



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

FACULTAD DE INFORMÁTICA Y ELECTRÓNICA

CARRERA DE INGENIERÍA EN ELECTRÓNICA Y AUTOMATIZACIÓN

**“CONSTRUCCIÓN DE UN SISTEMA PROTOTIPO DE
MONITOREO Y ALERTA EN TIEMPO REAL PARA EL USO DE
IMPLEMENTOS DE SEGURIDAD INDUSTRIAL EN CENTROS
DE TRABAJO MEDIANTE VISIÓN ARTIFICIAL”**

Trabajo de Titulación

Tipo: Dispositivo Tecnológico

Presentado para optar al grado académico de:

INGENIERO EN ELECTRÓNICA Y AUTOMATIZACIÓN

AUTORES: JIMSOP SANTIAGO PUCHA ALDAZ

JEFFERSON ERMEL GRANIZO TAPIA

DIRECTOR: ING. PABLO EDUARDO LOZADA YÁNEZ

Riobamba – Ecuador

2021

© 2021, Jefferson Ermel Granizo Tapia y Jimsop Santiago Pucha Aldaz

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

Nosotros, **JEFFERSON ERMEL GRANIZO TAPIA** y **JIMSOP SANTIAGO PUCHA ALDAZ**, declaramos que el presente trabajo de titulación es de nuestra autoría y los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autores asumimos la responsabilidad legal y académica de los contenidos de este trabajo de titulación; el patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

Riobamba, 9 de noviembre de 2021



Jefferson Ermel Granizo Tapia

060357277-7



Jimsop Santiago Pucha Aldaz

060410343-2

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO

FACULTAD DE INFORMÁTICA Y ELECTRÓNICA

CARRERA DE INGENIERÍA EN ELECTRÓNICA Y AUTOMATIZACIÓN

El Tribunal del Trabajo de Titulación certifica que: El trabajo de titulación: Tipo: Dispositivo Tecnológico, “**CONSTRUCCIÓN DE UN SISTEMA PROTOTIPO DE MONITOREO Y ALERTA EN TIEMPO REAL PARA EL USO DE IMPLEMENTOS DE SEGURIDAD INDUSTRIAL EN CENTROS DE TRABAJO MEDIANTE VISIÓN ARTIFICIAL**”, realizado por los señores **JEFFERSON ERMEL GRANIZO TAPIA** y **JIMSOP SANTIAGO PUCHA ALDAZ**, ha sido minuciosamente revisado por los Miembros del Tribunal del Trabajo de Titulación, el mismo que cumple con los requisitos científicos, técnicos, legales, en tal virtud el Tribunal Autoriza su presentación.

| | FIRMA | FECHA |
|-----------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|------------|
| Ing. José Luis Morales Gordón PRESIDENTE DEL TRIBUNAL | Jose Morales <small>Firmado digitalmente por Jose Morales Fecha: 2021.11.11 22:40:39 -0500'</small> | 2021-11-09 |
| Ing. Pablo Eduardo Lozada Yáñez DIRECTOR DE TRABAJO DE TITULACIÓN | Firmado digitalmente por PABLO EDUARDO LOZADA YANEZ | 2021-11-09 |
| Ing. Alberto Leopoldo Arellano Aucancela MIEMBRO DEL TRIBUNAL | ALBERTO LEOPOLDO ARELLANO AUCANCELA <small>Firmado digitalmente por ALBERTO LEOPOLDO ARELLANO AUCANCELA Fecha: 2021.11.10 16:17:13 -0500'</small> | 2021-11-09 |

DEDICATORIA

A mis padres, hermanas y mis amigos más cercanos, por haber estado para mí todo este tiempo.

JEFFERSON.

A mis padres y familiares que me apoyaron durante todo este proceso.

SANTIAGO.

AGRADECIMIENTO

Quiero agradecer a mis padres y hermanas por todo el apoyo que me brindaron a lo largo de esta etapa de mi vida.

JEFFERSON.

Quiero agradecer a Dios por permitirme vivir esta experiencia, y a mis padres quienes me apoyaron durante todo este proceso de esta etapa de mi vida.

SANTIAGO.

TABLA DE CONTENIDO

| | |
|-----------------------------|------|
| INDICE DE FIGURAS..... | x |
| INDICE DE GRÁFICOS..... | xii |
| INDICE DE ANEXOS | xiii |
| INDICE DE ABREVIATURAS..... | xiv |
| RESUMEN..... | xv |
| SUMMARY | xvi |
| INTRODUCCIÓN | 1 |

CAPÍTULO I

| | |
|------------------------------------------------------|----|
| 1. MARCO TEÓRICO | 6 |
| 1.1 Seguridad Industrial..... | 6 |
| 1.1.1 Generalidades | 6 |
| 1.1.2 Breve historia de la seguridad industrial..... | 7 |
| 1.1.3 Accidentes de Trabajo..... | 7 |
| 1.1.4 Factores..... | 8 |
| 1.2 Elementos de un Accidente | 8 |
| 1.2.1 Individuo | 8 |
| 1.2.2 Comportamiento humano..... | 9 |
| 1.3 Equipos de Protección Personal | 9 |
| 1.3.1 Análisis..... | 9 |
| 1.3.2 Ergonomía..... | 10 |
| 1.3.3 Factor Humano | 10 |
| 1.3.4 Clasificación de los EPP..... | 11 |
| 1.4 Medio Ambiente..... | 12 |
| 1.4.1 Lugar de trabajo | 12 |
| 1.4.2 Acústica..... | 13 |
| 1.4.3 Vibraciones | 14 |
| 1.4.4 Iluminación..... | 14 |
| 1.4.5 Sistema legislativo deficiente..... | 15 |
| 1.5 Visión Artificial..... | 16 |
| 1.5.1 Generalidades | 16 |
| 1.5.2 Procesamiento digital de imágenes | 16 |
| 1.5.3 Modelo RGB..... | 17 |
| 1.5.4 Modelo HSV..... | 18 |

| | | |
|--------|--------------------------------------------------------|----|
| 1.5.5 | <i>Detección de contornos</i> | 19 |
| 1.6 | Hardware | 20 |
| 1.6.1 | <i>Raspberry PI</i> | 20 |
| 1.7 | Redes Neuronales Convolucionales (CNN) | 20 |
| 1.8 | YOLOv3 | 21 |
| 1.9 | Yolov4-tiny | 22 |
| 1.10 | Fast R-CNN | 23 |
| 1.11 | Faster R-CNN | 24 |
| 1.12 | Software | 25 |
| 1.12.1 | <i>Python</i> | 25 |
| 1.12.2 | <i>Características de programación en Python</i> | 26 |
| 1.12.3 | <i>Open CV</i> | 26 |
| 1.13 | Google Colab | 26 |
| 1.14 | Darknet | 27 |

CAPITULO II

| | | |
|-------|----------------------------------------------------------|----|
| 2. | MARCO METODOLÓGICO | 27 |
| 2.1 | Método de Investigación | 27 |
| 2.2 | Requerimientos del Prototipo | 28 |
| 2.3 | Concepción General del Prototipo AZATHOTH | 28 |
| 2.4 | Diseño del prototipo AZATHOTH | 30 |
| 2.4.1 | <i>Etapa de entrenamiento</i> | 30 |
| 2.4.2 | <i>Etapa de adquisición de datos</i> | 31 |
| 2.4.3 | <i>Etapa de procesamiento de información</i> | 32 |
| 2.4.4 | <i>Etapa de alimentación</i> | 33 |
| 2.5 | Hardware del prototipo AZATHOTH | 34 |
| 2.5.1 | <i>Raspberry Pi</i> | 34 |
| 2.5.2 | <i>Dispositivo de captura</i> | 36 |
| 2.5.3 | <i>Interfaz</i> | 37 |
| 2.6 | Diagrama de Conexión | 38 |
| 2.7 | Elaboración del Case del Prototipo AZATHOTH | 39 |
| 2.8 | Software del Dispositivo | 41 |
| 2.8.1 | Requerimientos del software | 41 |
| 2.8.2 | Lenguaje de programación | 41 |
| 2.8.3 | Red neuronal convolucional | 42 |
| 2.9 | Programación | 44 |
| 2.9.1 | Programación en Google Colab | 44 |
| 2.9.2 | Programación en ID Python | 46 |

CAPITULO III

| | | |
|--------------|--------------------------------------------------------------|-----------|
| 3. | ANÁLISIS Y RESULTADOS..... | 49 |
| 3.1 | Método de Validación de Resultados | 50 |
| 3.1.1 | <i>Protocolo para evaluar el algoritmo de detección.....</i> | <i>51</i> |
| 3.2 | Pruebas de Campo..... | 52 |
| 3.2.1 | <i>Clase 1 – Casco de seguridad</i> | <i>52</i> |
| 3.2.2 | <i>Clase 2 – Botas de seguridad</i> | <i>56</i> |
| 3.2.3 | <i>Clase 3 – chaleco reflectivo</i> | <i>59</i> |
| 3.2.4 | <i>Clase 4 – Orejeras de seguridad.....</i> | <i>62</i> |
| 3.2.5 | <i>Clase 5 – Guantes de seguridad.....</i> | <i>65</i> |
| 3.2.6 | <i>Clase 6 – Persona.....</i> | <i>68</i> |
| 3.2.7 | <i>Integración de clases</i> | <i>72</i> |
| 3.3 | Resultados..... | 74 |

CAPITULO IV

| | | |
|-----------|-----------------------------------|-----------|
| 4. | EVALUACIÓN ECONÓMICA | 75 |
|-----------|-----------------------------------|-----------|

| | | |
|--|--------------------------|-----------|
| | CONCLUSIONES..... | 78 |
|--|--------------------------|-----------|

| | | |
|--|-----------------------------|-----------|
| | RECOMENDACIONES..... | 79 |
|--|-----------------------------|-----------|

BIBLIOGRAFIA

ANEXOS

INDICE DE TABLAS

| | | |
|-------------------|--------------------------------------------------------------------------|----|
| Tabla 1-2: | Simplificación de los nombres de cada equipo de protección personal..... | 31 |
| Tabla 2-2: | Características de módulos Raspberry..... | 34 |
| Tabla 3-2: | Características de los dispositivos de adquisición de imágenes..... | 36 |
| Tabla 4-2: | Módulos de interfaz..... | 37 |
| Tabla 5-2: | Comparación de tiempos de procesamiento entre redes neuronales..... | 44 |
| Tabla 1-3: | Datos obtenidos del procesamiento de imágenes de la Clase 1..... | 53 |
| Tabla 2-3: | Datos obtenidos del procesamiento de imágenes de la Clase 2..... | 56 |
| Tabla 3-3: | Datos obtenidos del procesamiento de imágenes de la Clase 3..... | 60 |
| Tabla 4-3: | Datos obtenidos del procesamiento de imágenes de la Clase 4..... | 63 |
| Tabla 5-3: | Datos obtenidos del procesamiento de imágenes de la Clase 5..... | 66 |
| Tabla 6-3: | Datos obtenidos del procesamiento de imágenes de la Clase 6..... | 69 |
| Tabla 7-3: | Resultados al integrar todas las clases..... | 72 |
| Tabla 8-3: | Resultados (Prototipo AZATHOTH vs Dataset Yolo COCO)..... | 74 |
| Tabla 1-4: | Análisis económico del prototipo AZATHOTH..... | 75 |
| Tabla 2-4: | Características (AIS / AZATHOTH)..... | 91 |

INDICE DE FIGURAS

| | | |
|---------------------|-----------------------------------------------------------------|----|
| Figura 1-1: | Modelo RGB representado como un cubo..... | 18 |
| Figura 2-1: | Modelo HVS representado por un cono..... | 19 |
| Figura 3-1: | Dilatación y erosión de imágenes | 19 |
| Figura 4-1: | Arquitectura de un Red Neuronal Convolutacional | 21 |
| Figura 5-1: | Modelo de Detección YOLOv3 | 21 |
| Figura 6-1: | Arquitectura R-CNN Rápida..... | 23 |
| Figura 7-1: | Faster R-CNN | 25 |
| Figura 1-2: | Concepción general del prototipo AZATHOTH..... | 29 |
| Figura 2-2: | Etapa de Entrenamiento AZATHOTH | 31 |
| Figura 3-2: | Etapa de adquisición de datos | 32 |
| Figura 4-2: | Etapa de procesamiento de información | 33 |
| Figura 5-2: | Etapa de alimentación del sistema | 34 |
| Figura 6-2: | Módulo Raspberry Pi 4 Modelo B | 36 |
| Figura 7-2: | Raspberry Pi Cámara V2 | 37 |
| Figura 8-2: | LCD Rpi de 3.5 Pulgadas (A) | 38 |
| Figura 9-2: | Conexiones del hardware del prototipo AZATHOTH..... | 39 |
| Figura 10-2: | Case del Prototipo AZATHOTH | 41 |
| Figura 11-2: | Entorno de trabajo de Python..... | 42 |
| Figura 12-2: | a) Frutas, b) Personas, c) Vehículos..... | 43 |
| Figura 13-2: | Diagrama de flujo para el proceso de entramiento..... | 45 |
| Figura 14-2: | Entrenamiento de la Red mediante Google Colab | 46 |
| Figura 15-2: | Diagrama de flujo del procesamiento de imágenes..... | 48 |
| Figura 16-2: | Imagen analizada mediante la red neuronal convolutacional | 49 |
| Figura 1-3: | Prototipo AZATHOTH..... | 50 |
| Figura 2-3: | Fórmulas para calcular valor mAP..... | 51 |
| Figura 3-3: | A) Casco de Seguridad, B) Casco de motocicleta, C) Gorra | 53 |
| Figura 4-3: | Detección Clase 1 | 55 |
| Figura 5-3: | Comparación de objetos en la Clase 1 | 55 |
| Figura 6-3: | A) Botas de Seguridad, B) Zapatillas, C) Zapatos | 56 |
| Figura 7-3: | Detección exitosa de la clase 2 | 58 |
| Figura 8-3: | Análisis de objeto de comparación de la clase 2..... | 59 |
| Figura 9-3: | A) Chaleco Reflectivo, B) Camisa, C) Buso Térmico | 59 |
| Figura 10-3: | Detección y etiquetado exitoso de la clase 3..... | 62 |

| | | |
|---------------------|----------------------------------------------------------------------|----|
| Figura 11-3: | Elemento de comparación de la clase 3 | 62 |
| Figura 12-3: | A) Orejeras de Seguridad, B) Auriculares, C) Orejeras de frio..... | 62 |
| Figura 13-3: | Detección y etiqueta de las orejetas de seguridad..... | 65 |
| Figura 14-3: | Detección de las orejetas de seguridad en una persona..... | 65 |
| Figura 15-3: | A) Guantes de Seguridad, B) Guantes sin dedos, C) Manos Humanas..... | 66 |
| Figura 16-3: | Comparación de objetos de la clase 5 | 68 |
| Figura 17-3: | Detección y etiquetado del guante de seguridad..... | 68 |
| Figura 18-3: | Elementos de comparación (Hombre/Mujer)..... | 69 |
| Figura 19-3: | Detección individual de la clase 6..... | 71 |
| Figura 20-3: | Detección grupal de la clase 6..... | 72 |
| Figura 21-3: | Detección y etiquetado correcto de todas las clases..... | 74 |

INDICE DE GRÁFICOS

| | | |
|---------------------|-------------------------------------------------------------------|----|
| Gráfica 1-1: | Comparación YOLOv4 vs Redes Neuronales de última Generación | 22 |
| Gráfica 1-3: | Resultados del prototipo AZATHOTH | 75 |

INDICE DE ANEXOS

- ANEXO A:** Especificaciones Raspberry Pi4
- ANEXO B:** Características Raspberry Pi Camera v2
- ANEXO C:** Características LCD rpi de 3.5 pulgadas
- ANEXO D:** Características prototipo P1x-Series
- ANEXO E:** Entrenamiento de la Red Neuronal mediante imágenes
- ANEXO F:** Pruebas de detección de clases.
- ANEXO G:** Ensamblaje del prototipo AZATHOTH
- ANEXO H:** Pruebas de detección a diferentes rangos de distancia
- ANEXO I:** Configuración del LCD

INDICE DE ABREVIATURAS

| | |
|-------------|------------------------------------|
| IA: | Inteligencia artificial. |
| VA: | Visión Artificial. |
| GPU: | Unidad de Procesamiento Grafico. |
| mAP: | Precisión media promedio. |
| AP: | Precisión media. |
| EPP: | Equipos de protección personal. |
| HMI: | Interfaz hombre – maquina. |
| PDI: | Procesamiento digital de imágenes. |
| DL: | Aprendizaje profundo. |
| CNN: | Redes neuronales convolucionales. |
| TPU: | Unidad de procesamiento tensorial. |
| VM: | Máquina virtual. |

RESUMEN

En el presente trabajo de titulación se realizó la implementación de un prototipo de monitoreo y alerta en tiempo real para el uso de implementos de seguridad industrial en centros de trabajo mediante visión artificial. El prototipo está conformado por 4 etapas: entrenamiento, alimentación, adquisición de datos y procesamiento. En la primera etapa se realiza el entrenamiento de la red neuronal convolucional Yolov4-tiny mediante la red pre entrenada Darknet para la detección de las distintas clases de estudio (casco, chaleco, orejeras, botas de seguridad, guante y personas) en la plataforma virtual Google Colab, en la segunda etapa se realiza el proceso de alimentación del prototipo mediante un adaptador de 12v 3A, el cual permite el correcto funcionamiento del prototipo, en la tercera etapa se realiza la adquisición de datos mediante el módulo Raspberry Pi Cámara V2 de 8Mpx que está diseñado para un mejor desempeño con el sistema embebido utilizado en la cuarta etapa, que se enfoca en el procesamiento de datos mediante la Raspberry PI 4 Modelo B de 8GB, el cual está programado mediante Python, el algoritmo de detección basado en Yolov4-tiny previamente entrenado para la detección de clases en equipos de protección personal, una vez identificadas las distintas clases en la imagen obtenida en la tercera etapa, el prototipo discernirá si los equipos de protección personal (EPPS) utilizados son correctos o se debe realizar una corrección en ellos, dando como resultado imágenes que se almacenarán tanto en el prototipo como en Google Drive para el posterior estudio de las mismas. Los resultados obtenidos mediante el prototipo realizado dieron como resultado una efectividad de la precisión media promedio por clase de estudio superior al 85% el cual es ampliamente más efectivo que el mostrado por el Dataset pre entrenado Yolo Coco que tiene un 57.9% de efectividad. Por los resultados obtenidos esta investigación puede ser susceptible a futuras modificaciones para trabajos futuros.

Palabras clave: <VISION ARTIFICIAL>, <RED NEURONAL>, <RED NEURONAL CONVOLUCIONAL YOLO>, <INTERNET>, <EQUIPOS DE PROTECCIÓN PERSONAL>, <SOFTWARE>, <GOOGLE COLAB>, <HARDWARE>.

SUMMARY

This research project aimed to develop a real-time monitoring and alert prototype for use in industrial safety implementation in work centers using artificial vision. Prototype development consisted of 4 stages: training, feeding, data acquisition, and processing. In the first stage, the Yolov4-tiny convolutional neural network was trained using the pre-trained Darknet network to detect the different research classifications (helmet, vest, earmuffs, safety boots, gloves, and people) on the virtual Google Colab platform. In the second stage, the process of feeding the prototype was carried out using a 12v 3A adapter, which ensured the correct functioning of the prototype. In the third stage, data acquisition is carried out using the Raspberry Pi module Camera V2, 8Mpx, which was designed for better performance with the embedded system used in the fourth stage, which focuses on data processing using the 8GB Raspberry PI 4 Model B, which was programmed using Python with the detection algorithm based on Yolov4- tiny previously trained for detecting classifications in personal protective equipment. Once the different categories have been identified in the images obtained in the third stage, the prototype will discern if the personal protective equipment (PPE) used is correct or a correction needs to be made, resulting in images that will be stored both in the prototype as well as on Google Drive for further study of same. The results obtained through the prototype demonstrated the effectiveness of average precision per study classification as greater than 85%, which is significantly more effective than that shown by the pre-trained Yolo Coco Dataset, which has 57.9% effectiveness. In conclusion, the results obtained suggest that further modifications as part of future research should be considered.

Keywords: <ARTIFICIAL VISION>, <NEURAL NETWORK>, <YOLO CONVOLUTIONAL NEURAL NETWORK>, <INTERNET>, <PERSONAL PROTECTIVE EQUIPMENT>, <SOFTWARE>, <GOOGLE COLAB>, <HARDWARE>.

INTRODUCCIÓN

A pesar de varias inspecciones de seguridad realizadas a lo largo de los años para garantizar el cumplimiento de las regulaciones y mantener condiciones de trabajo aceptables y seguras, la construcción sigue siendo una de las industrias más peligrosas, responsable de una gran parte del total de muertes de trabajadores. Un trabajador de la construcción tiene una probabilidad de 1 en 200 de morir en el trabajo durante una carrera de 45 años, principalmente debido a incendios, caídas y ser golpeado o atrapado entre objetos. Esto, en parte, puede atribuirse a cómo el control de la presencia y el uso adecuado de equipos de protección personal (EPP) por parte de los oficiales de seguridad se vuelve ineficaz cuando se inspeccionan áreas extensas y un número considerable de trabajadores (Mneymneh, Abbas and Khoury 2017).

El comportamiento inseguro de los trabajadores de la construcción es una de las principales razones que provocan accidentes en la construcción. Sin embargo, el enfoque de gestión existente para los comportamientos inseguros, p. Ej. La seguridad basada en el comportamiento (BBS) se basa principalmente en la observación y el registro manuales, que no solo consumen mucho tiempo y costos, sino que cubren de manera imposible todo un sitio de construcción o todos los trabajadores (Yu et al. 2017).

La gestión de la seguridad en la construcción ha sido un tema popular en la investigación y la práctica en los últimos años debido a las altas tasas de accidentes y muertes en la industria de la construcción. La complejidad y variabilidad de los sitios de construcción hace que la gestión de la seguridad sea más difícil de implementar que en otras industrias. Como tecnología prometedora, la visualización se ha explorado ampliamente para ayudar a la gestión de la seguridad en la construcción. Sin embargo, en la literatura no existe una revisión crítica integral de la tecnología de visualización en la gestión de la seguridad en la construcción (Guo, Yu and Skitmore 2017).

Se presenta un enfoque novedoso para integrar el seguimiento de la ubicación en tiempo real y las tecnologías de visualización de datos inmersivas en tres dimensiones en los entornos de formación y educación de los trabajadores de las construcciones existentes. El alcance se limita a las tareas de montaje de acero realizadas por trabajadores del hierro sindicalizados en un centro de formación interior. Se muestran los resultados del análisis y visualización de los datos recopilados de la sesión de entrenamiento. Se explica el potencial para evaluar y mejorar el desempeño en seguridad y productividad de los capacitadores y aprendices. Dado que estas tecnologías apenas se han utilizado como parte de las técnicas de formación y educación en

construcción existentes, se estudiaron las oportunidades, incluido el retorno de la inversión y los comentarios de los usuarios. Los resultados muestran que las prácticas inseguras en los entornos de formación de los trabajadores se pueden detectar y visualizar y, además, se puede medir indirectamente su eficacia en la formación (Teizer, Cheng and Fang 2013).

(Massiris, Delrieux and Fernández Muñoz 2018), propone utilizar visión artificial como alternativa cuantitativa para monitorear la utilización de equipo de protección personal. Se entrenó la red neuronal YOLO con la intención de detectar guantes, cascos, ropa de alta visibilidad y a los trabajadores con un dataset creado a partir de videos generados utilizando cámaras deportivas. Con el sistema entrenado, se presenta un análisis de caso in the open con un video grabado con cámara deportiva sujeta al casco de un trabajador metalúrgico en el sector de la construcción. Los resultados son promisorios y muestran que la estrategia planteada es adecuada para llegar a una solución implantable en ambientes de trabajo.

Con los estudios previos revisados y analizados se propone la implementación de UN SISTEMA PROTOTIPO DE MONITOREO Y ALERTA EN TIEMPO REAL PARA EL USO DE IMPLEMENTOS DE SEGURIDAD INDUSTRIAL EN CENTROS DE TRABAJO MEDIANTE VISIÓN ARTIFICIAL con el objetivo de evitar accidentes a trabajadores del área industrial.

FORMULACIÓN DEL PROBLEMA

¿Cómo se puede construir un sistema de monitoreo para el control del uso de implementos de seguridad industrial en centros de trabajo mediante visión artificial?

SISTEMATIZACIÓN DEL PROBLEMA

- ¿Cuál es la operatividad y el funcionamiento de los sistemas de monitoreo en tiempo real para el uso de implementos de seguridad industrial en centros de trabajo mediante visión artificial?
- ¿Cuáles son los requerimientos de funcionamiento en los sistemas de monitoreo en tiempo real para el uso de implementos de seguridad en centros de trabajo mediante visión artificial?
- ¿Cuál es el diseño para considerar la construcción de sistemas de monitoreo en tiempo real mediante visión artificial que cumplan los requerimientos necesarios para el control del uso de implementos de seguridad en centros de trabajo?
- ¿Cómo construir un sistema de monitoreo en tiempo real para el uso de implementos de seguridad en base a técnicas de monitoreo mediante visión artificial?

- ¿Cómo validar al sistema de monitoreo mediante pruebas en ambientes controlados para el monitoreo del correcto uso de implementos de seguridad con el fin de que cumpla con los requerimientos previamente establecidos?

JUSTIFICACIÓN TEÓRICA

En las empresas la falta de un monitoreo apropiado a sus máquinas proporciona una pérdida de tiempo en producción o funcionamiento de una parte necesaria en la empresa, generando costos por mal uso del equipo, o por fallas mecánicas. El monitoreo es primordial para poder tener un control. El conocimiento del tiempo de uso y el usuario nos brindan información necesaria para tener un buen plan de mantenimiento y tener una estadística tanto del personal autorizado para usarla, como los daños sufridos en ella.

Siendo una rama del aprendizaje automático, el aprendizaje profundo ha conseguido resolver problemas que hasta hace unos años resultaban muy difíciles o incluso imposibles para la visión artificial tradicional, superando así los límites del procesamiento digital de imágenes. Hoy en día, los modelos de aprendizaje profundo orientados a la visión artificial hacen uso de redes neuronales artificiales convolucionales, que permiten el desarrollo de sistemas capaces de clasificar, detectar, localizar y segmentar objetos. Para conseguir abordar estas tareas, es necesario obtener un conjunto de datos suficientemente grande para que los modelos sean capaces de aprender de ellos. Dicho aprendizaje de los datos es precisamente el cambio de paradigma que ha hecho triunfar al aprendizaje automático. Empleando las ventajas ofrecidas por el aprendizaje profundo, este trabajo refleja el desarrollo de un sistema de visión artificial capaz de detectar equipos de protección individual y que se puede integrar en una solución comercial (Bedoya Bedoya and Mecatrónico 2007).

La seguridad industrial en el concepto moderno significa más que una simple situación de seguridad física, una situación de bienestar personal, un ambiente de trabajo idóneo, una economía de costos importantes y una imagen de modernización y filosofía de vida humana en el marco de la actividad laboral contemporánea (Aguinaga Barragán 2007).

En tal sentido con el fin de contribuir al esfuerzo del uso de implementos de seguridad en centros de trabajo se presenta un prototipo el cual brindará información de monitoreo y alerta en tiempo real en un área específica mediante visión artificial, lo cual permitirá a la persona encargada de la seguridad en los trabajadores pueda recibir una alerta en caso de que un trabajador que no cumpla con las normas de seguridad.

En la universidad nova de Lisboa, se desarrolló un sistema móvil en tiempo real para reconocimiento de objetos. el sistema se ejecuta localmente en un dispositivo móvil. dispositivo. se diferencia de los enfoques más comunes para sistemas móviles, ya que no necesita conectarse a un servidor para procesar imágenes y clasificación. los métodos propuestos son eficientes y fiables, lo que significa que dan una rápida el usuario móvil y, al mismo tiempo, tienen razonables de detección/ identificación. la rápida retroalimentación promueve una mejor interacción y cooperación entre el usuario y la aplicación móvil, por consiguiente, también puede mejorar la calidad de la detección/ identificación (Estarita et al. 2017).

En la universidad de Firenzi, se realizó un proyecto de reconocimiento fiable para un robot de los objetos según el medio ambiente. El trabajo se realizó sobre el hecho de que los robots pueden observar a los seres humanos interactuando con los objetos en su entorno, y por lo tanto proporcionando numerosas señales no visuales a las identidades de esos objetos. se investigó un enfoque flexible de reconocimiento de objetos Que puede utilizar cualquier señal múltiple, si son visuales Claves intrínsecas al objeto o proporcionadas por la observación de un humano. Se notó la difícil cuestión de que múltiples pistas las cuales Pueden tener diferente peso en su asociación con un objeto Definición y deben tenerse en cuenta durante el reconocimiento (Aboutalib and Veloso 2007).

En la universidad de Texas tech, se desarrolló un algoritmo que permite a los robots identificar un Conjunto de objetos interesantes, utilizando la apariencia y contextual Cues extraídas de un pequeño número de imágenes para aprender eficientemente Modelos de estos objetos. Los robots aprenden el mapa del dominio y consideran Objetos que se mueven para ser interesantes, usando señales de movimiento para identificar Las regiones de imagen correspondientes. Modelos de objetos aprendidos automáticamente De estas regiones consisten en la disposición espacial del gradiente Características, modelos basados en gráficos de vecindades de características de gradiente, Modelos basados en partes de segmentos de imagen, distribuciones de color, Y modelos de mezcla de contexto local. Los modelos aprendidos se utilizan Para el reconocimiento de objetos en escenas novedosas basadas en la minimización de la energía Y un modelo generativo para la fusión de la información (Li, Sridharan and Meador 2013).

La inteligencia artificial es una disciplina académica relacionada con la teoría de la computación cuyo objetivo es emular algunas de las facultades intelectuales humanas en sistemas artificiales. Con inteligencia humana nos referimos típicamente a procesos de precepción sensorial (visión, audición, etc.) y a sus consiguientes procesos de reconocimiento de patrones, por lo que las aplicaciones más habituales de la inteligencia artificial son el tratamiento de datos y la identificación de sistemas (Nope, Loaiza and Caicedo 2008).

JUSTIFICACIÓN PRÁCTICA

Esta investigación se realiza debido a la existente necesidad de controlar la ausencia o el uso inadecuado de equipos de protección personal por parte de trabajadores, permitiendo reducir accidentes y enfermedades laborales.

El proceso de detección y reconocimiento se basa en los algoritmos de procesamiento de imágenes, las cuales serán detectadas mediante la ayuda de cámaras ubicadas estratégicamente por todo el área de trabajo en el que se requiera el uso obligatorio de indumentaria de seguridad, la cual detectara la presencia de las personas, mediante sensores de movimiento reduciendo así el tiempo de uso del sistema al mínimo posible ya que solo realizara la toma de imágenes cuando detecte movimientos en la zona, separándolas del entorno mediante procesamiento digital de la imagen para así poder distinguir en ellas el correcto uso de los implementos de seguridad, generando alertas e informes para así proceder con las acciones correspondientes.

Al sistema de monitoreo se accede mediante un HMI ubicado en la zona de monitoreo de la empresa, o mediante la web y en él se ve todos los reportes generados por el sistema, incluidos las imágenes analizadas del obrero que haya incurrido en el mal uso de su equipamiento de seguridad, una vez revisado los reportes del sistema se podrá proceder a la amonestación al trabajador

OBJETIVO GENERAL

Construir un sistema prototipo de monitoreo y alerta en tiempo real para el uso de implementos de seguridad en centros de trabajo mediante visión artificial.

OBJETIVOS ESPECÍFICOS

- Investigar cual es la operatividad y el funcionamiento de los sistemas de monitoreo y alerta en tiempo real mediante visión artificial enfocados a la seguridad en centros de trabajo.
- Determinar cuáles son los requerimientos para el funcionamiento de los sistemas de monitoreo y alerta en tiempo real para la seguridad en centros de trabajo.

- Realizar diseños para la construcción de sistemas de monitoreo y alerta que cumplan con los requerimientos de hardware y software necesario para el correcto control de los implementos de seguridad en centros de trabajo.
- Construir el prototipo de sistema en base a las técnicas de monitoreo de visión artificial.
- Evaluar el prototipo de monitoreo y alerta dedicado al control del uso de implementos de seguridad en centros de trabajo, con el fin de que cumpla con los requerimientos más rigurosos de seguridad industrial.

CAPÍTULO I

1. MARCO TEÓRICO

1.1 Seguridad Industrial

La seguridad industrial en el concepto moderno significa más que una simple situación de seguridad física, una situación de bienestar personal, un ambiente de trabajo idóneo, una economía de costos importantes y una imagen de modernización y filosofía de vida humana en el marco de la actividad laboral contemporánea (Cavassa 1991).

1.1.1 Generalidades

Durante mucho tiempo se ha dicho que la construcción es indiscutiblemente una de las industrias más peligrosas en la que existen muchas tareas y condiciones peligrosas que pueden presentar lesiones, riesgos y muertes para los trabajadores. Por ejemplo, los incendios, las caídas y los golpes o quedar atrapado entre objetos contribuyen a más del 50% del total de víctimas en el sector. Por lo tanto, se consideró necesario en los lugares de trabajo el uso y la adopción adecuados de equipos de seguridad, como equipos de protección personal (EPP), en particular cascos (Mneymneh, Abbas and Khoury 2017).

Si el accidente como resultado obedece a ciertos elementos dentro de un sistema de determinada estructura, el primer paso en la investigación consiste en el estudio del accidente y sus consecuencias. La pequeña y mediana empresa forman un vasto núcleo dentro del sistema industrial de la mayor parte de los países. Son auxiliares directos de las grandes empresas en la fabricación de productos especializados de alto valor añadido, o de difícil fabricación en serie, y antesala de las grandes empresas futuras (Cavassa 1991).

1.1.2 Breve historia de la seguridad industrial

En todos los tiempos y espacios, el hombre siempre ha estado en la búsqueda de su seguridad, los pueblos y sus organizaciones tratan por todos los medios de mejorar bajo condiciones propicias el desarrollo de sus actividades económicas para asegurar de esta manera el empleo, generando bienestar a sus habitantes y por tanto seguridad. Desde el hombre prehistórico que se guarecía en las cavernas para protegerse del ataque de las fieras y de los fenómenos naturales, hasta el habitante de las ciudades cosmopolitas y modernas, un largo proceso de conquistas paulatinas marca la historia de la seguridad. El nacimiento de la fuerza industrial (Revolución Industrial) y el de la seguridad en el trabajo no fue simultáneo. Recordemos cómo fue este proceso, que también tuvo que ver con los cambios sociales en los países en donde se llevó a cabo (Chamochumbi Barrueto 2014).

La revolución industrial marca el inicio de la seguridad industrial como consecuencia de la aparición de la fuerza del vapor y de la mecanización de la industria, lo que produjo el incremento de accidentes y enfermedades laborales. No obstante, el nacimiento de la fuerza industrial y el de la seguridad industrial no fueron simultáneos, Hola debido a la degradación y las condiciones de trabajo y de vida detestables. Es decir, en 1871 el cincuenta por ciento de los trabajadores moría antes de los veinte años, debido a los accidentes y a las pésimas condiciones de trabajo (Cavassa 1991).

1.1.3 Accidentes de Trabajo

La empresa actual inmersa en la globalización se enfrenta a grandes retos en la operatividad de sus maquinarias, equipos y herramientas, es esto un factor importante en la producción, ya que implica una secuencia matemática en el resultado del producto final. Es decir, la organización y un planeamiento correcto evitaría fallas en tomar las precauciones, prácticas razonables, fallas de administración, fallas de los trabajadores y fallas de la dirección. El número de accidentes de trabajo en todo el mundo se mantiene en una alarmante cifra de 270 millones de accidentes y de 160 millones de personas que contraen enfermedades profesionales de los cuales 2 millones culminan en muertes (Chamochumbi Barrueto 2014).

Según (Cavassa 1991), todo accidente es una combinación de riesgo físico y error humano. También se puede definir como un hecho en el cual ocurre o no la lesión de una persona, dañando o no a la propiedad; o solo se crea la posibilidad de tales efectos ocasionados por:

- El contacto de la persona con un objeto, sustancia u otra persona.
- Exposición del individuo a ciertos riesgos latentes.
- Movimientos de la misma persona.

1.1.4 Factores

Según (Cavassa 1991), se dijo que el accidente es la combinación de riesgos físicos y humanos; así, los factores que inciden en la producción del accidente son: técnicos y humanos.

- Factores humanos: Psicológicos, fisiológicos, sociológicos, económicos.
- Factores técnicos: organización.

Según (Cavassa 1991), estos factores causan el accidente una vez producido el disfuncionamiento en cualquiera de ellos. Los elementos cuyo disfuncionamiento origina el accidente o las enfermedades laborales se pueden agrupar en:

- El individuo (solo o en grupo).
- La tarea.
- Material y equipo.
- Medio ambiente o lugar de trabajo.
- Entorno.

1.2 Elementos de un Accidente

1.2.1 Individuo

Según (Chamocho Barrueto 2014), las causas de un accidente varían desde una simple distracción hasta la falta de equipos de protección personal adecuada al trabajo que se ejecuta, así tenemos:

- Causas sociales: Son variables psicológicas y materiales que tienen que ver con el bienestar mental y material del trabajador. Cuando una de estas falla entonces el trabajador entra en conflicto y puede accidentarse
- Causas de la dirección: Son de responsabilidad del empleador, la alta dirección tiene el deber moral y legal de darle al trabajador todas las protecciones del caso para que su labor sea productiva y a la vez segura.

- Causas del trabajador: Son directa responsabilidad del trabajador. La observación de las causas que producen accidentes y que podrían ser evitados por los trabajadores son simples, pero ineludibles.

1.2.2 Comportamiento humano

El comportamiento humano en una organización de trabajo más bien que espontaneo obedece a influencias estructurales. Se puede definir en términos de actividades, interacciones y sentimientos añadiéndose los conceptos de poder, autoridad y categoría, con lo que se consigues dos enfoques sociológicos, uno de carácter estructural – funcional y otro de interacción (Cavassa 1991).

El comportamiento del individuo aislado o en grupo puede ser formal o informal, según se aparte o no de las normas. Este último, si bien en determinados casos no es peligroso para la empresa, a nivel de seguridad va en contra de la organización y de sus componentes (Cavassa 1991).

El comportamiento individual diferente al comportamiento en grupo, esta función del entorno, las necesidades personales, las características genéticas y la formación del sujeto; más las motivaciones propias, cuando el individuo quiere alcanzar ciertos objetivos pudiendo, en caso negativo, sufrir insatisfacciones que se traducen en un sentimiento de frustración (Cavassa 1991).

1.3 Equipos de Protección Personal

Los equipos de protección personal son elementos de uso individual destinados a dar protección al trabajador frente a eventuales riesgos que puedan afectar su integridad durante el desarrollo de sus labores (Abrego, Molinos and Ruíz 2000).

1.3.1 Análisis

Se entenderá por Elemento de Protección Personal (EPP) a cualquier equipo destinado a ser llevado o sujetado por el trabajador para que le proteja de uno o varios riesgos que puedan amenazar su seguridad o su salud en el trabajo, así como cualquier complemento o accesorio destinado a tal fin.(Zorrilla 2012)

Los equipos de protección personal son elementos de uso individual destinados a dar protección al trabajador frente a eventuales riesgos que puedan afectar su integridad durante el desarrollo de sus labores. Es importante destacar que antes de decidir el uso de elementos de protección personal debieran agotarse las posibilidades de controlar el problema en su fuente de origen, debido a que ésta constituye la solución más efectiva. (Abrego, Molinos and Ruíz 2000)

1.3.2 Ergonomía

La ciencia de la ergonomía tiene poco más de 40 años. En su forma actual, se remonta a la Segunda Guerra Mundial y los años que siguieron (aunque tiene antecedentes reconocibles que se remontan mucho más atrás). Sus fundadores fueron un grupo de científicos británicos, que habían estado trabajando para las fuerzas armadas en varios proyectos relacionados con la eficiencia del combatiente. (El grupo incluía a anatomistas, fisiólogos, psicólogos e ingenieros). Creían que un enfoque científico multidisciplinario para el estudio de la eficiencia del trabajo podría ser igualmente relevante para la industria en tiempos de paz. Pero no parecía haber un nombre para esta área de investigación, por lo que tuvieron que inventar uno. Después de algunas deliberaciones, se decidieron por la ergonomía. La palabra se deriva del griego: ergos, trabajo; nomos, ley natural. (Pheasant 1991)

La ergonomía se ha definido como el estudio científico de la relación entre el hombre y su entorno laboral. En este sentido, el término entorno se toma para abarcar no solo el entorno ambiental en el que puede trabajar, sino también sus herramientas y materiales, sus métodos de trabajo y la organización de su trabajo, ya sea como individuo o dentro de un grupo de trabajo. Todos estos están relacionados con la naturaleza del hombre mismo; a sus habilidades, capacidades y limitaciones. Periféricas a la ergonomía, pero no consideradas en la actualidad como parte del campo, son las relaciones de un hombre con sus compañeros de trabajo, sus supervisores, su dirección y su familia. Por lo general, se considera que forman parte de las ciencias sociales, pero no pueden ignorarse, ya que pueden desempeñar un papel importante en la resolución de algunos problemas de ergonomía (Murrell 2012).

1.3.3 Factor Humano

Quienes utilizan equipos y dispositivos protectores deben conocer la necesidad de protección, los motivos por los cuales se utiliza en lugar (o, además) de otros métodos de control y las ventajas que se derivan de su empleo. Hay que explicar con claridad las consecuencias de la exposición sin protección y la forma en que el usuario puede detectar si el equipo no funciona correctamente.

Los usuarios deben recibir formación sobre métodos de inspección, ajuste, uso, mantenimiento y limpieza del equipo protector y deben conocer las limitaciones de dicho equipo, sobre todo en situaciones de emergencia. (Herrick 2001)

Según (Abrego, Molinos and Ruíz 2000) para que los elementos de protección personal resulten eficaces frente a los riesgos se deberá considerar lo siguiente:

- Entrega del protector a cada usuario.
- La responsabilidad de la empresa es proporcionar los EPP adecuados; la del trabajador es usarlos.
- El único EPP que sirve es aquel que ha sido seleccionado técnicamente y que el trabajador usa durante toda la exposición al riesgo.
- Capacitación respecto al riesgo que se está protegiendo. Es muy importante que los trabajadores conozcan los riesgos a que están expuestos para comprender la necesidad y conveniencia de utilizarlos.
- Responsabilidad de la línea de supervisión en el uso correcto y permanente de los EPP. Es fundamental la participación de los supervisores en el control del buen uso y mantenimiento de los elementos de protección personal. El supervisor debe dar el ejemplo utilizándolos cada vez que sea necesario.

1.3.4 Clasificación de los EPP

La utilización de un EPP o de una combinación de EPP contra uno o varios riesgos puede provocar una serie de molestias. Por consiguiente, a la hora de elegir un EPP apropiado, no solo hay que tener en cuenta el nivel de seguridad necesario, sino también la comodidad. Su selección deberá basarse en el estudio y la evaluación de los riesgos complejos presentes en el lugar de trabajo. Esto comprende la duración de la exposición a los riesgos, su frecuencia y la gravedad, las condiciones existentes en el trabajo y su entorno, el tipo de daños posibles para el trabajador y su constitución física. Solo son aptos para el uso los equipos de protección individual que se hallan en perfectas condiciones y pueden asegurar plenamente la función protectora prevista. (Zorrilla 2012)

Es importante enfatizar que cualquiera que sea el equipo de protección personal que se tenga que utilizar frente a un determinado riesgo, estos deben ser seleccionados por profesionales especializados y de acuerdo con las normas de calidad establecidas por el instituto Nacional de

Normalización (INN), o bien, provenientes de organismos reconocidos internacionalmente.(Abrego, Molinos and Ruíz 2000)

Según (Abrego, Molinos and Ruíz 2000), para describir los diferentes equipos se utilizará la siguiente clasificación:

- Protección de cráneo.
- Protección de ojos y cara.
- Protección de oído.
- Protección de las vías respiratorias.
- Protección de manos y brazos.
- Protección de pies y piernas.
- Cinturones de seguridad para trabajos de altura.
- Ropa protectora.

1.4 Medio Ambiente

1.4.1 Lugar de trabajo

Según (Spector, Vázquez and de la Pena 2002), la satisfacción laboral produce importantes aspectos vinculados con los empleados y las organizaciones, desde el desempeño laboral hasta la salud y la calidad de vida.

Para que una empresa contribuya con eficiencia a sus obligaciones con los empleados, accionistas y la sociedad, sus directivos deben generar una relación entre la organización y sus trabajadores que satisfaga las necesidades siempre cambiantes de ambas partes. Como mínimo la organización espera que los empleados desempeñen de manera confiable las tareas que se les asignen y con los estándares que se les fijan, que sigan las normas establecidas para administrar el sitio de trabajo. Con frecuencia, la dirección espera más: que tomen iniciativas, que se supervisen por sí mismos, que continúen desarrollando nuevas habilidades y que se respondan a las necesidades comerciales. Por su parte los empleados también esperan como mínimo que la organización les proporcione una paga equitativa, condiciones seguras de trabajo y un trato justo. Igual que la dirección, los empleados también esperan más, dependiendo de la magnitud de sus necesidades de seguridad, situación laboral y económica, entorno, retos, oportunidades, poder y responsabilidad. De esta forma la pretensión de las expectativas que tenga cada una de las partes

variará de organización a organización y de persona a persona, el reto es alinear los objetivos e intereses organizacionales con los particulares (Trujillo and Vargas 2017).

1.4.2 Acústica

El ruido es considerado actualmente como un grave problema de salud ocupacional que afecta notablemente la calidad de vida de los trabajadores y personal entrenado para arduas jornadas de trabajo en exposición permanente al mismo, evidentemente el grave problema que a ello concierne implica en primera instancia la marcada disminución de la capacidad auditiva en los individuos implicados, asociado en casos agravados, a sordera por destrucción de las células cocleares del oído interno. Indudablemente el único problema de salud ocupacional que implica el ruido en el medio laboral no incluye únicamente la salud auditiva, pues estudios múltiples desarrollados con el pasar de los años han demostrado que el deterioro cardiovascular, endocrino glandular y neurológico con predilección por las alteraciones en la percepción motora y la capacidad cognitivo-conductual son unas cuantas de las graves consecuencias que implica la exposición al mismo; uno de los más grandes agravantes se define como la exposición prolongada a altas frecuencias y mayor intensidad del ruido en jornadas superiores a ocho horas diarias de exposición, las cuáles, a todos los individuos, independiente de la edad o factores socioculturales asociados, afecta considerablemente (Martínez et al. 2012).

La disminución en la audición ocupacional se define como la disminución en la audición a la cual están expuestos los trabajadores, su severidad puede ir desde leve hasta la pérdida total de esta. La disminución en la audición de origen ocupacional puede ser de varios tipos: Conductiva, neurosensorial o una combinación de estas. La perforación del tímpano por trauma en cráneo, explosiones o elementos cortopunzantes, restos de metal son ejemplos de la hipoacusia conductiva. La exposición prolongada a ruido, la fístula perilinfática, trauma acústico por sustancias ototóxicas, y ruptura del oído interno y la membrana de la ventana redonda por trauma o barotrauma, son ejemplos de hipoacusia neurosensorial. Por lo general la hipoacusia es bilateral, aunque en ocasiones es unilateral (Kim 2010).

Dado que la exposición al ruido en el lugar de trabajo es continua, es importante para determinar la exposición a largo plazo los efectos de ruido persistente tienen sobre la pérdida de células ciliadas. La pequeña cantidad de daño de las células ciliadas causada por la exposición al ruido intermitente podría estar relacionado con los períodos de recuperación entre las fases de ruido. Varios estudios se han publicado en relación con el desplazamiento del umbral auditivo tras la exposición repetida al ruido durante un largo periodo de tiempo, sin embargo, sólo unos pocos

estudios han abordado la cuestión del umbral de los turnos producido después de la exposición repetida a ruidos durante un corto período de tiempo. Es necesario determinar si la exposición a un ruido pulsante es suficiente para iniciar el proceso de recuperación. Como consecuencia de sus efectos sobre la reducción de la exposición continua al ruido, el ruido induce formación de radicales libres puede ser un factor significativo en la disminución del flujo sanguíneo coclear, aún más formación de radicales libres y estrés oxidativo (Chou, Lai and Kuo 2009).

1.4.3 Vibraciones

Muchos factores pueden influir en los efectos de salud causados por la exposición humana a la vibración. Como ejemplo se puede mencionar el espectro de frecuencia de la vibración y la amplitud, la duración de la exposición diaria a las vibraciones y exposición a las vibraciones pre acumulado. Además, algunas personas pueden tener una mayor susceptibilidad a desarrollar otras enfermedades causadas por la exposición a las vibraciones (Osoreo 2017).

El estudio realizado a una población de trabajadores expuestos a las diversas vibraciones, debido a que operan equipos de mayor vibración como son: volquete, perforadora, excavadora y motoniveladoras, se determinó que la dirección en que inciden en el cuerpo humano las vibraciones de maquinarias en promedio no afecta la salud de los trabajadores, porque no llegan al mínimo establecido por ley que es el 0,5 m/s². El tiempo de exposición de 8 h de las vibraciones de maquinarias en cuerpo entero es de 0,5 m/s² (Valdiviezo Guzmán 2014).

En el caso de las vibraciones transmitidas al sistema mano-brazo pueden producirse alteraciones vasculares, neurológicas y musculoesquelético, de forma independiente o simultáneamente, conocido como síndrome de vibración mano-brazo. El principal problema es el fenómeno de Raynaud o enfermedad de los "dedos blancos", que es una disminución de los vasos periféricos y arterias, con el consecuente blanqueamiento de los extremos de los dedos, causados por la exposición a la vibración del sistema mano-brazo. Este fenómeno también se asocia por lo general con la exposición de los miembros al frío, cuando la actividad metabólica del cuerpo humano es menor (Osoreo 2017).

1.4.4 Iluminación

La iluminación industrial es uno de los factores ambientales de carácter micro climático que tiene como principal finalidad facilitar la visualización de los objetivos dentro de su contexto espacial,

de modo que, el trabajo se pueda realizar en unas condiciones aceptables de eficacia, comodidad y seguridad (Robledo 2014).

El ojo humano ha evolucionado a través de los tiempos, desde cuando usaba casi por completo la iluminación del aire libre y la luz brillante del día para una visión simple, de largo alcance. En la actualidad, el hombre vive y trabaja corrientemente en el interior de edificios y utiliza sus ojos con demasiada frecuencia y durante largas horas en condiciones de iluminación artificial inadecuadas y en trabajos delicados que exigen una constante acomodación. Un buen alumbrado puede hacer mucho para mejorar las condiciones de trabajo del ojo y aliviar el esfuerzo visual necesario para el ejercicio de labores visuales difíciles (Robledo 2014).

Según (Cavassa 1991), la seguridad de un ambiente de trabajo limpio, protegido y ordenado. Esto significa que la iluminación debe ser también adecuada. Aun después de que una planta u operación industrial ha sido bien iluminada, debe tenerse en cuenta la posibilidad de situaciones negativas inadvertidas; esto hace imprescindible una revisión periódica. Esta revisión puede dirigirse a los siguientes puntos:

- Cantidad de luz.
- Uniformidad.
- Sombra e iluminación localizada.
- Cubrimiento de focos y deslumbramiento.
- Tipo adecuado de reflectores.
- Mantenimiento.
- Operación de lámparas con voltaje específico.

1.4.5 Sistema legislativo deficiente

En muchos países existe por lo general, un aparato legal bastante completo sobre la prevención de accidentes y defensa del capital humano. Esto es apariencia, pues ¿Cómo se explica la permanente ocurrencia de accidentes, no solo de los que escapan a la previsión del individuo, sino por el contrario los debidos a causa simples y fáciles de detectar? Existen defectos en la aplicación práctica de las leyes. En otras palabras, un sistema permanente de control, mediante inspecciones permanentes realizadas por el aparato judicial, permitiría que lo planeado y cuestionado como medios de seguridad para el trabajador cristalice en efectos positivos, y se obligue a los patrones a cumplir con las normas de seguridad, a los técnicos a diseñar y fabricar equipo y herramientas

con la protección necesaria, y al propio personal a mantenerse preparado y actualizado en lo concerniente a la prevención de accidentes (Cavassa 1991).

1.5 Visión Artificial

Según (Sánchez 2002), visión por computador es una ciencia que estudia la interpretación de imágenes mediante computadoras digitales.

1.5.1 Generalidades

La visión por computador, también conocida bajo otras denominaciones como visión artificial o interpretación de imágenes, es una disciplina que persigue la deducción automática de la estructura y propiedad de un escenario o mundo tridimensional, posiblemente cambiante, a partir de una o varias imágenes captadas del mismo. Con respecto a las imágenes, estas pueden ser en blanco y negro o en color, capturadas por una única cámara o varias asociadas, o – incluso en un mayor grado de generalización del concepto de imagen –, proceder de sensores de naturaleza no visual como acústicos, térmicos, táctiles, etc. (Sánchez 2002).

Según (Alegre, Pajares and de la Escalera 2016), la visión artificial consiste básicamente en la deducción automática de la estructura y propiedades de un mundo tridimensional, posiblemente dinámico, a partir de una o varias imágenes bidimensionales de ese mundo. En esta área de conocimiento se añadan conceptos de la física del color, óptica, electrónica, geometría, algorítmica, sistemas de computación, etc.

Según (García 2007), la visión artificial también conocida como visión por computador, es una disciplina de la inteligencia artificial y su propósito general es programar un computador para que “entienda” una escena o las características de una imagen digital. Su campo de aplicación es muy amplio, desde la industria (control de calidad por visión, reconocimiento e identificación de formas, colores, materiales, etiquetas, códigos, etc.), en la medicina (detección precoz de cáncer mediante radiografías y ecografías), la seguridad (reconocimiento de matrículas, vehículos y personas), en la robótica y entretenimiento.

1.5.2 Procesamiento digital de imágenes

El procesamiento de imágenes es la primera etapa a la que se somete a las imágenes de la escena analizada con objeto de intentar determinar de forma automática su contenido. En este sentido hay que considerar el procesamiento de imágenes como el conjunto de técnicas que permiten transformar una imagen digital, con el fin de facilitar las etapas posteriores de segmentación y reconocimiento (Sánchez 2002).

El PDI utiliza un gran número de técnicas de las matemáticas puras y aplicadas. Las áreas utilizadas de las matemáticas son primordialmente el álgebra lineal, cálculo diferencial e integral en una y varias dimensiones, ecuaciones diferenciales ordinarias y parciales, ecuaciones integrales (transformadas integrales), probabilidad y estadística (Torres 1996).

Según (Torres 1996), las áreas de las matemáticas antes mencionadas se aplican de forma diversa y variada en los tres problemas distintos y fundamentales que son comunes en muchas de las áreas de aplicación del PDI:

- Restauración y reconstrucción de imágenes
- Reconocimiento de patrones
- Interpretación física.

1.5.3 Modelo RGB

El espacio RGB es el espacio de color más extendido y el que utilizan la gran mayoría de cámaras de video y fotográficas para construir una imagen de color. Y de ahí, su importancia en visión artificial ya que trabajar con el mismo espacio de color con el que trabaja la cámara con la que se capturan las imágenes permite evitar la alteración de las propiedades del color durante el proceso de segmentación, propia de los errores de conversión y transformación, y por otro lado conseguir una mayor velocidad de segmentación por ahorro de esas operaciones de conversión y redondeo (Gil Pablo AND Torres 2004).

Según (Gil Pablo AND Torres 2004), el espacio RGB se representa como un cubo dónde un color viene definido por la mezcla de valores de intensidad de tres colores primarios, rojo, verde y azul. Un color viene descrito por una tupla de 3- coordenadas en el cubo.

La Figura 1-1 describe el modelo RGB representado en un cubo.

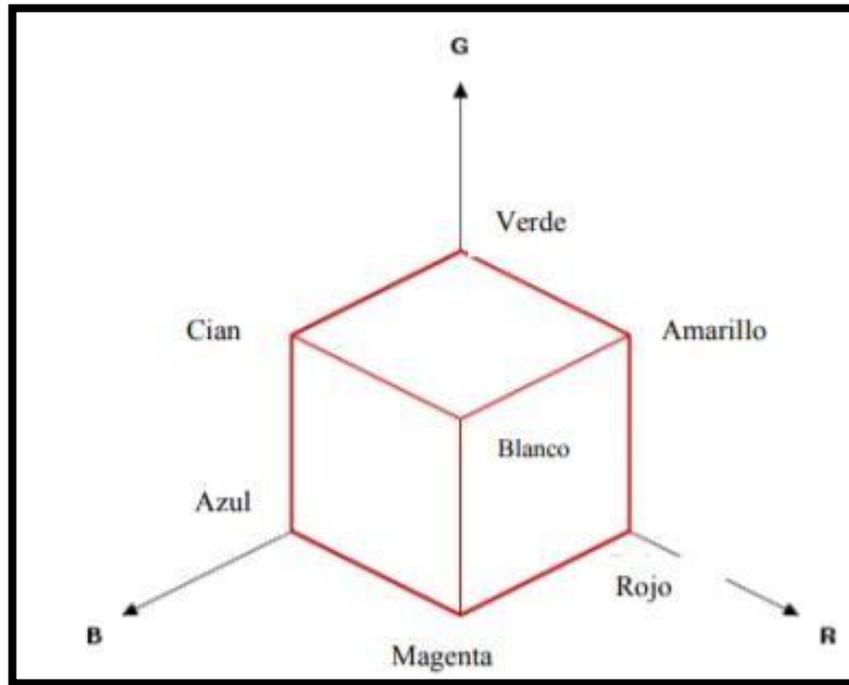


Figura 1-1: Modelo RGB representado como un cubo.

Fuente:(Gil Pablo AND Torres 2004)

1.5.4 Modelo HSV

Según (Smith 1978), el espacio HSV representa uno de los espacios de coordenadas más clásicos e intuitivos existentes en la literatura. Su interpretación geométrica viene determinada por un cono de base cuasi-hexagonal. Con esta representación del espacio de color, cada color trabaja con 3 componentes básicas: matiz, saturación y brillo.

Según (Smith 1978), el matiz, hHSV, hace referencia al valor de cromaticidad o clase de color. La saturación, sHSV, se refiere a las longitudes de onda que se suman a la frecuencia del color, y determina la cantidad de blanco que contiene un color. Contra menos saturado este un color más cantidad de blanco, y contra más saturado este un color menor cantidad de blanco. En definitiva, la saturación representa la pureza e intensidad de un color. Así, la falta de saturación viene dada por la generatriz en la representación del cono HSV. Esa falta de saturación representa la gama de grises desde el blanco hasta el negro. La luminancia, vHSV, se corresponde con la apreciación subjetiva de claridad y oscuridad.

La Figura 2-1 representa el ejemplo del modelo VHS en un cono en 3D.

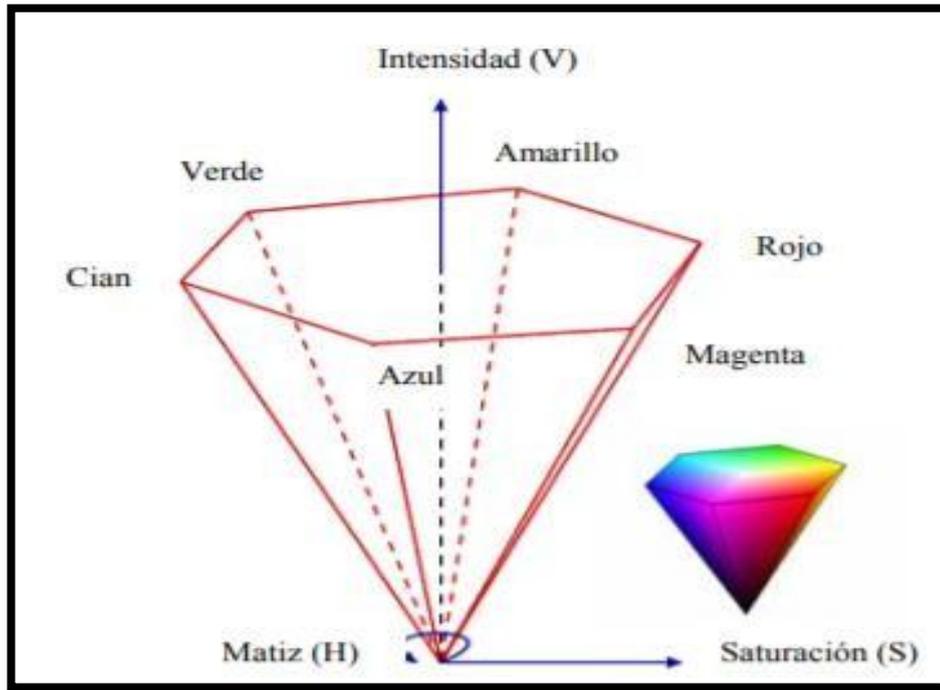


Figura 2-1: Modelo HVS representado por un cono.

Fuente: (Smith 1978)

1.5.5 *Detección de contornos*

Según (Alegre, Pajares and de la Escalera 2016), con las operaciones de erosión y dilatación es fácil obtener un contorno, de la parte exterior o interior al perímetro de una figura. Para ello tomaremos la imagen dilatada y le restaremos la original, en el caso del contorno exterior. Para la obtención del contorno interior restaremos de la imagen original la imagen erosionada. El contorno puede ser tanto o más marcado empleando EE de mayor tamaño tal como se ve contorno exterior del perímetro.

En la figura 3-1 se observa el método de detección de contornos aplicado a un objeto de estudio.

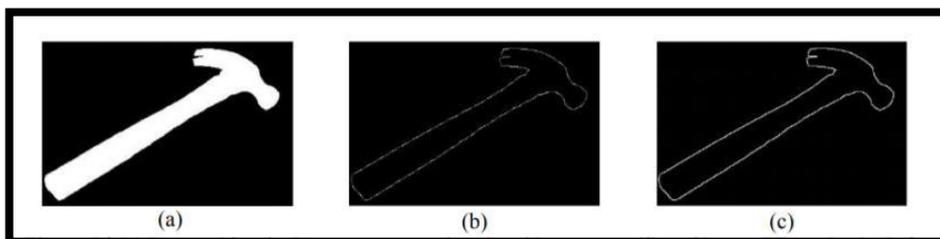


Figura 3-1: Dilatación y erosión de imágenes.

Fuente:(Alegre, Pajares and de la Escalera 2016)

1.6 Hardware

1.6.1 Raspberry PI

Según (Domínguez 2015) Raspberry Pi es un ordenador de placa reducida (SBC) de bajo coste, que se podría considerar como un ordenador de muy pequeño tamaño, comparable con el de una tarjeta de crédito, desarrollado en Reino Unido por la fundación Raspberry Pi, con el objetivo principal de incitar tanto a niños en sus colegios como a adultos a que aprendan sobre ordenadores y todo lo relacionado con ellos. La idea es tener una placa a la que poder conectar monitor, ratón y teclado y ayudar a personas de todas las edades a adentrarse en el mundo de la computación y la programación. La idea de desarrollar algo así surgió en 2006 cuando Eben Upton, Rob Mullins, Jack Lang and Alan Mycroft del laboratorio de informática de la Universidad de Cambridge empezaron a ver como había cambiado los conocimientos de los niños sobre la informática. En la década de 1990 la mayoría de los niños tenían mucha experiencia como programadores aficionados, en cambio en la década del 2000 solo eran capaces de realizar diseño web.

1.7 Redes Neuronales Convolucionales (CNN)

El aprendizaje profundo Deep Learning (DL) permite que modelos computacionales compuestos por varias capas de procesamiento puedan aprender representaciones sobre datos con múltiples niveles de abstracción y, mediante ese concepto, descubrir representaciones precisas de forma autónoma en grandes volúmenes de datos. El DL ha logrado recientemente grandes avances en el reconocimiento de imágenes y videos (Fang et al. 2018).

El propósito de CNN es extraer todas las características de una imagen y luego usar dichas características para detectar o clasificar los objetos en una imagen. Los parámetros de los filtros que se pueden aprender en estas capas; se ajustaran y optimizaran junto con los componentes de clasificación para minimizar el error de clasificación total (Massiris, Delrieux and Fernández Muñoz 2018).

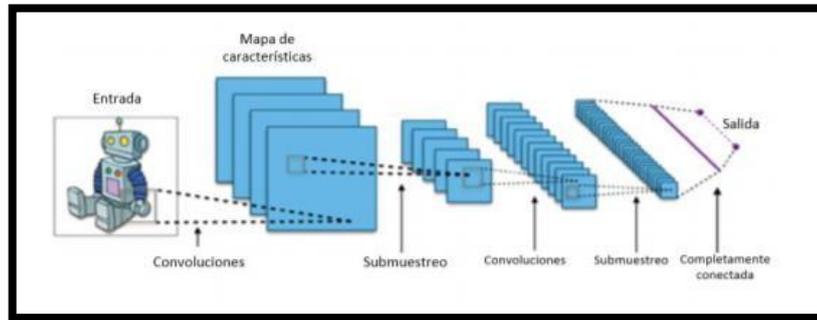


Figura 4-1: Arquitectura de un Red Neuronal Convolutional.

Fuente: (Pacheco 2017)

1.8 YOLOv3

La red YOLO-V3 (Redmon and Farhadi 2018) se desarrolla a partir de las redes YOLO (Redmon et al. 2016) y YOLO-V2 (Redmon and Farhadi 2017) En comparación con la red Faster R-CNN, la red YOLO transforma el problema de detección en un problema de regresión. No requiere una región de propuesta y genera las coordenadas del cuadro delimitador y las probabilidades de cada clase directamente a través de la regresión. Esto aumenta enormemente la velocidad de detección en comparación con Faster R-CNN (Tian et al. 2019).

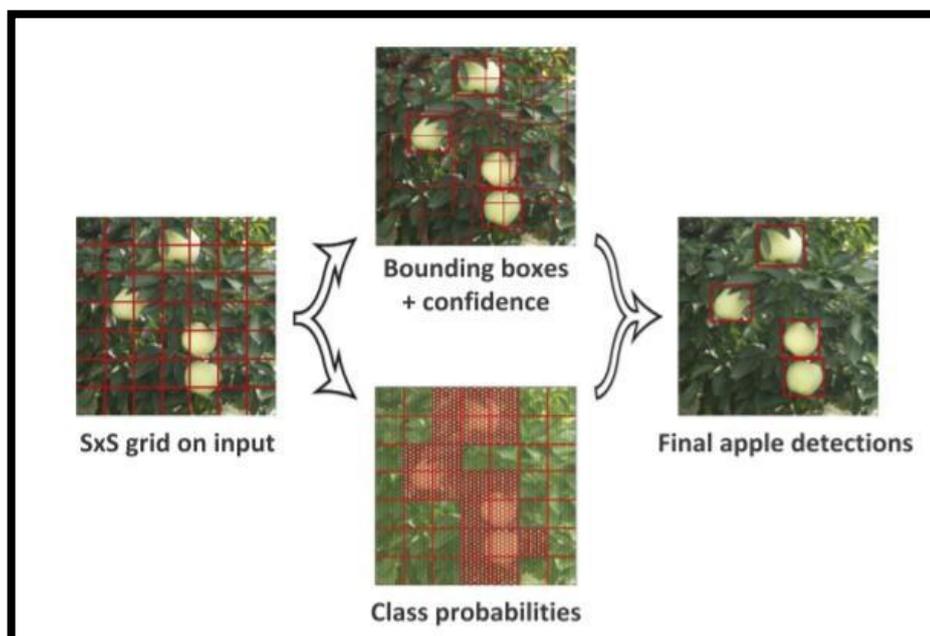
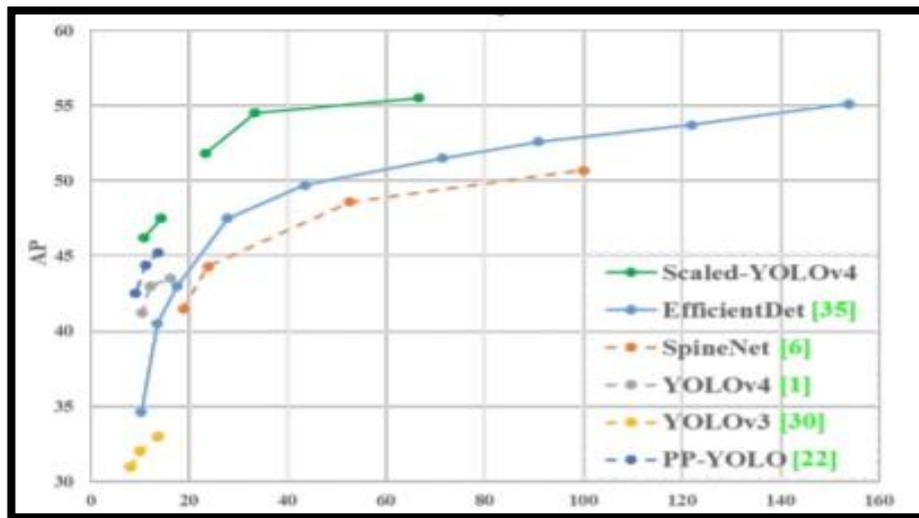


Figura 5-1: Modelo de Detección YOLOv3

Fuente: (Tian et al. 2019)

Según la figura 5-1. La red divide cada imagen en el conjunto de entrenamiento en cuadrículas $S \times S$ ($S = 7$). Si el centro de la verdad del terreno objetivo cae en una cuadrícula, entonces la cuadrícula es responsable de detectar el objetivo. Cada cuadrícula predice los cuadros delimitadores B y sus puntuaciones de confianza, así como las probabilidades condicionales de la clase C (Tian et al. 2019).

1.9 Yolov4-tiny



Gráfica 1-1: Comparación YOLOv4 vs Redes Neuronales de última Generación.

Fuente: (Wang, Bochkovskiy and Liao 2020)

El método Yolov4-tiny está diseñado en base al método Yolov4 para que tenga una velocidad más rápida de detección de objetos. La velocidad de detección de objetos para Yolov4-tiny puede alcanzar 371 fotogramas por segundo usando una GPU 108Ti con la precisión que satisface la demanda de la aplicación real. Aumenta enormemente la viabilidad de que el método de detección de objetos se implemente en sistemas integrados o dispositivos móviles (Jiang et al. 2020).

En método Yolov4-tiny utiliza la red CSPDarknet53-tiny como red troncal en lugar de la red CSPDarknet53 que se utiliza en el método Yolov4. La red CSPDarknet-tiny usa el módulo CSPBlock en la red parcial de etapas cruzadas en lugar del módulo ResBlock en la red residual. El módulo CSPBlock divide el mapa de características en dos partes y combina las dos partes por el borde residual de la etapa cruzada. Esto hace que el flujo de gradiente se pueda propagar en dos rutas de la red diferentes para aumentar la diferencia de correlación de la información de gradiente. El módulo CSPBlock puede mejorar la capacidad de aprendizaje de la red de

convolución de aprendizaje de la red de convolución en comparación con el módulo ResBlock. Aunque esto aumenta el cálculo en un 10% - 20%, mejorando la precisión (Jiang et al. 2020).

En la parte de la fusión de características, el método Yolov4-tiny utiliza una red piramidal de características para extraer mapas de características con diferentes escalas para aumentar la velocidad de detección de objetos, sin utilizar la red de agrupación de pirámides espaciales y agregación de rutas que se utilizan en el método Yolov4. Al mismo tiempo, el Yolov4-tiny utiliza dos mapas de características de escalas diferentes que son 13×13 y 26×26 para predecir resultados de detección. Suponiendo que el tamaño de la cifra de entrada es 416×416 y la clasificación de características es 80 (Jiang et al. 2020).

1.10 Fast R-CNN

Una red Fast R-CNN toma como entrada una imagen completa y un conjunto de propuestas de objetos. La red primero procesa la imagen completa con varias capas de agrupación convolucional (conv) y máxima para producir un mapa de características de conv. Luego, para cada propuesta de objeto, una capa de agrupación de región de interés (RoI) extrae un vector de características de longitud fija del mapa de características. Cada vector de características se alimenta en una secuencia de capas completamente conectadas (fc) que finalmente se ramifican en dos capas de salida hermanas: una que produce estimaciones de probabilidad softmax sobre K clases de objetos más una clase de "fondo" general y otra capa que genera cuatro -números valorados para cada una de las K clases de objetos. Cada conjunto de 4 valores codifica posiciones refinadas del cuadro delimitador para una de las clases K (Girshick 2015).

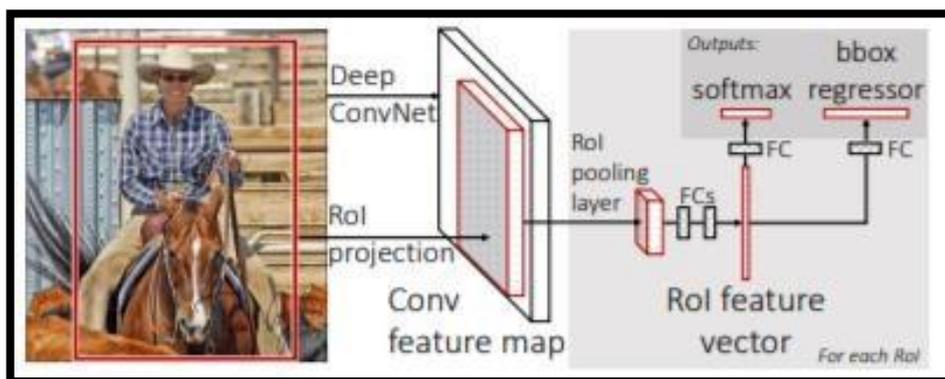


Figura 6-1: Arquitectura R-CNN Rápida.

Fuente:(Girshick 2015)

Según (Girshick 2015), una imagen de entrada y varias regiones de interés (RoI) se introducen en una red totalmente convolucional. Cada RoI se agrupa en un mapa de características de tamaño fijo y luego se asigna a un vector de características mediante capas completamente conectadas (FC).

La Figura 6-1 representa la arquitectura R-CNN rápida, la cual tiene sus vectores de salida mediante probabilidades y mediante compensaciones, aplicada en una imagen (Objeto de estudio).

Según (Girshick et al. 2014), el método de red convolucional basada en regiones (RCNN) logra una excelente precisión en la detección de objetos mediante el uso de una profunda ConvNet para clasificar las propuestas de objetos. R-CNN, sin embargo, tiene notables inconvenientes:

- La formación es un proceso de varias etapas. R-CNN primero afina una ConvNet en propuestas de objetos usando la pérdida de registro. Luego, adapta las SVM a las funciones de ConvNet. Estos SVM actúan como detectores de objetos, reemplazando el clasificador softmax aprendido por el ajuste fino. En la tercera etapa de entrenamiento, se aprenden los regresores de cuadro delimitador.
- La formación es cara en espacio y tiempo. Para el entrenamiento del regresor de caja delimitadora y SVM, las características se extraen de cada propuesta de objeto en cada imagen y se escriben en el disco. Con redes muy profundas, como VGG16, este proceso toma 2.5 GPU-días para las imágenes de 5k del conjunto de trenes VOC07. Estas funciones requieren cientos de gigabytes de almacenamiento.
- La detección de objetos es lenta. En el momento de la prueba, las características se extraen de cada propuesta de objeto en cada imagen de prueba. La detección con VGG16 toma 47 s / imagen (en una GPU).

1.11 Faster R-CNN

Según (Girshick 2015), el sistema de detección de objetos, llamado Faster R-CNN, está compuesto por dos módulos. El primer módulo es una red profunda totalmente convolucional que propone regiones, y el segundo módulo es el detector Fast R-CNN que utiliza las regiones propuestas.

Según (Chorowski et al. 2015), todo el sistema es una red única y unificada para la detección de objetos. Usando la terminología recientemente popular de redes neuronales con mecanismos de "atención", el módulo RPN le dice al módulo Fast R-CNN dónde buscar.

La Figura 7-1 representa el análisis mediante el método Faster R-CNN en cual analiza el objeto de estudio mediante regiones.

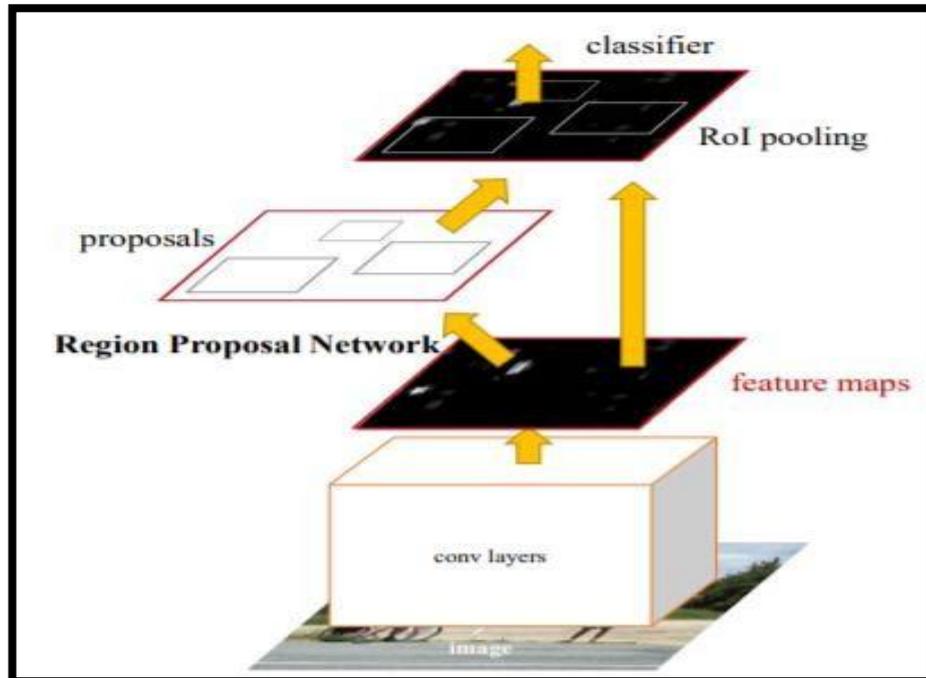


Figura 7-1: Faster R-CNN

Fuente: (Chorowski et al. 2015)

1.12 Software

1.12.1 Python

Según (Gutiérrez 2016), Python es un lenguaje interpretado, de alto nivel y enfocado principalmente a la legibilidad y facilidad de aprendizaje y uso.

Python es un lenguaje multiplataforma, lo que significa que puede usarse en multitud de sistemas distintos. Funciona en ordenadores con sistemas operativos Linux, BSD, Apple, Windows y muchos otros, pero también hay versiones para otros dispositivos, como terminales telefónicos inteligentes, etc. (Gutiérrez 2016).

Naturalmente, Python dispone, por medio del uso de bibliotecas, de herramientas para aprovechar las posibilidades concretas que le brinda cada plataforma, pero también es posible escribir

programas evitando el uso de esas bibliotecas específicas, de modo que esos programas funcionen indistintamente en cualquier ordenador (Gutiérrez 2016).

1.12.2 Características de programación en Python

Python es un software, y se distribuye bajo la licencia “Python Software Foundation License”. Esto, entre otras cosas, significa que se distribuye gratuitamente y no necesita del pago de licencias o royalties para su uso, ya sea privado o comercial (Gutiérrez 2016).

Según (van Rossum and Drake Jr 1995), Python permite que los programas se escriban de forma compacta y legible. Los programas escritos en Python suelen ser mucho más cortos que los programas equivalentes en C, C++ o Java, por varias razones:

- los tipos de datos de alto nivel le permiten expresar operaciones complejas en una sola declaración.
- La agrupación de sentencias se realiza mediante sangría en lugar de corchetes al principio y al final.
- No se necesitan declaraciones de variables o argumentos

1.12.3 Open CV

Según (Arévalo, González and Ambrosio 2004), la librería OpenCV está dirigida fundamentalmente a la visión por computador en tiempo real. Entre sus muchas áreas de aplicación destacarían: interacción hombre-máquina; segmentación y reconocimiento de objetos; reconocimiento de gestos; seguimiento del movimiento; estructura del movimiento; y robots móviles.

1.13 Google Colab

Google Colaboratory más comúnmente conocido como "Google Colab" o simplemente "Colab" es un proyecto de investigación para la creación de prototipos de modelos de aprendizaje automático en potentes opciones de hardware como GPU y TPU. Proporciona un entorno de portátil Jupyter sin servidor para el desarrollo interactivo. Google Colab es de uso gratuito como otros G Suite productos (Bisong 2019).

Los cuadernos colaboratorios se basan en Jupyter funcionan como un objeto de Google Docs: se pueden compartir y los usuarios pueden colaborar en el mismo cuaderno. Colaboraty proporciona tiempos de ejecución de Python 2 y 3 preconfiguradas con las bibliotecas esenciales de aprendizaje automático e inteligencia artificial, como TensorFlow, Matplotlib, keras. La máquina virtual bajo el tiempo de ejecución (VM) se desactiva después de un periodo de tiempo y se pierden todos los datos y configuraciones del usuario. Sin embargo, el portátil se conserva y también es posible transferir archivos desde el disco duro de la VM a la cuenta de Google Drive del usuario. Finalmente, este servicio de Google proporciona un tiempo de ejecución acelerado por GPU (Carneiro et al. 2018).

1.14 Darknet

"Darknet", coloquialmente, se refiere a una red distinta que admite sitios ocultos criptográficamente. Cuatro investigadores de Microsoft popularizaron el término en un artículo fundamental de 2002. Su caso de estudio fue la piratería digital y el desarrollo gradual de redes seguras de igual a igual como alternativas a los sitios de Internet que podrían censurarse fácilmente (Moore and Rid 2016).

La arquitectura Tor proporciona dos servicios: navegación anónima (propiedad 3) y alojamiento de intercambios de información anónimos (propiedad 5), a través de una pieza de software, el llamado "Navegador Tor". Aunque distintos, ambos servicios emplean aproximadamente los mismos protocolos y se basan en la misma infraestructura distribuida. Pero ahí es donde termina su dependencia mutua. No hay ningún requisito técnico para la navegación anónima y el alojamiento anónimo que se combinen (Moore and Rid 2016).

CAPÍTULO II

2. MARCO METODOLÓGICO

En el presente capítulo se describe el software y hardware que conforman el prototipo de monitoreo y alerta en tiempo real para el uso de implementos de seguridad industrial en centros de trabajo mediante visión artificial, a implementar, el cual a partir de este capítulo será denominado prototipo AZATHOTH.

2.1 Método de Investigación

Con la finalidad de convertir teorías generales en un elemento práctico, el presente trabajo de titulación fue elaborado mediante la utilización de dos métodos de investigación:

- Método investigativo.
- Método experimental.

El método investigativo se lo utilizo para la obtención de información (Temas, conceptos, características, etc.), con la finalidad de tener los conocimientos necesarios para la implementación del prototipo AZATHOTH.

Una vez implementado el prototipo AZATHOTH, se utilizó el método experimental para la realización de pruebas, manipulando las variables de estudio se obtuvo los resultados óptimos para el cumplimiento de los requerimientos del prototipo, los cuales a su vez ayudaron a validar el mismo.

2.2 Requerimientos del Prototipo

Para el diseño del prototipo AZATHOTH se definió los siguientes requerimientos con la finalidad de satisfacer los objetivos previamente propuestos:

- Seleccionar una tarjeta de desarrollo para el procesamiento de imágenes.
- Seleccionar un dispositivo para grabar videos con buena resolución.
- Recibir y procesar los datos de la captura de videos.
- Prototipo de fácil transportación e instalación.
- Capaz de almacenar las imágenes procesadas.
- Apto para alertar al usuario de manera visual y/o sonora.

El prototipo AZATHOTH está compuesto por elementos electrónicos previamente escogidos en base a sus características, los cuales son capaces de cumplir con los requerimientos antes mencionados.

2.3 Concepción General del Prototipo AZATHOTH

En la Figura 1-2 se muestra la concepción general del prototipo AZATHOTH el cual fue elaborado con componentes electrónicos, los cuales fueron seleccionados minuciosamente desde un estudio de ingeniería, con el fin de que cumpla con los objetivos planteados.

El diseño del prototipo AZATHOTH consta de un microprocesador Raspberry PI 4 Modelo B, el cual recibe imágenes mediante una cámara Raspberry Pi Cámara V2. Las imágenes son procesadas mediante visión artificial.

El proceso de detección y reconocimiento se basa en los algoritmos de procesamiento de imágenes, las cuales serán detectadas mediante la ayuda de una cámara ubicada estratégicamente en el área de trabajo en el que se requiera el uso obligatorio de indumentaria de seguridad, la cual detectara la presencia de las personas, mediante la toma de video en tiempo real para posteriormente extraer imágenes cuando detecte movimientos en la zona, separándolas del entorno mediante procesamiento digital de la imagen para así poder distinguir en ellas el correcto uso de los implementos de seguridad, generando alertas e informes para así proceder con las acciones correspondientes.

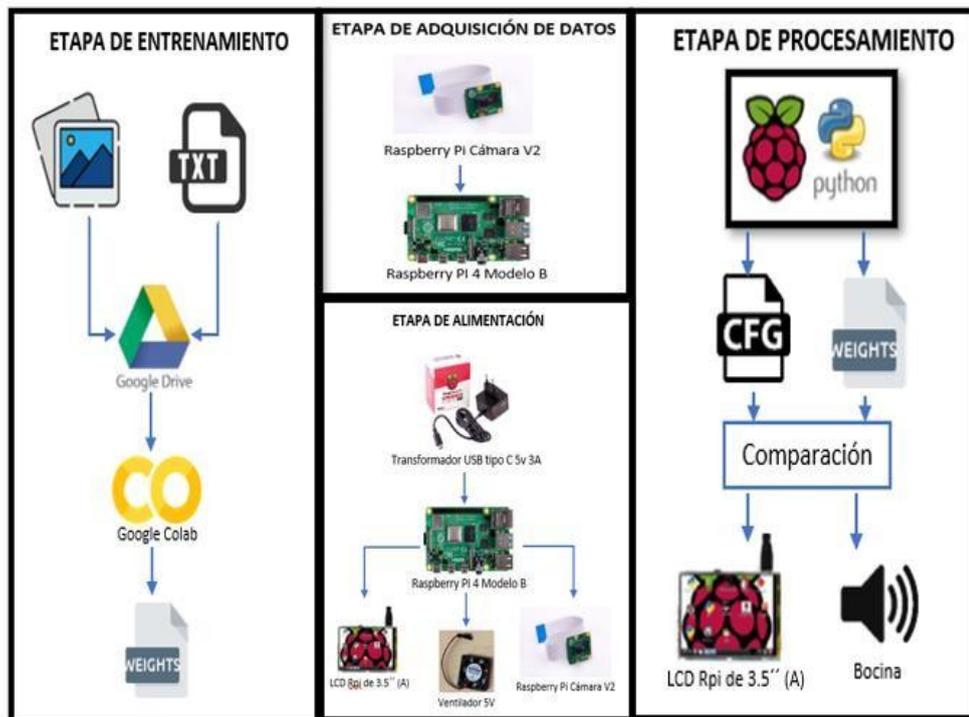


Figura 1-2: Concepción general del prototipo AZATHOTH

Realizado por: Pucha S, Granizo J, 2021.

2.4 Diseño del prototipo AZATHOTH

2.4.1 Etapa de entrenamiento

Según, (Redmon and Farhadi 2018) la cantidad mínima de imágenes necesarias para el entrenamiento por clase es de 400, para el mejor desempeño y optimización de la red neuronal YOLO, se recomienda incrementar el número de imágenes, tener en cuenta que el tiempo de entrenamiento es proporcional a la cantidad de imágenes que la red neuronal procesa.

La etapa de entrenamiento inicio con la recolección de 3000 imágenes (Objetos de estudio, en este caso EPP) en formato JPG y PNG, se utilizaron un total de 6 clases de objetos (Personas, casco, chaleco, guantes, orejeras y botas), el método utilizado en el entrenamiento de este sistema sugiere el uso de 500 imágenes por cada clase, en este caso se obtuvo un total de 3000 imágenes debido a que en cada imagen se encuentra más de una clase de los objetos de estudio.

Las imágenes empleadas fueron adaptadas a los formatos y colores con los que la plataforma YOLOv4-tiny debe ser entrenado, las cuales son, imágenes de dimensión 416*416 y color BRG, además el método requiere que se señale la ubicación espacial de la imagen en cada clase presente en la imagen para lo cual se utilizó archivos .TXT donde se detalla las coordenadas de cada clase existente en la imagen que posteriormente fueron almacenadas en Google drive para realizar la etapa de entrenamiento.

El entrenamiento se realizó en base a la arquitectura YOLOv4-tiny, el cual fue realizado en la plataforma online Google Colab, la cual brinda máquinas virtuales con características elevadas de procesamiento y GPU de manera gratuita enfocada para el estudio de inteligencia y visión artificial.

El entrenamiento tuvo una duración de 12 horas debido a que la plataforma permite que un usuario utilice las máquinas virtuales durante este tiempo, obteniendo al finalizar un archivo (.weight), el proceso realizado es el esquematizado en la Figura 2-2.

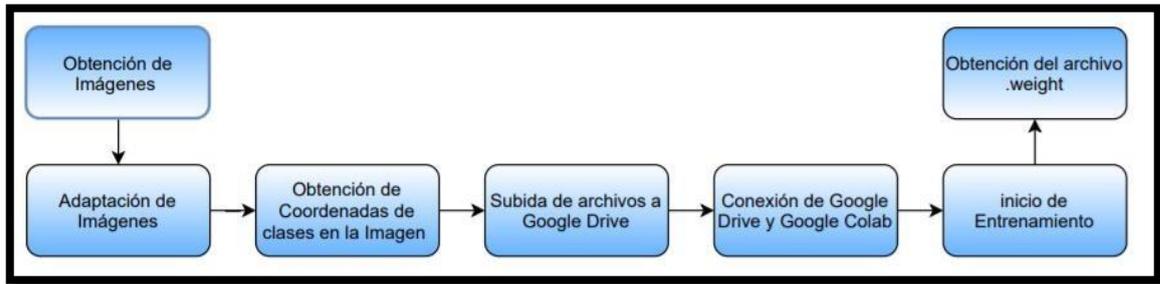


Figura 2-1: Etapa de Entrenamiento AZATHOTH.

Realizado por: Pucha S, Granizo J, 2021.

2.4.2 Etapa de adquisición de datos

Para la obtención de datos se analizaron 6 objetos de estudio, los cuales fueron sometidos a los métodos de visión artificial para cumplir con los objetivos antes propuestos, para un mejor tratamiento de estos conceptos a lo largo del texto se los denominó de la manera en la que nos muestra la Tabla 1-2.

Tabla 1-2: Simplificación de los nombres de cada equipo de protección personal.

| Denominación de clases | |
|------------------------|---------|
| Casco de Seguridad | Clase 1 |
| Botas de Seguridad | Clase 2 |
| Chaleco Reflectivo | Clase 3 |
| Orejeras de Seguridad | Clase 4 |
| Guantes de Seguridad | Clase 5 |
| Persona | Clase 6 |

Realizado por: Pucha S, Granizo J, 2021

El diagrama de bloques de la Figura 3-2 representa la etapa de adquisición de datos, la cual inicia detectando la presencia de una persona (Objeto de estudio), por medio del método de sustracción de imágenes que son obtenidas mediante el módulo Raspberry Pi Cámara V2, que se encuentra conectado al puerto especializado para dicho uso en la Raspberry PI 4 Modelo B.

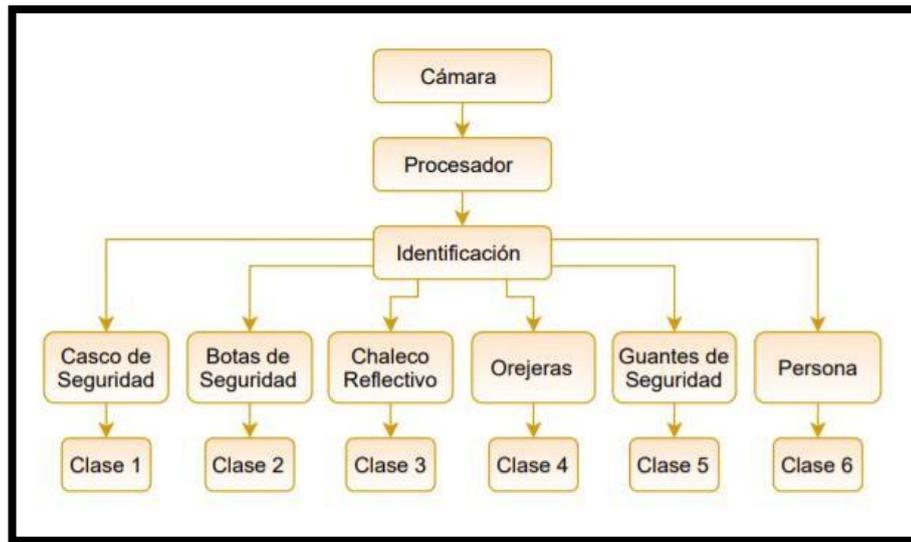


Figura 3-2: Etapa de adquisición de datos

Realizado por: Pucha S, Granizo J, 2021.

2.4.3 *Etapa de procesamiento de información.*

El diagrama de bloques representado en la Figura 4-2 se esquematiza la etapa de procesamiento de información (Imágenes), el cual inicia con la obtención del video en tiempo real a través de la Raspberry Pi Cámara V2 la cual está conectada al puerto especializado en el módulo Raspberry PI 4 Modelo B, la cual procesa Frame a Frame para realizar la sustracción de la imagen y posteriormente segmentarla y extraer sus características, las cuales serán utilizadas para realizar una comparación con el archivo obtenido en la etapa de entrenamiento, para la identificación de las EPP del objeto a estudio.

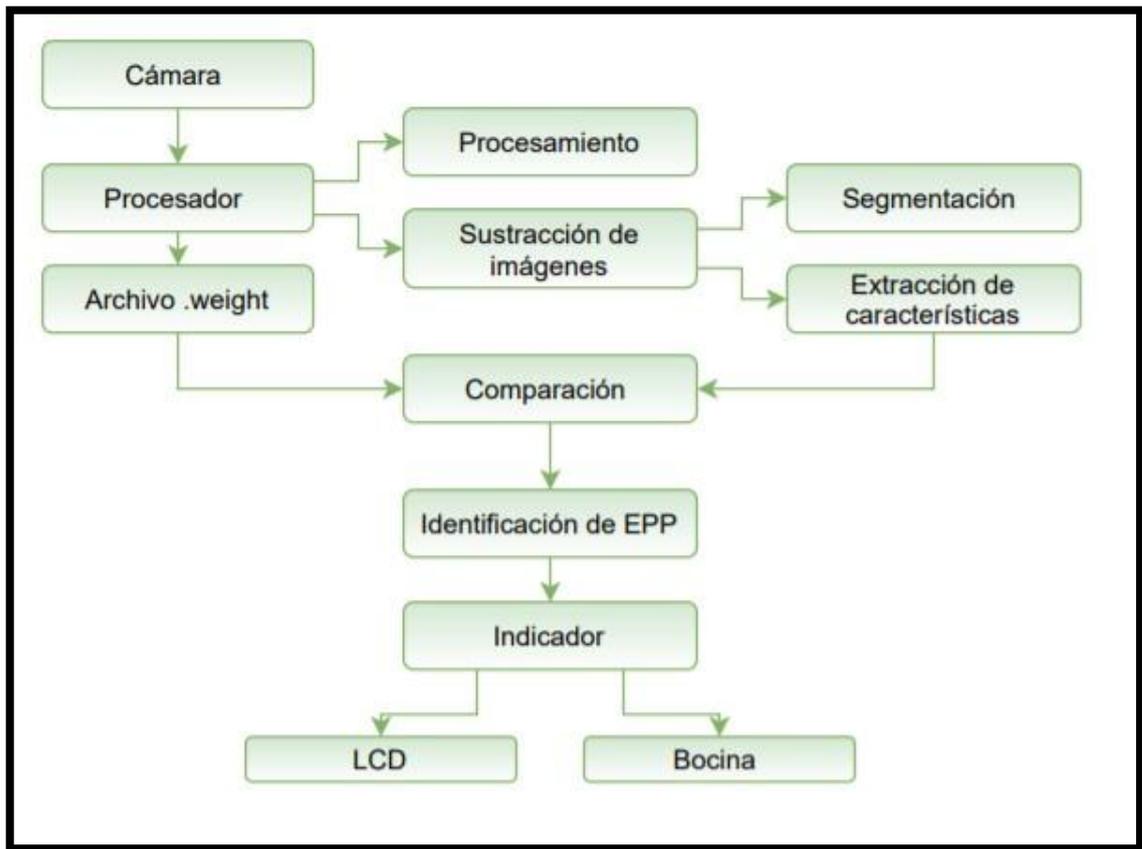


Figura 4-2: Etapa de procesamiento de información

Realizado por: Pucha S, Granizo J, 2021.

2.4.4 Etapa de alimentación

El diagrama de bloques representado en la Figura 5-2, se esquematiza la etapa de alimentación del sistema, en la cual tenemos 4 elementos (Raspberry PI 4 Modelo B, Raspberry Pi Cámara V2, LCD Rpi de 3.5 Pulgadas (A), 480x320, sistema de ventilación), los cuales serán alimentación mediante un adaptador de V a 3A.

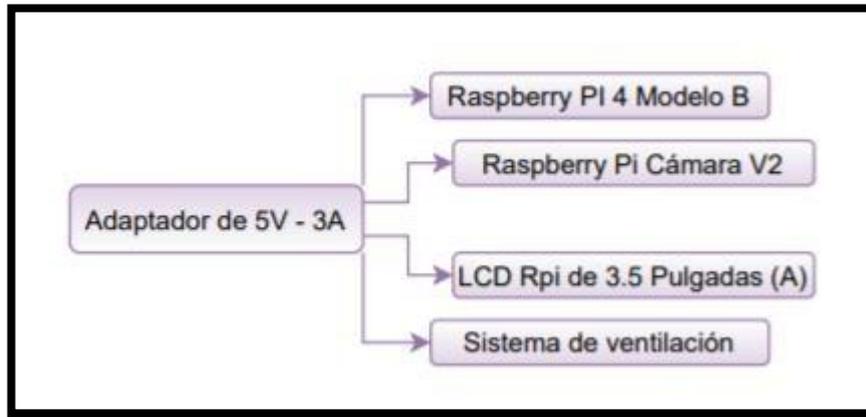


Figura 5-2: Etapa de alimentación del sistema

Realizado por: Pucha S, Granizo J, 2021.

2.5 Hardware del prototipo AZATHOTH

2.5.1 Raspberry Pi

En la Tabla 2-2 se presentan los 3 últimos modelos lanzados por la fundación Raspberry Pi (Raspberry PI 3 Modelo A+, Raspberry PI 3 Modelo B+, Raspberry PI 4 Modelo B), analizando las características principales de cada uno se seleccionó el hardware óptimo para la implementación el prototipo AZATHOTH.

Tabla 2-1: Características de módulos Raspberry

| Características | Raspberry PI 3 Modelo A+ | Raspberry PI 3 Modelo B+ | Raspberry PI 4 Modelo B |
|---------------------|-----------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------|
| Procesador | Broadcom BCM2837B0, Cortex-A-53 64bit SoC, 1.4GHz | Broadcom BCM2837B0, Cortex-A53 64-bit SoC, 1.4GHZ | Broadcom BCM2711, quad-core Cortex-A72 (ARM v8) 64-bit SoC, 1.5GHz |
| Memoria | 512MB LPDDR2SDRAM | 1GB LPDDR2 SDRAM | 2GB, 4GB, 8GB (dependiendo del modelo) |
| Conectividad | 2.4 GHz o 5 GHz IEEE 802.11.b/g/n/ac Wireless LAN, Bluetooth 4.2, BLE, 1 puerto USB 2.0 | 2.4 GHz o 5 GHz IEEE 802.11.b/g/n/ac Wireless LAN, Bluetooth 4.2, BLE, Gigabit Ethernet a través de USB 2.0, 4 puertos USB 2.0 | 2.4 GHz o 5.0 GHz IEEE 802.11.b/g/n/ac Wireless LAN, Bluetooth 5.0, BLE Gigabit Ethernet, 2 puertos USB 3.0, 2 puertos USB 2.0 |

| | | | |
|--------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Audio y Video | 1 HDMI de tamaño completo, Puerto de pantalla MIPI DSI, Puerto de cámara MIPI CSI, Salida de estéreo de 4 polos y puerto de video compuesto | 1 HDMI de tamaño completo, Puerto de pantalla MIPI DSI, Puerto de cámara MIPI CSI, Salida de estéreo de 4 polos y puerto de video compuesto | 2 puertos micro-HDMI (hasta 4kp60 compatible), Puerto de pantalla MIPI DSI de 2 carriles, Puerto de cámara MIPI CSI de 2 carriles, Puerto de video compuesto y audio estéreo de 4 polos |
| Multimedia | H.264, MPEG-4 decodificación (1080p30), H.264 codificación (1080p30), OpenGL ES 1.1, gráficos 2.0 | H.264, MPEG-4 decodificación (1080p30), H.264 codificación (1080p30), OpenGL ES 1.1, gráficos 2.0 | H.265 (decodificación 4kp60), H264 (decodificación 1080p60, codificación 1080p30), Gráficos OpenGL ES 3.0 |
| Soporte de tarjeta SD | Ranura para tarjeta micro-SD para cargar el sistema operativo y el almacenamiento de datos | Ranura para tarjeta micro-SD para cargar el sistema operativo y el almacenamiento de datos | Ranura para tarjeta micro-SD para cargar el sistema operativo y el almacenamiento de datos |
| Alimentación | 5 V / 2.5 A CC mediante conector micro USB, 5 V CC a través del encabezado GPIO | 5 V / 2.5 A CC mediante conector micro USB, 5 V CC a través del encabezado GPIO, Power over Ethernet (PoE) habilitado (Requiere PoE HAT separado) | 5 V CC a través del conector USB-C (mínimo 3A), 5 V CC a través del encabezado GPIO (mínimo 3A), Power over Ethernet (PoE) habilitado (Requiere PoE HAT separado) |
| Temperatura de funcionamiento | 0 – 50 °C | 0 – 50 °C | 0-50 grados C ambiente. |

Fuente: (Raspberry Pi, 2019)

Realizado por: Pucha S, Granizo J, 2021

Con las características presentadas en la Tabla 2-2 se puede deducir que los modelos Raspberry PI 3 Modelo B+ y Raspberry PI 4 Modelo B poseen características similares y a su vez superiores al modelo Raspberry PI 3 Modelo A+.

El módulo Raspberry PI 4 Modelo B posee mejoramientos en su procesador, conectividad y memoria el cual se lo puede observar en la Figura 6-2. En el Anexo A se muestra el datasheet del módulo de Raspberry seleccionado para la implementación del prototipo.

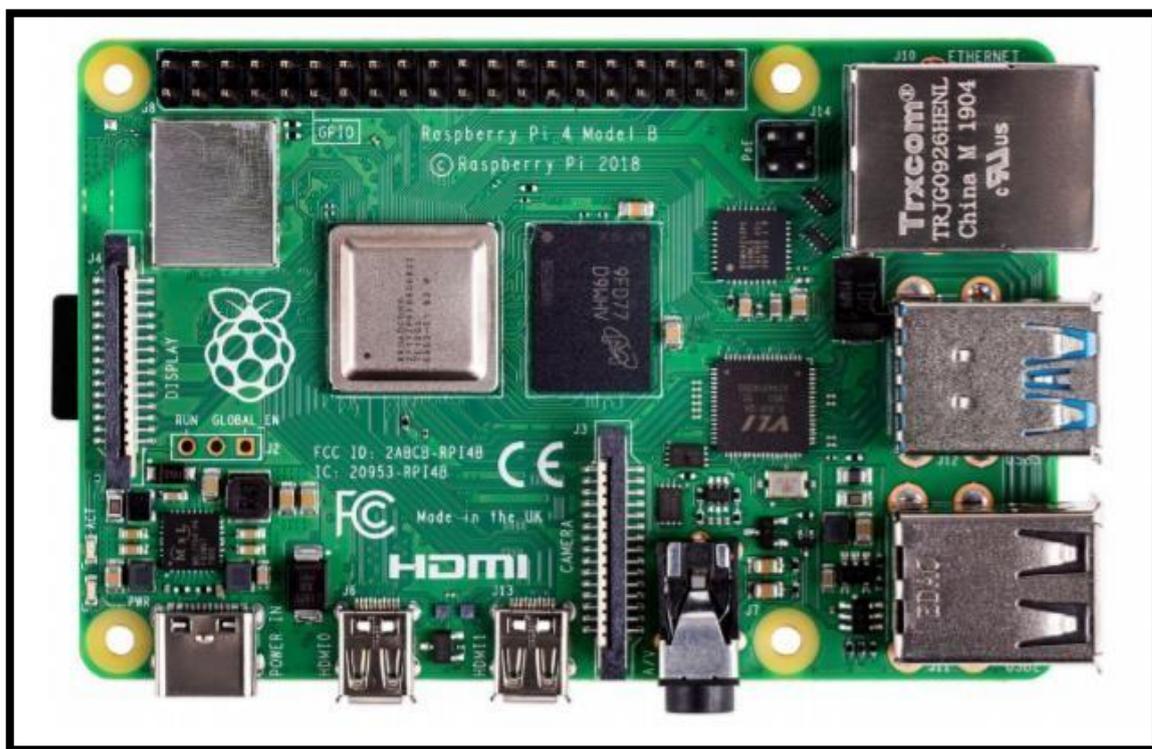


Figura 6-2: Módulo Raspberry Pi 4 Modelo B

Fuente: (Raspberry Pi, 2019)

2.5.2 Dispositivo de captura

En la tabla 3-2 se observa las características principales de 2 modelos de cámara que nos ofrece la empresa Raspberry los cuales son: Raspberry Pi Cámara V1.3 y Raspberry Pi Cámara V2.

Tabla 3-2: Características de los dispositivos de adquisición de imágenes.

| Características | Raspberry Pi Cámara V1.3 | Raspberry Pi Cámara V2 |
|----------------------------------|---------------------------------------------------------------|---------------------------------------------------------------|
| Resolución | Sensor Omnivisión 0V5647 5MP | Sensor Sony IMX219 de 8 MP |
| Resolución de Imagen Fija | 1080p HD a 30fps | Resolución de enfoque fijo. |
| Video | Admite 1080p A 30fps, 720p A 60fps y 640x480p Grabación 60/90 | Admite 1080p A 30fps, 720p A 60fps y 640x480p Grabación 60/90 |
| Audio | No cuenta con micrófono incorporado | No cuenta con micrófono incorporado |
| Interfaz | MIPI de 15 Pines | Conexión a puerto CSI |
| Tamaño | 20 x 25 x 9 mm | 25 x 24 x 9 mm |
| Peso | 3g | 3g |

Fuente: (Raspberry Pi, 2019)

Realizado por: Pucha S, Granizo J, 2021

Ambos modelos de LCD son compatibles con los módulos de Raspberry antes presentados, y para la implementación del prototipo AZATHOTH ambos son adecuados para utilizarlos, pero por el nivel de resolución nos inclinamos por el modelo Raspberry Pi Cámara V2, el cual nos brinda una resolución de 8 MP, adicionalmente posee una resolución de enfoque fija, en la Figura 7-2 se observa el elemento seleccionado, y en el Anexo B se encuentra el Datasheet del mismo.

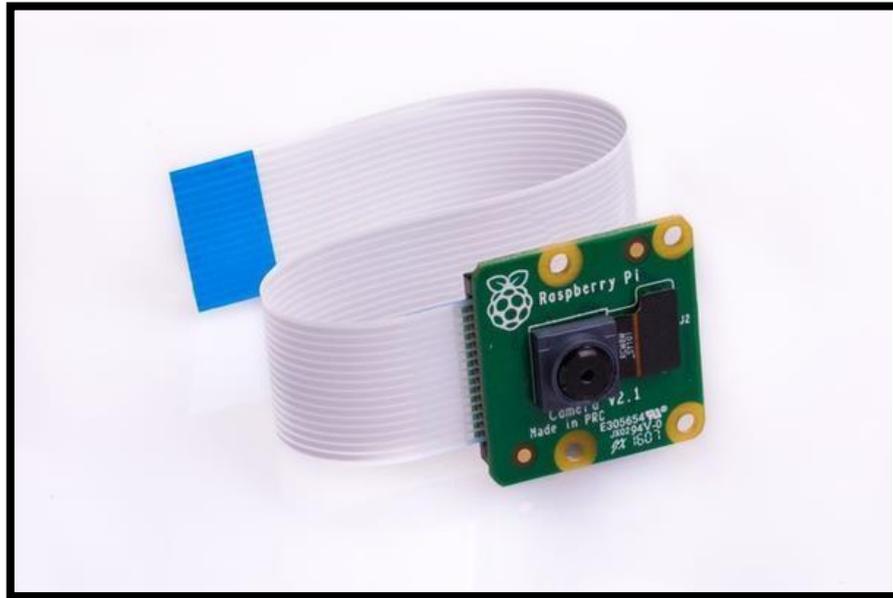


Figura 7-2: Raspberry Pi Cámara V2

Fuente: (Raspberry Pi, 2019)

2.5.3 Interfaz

En la Tabla 4-2, se presentan las características principales de dos tipos de LCD compatibles con los módulos Raspberry mencionados antes, los cuales son: LCD RPi 2.8 Pulgadas (A), 320x240 y LCD Rpi de 3.5 Pulgadas (A), 480x320.

Tabla 4-2: Módulos de interfaz

| Características | LCD RPi 2.8 Pulgadas (A), 320x240 | LCD Rpi de 3.5 Pulgadas (A), 480x320 |
|--------------------------------|--------------------------------------|-----------------------------------------|
| Tipo de LCD | TFT | TFT |
| Interfaz LCD | SPI | SPI |
| Tipo de Pantalla Táctil | Control Táctil Resistivo | Control Táctil Resistivo |
| Controlador de Pantalla Táctil | XPT2046 | XPT2046 |
| Colores | 65536 | 65536 |

| | | |
|-------------------------------------------|-------------------|-------------------|
| Iluminar desde el Fondo | LED | LED |
| Resolución | 320x240 (Píxeles) | 480x320 (Píxeles) |
| Relación de Aspecto | 4:3 | 8:5 |
| Consumo de Energía | TBD | TBD |
| Corriente de Luz de Fondo | TBD | TBD |
| Temperatura de Funcionamiento (°C) | TBD | TBD |

Fuente: (Raspberry Pi, 2019).

Realizado por: Pucha S, Granizo J, 2021.

Las dos pantallas LCD comparadas poseen características similares y son totalmente adecuadas para la implementación del prototipo AZATHOTH, en este caso nos inclinamos por el modelo LCD Rpi de 3.5 Pulgadas (A), 480x320 que se presenta en la Figura 8-2, únicamente por el tamaño, con la finalidad de tener una mejor experiencia con el usuario.

En el Anexo C encontraremos el Datasheet del LCD.



Figura 8-2: LCD Rpi de 3.5 Pulgadas (A)

Fuente: (Raspberry Pi, 2019)

2.6 Diagrama de Conexión

En la Figura 9-2 se presenta las conexiones del hardware del prototipo, en la cual se encuentra los elementos antes mencionados (Raspberry PI 4 Modelo B, Raspberry Pi Cámara V2, LCD Rpi de 3.5 Pulgadas (A), 480x320, sistema de ventilación).

El esquema fue realizado mediante la ayuda del software Fritzing el cual es de libre uso y especializado para el diseño electrónico, dicha plataforma dispone de los recursos de hardware que necesitamos para la implementación del prototipo AZATHOTH.

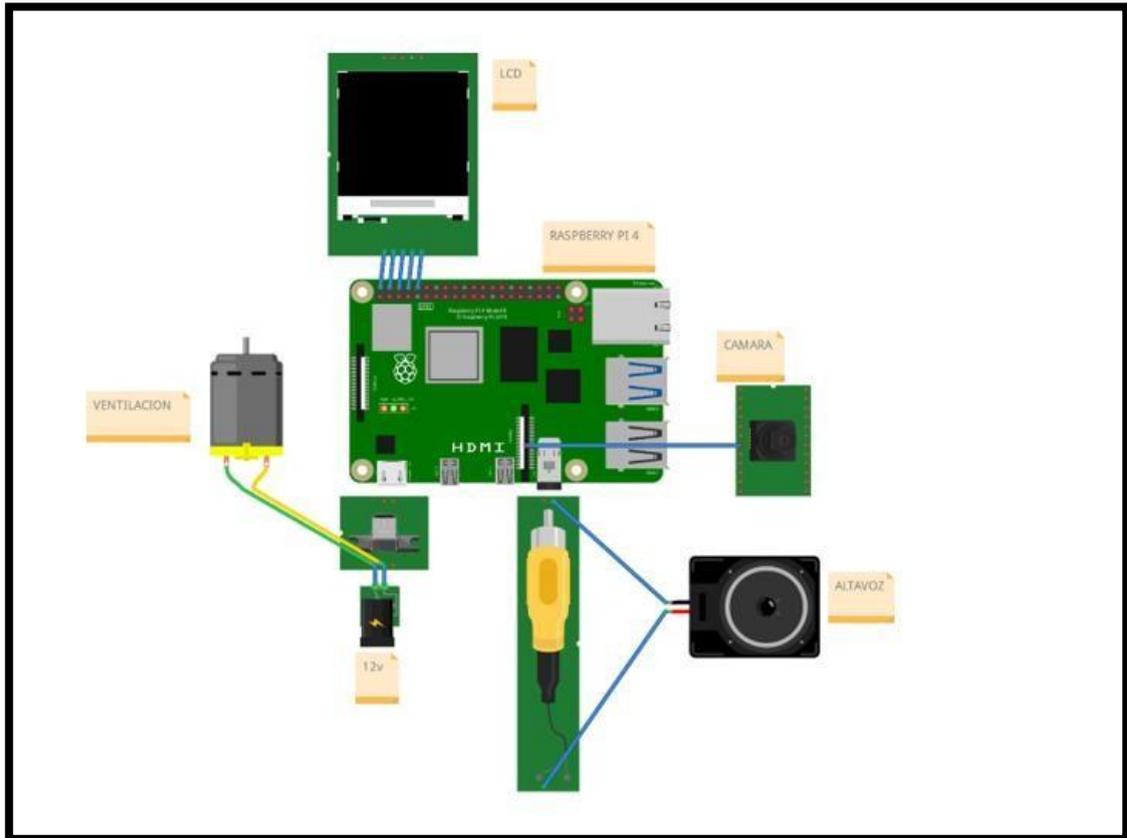


Figura 9-2: Conexiones del hardware del prototipo AZATHOTH

Realizado por: Pucha S, Granizo J, 2021.

2.7 Elaboración del Case del Prototipo AZATHOTH

Determinados los elementos electrónicos y su debida conexión se procedió a realizar la elaboración de un case (Carcasa para el prototipo) el cual en su interior tiene los siguientes elementos:

- Raspberry PI 4 Modelo B

- Raspberry Pi Cámara V2
- LCD Rpi de 3.5 Pulgadas (A)

El modelamiento del case para el prototipo AZATHOTH fue elaborado en el software SolidWorks.

El case tiene forma rectangular y posee las siguientes dimensiones:

- 61.98mm de ancho
- 122mm de largo
- 30.9mm de alto

Adicionalmente cuenta con orificios específicos para:

- Puertos de Raspberry PI 4 Modelo B
- Ventilación
- LCD

En la Figura 10-2 se observa el case elaborado con todas las características antes mencionadas.



Figura 10-2: Case del Prototipo AZATHOTH

Realizado por: Pucha S, Granizo J, 2021.

2.8 Software del Dispositivo.

2.8.1 *Requerimientos del software*

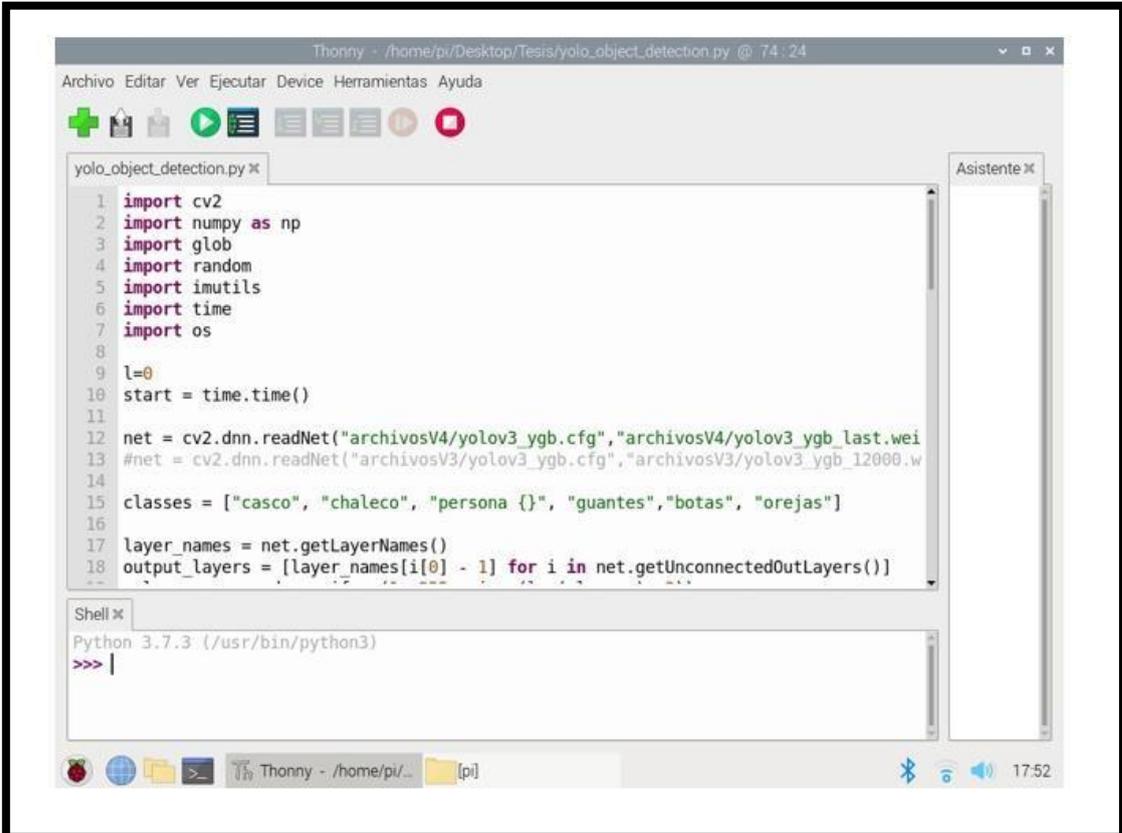
- Sistema Operativo Raspbian
- Lenguaje de programación Python.
- librerías OPEN CV, NumPy e imutil.
- Algoritmo para reconocimiento de objetos en imágenes.
- Emisión de una alarma por falta de EPP.

2.8.2 *Lenguaje de programación*

Por defecto el módulo (Raspberry PI 4 Modelo B) tiene preinstalado el lenguaje de programación Python 2.7, programa que es especializado para trabajos de visión artificial y el cual será utilizado para la implementación del prototipo.

Para la implementación del prototipo AZATHOTH, Python 2.7 se lo cambio por la versión de Python 3.0 la cual nos permite trabajar con las librerías más actuales de: OPEN CV, NumPy e imutil.

En la Figura 11-2 se observa la ventana de Python, en la cual se desarrolló el algoritmo de detección para el prototipo.



The image shows a screenshot of the Thonny Python IDE. The main window displays a Python script named 'yolo_object_detection.py'. The script includes the following code:

```
1 import cv2
2 import numpy as np
3 import glob
4 import random
5 import imutils
6 import time
7 import os
8
9 l=0
10 start = time.time()
11
12 net = cv2.dnn.readNet("archivosV4/yolov3_ygb.cfg", "archivosV4/yolov3_ygb_last.wei
13 #net = cv2.dnn.readNet("archivosV3/yolov3_ygb.cfg", "archivosV3/yolov3_ygb_12000.w
14
15 classes = ["casco", "chaleco", "persona {}]", "guantes", "botas", "orejas"]
16
17 layer_names = net.getLayerNames()
18 output_layers = [layer_names[i[0] - 1] for i in net.getUnconnectedOutLayers()]
```

Below the code editor is a shell window showing the Python 3.7.3 prompt: >>> |. The Thonny interface includes a menu bar (Archivo, Editar, Ver, Ejecutar, Device, Herramientas, Ayuda), a toolbar with icons for file operations and execution, and a status bar at the bottom showing the current directory and time (17:52).

Figura 11-2: Entorno de trabajo de Python

Realizado por: Pucha S, Granizo J, 2021.

2.8.3 Red neuronal convolucional

Tabla 5-2: Comparación de tiempos de procesamiento entre redes neuronales.

| Red Neuronal Convolutacional | Tiempo de Procesamiento | | |
|------------------------------|-------------------------|-----------|-----------|
| | Imagen A | Imagen B | Imagen C |
| Fast R-CNN | 41.64326 | 55.769754 | 83.81045 |
| Faster R-CNN | 38.53267 | 48.45637 | 73.88907 |
| YOLO | 1.929929 | 7.366887 | 19.389161 |

Fuente: (Huang, Pedoeem and Chen 2018)

Realizado por: Pucha S, Granizo J, 2021

Una vez analizados los valores mostrados en la Tabla 5-2, se observó que YOLO (You Only Look Once) tiene un tiempo de procesamiento superior a la red neuronal convolutacional Fast R-CNN y Faster R-CNN, por lo que la red neuronal (YOLO) será utilizada para la implementación del prototipo AZATHOTH.

2.9 Programación

2.9.1 Programación en Google Colab

En la Figura 13-2 se observa el diagrama de flujo para el proceso de entramiento de YOLO v4 Tiny (cuyo código fue genera generado y procesado en Google Colab utilizando como GPU una TESLA G4 de 16GB).

El proceso es detallado a continuación:

- Inicio
- Verificación del CUDA.
- Selección de GPU.
- Conexión con Google Drive
- Acceso al Dataset de Google Drive
- Clonado de la DARNET
- Clonación de la DARKNET para entrenamiento
- Generación del archivo train.txt
- Descarga de pesos pre entrenados
- Generación del archivo entrenado
- Fin

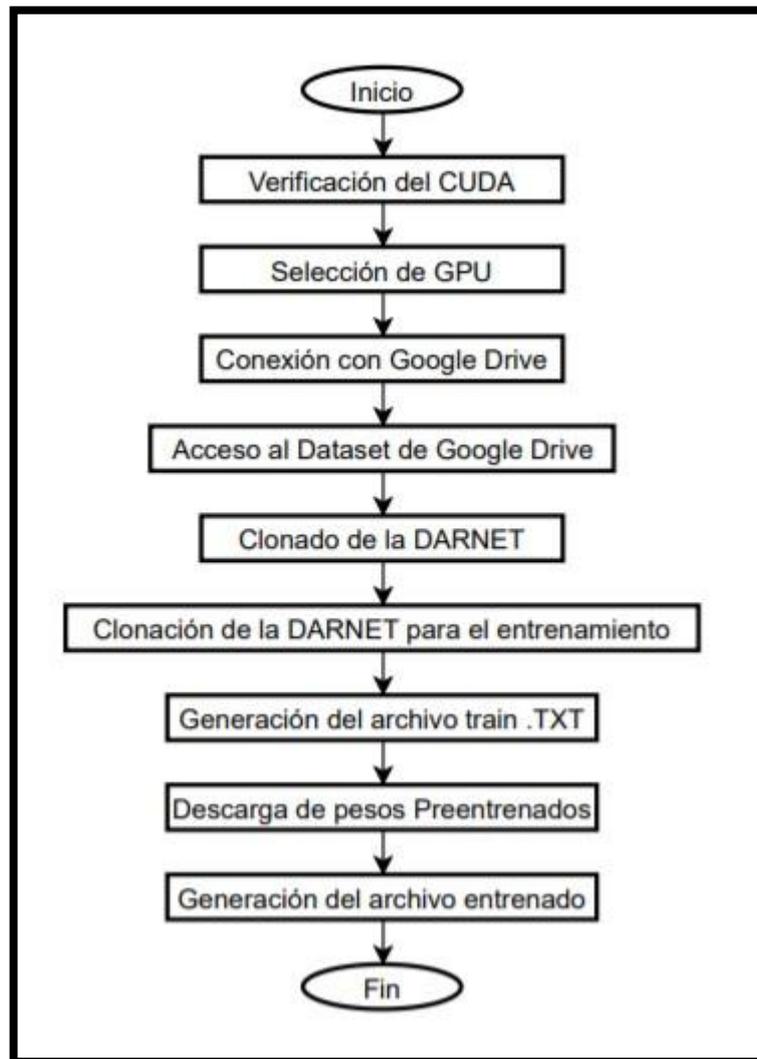


Figura 13-2: Diagrama de flujo para el proceso de entramiento.

Realizado por: Pucha S, Granizo J, 2021.

El entorno y la red usada para cumplir con el procedimiento fueron:

- Google colab: la cual es una máquina virtual que la empresa Google brinda a desarrollador para poder entrenar redes neuronales en máquinas de gama alta sin costo alguno.
- Darknet: es una red neuronal pre entrenada con capas de convolución y max pooling.

En esta etapa es necesario crear los archivos de etiquetado que se va a usar para el entrenamiento, para lo cual, se utilizara las 3000 imágenes generando así manualmente para cada una las coordenadas en las cuales está presente alguna clase de os objetos a estudio.

En la Figura 14-2 se muestra el proceso de entrenamiento de la red neuronal convolucional mediante la plataforma de Google Colab.

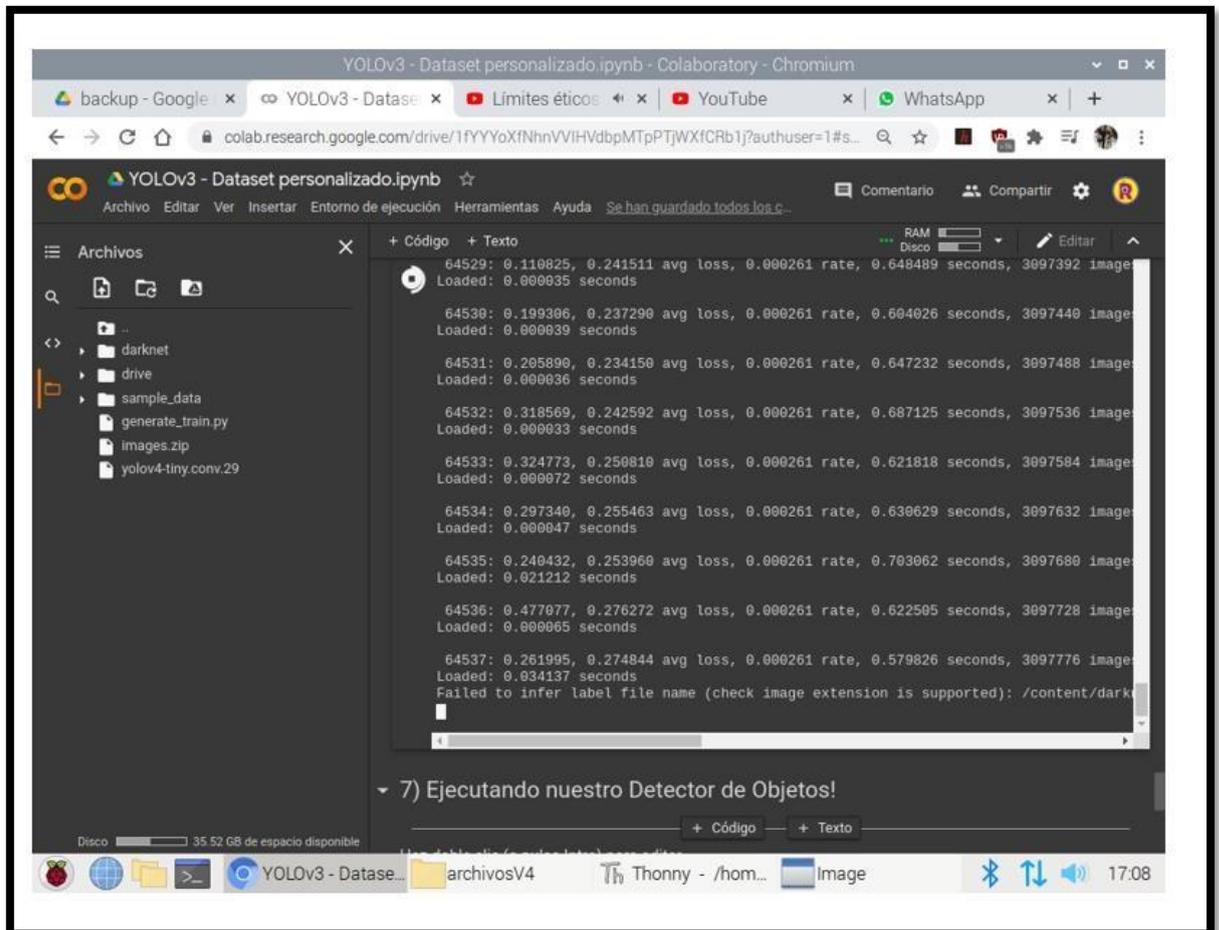


Figura 14-2: Entrenamiento de la Red mediante Google Colab.

Realizado por: Pucha S, Granizo J, 2021.

2.9.2 Programación en ID Python

En la figura 15-2 se presenta el diagrama de flujo del procesamiento de imágenes para la detección de EPP, el cual se detalla a continuación:

- Inicio
- Importación de librerías.
- Lectura de archivos entrenados
- Definición de clases
- Obtención de fotograma

- Redimensionamiento del Fotograma
- Obtención de datos del Fotograma
- Detección de objetos en el fotograma
- Obtención de coordenadas en objetos
- Señalización de objetos encontrados
- Conteo de clases
- Generación de indicadores en la bocina y la LCD
- Generación de la imagen señalizada
- Subida del archivo de EXCEL a Google Drive
- FIN

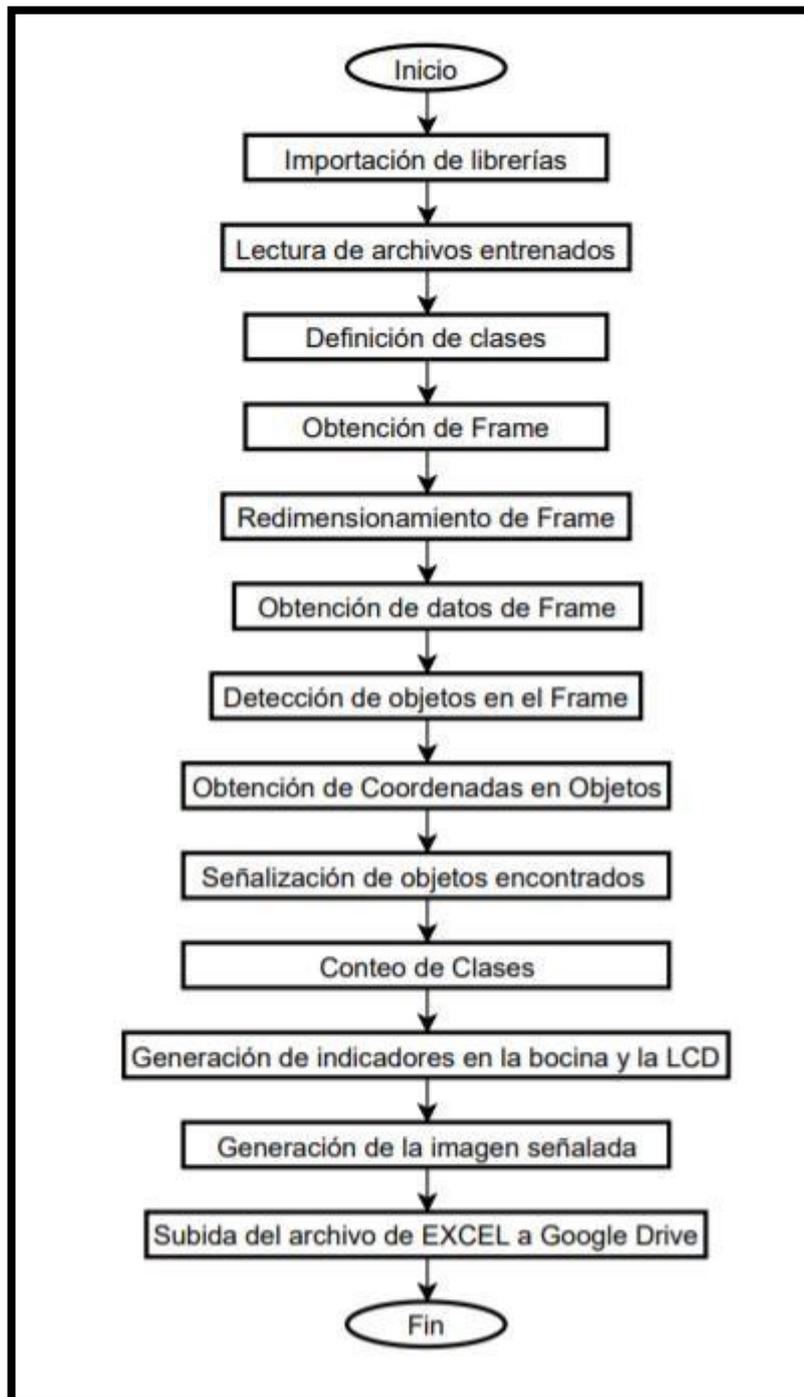


Figura 15-2: Diagrama de flujo del procesamiento de imágenes

Realizado por: Pucha S, Granizo J, 2021.

Para el cumplimiento de este procedimiento se usó las librerías:

- OPEN CV: es una librería de Python especializada en visión artificial que posee algoritmos capaces de simplificar el tratamiento y análisis de imágenes.

- **NUMPY:** es una librería la cual permite trabajar más eficientemente con arreglos como vectores y matrices.
- **IMUTILS:** es una librería generada para el procesamiento de imágenes que permite el re-escalamiento conversión y otras funciones de una manera más sencilla.
- **RANDOM:** esta librería es utilizada para genera número aleatorios.
- **TIME:** esta librería es usada para obtener el tiempo real de la máquina y saber cuál es el tiempo consumido en el procesamiento del fotograma.
- **Pygame:** es una librería que permite que permite la reproducción de archivos en formato mp3.

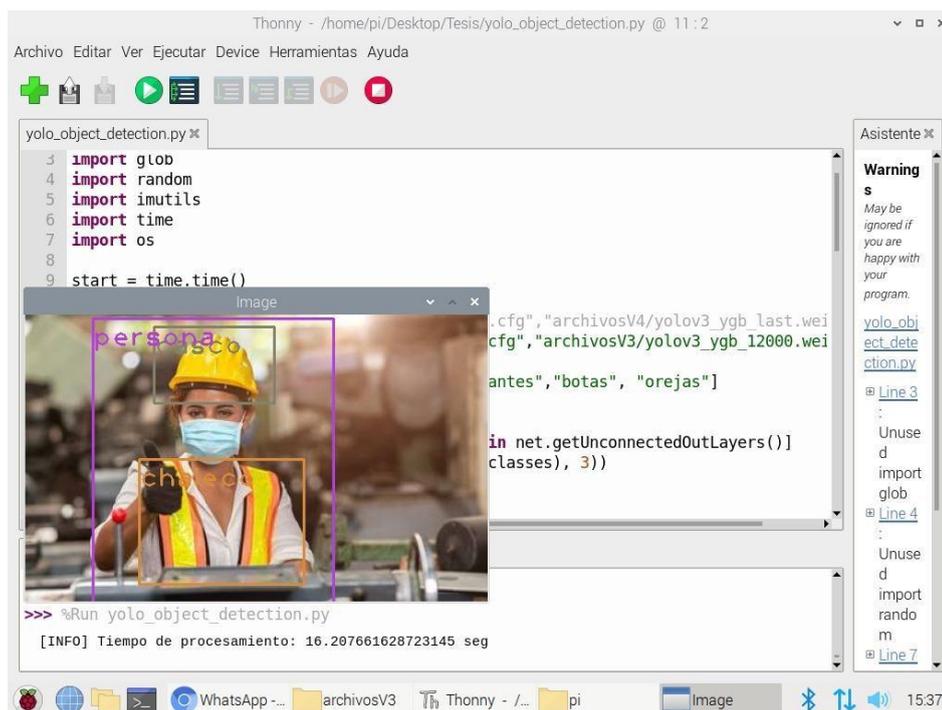


Figura 16-2: Imagen analizada mediante la red neuronal convolucional.

Realizado por: Santiago P, Jefferson G

En la Figura 16-2 se observa una imagen analizada mediante el entrenamiento de la red neuronal convolucional, las pruebas en tiempo real se realizarán cuando el prototipo cumpla todos los requerimientos establecidos.

CAPÍTULO III

3. ANÁLISIS Y RESULTADOS

En el presente capítulo se evidencian las pruebas a las cuales el prototipo AZATHOTH fue sometido para su debida validación, los resultados obtenidos a través de las pruebas nos

permitieron evaluar el correcto funcionamiento del prototipo con el fin de realizar y cumplir eficazmente los requerimientos y objetivos establecidos los capítulos anteriores.

En la Figura 1-3 se muestra el prototipo AZATHOTH en su versión final.



Figura 1-3: Prototipo AZATHOTH

Realizado por: Pucha S, Granizo J, 2021.

3.1 Método de Validación de Resultados

Para la validación de los resultados obtenidos por nuestro prototipo AZATHOTH se utilizó el método experimental. Mediante el cálculo de la variable mAP (Precisión media media) para cada clase, con el objetivo de obtener una precisión de detección adecuada para la red neuronal convolucional obtenida.

Para el cálculo de esta variable se debe considerar el siguiente parámetro:

- La variable (Precisión) nos permite obtener un valor porcentual de las predicciones correctas en una clase.

Para calcular el valor porcentual de mAP se utilizarán las siguientes formulas:

| | |
|---------------------------------|-----------------------------|
| $Precision = \frac{TP}{TP+FP}$ | $Recall = \frac{TP}{TP+FN}$ |
| $AP = \frac{\sum Precision}{N}$ | $mAP = \frac{\sum AP}{NC}$ |

Figura 2-3: Fórmulas para calcular valor mAP.

Realizado por: Pucha S, Granizo J, 2021.

En la Figura 2-3 se observa las fórmulas que se utilizarán para obtener el valor mAP, dichas fórmulas son usadas específicamente para cuando se procesa información sobre detección de objetos.

Para la utilización de la ecuación 1 hasta la Ecuación 4 debemos tener en cuenta la consideración de la información obtenida de la siguiente manera:

- TP es un verdadero positivo.
- FP es un falso positivo.
- FN es un falso negativo.
- N es el número de imágenes analizadas.
- NC es el número de clases evaluadas.

El valor mAP referencial utilizado para la validación de nuestros resultados en cada clase será de 57.9%. El valor mAP referencial es obtenido mediante el estudio: YOLO-LITE: A Real-Time Object Detection Algorithm Optimized for Non-GPU Computers (Huang, Pedoem and Chen 2018), el cual realiza un análisis de detección de objetos en dispositivos portátiles, por lo que se espera que los valores de mAP que se obtengan en cada clase de estudio por el prototipo AZATHOTH supere al valor referencial.

3.1.1 Protocolo para evaluar el algoritmo de detección

Teniendo en cuenta que la apertura focal del módulo utilizado es de $F = 2.35\text{mm}$, y basándonos en el campo de visión (horizontal y vertical) de la misma obtenemos que la distancia máxima para que el módulo pueda reconocer, un objeto es de 2 metros, por lo cual hemos implementado

pruebas a distancias menores a la indicada, todos los valores antes mencionados se encuentran en el Datasheet del módulo utilizado.

Para la ejecución del algoritmo de detección se realizó el siguiente proceso:

- Abrir el algoritmo de detección de EPP.
- Ejecutar del algoritmo.
- Colocar el objeto de prueba y sus similares al frente de la cámara.
- Tomar imágenes a distintas distancias (50 cm, 100 cm, 150cm)
- Obtención de resultados obtenidos por el algoritmo de detección.

La obtención de datos se la realizo en dos etapas:

- Pruebas individuales a cada clase (Casco, Guantes, Orejeras, Botas, Chaleco, Persona)
- Integración de todas las EPP en una persona.

Para las pruebas individuales de cada clase se introdujo 2 objetos de similar forma, pero diferente funcionalidad, se obtuvo 30 muestras para cada clase, las cuales fueron tomadas a diferentes rangos de distancia (50 cm, 100 cm, 150 cm), la posición de los objetos evaluados en cada clase fueron los siguientes:

- Vista frontal
- Vista posterior
- Lateral derecho
- Lateral izquierdo

Para la segunda etapa, 5 clases evaluadas son EPP y la restante es una persona, se procedió a integrar todas las clases un una solo imagen con el fin de evaluar todos los elementos en conjunto, se obtuvo un total de 30 muestras las cuales fueron tomadas a diferentes rangos de distancia (50 cm, 100 cm, 150 cm).

3.2 Pruebas de Campo

3.2.1 Clase 1 – Casco de seguridad

Para la comparación y tratamiento de imágenes de la primera clase se integró 2 objetos como lo muestra la Figura 3-3.



Figura 3-1 A) Casco de Seguridad, B) Casco de motocicleta, C) Gorra

Realizado por: Pucha S, Granizo J, 2021.

Todas las imágenes obtenidas fueron diferentes, en base a la Figura 3-3 se realizó las siguientes combinaciones:

- A, B y C (Lateral Derecho)
- A, B y C (Lateral Izquierdo)
- A, B y C (Vista Frontal)
- A, B y C (Vista Posterior)

Obteniendo un total de 30 muestras, 10 imágenes por cada rango de distancia (50 cm, 100 cm, 150 cm).

Tabla 1-3: Datos obtenidos del procesamiento de imágenes de la Clase 1.

| CLASE 1 – CASCO DE SEGURIDAD | | | | | | | |
|------------------------------|-------------------------|-----------------------|----------------------|----|----|-------------------------------------------------------|--------------------|
| # | Tiempo de Procesamiento | Distancia de Medición | Detección del Objeto | | | El algoritmo detecto y etiqueto solamente el objeto A | Valor de Precisión |
| | | | A | B | C | | |
| 1 | 16.50 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 2 | 16.31 s | 50 cm | SI | SI | NO | Correcto | 0.67 |
| 3 | 16.28 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 4 | 16.29 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 5 | 16.33 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 6 | 16.32 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 7 | 11.28 s | 50 cm | SI | NO | SI | Correcto | 0.67 |
| 8 | 16.30 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 9 | 11.26 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 10 | 16.30 s | 50 cm | SI | SI | SI | Correcto | 1 |

| | | | | | | | |
|----|---------|--------|----|----|----|------------|------|
| 11 | 16.32 s | 100 cm | NO | SI | NO | Incorrecto | 0.33 |
| 12 | 16.32 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 13 | 11.26 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 14 | 16.30 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 15 | 16.31 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 16 | 16.32 s | 100 cm | NO | SI | SI | Incorrecto | 0.67 |
| 17 | 16.30 s | 100 cm | SI | NO | SI | Correcto | 0.67 |
| 18 | 11.31 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 19 | 16.29 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 20 | 16.32 s | 100 cm | NO | SI | SI | Incorrecto | 0.67 |
| 21 | 11.31 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 22 | 16.29 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 23 | 16.31 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 24 | 16.31 s | 150 cm | NO | SI | SI | Incorrecto | 0.67 |
| 25 | 11.33 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 26 | 16.30 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 27 | 16.31 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 28 | 11.29 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 29 | 11.28 s | 150 cm | SI | NO | NO | Correcto | 0.33 |
| 30 | 16.31 s | 150 cm | SI | SI | SI | Correcto | 1 |

Realizado por: Pucha S, Granizo J, 2021.

La Tabla 1-3 presenta los resultados de la clase 1 (Casco de Seguridad) en sus diferentes rangos de distancia, los valores individuales de cada imagen procesada obtenidos son:

- Tiempo de procesamiento.
- Distancia de medición.
- Detección exitosa del objeto: A, B y C.
- Etiquetado exitoso solo del objeto A (Casco).
- Valor de precisión en la predicción de los objetos A, B y C en cada imagen.

De los datos obtenidos anteriormente se procedió a obtener el valor promedio del tiempo de procesamiento y se utilizó el valor de precisión de cada imagen analizada para obtener el valor de mAP de la clase 1, los resultados fueron los siguientes:

- El tiempo promedio de procesamiento para la clase 1 es de 15.894 s
- El mAP para la clase 1 es de 88.933%

En la Figura 4-3 se observa que en la imagen procesada detecta y etiqueta correctamente el objeto de estudio, en la Figura 5-3 se observa la comparación de objetos de en la cual fue detectada y etiquetada correctamente el objeto de estudio y su similar de comparación fue omitido.



Figura 4-3: Detección Clase 1.

Realizado por: Pucha S, Granizo J, 2021.



Figura 5-3: Comparación de objetos en la Clase 1.

Realizado por: Pucha S, Granizo J, 2021.

3.2.2 Clase 2 – Botas de seguridad

Para la comparación y tratamiento de imágenes de la segunda clase se integró 2 objetos como lo muestra la Figura 6-3.



Figura 6-3: A) Botas de Seguridad, B) Zapatillas, C) Zapatos

Realizado por: Pucha S, Granizo J, 2021.

Todas las imágenes obtenidas fueron diferentes, en base a la Figura 6-3 se realizó las siguientes combinaciones:

- A, B y C (Lateral Derecho)
- A, B y C (Lateral Izquierdo)
- A, B y C (Vista Frontal)
- A, B y C (Vista Posterior)

Obteniendo un total de 30 muestras, 10 imágenes por cada rango de distancia (50 cm, 100 cm, 150 cm).

Tabla 2-3: Datos obtenidos del procesamiento de imágenes de la Clase 2.

| CLASE 2 – BOTAS DE SEGURIDAD | | | | | | | |
|------------------------------|-------------------------|-----------------------|----------------------|----|----|-------------------------------------------------------|--------------------|
| # | Tiempo de Procesamiento | Distancia de Medición | Detección del Objeto | | | El algoritmo detecto y etiqueta solamente el objeto A | Valor de Precisión |
| | | | A | B | C | | |
| 1 | 16.27 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 2 | 16.29 s | 50 cm | SI | SI | No | Correcto | 0.67 |

| | | | | | | | |
|----|---------|--------|----|----|----|------------|------|
| 3 | 16.31 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 4 | 16.28 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 5 | 16.33 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 6 | 16.30 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 7 | 16.29 s | 50 cm | NO | NO | SI | Incorrecto | 0.33 |
| 8 | 16.30 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 9 | 16.32 s | 50 cm | NO | SI | SI | Incorrecto | 0.67 |
| 10 | 16.31 s | 50 cm | NO | SI | SI | Incorrecto | 0.67 |
| 11 | 16.28 s | 100 cm | SI | SI | No | Correcto | 0.67 |
| 12 | 16.32 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 13 | 16.29 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 14 | 16.31 s | 100 cm | SI | SI | NO | Correcto | 0.67 |
| 15 | 16.28 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 16 | 11.33 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 17 | 11.29 s | 100 cm | SI | No | SI | Correcto | 0.67 |
| 18 | 16.33 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 19 | 16.29 s | 100 cm | NO | SI | SI | Incorrecto | 0.67 |
| 20 | 16.30 s | 100 cm | NO | SI | SI | Incorrecto | 0.67 |
| 21 | 16.31 s | 150 cm | NO | SI | SI | Incorrecto | 0.67 |
| 22 | 16.33 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 23 | 16.28 s | 150 cm | SI | NO | SI | Correcto | 0.67 |
| 24 | 16.29 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 25 | 11.33 s | 150 cm | SI | NO | NO | Correcto | 0.33 |
| 26 | 11.28 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 27 | 11.32 s | 150 cm | NO | SI | SI | Incorrecto | 0.67 |
| 28 | 16.30 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 29 | 11.29 s | 150 cm | SI | NO | NO | Correcto | 0.33 |
| 30 | 11.28 s | 150 cm | NO | SI | SI | Incorrecto | 0.67 |

Realizado por: Pucha S, Granizo J, 2021.

La Tabla 2-3 presenta los resultados de la clase 2 (Botas de Seguridad) en sus diferentes rangos de distancia, los valores individuales de cada imagen procesada obtenidos son:

- Tiempo de procesamiento.
- Distancia de medición.
- Detección exitosa del objeto: A, B y C.
- Etiquetado exitoso solo del objeto A (Botas de Seguridad).
- Valor de precisión en la predicción de los objetos A, B y C en cada imagen.

De los datos obtenidos anteriormente se procedió a obtener el valor promedio del tiempo de procesamiento y se utilizó el valor de precisión de cada imagen analizada para obtener el valor de mAP de la clase 2, los resultados fueron los siguientes:

- El tiempo promedio de procesamiento para la clase 2 es de 15.301 s
- El mAP para la clase 2 es de 80.1%

En la Figura 7-3 se observa la correcta detección y etiquetado del objeto de estudio por medio del algoritmo de detección, en la Figura 8-3 se observa el análisis de un objeto comparativo el cual es reconocido, pero no etiquetado.



Figura 7-3: Detección exitosa de la clase 2.

Realizado por: Pucha S, Granizo J, 2021.



Figura 8-3: Análisis de objeto de comparación de la clase 2.

Realizado por: Pucha S, Granizo J, 2021.

3.2.3 Clase 3 – Chaleco reflectivo

Para la comparación y tratamiento de imágenes de la tercera clase se integró 2 objetos como lo muestra la Figura 9-3.



Figura 9-3: A) Chaleco Reflectivo, B) Camisa, C) Buso Térmico

Realizado por: Pucha S, Granizo J, 2021.

Todas las imágenes obtenidas fueron diferentes, en base a la Figura 9-3 se realizó las siguientes combinaciones:

- A, B y C (Lateral Derecho)
- A, B y C (Lateral Izquierdo)
- A, B y C (Vista Frontal)
- A, B y C (Vista Posterior)

Obteniendo un total de 30 muestras, 10 imágenes por cada rango de distancia (50 cm, 100 cm, 150 cm).

Tabla 3-1: Datos obtenidos del procesamiento de imágenes de la Clase 3.

| CLASE 3 – CHALECO REFLECTIVO | | | | | | | |
|------------------------------|-------------------------|-----------------------|----------------------|----|----|-------------------------------------------------------|--------------------|
| # | Tiempo de Procesamiento | Distancia de Medición | Detección del Objeto | | | El algoritmo detecto y etiqueto solamente el objeto A | Valor de Precisión |
| | | | A | B | C | | |
| 1 | 16.32 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 2 | 16.31 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 3 | 16.30 s | 50 cm | SI | SI | NO | Correcto | 0.67 |
| 4 | 16.31 s | 50 cm | SI | NO | NO | Correcto | 0.33 |
| 5 | 16.29 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 6 | 16.32 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 7 | 16.32 s | 50 cm | NO | NO | SI | Incorrecto | 0.33 |
| 8 | 11.27 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 9 | 16.32 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 10 | 16.31 s | 50 cm | NO | SI | SI | Incorrecto | 0.67 |
| 11 | 11.46 s | 100 cm | SI | NO | SI | Correcto | 0.67 |
| 12 | 11.26 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 13 | 11.28 s | 100 cm | NO | SI | SI | Incorrecto | 0.67 |
| 14 | 11.28 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 15 | 11.27 s | 100 cm | SI | SI | NO | Correcto | 0.67 |
| 16 | 11.27 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 17 | 16.32 s | 100 cm | SI | SI | NO | Correcto | 0.67 |
| 18 | 16.28 s | 100 cm | NO | NO | SI | Incorrecto | 0.33 |
| 19 | 11.46 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 20 | 11.26 s | 100 cm | SI | NO | SI | Correcto | 0.67 |
| 21 | 16.32 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 22 | 16.31 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 23 | 16.30 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 24 | 11.32 s | 150 cm | NO | NO | SI | Incorrecto | 0.33 |
| 25 | 16.28 s | 150 cm | SI | NO | SI | Correcto | 0.67 |
| 26 | 16.27 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 27 | 16.30 s | 150 cm | SI | SI | NO | Correcto | 0.67 |

| | | | | | | | |
|----|---------|--------|----|----|----|------------|------|
| 28 | 16.29 s | 150 cm | NO | SI | SI | Incorrecto | 0.67 |
| 29 | 11.28 s | 150 cm | NO | SI | SI | Incorrecto | 0.67 |
| 30 | 16.32 s | 150 cm | SI | SI | SI | Correcto | 1 |

Realizado por: Pucha S, Granizo J, 2021.

La Tabla 3-3 presenta los resultados de la clase 3 (Chaleco Reflectivo) en sus diferentes rangos de distancia, los valores individuales de cada imagen procesada obtenidos son:

- Tiempo de procesamiento.
- Distancia de medición.
- Detección exitosa del objeto: A, B y C.
- Etiquetado exitoso solo del objeto A (Casco).
- Valor de precisión en la predicción de los objetos A, B y C en cada imagen.

De los datos obtenidos anteriormente se procedió a obtener el valor promedio del tiempo de procesamiento y se utilizó el valor de precisión de cada imagen analizada para obtener el valor de mAP de la clase 3, los resultados fueron los siguientes:

- El tiempo promedio de procesamiento para la clase 3 es de 14.64 s
- El mAP para la clase 3 es de 79.97%

En la Figura 10-3 se observa el correcto reconocimiento y etiquetado de la clase 3 – Chaleco reflectivo, en la Figura 11-3 se observa que el algoritmo reconoció, pero no etiquetó el objeto de estudio.

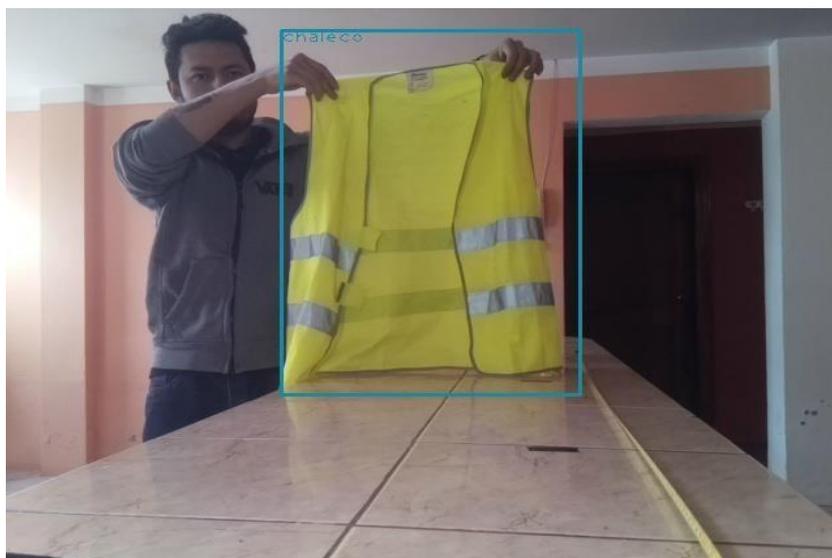


Figura 10-3: Detección y etiquetado exitoso de la clase 3.

Realizado por: Pucha S, Granizo J, 2021.



Figura 11-3: Elemento de comparación de la clase 3.

Realizado por: Pucha S, Granizo J, 2021.

3.2.4 Clase 4 – Orejeras de seguridad

Para la comparación y tratamiento de imágenes de la cuarta clase se integró 2 objetos como lo muestra la Figura 12-3.



Figura 12-3: A) Orejeras de Seguridad, B) Auriculares, C) Orejeras de frio.

Realizado por: Pucha S, Granizo J, 2021.

Todas las imágenes obtenidas fueron diferentes, en base a la Figura 12-3 se realizó las siguientes combinaciones:

- A, B y C (Lateral Derecho)
- A, B y C (Lateral Izquierdo)
- A, B y C (Vista Frontal)
- A, B y C (Vista Posterior)

Obteniendo un total de 30 muestras, 10 imágenes por cada rango de distancia (50 cm, 100 cm, 150 cm).

Tabla 4-3: Datos obtenidos del procesamiento de imágenes de la Clase 4.

| CLASE 4 – OREJERAS DE SEGURIDAD | | | | | | | |
|---------------------------------|-------------------------|-----------------------|----------------------|----|----|-------------------------------------------------------|--------------------|
| # | Tiempo de Procesamiento | Distancia de Medición | Detección del Objeto | | | El algoritmo detecto y etiqueto solamente el objeto A | Valor de Precisión |
| | | | A | B | C | | |
| 1 | 16.33 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 2 | 16.30 s | 50 cm | SI | SI | No | Correcto | 0.67 |
| 3 | 16.27 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 4 | 16.39 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 5 | 11.32 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 6 | 16.31 s | 50 cm | SI | NO | SI | Correcto | 0.67 |
| 7 | 11.29 s | 50 cm | SI | NO | SI | Correcto | 0.67 |
| 8 | 11.30 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 9 | 16.33 s | 50 cm | NO | SI | SI | Incorrecto | 0.67 |
| 10 | 16.28 s | 50 cm | SI | SI | NO | Correcto | 0.67 |
| 11 | 16.27 s | 100 cm | NO | SI | SI | Incorrecto | 0.67 |
| 12 | 16.33 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 13 | 11.29 s | 100 cm | SI | SI | NO | Correcto | 0.67 |
| 14 | 16.33 s | 100 cm | SI | NO | SI | Correcto | 0.67 |
| 15 | 11.28 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 16 | 16.33 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 17 | 16.33 s | 100 cm | NO | NO | SI | Incorrecto | 0.33 |
| 18 | 16.27 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 19 | 11.29 s | 100 cm | SI | NO | SI | Correcto | 0.67 |
| 20 | 16.32 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 21 | 16.30 s | 150 cm | NO | SI | SI | Incorrecto | 0.67 |
| 22 | 16.31 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 23 | 16.29 s | 150 cm | NO | SI | SI | Incorrecto | 0.67 |
| 24 | 16.29 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 25 | 16.33 s | 150 cm | SI | SI | SI | Correcto | 1 |

| | | | | | | | |
|----|---------|--------|----|----|----|------------|------|
| 26 | 16.28 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 27 | 16.32 s | 150 cm | NO | SI | SI | Incorrecto | 0.67 |
| 28 | 11.30 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 29 | 11.28 s | 150 cm | SI | NO | SI | Correcto | 0.67 |
| 30 | 16.28 s | 150 cm | NO | SI | NO | Incorrecto | 0.33 |

Realizado por: Pucha S, Granizo J, 2021.

La Tabla 4-3 presenta los resultados de la clase 4 (Orejeras de Seguridad) en sus diferentes rangos de distancia, los valores individuales de cada imagen procesada obtenidos son:

- Tiempo de procesamiento.
- Distancia de medición.
- Detección exitosa del objeto: A, B y C.
- Etiquetado exitoso solo del objeto A (Orejeras de Seguridad).
- Valor de precisión en la predicción de los objetos A, B y C en cada imagen.

De los datos obtenidos anteriormente se procedió a obtener el valor promedio del tiempo de procesamiento y se utilizó el valor de precisión de cada imagen analizada para obtener el valor de mAP de la clase 4, los resultados fueron los siguientes:

- El tiempo promedio de procesamiento para la clase 4 es de 14.971 s
- El mAP para la clase 4 es de 81.23%

En la Figura 13-3 y Figura 14-3 se observa que el algoritmo de detección identifico y etiqueto correctamente el objeto de estudio de la clase 4.



Figura 13-3: Detección y etiqueta de las orejetas de seguridad.

Realizado por: Pucha S, Granizo J, 2021.

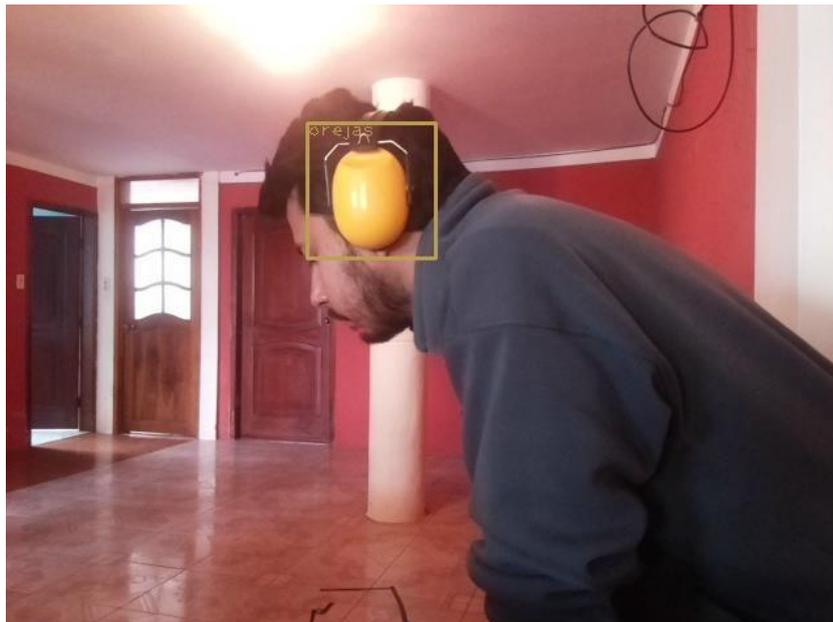


Figura 14-3: Detección de las orejetas de seguridad en una persona.

Realizado por: Pucha S, Granizo J, 2021.

3.2.5 Clase 5 – Guantes de seguridad

Para la comparación y tratamiento de imágenes de la quinta clase se integró 2 objetos como lo muestra la Figura 15-3.



Figura 15-3: A) Guantes de Seguridad, B) Guantes sin dedos, C) Manos Humanas.

Realizado por: Pucha S, Granizo J, 2021.

Todas las imágenes obtenidas fueron diferentes, en base a la Figura 15-3 se realizó las siguientes combinaciones:

- A, B y C (Lateral Derecho)
- A, B y C (Lateral Izquierdo)
- A, B y C (Vista Frontal)
- A, B y C (Vista Posterior)

Obteniendo un total de 30 muestras, 10 imágenes por cada rango de distancia (50 cm, 100 cm, 150 cm).

Tabla 5-3: Datos obtenidos del procesamiento de imágenes de la Clase 5.

| CLASE 5 – GUANTES DE SEGURIDAD | | | | | | | |
|--------------------------------|-------------------------|-----------------------|----------------------|----|----|-------------------------------------------------------|--------------------|
| # | Tiempo de Procesamiento | Distancia de Medición | Detección del Objeto | | | El algoritmo detecto y etiqueto solamente el objeto A | Valor de Precisión |
| | | | A | B | C | | |
| 1 | 11.30 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 2 | 11.27 s | 50 cm | SI | SI | No | Correcto | 0.67 |
| 3 | 16.29 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 4 | 16.30 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 5 | 16.29 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 6 | 11.30 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 7 | 16.29 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 8 | 16.33 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 9 | 16.28 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 10 | 11.31 s | 50 cm | SI | SI | SI | Correcto | 1 |
| 11 | 16.30 s | 100 cm | SI | SI | SI | Incorrecto | 1 |
| 12 | 16.32 s | 100 cm | SI | NO | SI | Correcto | 0.67 |

| | | | | | | | |
|----|---------|--------|----|----|----|------------|------|
| 13 | 16.32 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 14 | 16.31 s | 100 cm | NO | SI | SI | Correcto | 0.67 |
| 15 | 16.32 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 16 | 16.33 s | 100 cm | NO | SI | SI | Incorrecto | 0.67 |
| 17 | 16.29 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 18 | 16.31 s | 100 cm | SI | SI | NO | Correcto | 0.67 |
| 19 | 16.31 s | 100 cm | SI | SI | SI | Correcto | 1 |
| 20 | 11.30 s | 100 cm | SI | SI | SI | Incorrecto | 1 |
| 21 | 11.30 s | 150 cm | SI | SI | NO | Correcto | 0.67 |
| 22 | 11.32 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 23 | 11.32 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 24 | 16.29 s | 150 cm | SI | NO | NO | Incorrecto | 0.33 |
| 25 | 16.33 s | 150 cm | SI | NO | SI | Correcto | 0.67 |
| 26 | 16.28 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 27 | 16.32 s | 150 cm | NO | SI | SI | Correcto | 0.67 |
| 28 | 16.30 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 29 | 16.29 s | 150 cm | SI | SI | SI | Correcto | 1 |
| 30 | 11.32 s | 150 cm | SI | SI | NO | Correcto | 0.67 |

Realizado por: Pucha S, Granizo J, 2021.

La Tabla 5-3 presenta los resultados de la clase 5 (Guantes de Seguridad) en sus diferentes rangos de distancia, los valores individuales de cada imagen procesada obtenidos son:

- Tiempo de procesamiento.
- Distancia de medición.
- Detección exitosa del objeto: A, B y C.
- Etiquetado exitoso solo del objeto A (Guantes de Seguridad).
- Valor de precisión en la predicción de los objetos A, B y C en cada imagen.

De los datos obtenidos anteriormente se procedió a obtener el valor promedio del tiempo de procesamiento y se utilizó el valor de precisión de cada imagen analizada para obtener el valor de mAP de la clase 5, los resultados fueron los siguientes:

- El tiempo promedio de procesamiento para la clase 5 es de 14.805 s
- El mAP para la clase 5 es de 87.867%

En la Figura 16-3 se observa la comparación de 2 objetos de estudios en la cual el elemento de la clase 5 es reconocido y etiquetado correctamente, en la Figura 17-3 se observa la correcta detección y etiquetado del guante de seguridad.



Figura 16-3: Comparación de objetos de la clase 5.

Realizado por: Pucha S, Granizo J, 2021.



Figura 17-3: Detección y etiquetado del guante de seguridad.

Realizado por: Pucha S, Granizo J, 2021.

3.2.6 Clase 6 – Persona

Para la comparación y tratamiento de imágenes de la sexta clase se realizó la comparación entre dos objetos de estudio como lo muestra la Figura 18-3.



Figura 18-3: Elementos de comparación (Hombre/Mujer).

Realizado por: Pucha S, Granizo J, 2021.

Todas las imágenes obtenidas fueron diferentes, en base a la Figura 18-3 se realizó las siguientes combinaciones:

- A y B (Lateral Derecho)
- A y B (Lateral Izquierdo)
- A y B (Vista Frontal)
- A y B (Vista Posterior)

Obteniendo un total de 30 muestras, 10 imágenes por cada rango de distancia (50 cm, 100 cm, 150 cm).

Tabla 6-3: Datos obtenidos del procesamiento de imágenes de la Clase 6.

| CLASE 6 – PERSONA | | | | | | |
|-------------------|-------------------------|-----------------------|----------------------|----|-----------------------------------------------------------------|--------------------|
| # | Tiempo de Procesamiento | Distancia de Medición | Detección del Objeto | | El algoritmo detecto y etiqueto solo el objeto (A o B), (A y B) | Valor de Precisión |
| | | | A | B | | |
| 1 | 16.30 s | 50 cm | SI | SI | Correcto | 1 |
| 2 | 16.32 s | 50 cm | SI | NO | Correcto | 0.5 |

| | | | | | | |
|----|---------|--------|----|----|----------|-----|
| 3 | 16.32 s | 50 cm | SI | SI | Correcto | 1 |
| 4 | 16.31 s | 50 cm | NO | SI | Correcto | 0.5 |
| 5 | 16.32 s | 50 cm | SI | SI | Correcto | 1 |
| 6 | 16.33 s | 50 cm | SI | SI | Correcto | 1 |
| 7 | 16.29 s | 50 cm | SI | SI | Correcto | 1 |
| 8 | 16.31 s | 50 cm | SI | SI | Correcto | 1 |
| 9 | 16.31 s | 50 cm | SI | SI | Correcto | 1 |
| 10 | 11.30 s | 50 cm | SI | SI | Correcto | 1 |
| 11 | 11.29 s | 100 cm | SI | SI | Correcto | 1 |
| 12 | 16.32 s | 100 cm | SI | SI | Correcto | 1 |
| 13 | 11.27 s | 100 cm | SI | SI | Correcto | 1 |
| 14 | 16.30 s | 100 cm | SI | SI | Correcto | 1 |
| 15 | 11.28 s | 100 cm | SI | SI | Correcto | 1 |
| 16 | 16.28 s | 100 cm | SI | SI | Correcto | 1 |
| 17 | 16.29 s | 100 cm | SI | SI | Correcto | 1 |
| 18 | 11.31 s | 100 cm | SI | SI | Correcto | 1 |
| 19 | 11.27 s | 100 cm | SI | SI | Correcto | 1 |
| 20 | 16.31 s | 100 cm | NO | SI | Correcto | 0.5 |
| 21 | 16.29 s | 150 cm | SI | SI | Correcto | 1 |
| 22 | 16.32 s | 150 cm | SI | SI | Correcto | 1 |
| 23 | 16.28 s | 150 cm | SI | SI | Correcto | 1 |
| 24 | 16.30 s | 150 cm | SI | SI | Correcto | 1 |
| 25 | 16.31 s | 150 cm | SI | NO | Correcto | 0.5 |
| 26 | 11.28 s | 150 cm | SI | SI | Correcto | 1 |
| 27 | 16.31 s | 150 cm | NO | SI | Correcto | 0.5 |
| 28 | 11.30 s | 150 cm | SI | SI | Correcto | 1 |
| 29 | 11.27 s | 150 cm | SI | SI | Correcto | 1 |
| 30 | 16.32 s | 150 cm | SI | SI | Correcto | 1 |

Realizado por: Pucha S, Granizo J, 2021.

La Tabla 6-3 presenta los resultados de la clase 6 (Persona) en sus diferentes rangos de distancia, los valores individuales de cada imagen procesada obtenidos son:

- Tiempo de procesamiento.
- Distancia de medición.
- Detección exitosa del objeto: A y B.
- Etiquetado exitoso del objeto A y B.
- Valor de precisión en la predicción de los objetos A y B en cada imagen.

De los datos obtenidos anteriormente se procedió a obtener el valor promedio del tiempo de procesamiento y se utilizó el valor de precisión de cada imagen analizada para obtener el valor de mAP de la clase 6, los resultados fueron los siguientes:

- El tiempo promedio de procesamiento para la clase 6 es de 14.800 s
- El mAP para la clase 6 es de 91.67%

En la Figura 19-3 y Figura 20-3 se observa la detección y correcto etiquetado de la clase 6.



Figura 19-3: Detección individual de la clase 6.

Realizado por: Pucha S, Granizo J, 2021.



Figura 20-3: Detección grupal de la clase 6.

Realizado por: Pucha S, Granizo J, 2021.

3.2.7 Integración de clases

En la etapa 2 se procedió a integrar todas las clases en un solo análisis, obteniendo 30 imágenes diferentes, a diferentes rangos de distancia, las combinaciones realizadas para la obtención de información fue la siguiente:

- Lateral Derecho
- Lateral Izquierdo
- Vista Frontal
- Vista Posterior

Tabla 7-3: Resultados al integrar todas las clases.

| Integración de clases | | | | | | | | | | |
|-----------------------|-------------------------|-----------------------|-----------------------|----|----|----|----|----|--------------------------------------------------------------|--------------------|
| # | Tiempo de Procesamiento | Distancia de Medición | Detección de la Clase | | | | | | El algoritmo detecto y etiqueto a la clase 1, 2, 3, 4, 5 y 6 | Valor de Precisión |
| | | | 1 | 2 | 3 | 4 | 5 | 6 | | |
| 1 | 16.33 s | 50 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 2 | 16.28 s | 50 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 3 | 16.30 s | 50 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 4 | 16.31 s | 50 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |

| | | | | | | | | | | |
|----|---------|--------|----|----|----|----|----|----|------------|------|
| 5 | 16.30 s | 50 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 6 | 16.32 s | 50 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 7 | 16.27 s | 50 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 8 | 11.29 s | 50 cm | SI | SI | SI | NO | SI | SI | Incorrecto | 0.83 |
| 9 | 11.28 s | 50 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 10 | 11.30 s | 50 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 11 | 16.30 s | 100 cm | SI | SI | SI | SI | SI | NO | Incorrecto | 0.83 |
| 12 | 11.30 s | 100 cm | SI | NO | SI | SI | SI | SI | Incorrecto | 0.83 |
| 13 | 16.31 s | 100 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 14 | 16.28 s | 100 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 15 | 16.28 s | 100 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 16 | 16.33 s | 100 cm | NO | SI | SI | SI | SI | SI | Incorrecto | 0.83 |
| 17 | 16.29 s | 100 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 18 | 11.30 s | 100 cm | SI | NO | SI | NO | NO | SI | Incorrecto | 0.5 |
| 19 | 16.30 s | 100 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 20 | 11.28 s | 100 cm | SI | NO | SI | SI | SI | SI | Incorrecto | 0.83 |
| 21 | 11.27 s | 150 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 22 | 16.30 s | 150 cm | SI | SI | SI | NO | SI | SI | Incorrecto | 0.83 |
| 23 | 16.28 s | 150 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 24 | 16.32 s | 150 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 25 | 11.30 s | 150 cm | SI | SI | NO | SI | SI | SI | Incorrecto | 0.83 |
| 26 | 16.30 s | 150 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 27 | 11.27 s | 150 cm | SI | NO | SI | SI | SI | SI | Incorrecto | 0.83 |
| 28 | 11.30 s | 150 cm | SI | SI | SI | SI | SI | SI | Correcto | 1 |
| 29 | 16.31 s | 150 cm | NO | SI | SI | SI | NO | SI | Incorrecto | 0.67 |
| 30 | 11.27 s | 150 cm | SI | SI | SI | NO | SI | NO | Incorrecto | 0.67 |

Realizado por: Pucha S, Granizo J, 2021.

En la Figura 21-3 se observa la integración de las clases: 1, 2, 3, 4, 5 y 6, en conjunto las cuales son detectadas y etiquetadas correctamente.

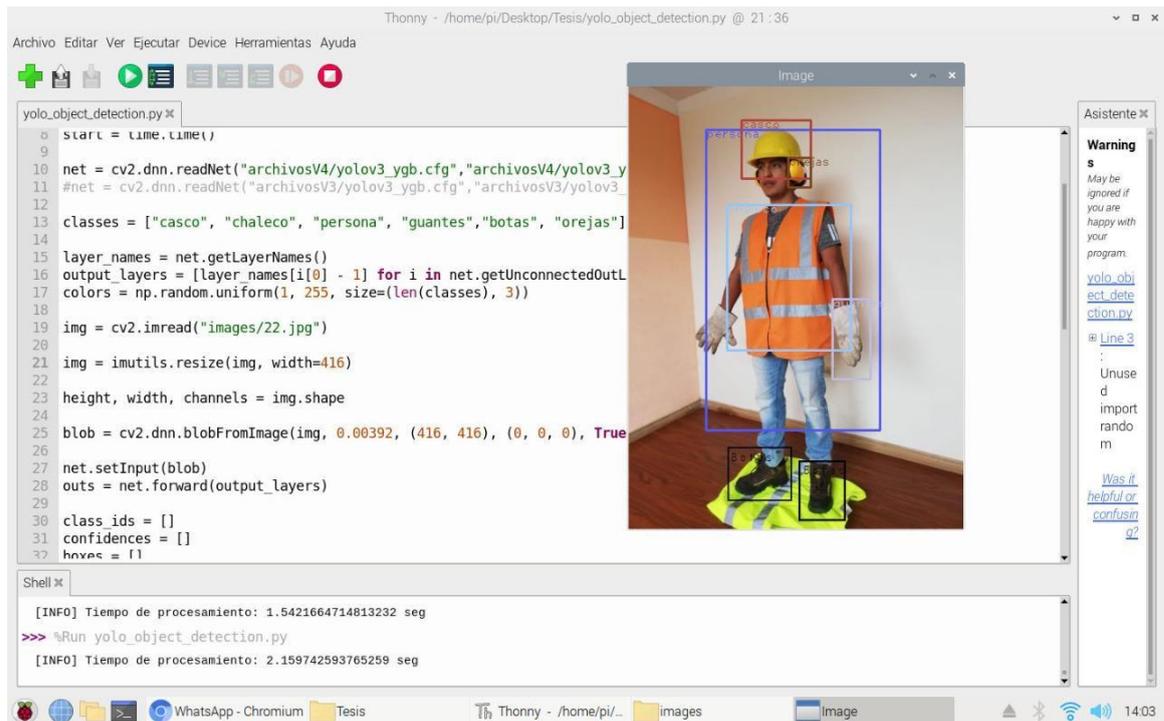


Figura 21-3: Detección y etiquetado correcto de todas las clases.

Realizado por: Pucha S, Granizo J, 2021.

3.3 Resultados

En la Tabla 8-3 se evidencia los resultados obtenidos en:

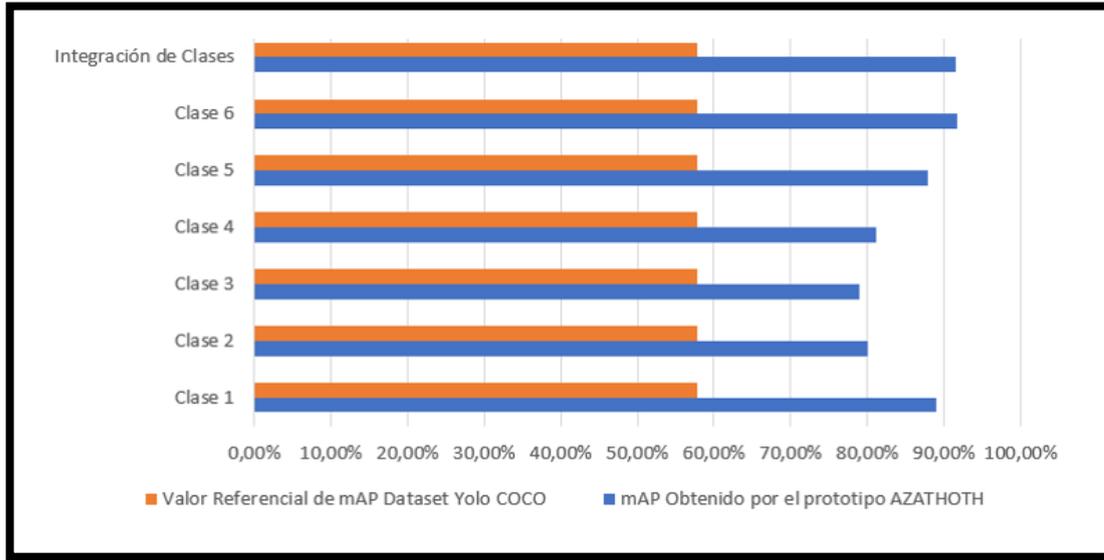
- Pruebas individuales de cada clase.
- Pruebas en la integración de clases.
- El porcentaje de mAP del Dataset Yolo COCO

Tabla 8-3: Resultados (Prototipo AZATHOTH vs Dataset Yolo COCO)

| | mAP Obtenido por el prototipo AZATHOTH | Valor Referencial de mAP Dataset Yolo COCO |
|-----------------------|----------------------------------------|--------------------------------------------|
| Clase 1 | 88.93 % | 57.9% |
| Clase 2 | 80.1 % | 57.9% |
| Clase 3 | 78.97 % | 57.9% |
| Clase 4 | 81.23 % | 57.9% |
| Clase 5 | 87.87 % | 57.9% |
| Clase 6 | 91.67 % | 57.9% |
| Integración de Clases | 91.6 % | 57.9% |

Realizado por: Pucha S, Granizo J, 2021.

Como se observa en la Tabla 8-3 los resultados obtenidos por el prototipo AZATHOTH en valores de mAP son superiores o los valores referenciales de mAP del Dataset Yolo COCO, resultando un prototipo con un alto porcentaje de mAP lo cual nos ayuda a cumplir eficazmente con los objetivos y requerimientos propuestos para el prototipo.



Gráfica 1-3: Resultados del prototipo AZATHOTH

Realizado por: Pucha S, Granizo J, 2021.

La Gráfica 3-1 muestra los valores porcentuales del mAP del dataset Yolo COCO vs los resultados obtenidos mediante la red neuronal entrenada en el prototipo AZATHOTH.

CAPÍTULO IV

4. EVALUACIÓN ECONÓMICA

Los elementos utilizados para la implementación del prototipo AZATHOTH se detallan en la Tabla 1-4, los precios establecidos son medidos en dólares americanos, además que incluyen sus debidos impuestos.

Tabla 1-4: Análisis económico del prototipo AZATHOTH

| Cantidad | Elemento | Precio unitario | Precio total |
|----------|-----------------------------|-----------------|--------------|
| 1 | LCD Rpi de 3.5 Pulgadas (A) | \$70 | \$70 |

| | | | |
|---|--------------------------------------|-------|-------|
| 1 | Raspberry Pi Cámara V2 | \$50 | \$50 |
| 1 | Raspberry PI 4 Modelo B Kit Completo | \$190 | \$190 |
| 1 | Elementos de la estructura | \$20 | \$20 |
| 1 | Elementos electrónicos | \$60 | \$60 |
| 1 | Altavoz Genius | \$15 | \$15 |
| | Mano de obra de Ingeniería | \$300 | \$300 |
| | | TOTAL | \$705 |

Realizado por: Pucha S, Granizo J, 2021.

Se utilizó como elemento de comparación en precios y características de funcionamiento un prototipo de reconocimiento de objetos mediante visión artificial.

El prototipo P1x-Series de la empresa AIS (VISION SYSTEMS) posee funciones de visión artificial avanzadas aplicadas a la industria y a dispositivos autónomos, tiene un valor en el mercado internacional de \$1350 su Datasheet se encuentra en el Anexo D, la Tabla 2-4 presenta las principales características de los prototipos comparados.

Tabla 2-4: Características (AIS / AZATHOTH)

| Característica | P1x-Series | AZATHOTH |
|--------------------------|------------------------------------|--------------------------------------------------------------------|
| Resolución en imagen | VGA (640×480) - 1.3 MP (1280×1024) | Sensor Sony IMX219 de 8 MP |
| Procesador | 660 MHz DPS | Broadcom BCM2711, quad-core Cortex-A72 (ARM v8) 64-bit SoC, 1.5GHz |
| Conexión | Serial y Ethernet | Serial y Ethernet |
| Distancias de trabajo | 70mm – 600mm | Hasta los 1500mm |
| Lentes e iluminadores | Intercambiables | Intercambiables |
| Iluminación | 9 niveles de iluminación | Trabaja bajo un ambiente de iluminación controlado |
| Software | IMPACT LITE | Software libre |
| Grado de Protección | IP67 | Ninguna |
| FPS | 120 fps (VGA) – 60 fps (1.3MP) | 0.09 fps |
| Almacenamiento | 256MB | 8GB |
| Alimentación | 10 – 30 Vdc | 12v - 3 A |
| Dimensiones | 95 x 54 x 43 mm | 61.98 x 122 x 30.9 mm |
| Material de la caja | Aluminio | Filamento PLA |
| Temperatura de operación | 0 – 50° | 0 – 50° |
| Certificación | CE, CSA | Ninguna |
| Precio | \$1350 | \$705 |

Realizado por: Pucha S, Granizo J, 2021.

Para recompensar la falta de fps el prototipo AZATHOTH realiza la detección de clases únicamente cuando detecta movimiento en el área asignada, y está diseñado para actuar óptimamente para permitir el ingreso a zonas designadas.

El costo final de prototipo AZATHOTH es de \$705 que representa el 52.22% una vez comparado con el precio de mercado del prototipo P1x-Series de la empresa AIS (VISION SYSTEMS), demostrado así que el prototipo AZATHOTH supera en varios aspectos al elemento de comparación y es de bajo costo.

CONCLUSIONES

Los sistemas de monitoreo y alerta mediante visión artificial enfocados en el área industrial son una tecnología emergente y que se encuentra en desarrollo, con un objetivo claro, el de ayudar a reducir de manera significativa los accidentes laborales presentes en la industria. En la actualidad se puede evidenciar varios sistemas que ayudan en la detección de personas u objetos (EPPS) en el área industrial que sirven para un control más riguroso de las mismas, presentes en un área determinada de la industria.

Para el cumplimiento de los requerimientos previamente establecidos para el prototipo AZATHOTH, se realizó un estudio a fondo para encontrar los elementos de Hardware y Software necesarios para su correcto funcionamiento y accesibilidad económica, enfocándonos en una relación calidad-precio lo suficientemente buena para que el prototipo pueda competir con sus semejantes.

La estructura que encapsula a los elementos que conforman al prototipo AZATHOTH fue diseñada mediante el Software SolidWorks, además la adquisición de datos se realiza mediante la Raspberry Pi Cámara V2 de 8Mpx, para su posterior procesamiento en una Raspberry PI 4 Modelo B de 8Gb mediante Yolo V4 Tiny programado en Python.

El prototipo AZATHOTH es capaz de reconocer hasta 6 tipos de objetos, entre los cuales se encuentran personas y EPPS (Casco, Chaleco, Guantes, Orejeras y Botas), gracias al uso de la red neuronal convolucional Yolo V4 Tiny, que permite un mejor desempeño del prototipo, el mismo que puede alertar del mal uso de las EPPS mediante su pantalla o sus altavoces.

Al realizar las pruebas de efectividad del algoritmo de detección de EPPS AZATHOTH se logró evidenciar los siguientes resultados: un porcentaje de precisión media (mAP) de 88.93% en la Clase1 (Casco), 80.01% Clase 2 (Botas de Seguridad), 78.97 Clase 3 (Chaleco Reflectivo), 81.23% Clase 4 (Orejeras), 87.87% Clase 5(Guante de Seguridad), 91.67% Clase 6 (Personas) y un 91.6% en la Integración de Clases, en comparación al Dataset Yolo COCO con un mAP de 57.9% de efectividad, el prototipo AZATHOTH es capaz de detectar con gran precisión y a un bajo costo las clases previamente definidas.

RECOMENDACIONES

Para un rendimiento óptimo de la red neuronal aplicada al prototipo AZATHOTH, esta debe ser implementada en un equipo de cómputo con una alta capacidad de procesamiento gráfico, elevando así el valor económico del prototipo, pero también su efectividad y fidelidad.

Utilizar un sistema embebido diferente al utilizado en el prototipo AZATHOTH, que conste con mayor capacidad de GPU para dotar así de mayor rapidez en procesamiento de imágenes, una de dichas alternativas es la Nvidia Jetson Nano, sistema embebido enfocado en el procesamiento de imágenes.

Para tener una mejor tasa de FPS en el mismo sistema embebido (Raspberry PI 4 Modelo B), se puede aumentar su capacidad de GPU mediante la implementación del Neural Computer Stick de INTEL, aumentando también el valor económico del prototipo.

Implementar una pantalla de dimensiones mayores a la usada en el prototipo (LCD Rpi de 3.5 Pulgadas (A)), para así poder visualizar mejor las indicaciones y alertas emitidas en ella.

Para el correcto funcionamiento del dispositivo AZATHOTH se debe tener en cuenta una adecuada iluminación en el área en la que va a trabajar el prototipo ya que este es sensible a cambios de iluminación.

Adaptar la red neuronal del prototipo AZATHOTH para la detección de diferentes EPPS (adicionales a las analizadas) para la diversificación del dispositivo en diferentes industrias, además de implementar un algoritmo extra para la diferenciación de personas, cubriendo así un mayor número de necesidades que se pueden presentar en el ámbito laboral.

BIBLIOGRAFIA

ABOUTALIB, S; & VELOSO, M. Towards using multiple cues for robust object recognition. *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*. 2007, pp. 1–8.

ABREGO, M; et al. *Equipos de protección personal*. S.l.: ACHS. 2000.

AGUINAGA BARRAGÁN, A. Prototipo para el control de calidad de productos industriales por visión artificial. 2007. S.l.: CIBIM.

ALEGRE, E., PAJARES, G. & DE LA ESCALERA, A. “Conceptos y métodos en visión por computador”. *España: Grupo de Visión del Comité Español de Automática (CEA)*, 2016.

ARÉVALO, V., GONZÁLEZ, J. and AMBROSIO, G. “La Librería de Visión Artificial OpenCV. Aplicación a la Docencia e Investigación”. *Base Informática*, vol. 40, 2004, pp 61–66.

BEDOYA BEDOYA, N., & MECATRÓNICO, I. 2007. Diseño de dispositivo monitoreo para supervisar operación de un montacarga. *instname:Universidad Autónoma de Occidente* [en línea]. 2007. [Consulta: 23 November 2020]. Disponible en: <http://hdl.handle.net/10614/6336>.

BISONG, E. “Google Colaboratory”, Building Machine Learning and Deep Learning Models on Google Cloud Platform: A Comprehensive Guide for Beginners [en línea], 2019, Berkeley, CA: Apress, pp. 59–64. ISBN 978-1-4842-4470-8. Disponible en: https://doi.org/10.1007/978-1-4842-4470-8_7.

CARNEIRO, T., NÓBREGA, R.V.M. da, NEPOMUCENO, T., BIAN, G., ALBUQUERQUE, V.H.C. de & FILHO, P.P.R. “Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications”. *IEEE Access*, vol. 6, 2018, pp. 61677–61685. ISSN 2169-3536. DOI 10.1109/ACCESS.2018.2874767.

CAVASSA, C.R. *Seguridad industrial: un enfoque integral*. S.l.: Editorial Limusa, 1991.

CHAMOCHUMBI BARRUETO, C.M. Seguridad e higiene industrial. S.l.: Universidad Inca Garcilaso de la Vega. ISBN 6124050633. 2014.

CHOROWSKI, J.K., BAHDANAU, D., SERDYUK, D., CHO, K. & BENGIO, Y.

“Attention-based models for speech recognition”. *Advances in neural information processing systems*. 2015. S.l.: s.n., pp. 577–585.

CHOU, Y.-F., LAI, J.-S. & KUO, H.-W. “Effects of shift work on noise-induced hearing loss”.

Noise and Health, vol. 11, n° 45 (2009), pp. 185. ISSN 1463-1741.

DESHPANDE, H., SINGH, A. & HERUNDE, H. “Comparative analysis on YOLO object detection with OpenCV”. *International Journal of Research in Industrial Engineering* [en línea],

2020, vol. 9, no. 1, pp. 46–64. ISSN 2717-2937. DOI 10.22105/riej.2020.226863.1130.

Disponible en: http://www.riejournal.com/article_106905.html.

DOMÍNGUEZ, C.G. *Aplicaciones orientadas a la domótica con Raspberry Pi*. Sevilla: Universidad de Sevilla, 2015.

ESTARITA, J., JIM, A., BROCHERO, J., ESCOBAR, H. & MORENO, S. “Sistema de

Reconocimiento de objetos en tiempo real”. *Investigación y desarrollo en TIC*, vol. 8, n° 2 (2017), pp. 41–45. ISSN 2216-1570.

FANG, Q., LI, H., LUO, X., DING, L., LUO, H., ROSE, T.M. & AN, W. “Detecting non-hardhat-use by a deep learning method from far-field surveillance videos”. *Automation in*

Construction [en línea], 2018, vol. 85, pp. 1–9. ISSN 0926-5805. DOI <https://doi.org/10.1016/j.autcon.2017.09.018>.

Disponible en: <https://www.sciencedirect.com/science/article/pii/S0926580517304429>.

GARCÍA, I. “Visión artificial y procesamiento digital de imágenes”. *REVISTA CIENTÍFICA*

AXIOMA, vol. 1, n° 3 (2007), pp. 40–41. ISSN 2550-6684.

GIL PABLO & TORRES, F.A.N.D.O.Z.F.G. *Detección de objetos por segmentación*

multinivel combinada de espacios de color, 2004.

GIRSHICK, R., DONAHUE, J., DARRELL, T. & MALIK, J. “Rich feature hierarchies for accurate object detection and semantic segmentation”. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014. pp. 580–587.

GIRSHICK, R.B. “Fast R-CNN”. CoRR [en línea], 2015, vol. abs/1504.08083. Disponible en: <http://arxiv.org/abs/1504.08083>.

GUO, H., YU, Y. & SKITMORE, M. “Visualization technology-based construction safety management: A review”. *Automation in Construction* [en línea], 2017, vol. 73, pp. 135–144. ISSN 0926-5805. DOI <https://doi.org/10.1016/j.autcon.2016.10.004>. Disponible en: <http://www.sciencedirect.com/science/article/pii/S092658051630293X>.

GUTIÉRREZ, Á.P.H. *Python paso a paso*. S.l.: Grupo Editorial RA-MA, 2016.

HERRICK, R.F. *DE LA PROTECCION PERSONAL. Enciclopedia de salud y seguridad en el trabajo*, 2001.

HUANG, R., PEDOEEM, J. & CHEN, C. “YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers”. *2018 IEEE International Conference on Big Data (Big Data)*. S.l.: IEEE, pp. 2503–2510. ISBN 1538650355.

JIANG, Z., ZHAO, L., LI, S. & JIA, Y. “Real-time object detection method based on improved YOLOv4-tiny”. [en línea], 2020, [Consulta: 8 February 2021]. Disponible en: <http://arxiv.org/abs/2011.04244>.

KIM, K.S. “Occupational hearing loss in Korea”. *Journal of Korean medical science*, 2010, vol. 25, no. Suppl, pp. S62–S69. ISSN 1011-8934.

LI, X., SRIDHARAN, M. & MEADOR, C. “Learning visual object models on a robot using context and appearance cues”. *AAMAS*. 2013. S.l.: Citeseer, pp. 1167–1168.

MARTÍNEZ, M.G., GARCÍA, J.J.J., CEBALLOS, L.Y., VALENCIA, A.M., ZAPATA, M.A.V. & TRESPALACIOS, E.M.V. “Ruido industrial: efectos en la salud de los trabajadores expuestos”. *Revista CES Salud Pública*, vol. 3, n° 2 (2012), pp. 174–183. ISSN 2145-9932.

MASSIRIS, M., DELRIEUX, C. & FERNÁNDEZ MUÑOZ, J.Á. Detección de equipos de protección personal mediante red neuronal convolucional YOLO. *XXXIX Jornadas de Automática*. S.l.: Área de Ingeniería de Sistemas y Automática, Universidad de Extremadura. 2018. pp. 1022–1029. ISBN 8409044609.

MNEYMNEH, B.E., ABBAS, M. & KHOURY, H. “Automated Hardhat Detection for Construction Safety Applications”. *Procedia Engineering* [en línea], 2017, vol. 196, pp. 895–902. ISSN 1877-7058. DOI <https://doi.org/10.1016/j.proeng.2017.08.022>. Disponible en: <http://www.sciencedirect.com/science/article/pii/S1877705817331430>.

MOORE, D. & RID, T. “Cryptopolitik and the Darknet”. *Survival* [en línea], 2016, vol. 58, nº 1, pp. 7–38. DOI 10.1080/00396338.2016.1142085. Disponible en: <https://doi.org/10.1080/00396338.2016.1142085>.

MURRELL, K. *Ergonomics: Man in his working environment*. S.l.: Springer Science & Business Media, 2012. ISBN 9400958781.

NOPE, S.E., LOAIZA, H. & CAICEDO, E. “Estudio Comparativo de Técnicas para el Reconocimiento de gestos por Visión Artificial”. *Avances en Sistemas e Informática*, vol. 5, nº 3 (2008), pp. 127–134.

OSORES, N. Investigación de la vibración de cuerpo entero en una planta industrial: una comprobación acerca de la comodidad y la salud. 2017.

PACHECO, M.A. Identificación de sistemas no lineales con redes neuronales convolucionales. Ciudad de México: CENTRO DE INVESTIGACION Y DE ESTUDIOS AVANZADOS, 2017.

PHEASANT, S. “Introduction Ergonomics”. *Work and Health* [en línea], 1991, London: Macmillan Education UK, pp. 3–23. ISBN 978-1-349-21671-0. Disponible en: https://doi.org/10.1007/978-1-349-21671-0_2.

REDMON, J., DIVVALA, S., GIRSHICK, R. & FARHADI, A. “You Only Look Once: Unified, Real-Time Object Detection”. *IEEE Conference on Computer Vision and Pattern*

Recognition (CVPR). 2016. S.l.: s.n., pp. 779–788. ISBN 1063-6919. DOI 10.1109/CVPR.2016.91.

REDMON, J. & FARHADI, A. “YOLO9000: Better, Faster, Stronger”. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. S.l.: s.n., pp. 6517–6525. ISBN 1063-6919. DOI 10.1109/CVPR.2017.690.

REDMON, J. & FARHADI, A. “Yolov3: An incremental improvement”. *arXiv preprint arXiv:1804.02767*, 2018.

ROBLEDO, F.H. *Riesgos físicos II: iluminación*. S.l.: Ecoe Ediciones, 2014, ISBN 9587711033.

SÁNCHEZ, J.A.S. *Avances en robótica y visión por computador*. S.l.: Univ de Castilla La Mancha, 2002, ISBN 8484271994.

SMITH, A.R. “Color Gamut Transform Pairs”. *SIGGRAPH Comput. Graph.* [en línea], vol. 12, n° 3 (1978), pp. 12–19. ISSN 0097-8930. DOI 10.1145/965139.807361. Disponible en: <https://doi.org/10.1145/965139.807361>.

SPECTOR, P.E., VÁZQUEZ, A.A. & DE LA PENA, M.F. *Psicología industrial y organizacional: investigación y práctica*. S.l.: Editorial El Manual Moderno, 2002, ISBN 9684269757.

TEIZER, J., CHENG, T. & FANG, Y. “Location tracking and data visualization technology to advance construction ironworkers’ education and training in safety and productivity”. *Automation in Construction* [en línea], 2013, vol. 35, pp. 53–68. ISSN 0926-5805. DOI <https://doi.org/10.1016/j.autcon.2013.03.004>. Disponible en: <http://www.sciencedirect.com/science/article/pii/S0926580513000320>.

TIAN, Y., YANG, G., WANG, Z., WANG, H., LI, E. & LIANG, Z. “Apple detection during different growth stages in orchards using the improved YOLO-V3 model”. *Computers and electronics in agriculture*, vol. 157, 2019, pp. 417–426. ISSN 0168-1699.

TORRES, A.D. “Procesamiento digital de imágenes”. *Perfiles Educativos*, nº 72 (1996). ISSN 0185-2698.

TRUJILLO, M.G.S. & VARGAS, M.D.L.E.G. “Satisfacción Laboral en los Entornos de Trabajo. Una exploración cualitativa para su estudio”. *Scientia et technica*, vol. 22, nº 2 (2017), pp. 161–166. ISSN 0122-1701.

VALDIVIEZO GUZMÁN, L.A. *Estudio de la vibración producida por maquinaria minera en la salud de los trabajadores en la Unidad Minera Breapampa*, 2014.

VAN ROSSUM, G. & DRAKE JR, F.L. *Python tutorial*. S.l.: Centrum voor Wiskunde en Informatica Amsterdam, 1995.

WANG, C.-Y., BOCHKOVSKIY, A. & LIAO, H.-Y.M. “Scaled-YOLOv4: Scaling Cross Stage Partial Network”. *arXiv preprint arXiv:2011.08036*, 2020.

YU, Y., GUO, H., DING, Q., LI, H. & SKITMORE, M. “An experimental study of real-time identification of construction workers’ unsafe behaviors”. *Automation in Construction* [en línea], 2017, vol. 82, pp. 193–206. ISSN 0926-5805. DOI <https://doi.org/10.1016/j.autcon.2017.05.002>. Disponible en: <http://www.sciencedirect.com/science/article/pii/S0926580517304223>.

ZORRILLA, S.P. *Elementos de Protección Personal*. 2012. S.l.: Córdoba.

ANEXOS

ANEXO A: Especificaciones Raspberry Pi4

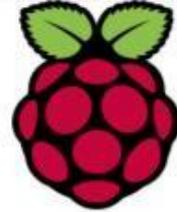
Specification

| | |
|-----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Processor: | Broadcom BCM2711, quad-core Cortex-A72 (ARM v8) 64-bit SoC @ 1.5GHz |
| Memory: | 1GB, 2GB, 4GB or 8GB LPDDR4 (depending on model) with on-die ECC |
| Connectivity: | 2.4 GHz and 5.0 GHz IEEE 802.11b/g/n/ac wireless LAN, Bluetooth 5.0, BLE Gigabit Ethernet 2 × USB 3.0 ports 2 × USB 2.0 ports. |
| GPIO: | Standard 40-pin GPIO header (fully backwards-compatible with previous boards) |
| Video & sound: | 2 × micro HDMI ports (up to 4Kp60 supported) 2-lane MIPI DSI display port 2-lane MIPI CSI camera port 4-pole stereo audio and composite video port |
| Multimedia: | H.265 (4Kp60 decode); H.264 (1080p60 decode, 1080p30 encode); OpenGL ES, 3.0 graphics |
| SD card support: | Micro SD card slot for loading operating system and data storage |
| Input power: | 5V DC via USB-C connector (minimum 3A) 5V DC via GPIO header (minimum 3A) Power over Ethernet (PoE)-enabled (requires separate PoE HAT) |
| Environment: | Operating temperature 0–50°C |
| Compliance: | For a full list of local and regional product approvals, please visit https://www.raspberrypi.org/documentation/hardware/raspberrypi/conformity.md |
| Production lifetime: | The Raspberry Pi 4 Model B will remain in production until at least January 2026. |

ANEXO B: Características Raspberry Pi Camera v2

Raspberry Pi Camera v2

Part number: RPI 8MP CAMERA BOARD



- 8 megapixel camera capable of taking photographs of 3280 x 2464 pixels
- Capture video at 1080p30, 720p60 and 640x480p90 resolutions
- All software is supported within the latest version of Raspbian Operating System

The Camera v2 is the new official camera board released by the Raspberry Pi foundation.

The Raspberry Pi Camera Module v2 is a high quality 8 megapixel Sony IMX219 image sensor custom designed add-on board for Raspberry Pi, featuring a fixed focus lens. It's capable of 3280 x 2464 pixel static images, and also supports 1080p30, 720p60 and 640x480p60/90 video. It attaches to Pi by way of one of the small sockets on the board upper surface and uses the dedicated CSI interface, designed especially for interfacing to cameras.

- 8 megapixel native resolution sensor-capable of 3280 x 2464 pixel static images
- Supports 1080p30, 720p60 and 640x480p90 video
- Camera is supported in the latest version of Raspbian, Raspberry Pi's preferred operating system

The board itself is tiny, at around 25mm x 23mm x 9mm. It also weighs just over 3g, making it perfect for mobile or other applications where size and weight are important. It connects to Raspberry Pi by way of a short ribbon cable.

The high quality Sony IMX219 image sensor itself has a native resolution of 8 megapixel, and has a fixed focus lens on-board. In terms of still images, the camera is capable of 3280 x 2464 pixel static images, and also supports 1080p30, 720p60 and 640x480p90 video.

Applications

- CCTV security camera
- motion detection
- time lapse photography



3.5inch TFT Touch Shield User Manual

OVERVIEW

This product is 3.5-inch resistive screen module with a resolution 320x480. It has an internal controller allowing to draw on the display geometric figures, text and pictures by using correspondent functions available via SPI communication interface.

STM32 and Arduino examples are provided for easy porting.

PARAMETERS

LCD Type: TFT

LCD Interface: SPI

LCD Controller: ILI9486

Touch Screen Controller: XPT2046

Touch Screen Type: Resistive

Resolution: 320x480 (Pixel)

Color Gradation Exponent: 65536

INTERFACE

| Symbol | Arduino PIN | STM32 PIN | Description |
|---------|-------------|-----------|-----------------|
| 5V | 5V | 5V | 5V power input |
| GND | GND | GND | Ground |
| SCLK | D13 | PA5 | SPI clock |
| MISO | D12 | PA6 | SPI data input |
| MOSI | D11 | PA7 | SPI data output |
| LCD_CS | D10 | PB6 | LCD chip select |
| LCD_BL | D9 | PC7 | LCD back light |
| LCD_RST | D8 | PA9 | LCD reset |

ANEXO D: Características prototipo P1x-Series

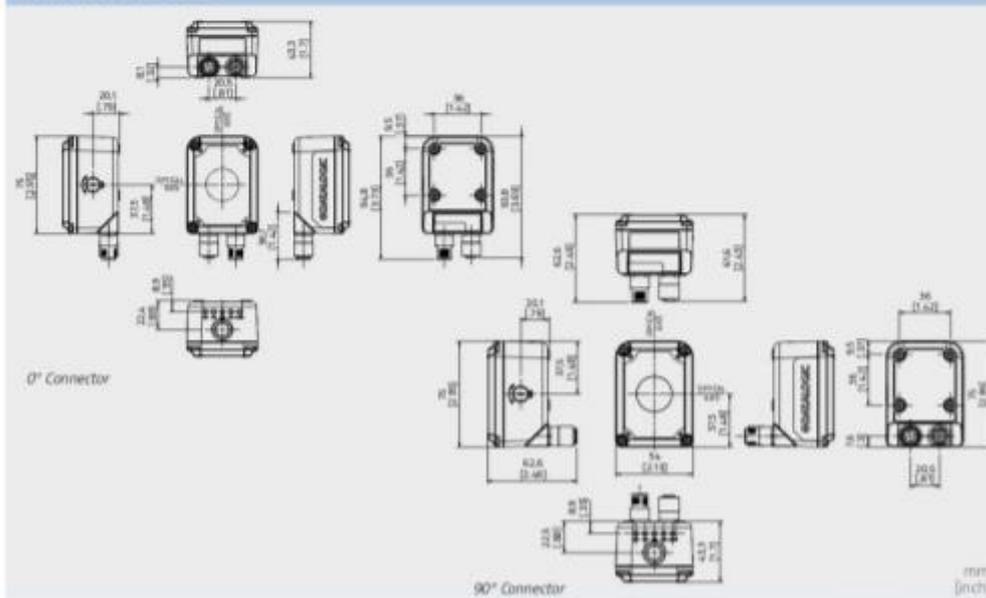
P-Series



DATOS TÉCNICOS

| | P10 | P11 | P15 | P16 |
|-----------------------|---------------------------------------------------------------------------------|---------------------------|--------------------------------|---------------------------|
| Resolution | VGA (640 x 480 pixels) | | SXGA (1280 x 1024 pixels) | |
| Imager type | Gray-scale CMOS global shutter | Color CMOS global shutter | Gray-scale CMOS global shutter | Color CMOS global shutter |
| Pixel size | 5.3 µm | | | |
| Imager size | 1.4" | | 1.18" | |
| Frame rate | 120 fps | | 60 fps | |
| Processor | 660 MHz DSP | | | |
| Dedicated I/O | 1 I/O | | | |
| Configurable I/O | 1 IN / 3 OUT | | | |
| I/O type | Isolated I/O I/O | | | |
| External button | Non-isolated I/O I/O I/O | | | |
| Ethernet | Yes | | | |
| RS-232 serial | 10/100 Mbit/s | | | |
| RAM | Yes | | | |
| Storage | 256 MB | | | |
| Lens options | 256 MB | | | |
| Illuminator options | 6mm / 8mm / 12mm / 16mm / 25 mm | | | |
| Power requirements | Wide angle LEDs: Red, White, IR, Blue, Red with Polarizer, White with Polarizer | | | |
| Protection class | Multi-chain 50 LEDs: Red, White, Blue | | | |
| Dimensions | 10 - 30 Wc | | | |
| Case material | IP67 | | | |
| Operating temperature | 85 x 54 x 43 mm (3.7 x 2.1 x 1.7 in.) - Connector @ 0° | | | |
| Storage temperature | 75 x 56 x 62 mm (3.0 x 2.1 x 2.4 in.) - Connector @ 90° | | | |
| Humidity | Aluminum, Plastic protective window cover | | | |
| Certifications | 0 to 90 °C (32 to 122 °F) | | | |
| | -20 to 70 °C (-4 to 158 °F) | | | |
| | 90% non-condensing | | | |
| | IE, CSA | | | |

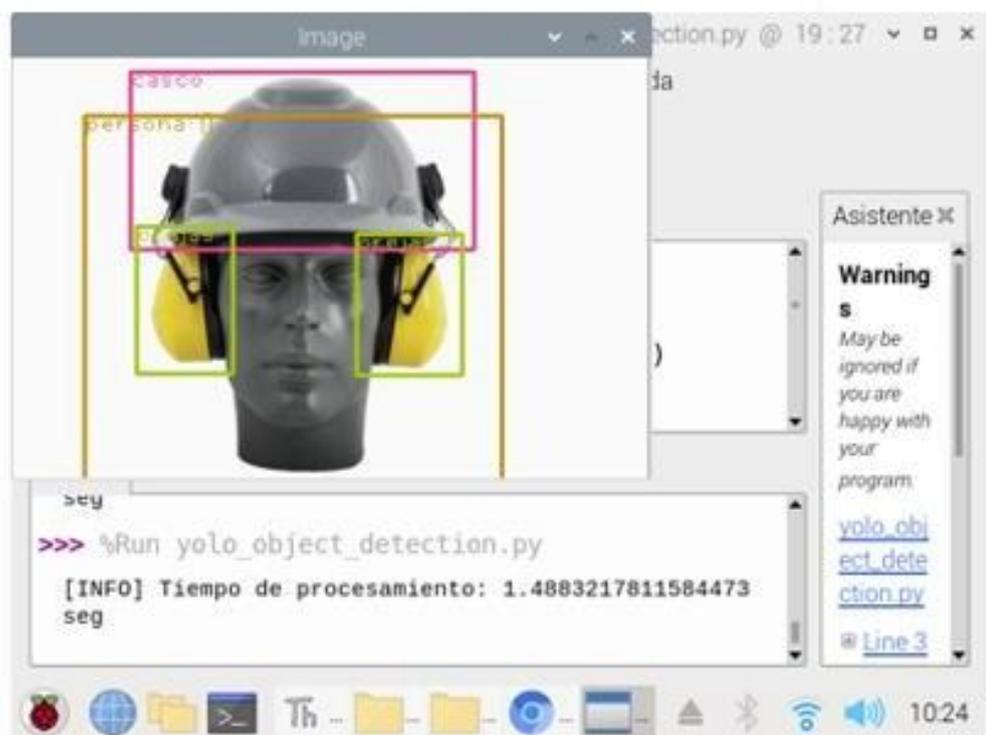
DIMENSIONES



Engineering peace of mind
Machine Vision - ISO Verification - Serialization

AIS Vision Systems
(+34) 931 003 035
info@aisvision.com
www.aisvision.com

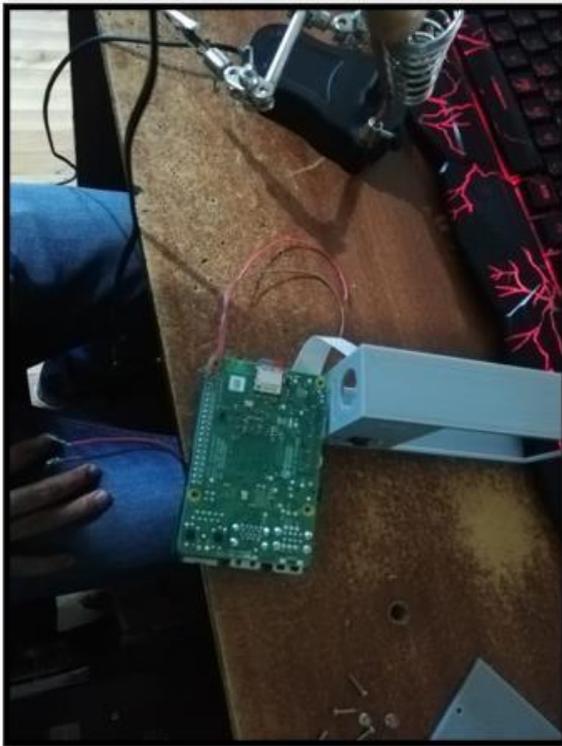
ANEXO E: Entrenamiento de la Red Neuronal mediante imágenes.



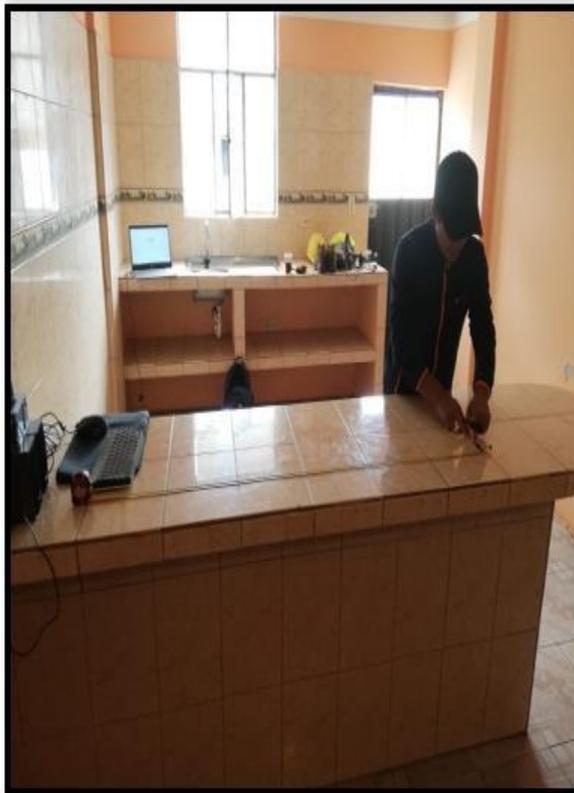
ANEXO F: Pruebas de detección de clases.



ANEXO G: Ensamblaje del prototipo AZATHOTH



ANEXO H: Pruebas de detección a diferentes rangos de distancia



ANEXO I: Configuración del LCD

