UNIVERSIDAD NACIONAL DE CHIMBORAZO
FACULTAD DE INGENIERIA
CARRERA DE INGENIERIA EN TELECOMUNICACIONES

Desarrollo de un sistema aplicando visión artificial para el análisis de comportamientos sospechosos y la prevención de atracos armados en la compañía de telecomunicaciones Global Cell.

Trabajo de Titulación para optar al título de:
Ingeniero en Telecomunicaciones

Autor:
Pungaña Villacrés, Augusto Luciano

Tutor:
PhD. Leonardo Fabian Rentería Bustamante.

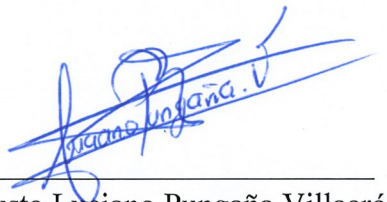
Riobamba, Ecuador. 2026

DECLARATORIA DE AUTORÍA

Yo, Augusto Luciano Pungaña Villacrés, con cédula de ciudadanía 180515684-9, autor del trabajo de investigación titulado: **"Desarrollo de un sistema aplicando visión artificial para el análisis de comportamientos sospechosos y la prevención de atracos armados en la compañía de telecomunicaciones Global Cell"**, certifico que la producción, ideas, opiniones, criterios, contenidos y conclusiones expuestas son de mí exclusiva responsabilidad.

Asimismo, cedo a la Universidad Nacional de Chimborazo, en forma no exclusiva, los derechos para su uso, comunicación pública, distribución, divulgación y/o reproducción total o parcial, por medio físico o digital; en esta cesión se entiende que el cesionario no podrá obtener beneficios económicos. La posible reclamación de terceros respecto de los derechos de autor (a) de la obra referida, será de mi entera responsabilidad; librando a la Universidad Nacional de Chimborazo de posibles obligaciones.

En Riobamba, 04 de junio de 2026.



Augusto Luciano Pungaña Villacrés
C.I: 180515684-9

DICTAMEN FAVORABLE DEL PROFESOR TUTOR

Quien suscribe, **PhD. Leonardo Fabian Renteria Bustamante** catedrático adscrito a la **Facultad de Ingeniería**, por medio del presente documento certifico haber asesorado y revisado el desarrollo del trabajo de investigación titulado: **"Desarrollo de un sistema aplicando visión artificial para el análisis de comportamientos sospechosos y la prevención de atracos armados en la compañía de telecomunicaciones Global Cell"**, bajo la autoría de **Augusto Luciano Pungaña Villacrés**; por lo que se autoriza ejecutar los trámites legales para su sustentación.

Es todo cuanto informar en honor a la verdad; en Riobamba, a los 4 días del mes de junio de 2026.

PhD. Leonardo Fabian Rentería Bustamante

C.I: 1104064132

CERTIFICADO DE LOS MIEMBROS DEL TRIBUNAL

Quienes suscribimos, catedráticos designados Miembros del Tribunal de Grado para la evaluación del trabajo de investigación **Desarrollo de un sistema aplicando visión artificial para el análisis de comportamientos sospechosos y la prevención de atracos armados en la compañía de telecomunicaciones Global Cell**, presentado por Augusto Luciano Pungaña Villacrés, con cédula de identidad número 180515684-9, bajo la tutoría de PhD. Leonardo Fabián Rentería Bustamante; certificamos que recomendamos la **APROBACIÓN** de este con fines de titulación. Previamente se ha evaluado el trabajo de investigación y escuchada la sustentación por parte de su autor; no teniendo más nada que observar.

De conformidad a la normativa aplicable firmamos, en Riobamba 29 de junio de 2026.

Luis Santillán, Mgs.
PRESIDENTE DEL TRIBUNAL DE GRADO

Alejandra Pozo, Mgs.
MIEMBRO DEL TRIBUNAL DE GRADO

Klever Torres, PhD.
MIEMBRO DEL TRIBUNAL DE GRADO

CERTIFICACIÓN

Que, **Pungaña Villacrés Augusto Luciano** con CC: **180515684-9**, estudiante de la Carrera **Ingeniería**, Facultad de **Telecomunicaciones**; ha trabajado bajo mi tutoría el trabajo de investigación titulado "**DESARROLLO DE UN SISTEMA APLICANDO VISIÓN ARTIFICIAL PARA EL ANÁLISIS DE COMPORTAMIENTOS SOSPECHOSOS Y LA PREVENCIÓN DE ATRACOS ARMADOS EN LA COMPAÑÍA DE TELECOMUNICACIONES GLOBAL CELL**", cumple con el 10%, de acuerdo al reporte del sistema Anti plagio Compilatio Magister+, porcentaje aceptado de acuerdo a la reglamentación institucional, por consiguiente autorizo continuar con el proceso.

Riobamba, 24 de junio de 2026

Validar únicamente en FirmaEC
Firmado electrónicamente por:
LEONARDO FABIAN
RENTERIA BUSTAMANTE

PhD. Leonardo Fabian Renteria Bustamante
TUTOR(A)

DEDICATORIA

Doy gracias a Dios, en los momentos que creía estar derrotado me dio el aliento suficiente para no darme por vencido, me resguardo en todo lugar de las situaciones complicadas en las que me encontraba, es por ello, que la sabiduría y conocimiento que he conseguido a lo largo de esta travesía es gracias a él y en su honor.

A mis padres, Augusto Pungaña, Olga Villacrés quienes, con sus oraciones llegaron a cada rincón de mi vida, su amor incondicional, sus sabios consejos, la enseñanza de los valores con la que me inculcaron, lograron formar parte de la persona en la que me he convertido, en los días más tétricos su apoyo y cariño incondicional siempre fueron motivo para mantenerme en pie de lucha.

A mis hermanos, Mayra, Jenny, Johanna, Nicolás, que siempre me aconsejan en el día a día sin importar donde me encuentre o en la situación que me involucre, su apoyo, sus charlas y el cariño que me demuestran me dan valor para encontrar otro motivo por el cual no me he dado por vencido, por ser mi refugio cuando he sentido decaer en este ciclo llamado vida.

A mis cuñados, Fabián y Saulo, por ser ese ejemplo de apoyo y consejeros, por saber que en ustedes también he encontrado un hogar, siendo un ejemplo de buenos padres y el reflejo de lo que se anhela para llegar al éxito.

A mis sobrinitos Alan y Gaia, que desde que llegaron dieron luz de pureza a toda la familia en especial a sus abuelitos, estoy agradecido con la vida porque han sido más que una etapa de madurez en mi vida, sino un regalo del saber que todos necesitamos seguir soñando aún sin importar la edad en la que nos encontremos.

A mi familia, que siempre confiaron en mí, les dedico este trabajo que es el reflejo de perseverancia, dedicación y esfuerzo, gracias por todo el apoyo incondicional que me bridaron.

Augusto Luciano Pungaña Villacrés

AGRADECIMIENTO

Agradezco primeramente a Dios por permitirme vivir el día a día y mantenerme con salud para alcanzar mis sueños, por siempre saber que puedo recibir el perdón sin importar los días malos que tenga, por la sabiduría y fortaleza que me brinda y así poder culminar esta etapa con éxito.

A mi madre Olga Villacrés, por ser el pilar fundamental en mi vida, recuerdo cada una de sus palabras cuando estaba dándome por vencido y que sin su apoyo moral y consejos no hubiese llegado al objetivo, ni la vida misma me alcanzará para agradecerle, por tanto.

A mi padre por forjar el carácter que tengo, por ser palabras de aliento y sintonía cuando lo necesité.

A mis hermanos, infinitamente por todo, por la compañía, por escucharme, por comprenderme y por aconsejarme en todo momento.

A mi familia y amigos, los que estuvieron ahí para aportar con un consejo, una sonrisa y por demostrar que en la vida uno se rodea de lo que da.

De manera especial, agradezco al Ingeniero Leonardo Rentería, tutor de esta tesis, por su dedicación, orientación y valiosas recomendaciones que fueron fundamentales para la culminación de este trabajo.

Augusto Luciano Pungaña Villacrés

ÍNDICE GENERAL

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

RESUMEN

La presente investigación aborda el desarrollo de un sistema de monitoreo y seguridad inteligente, fundamentado en el uso de visión artificial. Este sistema está concebido para la identificación automatizada de conductas consideradas sospechosas en cuanto a gestos corporales, además de la detección de prendas y accesorios que pueden opacar la visibilidad de una persona, todo ello mediante escenarios simulados. La implementación incorpora herramientas avanzadas de aprendizaje profundo, tales como Yolov11, Yolo pose, destinadas al seguimiento de individuos y a la estimación de posturas. Así mismo, se utiliza la biblioteca OpenCV para el procesamiento de video multicanal, estableciendo una interfaz de comunicación con la API de Telegram para la transmisión remota de notificaciones y alertas. No obstante, el uso de cámaras de video vigilancia MI 360 a 1080 de la compañía Xiaomi, un computador para su respectivo desarrollo y un móvil para el envío de alertas. La eficacia del sistema fue corroborada mediante evaluaciones de precisión en la detección de accesorios como gafas, gorras, mascarillas, etc. y en el reconocimiento de posturas indicativas de amenaza como manos en la cabeza o elevadas.

El proyecto se llevó a cabo en tres etapas. En primer lugar, tras una exhaustiva revisión del estado del arte en sistemas de video vigilancia, se seleccionaron modelos de detección de objetos, ponderando su equilibrio entre la velocidad de inferencia y la precisión. En segundo lugar, se procedió a la programación del sistema en Python utilizando la interfaz de Visual Studio Code. La activación de alarmas se produce ante la presencia de tres o más objetos considerados sospechosos y posturas evaluadas como críticas.

Se determina que la integración de redes neuronales convolucionales para la vigilancia representa una solución tecnológica sólida y adaptable. Se sugiere que las futuras generaciones incorporen el empleo de hardware especializado para aceleración como unidades de procesamiento gráfico dedicadas como por ejemplo NVIDIA Nano Jetson y la inclusión de algoritmos de identificación facial para perfeccionar el seguimiento de los individuos detectados, lo que optimizará los tiempos de respuesta en los sistemas de seguridad pública.

Palabras claves: YOLO, Python, visión artificial, detección de poses, seguridad ciudadana, monitoreo en tiempo real.

ABSTRACT

This research addresses the development of an intelligent monitoring and security system based on computer vision. The system is designed to automate the identification of suspicious behaviors related to body gestures, as well as the detection of garments and accessories that may obscure a person's visibility, all within simulated scenarios. The implementation incorporates advanced deep learning tools, such as YOLOv11 and YOLO Pose, for individual tracking and posture estimation. Furthermore, the OpenCV library is used for multi-channel video processing, and a communication interface with the Telegram API is established for remote transmission of notifications and alerts. The hardware setup included Xiaomi MI 360 1080p surveillance cameras, a development computer, and a mobile device for receiving alerts. The system's effectiveness was validated through precision evaluations for detecting accessories such as glasses, caps, and masks, and recognizing threat-indicative postures such as hands on head or raised.

The project was executed in three stages. First, following an exhaustive state-of-the-art review of video surveillance systems, object detection models were selected based on the balance between inference speed and accuracy. Second, the system was programmed in Python using the Visual Studio Code interface. Alarms are triggered by the presence of three or more suspicious objects or postures evaluated as critical.

It is concluded that integrating Convolutional Neural Networks into surveillance represents a robust and adaptable technological solution. Future iterations are encouraged to incorporate specialized hardware for acceleration, such as NVIDIA Jetson Nano-based GPUs, and to include facial identification algorithms to refine individual tracking, thereby optimizing response times in public security systems.

Keywords: YOLO, Python, computer vision, pose detection, public safety, real-time monitoring.

Validar únicamente en FirmaEC.
Firmado electrónicamente por:
ANA ELIZABETH MALDONADO LEON

Reviewed by:
Ms.C. Ana Maldonado León
ENGLISH PROFESSOR
C.I.0601975980

CAPÍTULO I

1. INTRODUCCIÓN

1.1. Antecedentes

Según la Convención de las Naciones Unidas contra la Delincuencia Organizada Transnacional (UNTOC), un Grupo Delictivo Organizado. *Se define como un grupo estructurado de tres o más personas, existente durante un periodo de tiempo, que actúa concertadamente con el propósito de cometer delitos graves con la finalidad de obtener, directa o indirectamente, un beneficio económico o material [1].*

Es importante considerar teorías y modelos previos que abordan el fenómeno de la delincuencia y su relación con el uso de tecnología para combatirla. Por ejemplo, se han propuesto modelos que analizan y procesan imágenes cómo el uso de sistemas de vigilancia convencionales y otros mejor desarrollados con visión artificial, que influyen en la reducción del crimen organizado y con el aprendizaje automático que pueden optimizar estos sistemas.

En los últimos años, Ecuador ha enfrentado un aumento en la actividad de grupos delictivos relacionados con el narcotráfico, el contrabando, la violencia urbana. Según datos del Observatorio de Seguridad y Convivencia Ciudadana, en 2022, la tasa de homicidios en Ecuador fue de aproximadamente 14,8 por cada 100,000 habitantes, aproximándose al promedio regional de 15,4 pero aún inferior a países como Brasil y Venezuela, que presentan tasas superiores a 30 por cada 100,000 habitantes [2].

La compañía Global Cell, reconocida por su amplia oferta de servicios y su compromiso con la innovación tecnológica, confronta el reto de asegurar sus activos, así como la seguridad de su personal y clientela. En este contexto, el desarrollo de sistemas que incorporen tecnologías avanzadas se presenta como una solución viable e indispensable.

La visión artificial, llamada una disciplina científica que incluye métodos para adquirir, procesar y analizar imágenes del mundo real con el fin de producir información que pueda ser tratada por una máquina. Es una aplicación de la teórica clásica de matrices a dimensiones considerablemente grandes. Su objetivo principal es la deducción automática de la estructura y propiedades de un mundo tridimensional a partir de una o varias imágenes bidimensionales [3].

Utilizada en una variedad de campos y aplicaciones, desde la mejora de imágenes y la restauración hasta el reconocimiento de patrones. Este campo abarca métodos como la segmentación, la filtración y la transformación de imágenes, y es fundamental en áreas como la visión por computadora y el análisis médico entre las más importantes.

El objetivo radica en el diseño e implementación de un sistema basado en visión artificial que proteja a la empresa de telecomunicaciones Global Cell. Este sistema busca identificar

comportamientos sospechosos para prevenir robos, empleando cámaras de vigilancia y algoritmos avanzados de procesamiento de imágenes. La finalidad es establecer un entorno más seguro para el personal, tanto en las operaciones internas como en las actividades externas de la compañía.

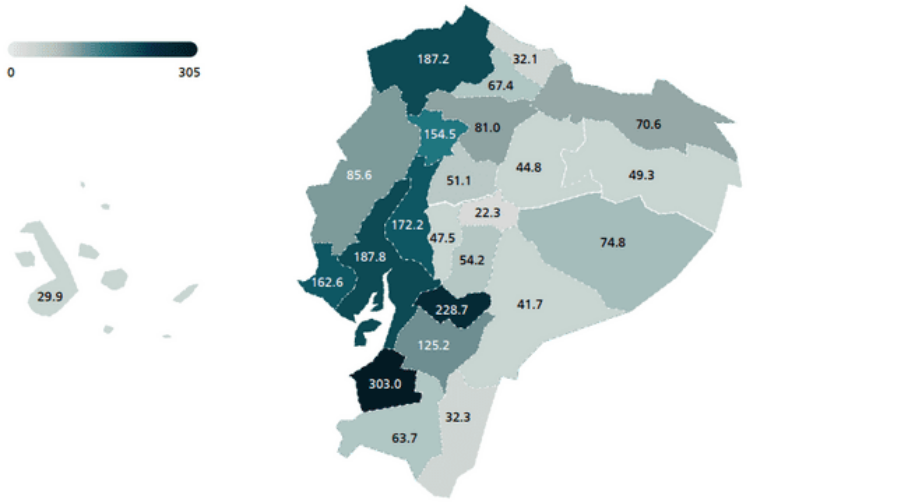
En este contexto, se establece el objetivo general que aborda el desarrollo del sistema propuesto para mejorar la seguridad de Global Cell. Los objetivos específicos engloban la elaboración y verificación del software para el procesamiento eficiente de imágenes mediante frames y el entrenamiento del modelo basado en el aprendizaje automático para identificar comportamientos extraños. En este documento, se expondrán las bases teóricas que respaldan esta investigación, además de la metodología utilizada para la creación del sistema previamente mencionado. Se anhela que los resultados logrados no solo ayuden a incrementar la seguridad en Global Cell, sino que también promuevan la mejora de la seguridad en la empresa.

Durante el desarrollo del proyecto se llevó a cabo un estudio de los distintos algoritmos usados en Inteligencia artificial, así mismo los más frecuentes y efectivos en visión artificial, además de una variedad de tecnologías existentes en el mercado que puedan ser utilizadas en el desarrollo del proyecto. Del mismo modo se implementó la parte electrónica del entorno a vigilar integrando un computador, software, cámaras y un teléfono para la notificación de los mensajes de alerta, así como el sistema de control de la nube nativa de las cámaras en donde se almacenan los videos grabados durante 48 horas para su respectiva revisión, entre otros. Se desarrolló el firmware, las interfaces de comunicación entre el usuario y el sistema. Además, programas de software que permiten la interacción de todos los elementos del sistema. Finalmente, se llevó a cabo pruebas de funcionamiento y validación del proyecto.

1.2. Planteamiento del problema

El dominio de ciertos territorios a nivel local por parte del narcotráfico, especialmente en zonas clave, se muestra mediante el cobro de extorsiones o "vacunas", lo que a su vez origina los delitos cometidos por grupos organizados contra negocios de todo tamaño. Los casos de extorsión en todo el país se concentran en las mismas áreas donde ocurren otros tipos de crímenes. La provincia de El Oro tiene la mayor cantidad de estos incidentes, y Santa Elena es la que más ha crecido en un año, con un aumento del 83% entre 2023 y 2024.

La concentración de la extorsión en la zona costera muestra una consolidación del crimen organizado en este espacio geográfico. Sin embargo, las altas tasas registradas en el resto del territorio nacional demuestran la presencia de organizaciones criminales con control territorial en todo el país. Así mismo, se observa una creciente importancia de la Amazonía como un enclave estratégico para el crimen organizado. De hecho, entre 2023 y 2024, Sucumbíos registró un incremento de este fenómeno en un 61% aproximadamente, seguido de 59% de incremento en Morona, 52% en Zamora y 23% en Orellana, datos que se pueden ser corroborados en la figura 1 [1].



Fuente: Fiscalía General del Estado, 2024b.
Elaboración: Observatorio Ecuatoriano de Crimen Organizado, 2025.

Figura 1. Tasa de extorsión por provincia en el Ecuador [1].

La violencia no solo se expresa por medio de asaltos y homicidios, sino también a través de otros crímenes, como hurtos y ataques armados. En este escenario, compañías de telecomunicaciones, como Global Cell ubicada en la ciudad de Ambato, se han transformado en blancos atractivos para criminales gracias a sus activos de gran valor y la constante presencia de clientes, tanto niños, jóvenes, adultos y adultos mayores. Estos crímenes no solo constituyen un peligro para la protección física de trabajadores y clientes, sino que también ocasionan pérdidas económicas considerables y perjudican la confianza de sus clientes.

En esta situación, surge la necesidad apremiante de implementar soluciones innovadoras que permitan la identificación de comportamientos sospechosos en individuos de cualquier género y edad, con el propósito de disuadir actividades delictivas. La visión artificial se postula como una herramienta prometedora en esta lucha contra la criminalidad [3]. Sin embargo, el mecanismo propuesto enfrenta desafíos técnicos y operativos que requieren atención.

El presente análisis busca responder a la siguiente pregunta: ¿Cuáles son los atributos fundamentales de un sistema de seguridad que incorpora visión artificial, a fin de posibilitar la detección efectiva de conductas sospechosas y la prevención de robos a mano armada en Global Cell, considerando aspectos técnicos, operativos y criterios de evaluación?

El enfoque de esta investigación es crear e instalar un sistema que, usando visión artificial, pueda detectar rápidamente actividades sospechosas en lugares controlados o simulados. Este estudio no busca probar si se pueden mitigar atracos automáticamente en tiempo real en situaciones muy peligrosas, sino más bien ver si el sistema puede reconocer conductas extrañas y objetos que opacan la visibilidad de los usuarios en escenarios de prueba. Para comprobar si funciona el prototipo, se realizarán varias pruebas piloto controladas de dos a cuatro semanas. Esta forma de trabajar nos permitirá recoger datos importantes para mejorar el sistema técnicamente con revisiones semanales, lo que garantiza que el modelo sea sólido antes de pensar en usarlo en negocios reales a largo plazo.

Mediante un análisis minucioso de los requisitos técnicos y operativos, se propuso un modelo denominado “Desarrollo de un sistema aplicando visión artificial para el análisis de comportamientos sospechosos y la prevención de atracos armados en la compañía de telecomunicaciones Global Cell” que no solo contribuya a la disminución de riesgos, sino que también promueva la confianza entre los usuarios.

1.3. Justificación

Dado que, en los últimos años, en el primer semestre de 2016, hubo 14.596 delitos denunciados, 480 de los cuales tuvieron participaciones jóvenes que cometieron asaltos o robos a vehículos privados, taxis y locales comerciales según estudios preliminares. El mayor problema yace en la zona costera como Guayaquil puesto que, está relacionado con las pandillas juveniles, se calcula que existe alrededor de 200 aproximadamente, integradas por cerca de 70.000 niños y jóvenes entre los 12 y los 25 años que posteriormente cruzan la zona de la sierra ecuatoriana a cometer actividades delictivas [4] A pesar de que la información estadística que se examinó es de 2016, se eligió porque ofrece una base sólida y académicamente reconocida sobre la cantidad de incidentes, lo que posibilita una comparación objetiva para crear el algoritmo de detección actual.

Actualmente, los sistemas de videovigilancia tradicionales operan de manera pasiva, su eficacia depende de la intervención de un operador humano quien, a causa de la fatiga o la sobrecarga de información, podría omitir la detección de comportamientos críticos. Esta investigación se justifica por la imperativa necesidad de transformar la vigilancia en un procedimiento proactivo. Al identificar en tiempo real gestos sospechosos y el empleo de elementos que disimulan la identidad, tales como gafas, mascarillas o gorras etc.

La presente investigación se basa en la aplicación de algoritmos de vanguardia. La utilización de YOLOv11 y YOLO Pose constituye una mejora considerable en comparación con los métodos convencionales de detección de movimiento, al posibilitar una interpretación contextual del entorno (estimación de la postura) [5]. Asimismo, la optimización lograda mediante el procesamiento paralelo y la lógica de omisión de fotogramas avala la factibilidad técnica del sistema, evidenciando la capacidad de realizar análisis complejos de visión artificial sin la necesidad de infraestructuras informáticas inalcanzables ni la saturación del ancho de banda de la red.

El estudio propone una solución concreta a la escasez de personal en los centros de monitoreo. La integración con la interfaz de programación de aplicaciones (API) de Telegram universaliza el acceso a la seguridad de alto nivel, posibilitando la recepción instantánea de alertas en dispositivos móviles. Esta característica posiciona al sistema como una herramienta adaptable, no solo para la seguridad pública, sino también para contextos privados o comerciales como lo es la implementación de un sistema basado en visión artificial que proteja a la empresa de telecomunicaciones Global Cell en Ambato, puesto que, se requiere una supervisión automatizada con bajos recursos operativos.

En el ámbito académico, se presenta un caso de análisis referente a la confluencia de redes neuronales convolucionales y sistemas de comunicación a distancia. Ofrece información empírica acerca de la efectividad en la detección de objetos y poses que se basan en alertas, estableciendo así una base para investigaciones futuras que persigan incorporar hardware especializado (tal como NVIDIA Jetson) o algoritmos de reconocimiento facial en contextos de seguridad inteligente.

1.4. Objetivos

1.4.1. Objetivo general

Desarrollar un sistema aplicando visión artificial para el análisis de comportamientos sospechosos y la prevención de atracos armados en la compañía de telecomunicaciones Global Cell, con un periodo de implementación de 6 meses aproximadamente.

1.4.2. Objetivos específicos

- Examinar y evaluar los diferentes tipos de algoritmos basados en visión artificial que puedan utilizarse para identificar conductas irregulares en ambientes comerciales, reconociendo sus beneficios, restricciones y requerimientos técnicos requeridos para su aplicación en Global Cell.

- Elaborar y poner en marcha un prototipo operativo del sistema que incorpore cámaras de seguridad y algoritmos de procesamiento de imágenes, posibilitando la supervisión en tiempo real y la creación de alertas frente a circunstancias inusuales.

- Verificar la eficacia del sistema implementado a través de ensayos piloto en los locales de Global Cell, evaluando su habilidad para identificar conductas sospechosas y su influencia en la prevención de actos delictivos.

CAPÍTULO II

2. MARCO TEÓRICO

A continuación, se establecen los fundamentos necesarios para el desarrollo del proyecto titulado “Desarrollo de un sistema aplicando visión artificial para el análisis de comportamientos sospechosos y la prevención de atracos armados en la compañía de telecomunicaciones Global Cell.”. Este proyecto tiene como objetivo la implementación de un sistema de seguridad con el uso de visión artificial, el cual consta de un sistema de seguridad que brinde mayor independencia y eficiencia en la detección de accesorios y comportamientos inusuales en la tienda tecnológica Global Cell.

2.1. Estado del arte

El ideal de implementar un sistema autónomo con visión artificial no es cuestión de algo único o quizá nuevo, pero su viabilidad práctica ha crecido exponencialmente con los recientes avances en hardware y software, mediante el desarrollo de la tecnología de punta de países que hoy en día son potencia mundial. Los primeros indicios de la cámara de videovigilancia que conocemos hoy en día fueron una versión mecánica , desarrollada en junio de 1927 por el físico ruso Léon Theremin para proteger el Kremlin. El ingeniero eléctrico alemán y pionero de la televisión alemana, Walter Bruch , es ampliamente reconocido como el inventor de la cámara de videovigilancia actual [6]. En 1942, durante la Segunda Guerra Mundial, se utilizaron en Alemania para monitorear los cohetes V-2.

En la década de 1970, los sistemas de videovigilancia comenzaron a implementarse en espacios públicos y comerciales. Empresas, bancos y autoridades municipales se percataron del potencial de la vigilancia ambiental para disuadir el delito y mejorar la seguridad pública [7]. Durante esta década, las cámaras monocromáticas conectadas a través de sistemas cableados se establecieron como el estándar. Aunque voluminosas según las normas contemporáneas, estas instalaciones posibilitaban la vigilancia elemental en establecimientos comerciales, áreas de estacionamiento y edificaciones gubernamentales.

Proyectos posteriores comenzaron a incorporar hardware más especializado y algoritmos más sofisticados. Un ejemplo concreto es el trabajo realizado en la Universidad politécnica salesiana sede cuenca, que implementó un sistema de videovigilancia basado en técnicas de reconocimiento facial, visión artificial y algoritmos inteligentes montados sobre sistemas embebidos [8]. Este dispositivo, que utiliza tecnología de visión artificial y cámaras de alta calidad, permitió alcanzar un desempeño de 84.69% de aciertos en las pruebas realizadas demostrando el potencial encaminado a futuras implementaciones.

Implementaciones y estudios recientes han demostrado una marcada preferencia por el desarrollo de sistemas inteligentes de bajo costo que capitalizan el potencial de la Inteligencia Artificial. Los investigadores emplean ordenadores de placa única, como la Jetson Nano o la Raspberry Pi, en calidad de unidad central de procesamiento. Estos están combinados con cámaras web convencionales y algoritmos avanzados de visión artificial,

tales como las diversas iteraciones de YOLO. Los sistemas propuestos tienen como objetivo la democratización tecnológica, fomentando su accesibilidad. Las tendencias más actuales en proyectos de esta índole abarcan:

Proyectos como el sistema de vigilancia inteligente implementado en el local comercial "El chaval" que utiliza modelos de yolov5 para implementar un sistema de vigilancia inteligente mediante el uso de visión artificial y reconocimiento de lenguaje natural que incorpora mecanismos de atención para enfocarse en las partes más relevantes dentro del entorno del local, además de enviar una alerta cuando se detectan palabras especificas con el que rigorosamente el modelo fue entrenado [9]. Durante varias pruebas de campo, su eficiencia total fue del 75%, logrando así dar por hecho que el sistema se adapta de acuerdo con los mecanismos programados, logrando además de ser más eficiente que los mecanismos comunes.

Entre otro avance relevante que se logró en la universidad nacional de Chimborazo, se encuentra el denominado sistema inteligente basado en visión artificial para el monitoreo y prevención de robos en tiempo real en la farmacia "Pharmatodo", el cual tras su desarrollo buscó evaluar el rendimiento del sistema mediante pruebas prácticas dentro del área mencionada, comparando su precisión, sensibilidad y detección de objetos [10].De los casos analizados, el 57.0% fueron detectados como verdaderos positivos, lo cual muestra una capacidad efectiva del sistema en eventos relevantes.

No obstante, cabe recalcar que ninguna interfaz o patente es perfecta para todas las situaciones, la investigación futura se dirige hacia sistemas de videovigilancia avanzada y detección de postura. Un trabajo presentado en 2023 describe un sistema de control de posturas y velocidad basado en el procesamiento de imágenes que permiten la detección y control de posturas en el entrenamiento de los fundamentos básicos del patinaje [11]. Esto ayudándonos a ser más eficientes y tener mayor enfoque en cuanto a los objetivos planteados. En el proyecto se implementa el sistema utilizando la técnica, el algoritmo estudiado y acorde a las necesidades para detectar la postura, posterior de ser detectado y procesado mediante una cámara de dron en tiempo real. La Tabla 1 resume algunos de los proyectos más relevantes en el campo.

Tabla 1. Revisión de sistemas implementados con visión artificial.

Autor / Año	Objetivo	Algoritmo	Hardware	Métricas	Resultado	Limitación
Romero Mogrovejo, 2018 [12].	Detección de armas de fuego	CNN, TensorFlow, YOLO	Cámaras de seguridad, computador	Recall, Precisión	Recall 0.9097, Precisión 0.9093	**Tipo de cámaras.**
Gutiérrez Meneses, 2022 [8].	Reconocimiento facial	CNN, Clasificador de Haar	Raspberry Pi 4, cámara IP, móvil	Porcentaje de aciertos	84.69% de aciertos	**Entrenamiento y enlace de Raspberry.**
Lovato Granizo, 2024 [9].	Detección de objetos y voz	YOLOv5, Google Speech Recognition	Arduino nano, Jetson nano, cámara IP	Eficiencia total	75% de eficiencia	**YoloV5 mantiene librerías antiguas.**

Villa Yuquilema, 2024 [10].	Detección de armas y gestos	YOLOv5, Haar Cascade, CNN	Cámaras, computador, móvil	Verdaderos positivos	57.0% de detección	**CPU y GPU de alto rendimiento.**
González & Muñoz, 2023 [11]	Control de posturas y velocidad	Mediapipe, CNN	Arduino Uno, dron DJI Tello, PC	Mediana de postura	Mejora de 3 a 4 (escala de postura)	**Características del dron.**

2.2. Términos y definiciones

2.2.1. Inseguridad en el Ecuador

La preocupación principal de la sociedad ecuatoriana en estos tiempos se centra en la inseguridad, un fenómeno que afecta directamente la tranquilidad de los ciudadanos. En los últimos años, diversas formas de delincuencia, como los asaltos, robos y fraudes, han aumentado, generando un clima de inseguridad generalizado. Este problema es alimentado por la actividad de asaltantes, bandas urbanas y grupos del crimen organizado, lo que representa un desafío significativo para la seguridad pública en el país [10]
Esta situación perjudica directamente a comerciantes y empresarios, ocasionando perjuicios económicos, así como un impacto psicológico en las víctimas de los robos y un ambiente de inseguridad en el sector comercial. Entre enero y junio en ecuador se registraron 2.920 robos a empresas y a locales comerciales, el pico más alto fue en marzo con 584 robos [13]

2.2.2. Impacto psicológico en empleados

De igual manera las consecuencias psicológicas pueden ser peores que sufrido haber una agresión. La persona afectada pude quedar con secuelas como por ejemplo miedo, ansiedad y estrés postraumático. Un robo es un evento estresante que puede tener consecuencias psicologías a largo plazo. La mayoría de las víctimas pueden sufrir una variedad de síntomas como por ejemplo depresión, miedo y culpa. También pueden estar estresados e irritables y poseer dificultad para dormir o concentrarse [14].

2.2.3. Impacto económico de la inseguridad

La inseguridad tiene repercusiones económicas significativas, ya que provoca la paralización de negocios, la interrupción de la producción y el comercio, y conlleva a una disminución en el empleo y los ingresos para los comerciantes y empresas [10].
Según el informe Top Risks 2024, Ecuador terminó el año 2023 como uno de los países más afectados por la inseguridad, desde junio de 2023 se evidenció una caída en ventas en locales comerciales entre -0.9% y -1.5% mensual durante el segundo trimestre del año. Actualmente en 2024 el comercio y el turismo proveen un descenso de ventas de hasta un 20% esto debido a los hechos delictivos que sufrió el país, como consecuencia se obtiene menos ingresos[15]

2.2.4. Comportamientos sospechosos

El análisis de comportamientos sospechosos implica identificar patrones que podrían predecir actos delictivos. Un estudio realizado por Mohammad en el año 2015 utilizó

algoritmos de aprendizaje automático para clasificar comportamientos en vídeos de vigilancia [16]. Este enfoque permite a los sistemas no solo detectar intrusiones, sino también anticipar acciones que podrían llevar a un crimen, como merodear de manera sospechosa, ver figura 2.

Figura 2. Comportamientos de inseguridad en una persona [17].

2.2.5. Conductas para considerarse

Las conductas que podrían catalogarse como sospechosas se vinculan principalmente con la expresión verbal y no verbal del individuo, por lo que se recomienda prestar atención, escuchar y observar.

Auditivas

- Incapacidad de dar respuestas.
- Renuncia a dar respuestas.
- Respuesta a una pregunta con otra pregunta. Etc.

Visuales

- Enrojece.
- Se pone blanco/palidece más peligro.
- Tiembla claramente. Etc.

2.2.6. Detección de atracos armados

La detección temprana de atracos armados es crucial para minimizar daños tanto a personas como a bienes materiales. Indagaciones previas han demostrado que los sistemas basados en visión artificial pueden identificar situaciones potencialmente peligrosas mediante el análisis del comportamiento humano y la detección de armas, ver figura 3. Estos sistemas utilizan técnicas avanzadas como redes neuronales profundas para clasificar imágenes y reconocer patrones asociados con situaciones de riesgo [14].

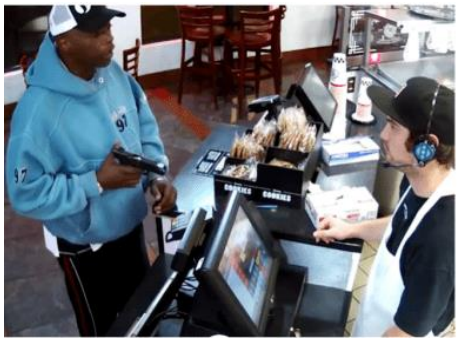


Figura 3. Atracos armados captados en tiempo real [12].

2.2.7. Sistemas de seguridad

Un sistema de seguridad puede ser conceptualizado como el conjunto de dispositivos y elementos indispensables para asegurar la protección requerida a individuos y propiedades materiales, presentes en una ubicación específica, contra amenazas externas, varían significativamente en función de los requisitos del usuario, las características del sitio a salvaguardar y el presupuesto asignado. En el mercado tecnológico se halla una amplia gama de componentes con especificaciones técnicas diversas.

2.2.8. Clasificación de los sistemas de seguridad

Este tipo de sistemas pueden encontrarse ubicados en cualquier edificación. La existencia de un sistema de seguridad electrónica en una instalación determinada puede ser obligatoria por ley en algunos casos como, sucursales bancarias, hospitales, residencias, aeropuertos, garajes, cárceles o puede ser opcional como los que se mencionan en la tabla 2, pueden instalarse simplemente por recomendación [18].

Tabla 2. Clasificación de los sistemas de seguridad [18].

Seguridad de las instalaciones		**Seguridad Privada**	
Sistemas de seguridad contra incendios.	Sistemas de detección de gas.	Sistemas antirrobo e intrusión.	Circuito cerrado de televisión (CCTV).

- **Sistemas de seguridad contra incendio**

Tienen la finalidad de localizar un incendio lo más tempranamente posible y dar aviso de este, evitando que las llamas se propaguen y minimizando al máximo los daños que puedan producirse sobre las personas, bienes o inmuebles [18].

- **Sistemas de detección de gas**

Los sistemas de detección electrónica de gas están diseñados para notificar al personal de una instalación sobre varias condiciones peligrosas:

- Riesgo de explosión por acumulación de gases o vapores inflamables.
- Riesgo de intoxicación por presencia de gases o vapores tóxicos [18].

- **Sistemas antirrobo e intrusión**

Se denomina sistema de seguridad electrónica anti-intrusión al conjunto de equipos y componentes capaces de administrar una o varias de las siguientes funciones:

- Intrusión. - Los dispositivos anti-intrusión advierten cualquier intento de irrupción o allanamiento en un determinado perímetro o recinto.
- Robo o atraco. - Los dispositivos antirrobo previenen los ataques contra personas, bienes e inmuebles [18].

Figura 4. Alarma como sistema antirrobo [18].

2.2.9. Elementos de un sistema de seguridad

A continuación, se detalla en la Figura 5, los elementos que embarcan los sistemas de seguridad.

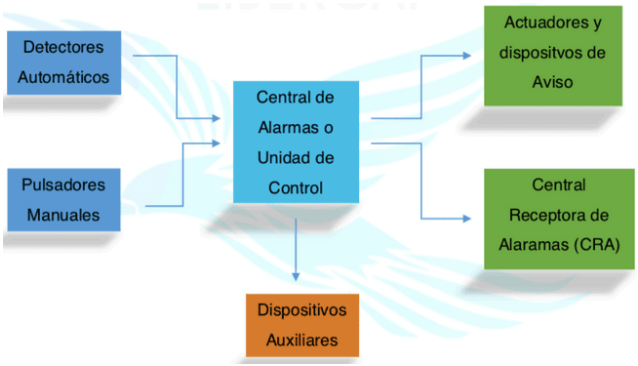


Figura 5. Bloques que componen un sistema de seguridad [18].

- **Sistemas de alarma**

Este tipo de sistema se caracteriza principalmente por su activación ante la sospecha de intrusión de individuos. Su beneficio radica en la pronta notificación al detectarse una presencia no deseada y en la protección de activos.

- **Sistemas de cámaras de seguridad**

Visualiza el espacio a resguardar. Se encuentra a la vista del personal de seguridad encargado. Ya sea en un cuarto de seguridad o mediante dispositivos remotos como pc, teléfono, tableta. Las cámaras se ubican estratégicamente para tener un ángulo adecuado y prevenir eventos considerables de sospecha. Su ventaja es la supervisión directa de los espacios escogidos [19].

2.2.10. Cámaras IP

Las cámaras de red tienen direcciones IP como cualquier otro dispositivo de red y se pueden instalar con pocos gastos en cualquier parte de la red, siendo controlada centralmente por

medio de software. Las cámaras IP pueden ofrecer una resolución hasta 16 veces superior y excelentes capacidades de zoom digital para cubrir un área más amplia [18].
En aplicaciones de videovigilancia IP, la tecnología inalámbrica representa un método flexible, económico y esencial para la instalación de cámaras, particularmente en sistemas IP que abarcan extensas áreas, como los sistemas de vigilancia para estacionamientos. Constituye una alternativa sobresaliente al prescindir de la necesidad de emplear cables de comunicación.

Figura 6. Modelos de cámaras IP [18].

2.2.11. Aplicaciones móviles

Representa un sistema contemporáneo en el mercado. Mediante la tecnología, el personal de seguridad, se podrá supervisar los acontecimientos en la empresa a través de un dispositivo inteligente.

2.2.12. Inteligencia artificial

Pertenece a la rama de la ciencia que se encarga del estudio de la inteligencia en elementos artificiales, es decir, crear sistemas y máquinas que realicen tareas parecidas y similares a las del ser humano [11]. Los algoritmos de mayor relevancia en el reconocimiento de imágenes incluyen redes neuronales, sistemas difusos e inteligencia de enjambre, conforme se ilustra en la figura 7.

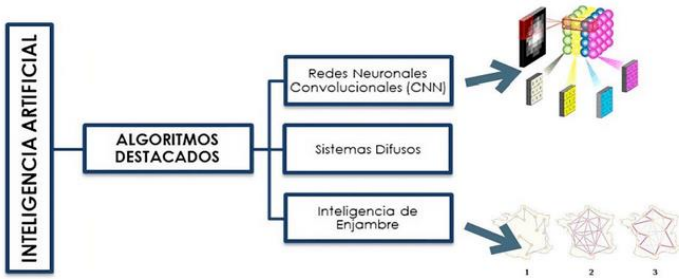


Figura 7. Algoritmos de IA para el reconocimiento de imágenes [11].

2.2.13. Visión artificial

Es una disciplina que se ocupa de cómo las computadoras pueden ser programadas para entender imágenes y vídeos. Esto incluye técnicas como la detección de objetos, el reconocimiento facial y el seguimiento de movimientos, ya que, combina algoritmos y

modelos matemáticos para extraer información útil de imágenes digitales [20]. En la Figura 8 se presenta paso a paso el proceso que se ejecuta en la visión artificial.

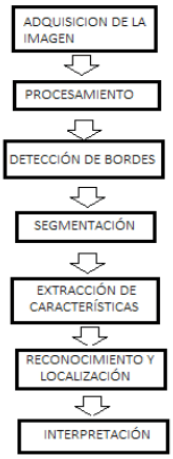


Figura 8. Procesos de la visión artificial al analizar una imagen [20].

2.2.14. Etapas del procesamiento de imágenes con visión artificial

Las etapas que involucra el procesamiento de imágenes se muestran en la tabla 3. La Visión Artificial puede ser definida como los procesos de obtención, caracterización e interpretación de información de imágenes tomadas desde un mundo tridimensional a partir de imágenes bidimensionales [21]. Estos pueden ser subdivididos en seis procesos como se muestra a continuación.

Tabla 3. Procesos de la visión artificial [22].

Procesos	Nivel de Visión	Entrada	Salida	Área
1. Captura	Bajo	Imagen	Imagen	Procesamiento de imágenes
2. Preprocesamiento				
3. Segmentación	Medio	Imagen	Grupos de píxeles en bruto (objetos o regiones)	Análisis de Imágenes
4. Descripción		Objetos o regiones	Información cuantitativa de los objetos o regiones	
5. Reconocimiento (clasificación)		Información cuantitativa	Objetos clasificados en categorías	
6. Interpretación	Alto	Objetos clasificados en categorías	Comprensión de la escena	Visión por Computador

El propósito primordial del procesamiento de imágenes mediante visión artificial es discernir y comprender el contenido visual de la imagen con el fin de ejecutar una labor específica.

2.2.14.1 Red neuronal convolucional (CNN)

Una red neuronal convolucional (CNN) es un tipo de red neuronal artificial que ha demostrado un gran éxito en el procesamiento de imágenes y en la visión por computadora. Están compuestas por capas de convolución, que aplican filtros a la imagen de entrada para extraer características relevantes, seguidas de capas de agrupación que reducen la dimensionalidad [23]

Han sido implementadas con resultados favorables en una variedad de aplicaciones, tales como la distinción de objetos, la segmentación semántica y la detección facial, etc. Este algoritmo para ser eficaz necesita de un requerimiento estricto en cuanto a hardware para su operación, por tal motivo está considerado uno de los más costosos en cuanto a ejecución.

2.2.14.2 Haar cascade Trainer

Haar Cascade Trainer es una herramienta utilizada para entrenar clasificadores de objetos basados en características de Haar. El entrenamiento implica proporcionar al algoritmo una serie de imágenes positivas que contienen el objeto de interés y una serie de imágenes negativas que no lo contienen [24].
Se emplea habitualmente en la identificación de objetos dentro de imágenes fijas, tales como rostros, ojos y vehículos, entre otros. Este método se fundamenta en las características de Haar y puede resultar efectivo para objetos que poseen atributos visuales notables y claramente delimitados.

2.2.14.3 You only look once (YOLO)

Es un algoritmo que detecta objetos en imágenes y vídeos en tiempo real. Utiliza una única red neuronal convolucional para detectar y clasificar objetos sin requerir una región de interés previa, como lo hacen otros algoritmos [11]. El proceso de detección consiste en dividir la imagen en una cuadrícula de celdas, es capaz de detectar varios objetos y asignarles las clases correspondientes en una sola pasada[11]. Aunque YOLO es rápido y eficiente, puede tener dificultades para detectar objetos pequeños o con formas complicadas, lo que puede afectar su precisión, depende además de la conectividad y la velocidad de procesamiento de la red.
Al ser un algoritmo no tan estricto en cuanto a ejecución y procesamiento de un modelo, el consumo de recursos es regular y se puede optimizar con limitantes dentro del script, esto ayuda a reducir el uso de la GPU, RAM y memoria, además de gestionar los famosos cuellos de botella a diferencia del algoritmo CNN.

2.2.15. Clasificación de algoritmos de postura

A continuación, se detallan los algoritmos utilizados en estimación de pose.

2.2.15.1 OpenPose

OpenPose es un algoritmo que se basa en una red neuronal convolucional (CNN) el cual detecta hasta 24 puntos claves del cuerpo humano como se puede ver en la figura 9, como las 29 articulaciones, los huesos, manos y puntos claves faciales [11]. Estos puntos se utilizan para rastrear y estimar el movimiento y la posición del cuerpo en tiempo real. El algoritmo utiliza técnicas de aprendizaje profundo y de visión por computadora para mejorar la precisión y la velocidad del seguimiento [11].

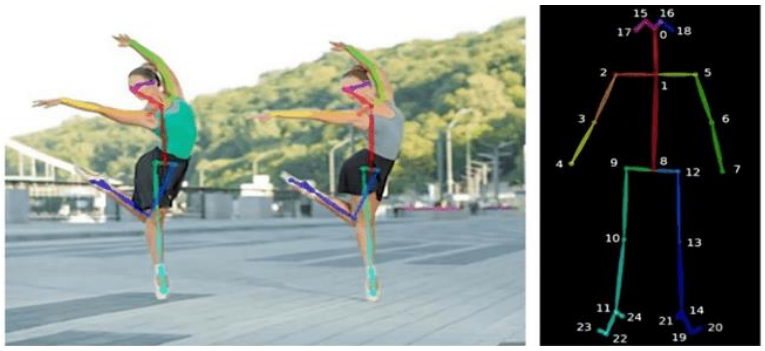


Figura 9. Puntos clave de OpenPose [11].

La finalidad de estos puntos es el seguimiento y la estimación en tiempo real del movimiento y la postura corporal. Emplea técnicas de aprendizaje profundo y visión por computadora para optimizar la precisión y la celeridad del seguimiento.

2.2.15.2 PoseNet

PoseNet es una herramienta desarrollada en colaboración con Google y TensorFlow que permite detectar puntos clave en imágenes de una o varias personas. Actualmente, es capaz de detectar hasta 17 puntos clave como se muestra en la figura 10. Es un modelo pre entrenado que puede ser utilizado de manera gratuita desde el navegador o en dispositivos móviles. PoseNet funciona en dos fases para realizar su estimación [11].

- Proporciona una imagen RGB como entrada a una red neuronal convolucional.
- Emplea un algoritmo decodificador para descifrar la estimación de pose.

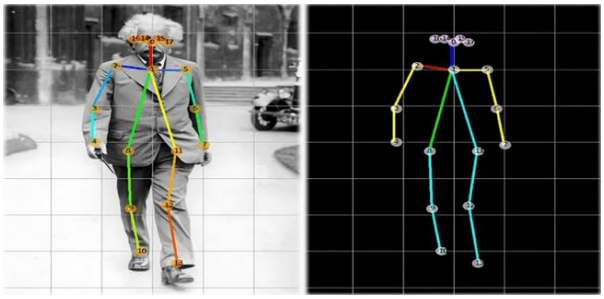


Figura 10. Puntos clave de OpenPose [11].

2.2.15.3 Blaze pose (MediaPipe pose)

Es un modelo de estimación de poses que se encuentra dentro de la biblioteca MediaPipe, fue desarrollado por Google y ofrece soluciones de alta fidelidad de visión por computadora y aprendizaje profundo en tiempo real [11]. Además, utiliza una red neuronal convolucional (CNN) que tiene una arquitectura de red basada en el modelo EfficentNet. Esta red es una arquitectura de red escalable que se puede ajustar a diferentes tamaños y resoluciones de la imagen. Utiliza una entrada de imagen RGB y devuelve una lista de 33 puntos clave de pose en 3D del cuerpo humano a partir de imágenes o videos como se aprecia en la figura 11.

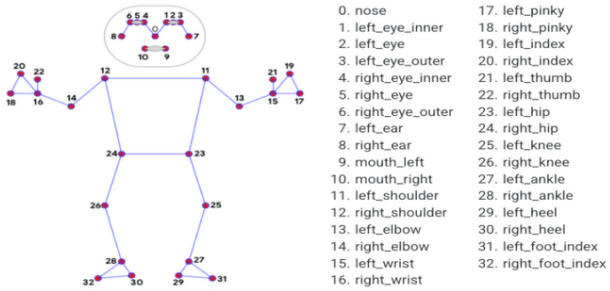


Figura 11. Puntos clave de BlazePose [11].

2.2.15.4 YoloPose v11

Los modelos de aprendizaje profundo como YOLO11 pueden identificar, localizar y dibujar puntos clave en un objeto o persona. En el caso de los objetos, estos puntos clave pueden incluir esquinas, bordes o marcas superficiales distintivas, mientras que, en el caso de las personas, estos puntos clave representan articulaciones importantes como el codo, la rodilla o el hombro [25].

Los puntos clave representan articulaciones fundamentales como el codo, la rodilla o el hombro, la detección de objetos determina su ubicación en una imagen mediante la delimitación con un recuadro, la estimación de la postura excede esta capacidad al pronosticar las ubicaciones precisas de los puntos clave en el individuo, figura 12.

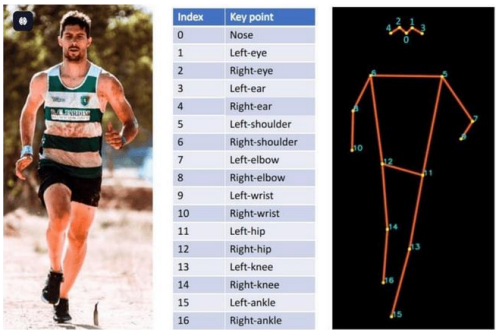


Figura 12. Puntos clave de YoloPose [26].

Actualmente, es capaz de identificar hasta 17 puntos clave. Se trata de un modelo pre entrenado de acceso gratuito como se observa en la figura 13, utilizable mediante navegador web o en dispositivos móviles.

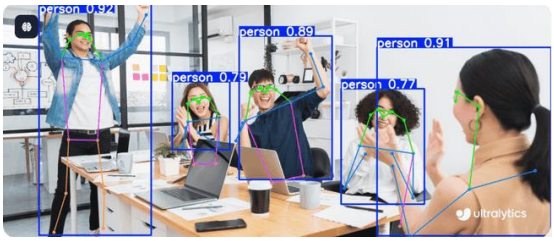


Figura 13. Uso de YOLOv11 para detectar y estimar las poses de varios individuos [25].

Mientras que la detección de objetos identifica su ubicación en una imagen mediante la delimitación con un recuadro, la estimación de la postura supera esto al pronosticar las ubicaciones precisas de los puntos clave en el objeto.

2.2.16. Métricas de detección en visión artificial

Para que un sistema pueda detectar objetos en una imagen, es necesario entrenar una red con varios ejemplos diferentes. Así, el sistema aprende a identificar los objetos y cómo se ubican entre sí. Los servicios que etiquetan imágenes profesionalmente pueden ser de gran ayuda para esto. Para un software que detecta objetos, algunas de las medidas más importantes son las siguientes:

- **Intersección sobre unión (IoU):**

El IoU mide cuánto se superponen dos recuadros, uno que el modelo predijo y otro que es el real. Se basa en el índice de Jaccard [27].

- **Precisión y Recall:**

La precisión se refiere a qué tan bien el modelo identifica correctamente lo que busca, mientras que la exhaustividad o recall se refiere a qué tan bien el modelo encuentra todos los elementos correctos [27].

- **Precisión media (AP):**

La AP combina la precisión, la exhaustividad y la confianza del modelo para cada cosa detectada. Se calcula para cada tipo de objeto y resume el rendimiento del modelo en un solo número [27].

- **Precisión media promedio (mAP)**

La Precisión Media Promedio (mAP), utilizada especialmente en situaciones en donde se tienen varias categorías, se determina calculando el promedio de la Precisión Promedio (AP) de cada clase. Esta medida evalúa la exactitud y la capacidad de encontrar objetos, considerando diferentes niveles de superposición (IoU) y tipos de objetos, lo que significa que un mAP mayor demuestra que el modelo funciona mejor en general [27].

- **F1 Score:**

La puntuación F1 es un promedio que equilibra la precisión y la exhaustividad [27].

2.2.17. Matriz de confusión para modelos de detección de objetos

A continuación, se muestra la matriz de confusión con sus respectivos conceptos a la hora de detectar objetos, ver tabla 4.

Tabla 4. Matriz de confusión para modelos de detección de objetos [27].

TIPO	CONCEPTO
Verdadero Positivo (VP)	El sistema de detección de objetos acierta al encontrar y ubicar objetos, y el área de superposición entre lo que predice y lo real es igual o mayor a un valor establecido.
Falso positivo (FP)	El sistema se equivoca al señalar un objeto que no existe en la realidad o cuando el área de superposición es menor al valor fijado.
Falso negativo (FN)	El sistema no ve un objeto que sí está presente en la realidad, es decir, lo pasa por alto.
Verdadero Negativo (VN)	Este concepto no se usa en la detección de objetos, ya que el propósito es encontrar e identificar objetos, no confirmar su ausencia.

Los conceptos mencionados se emplean comúnmente para determinar las mediciones fundamentales en la detección de objetos, como la exactitud, la cobertura, el puntaje F1 y la IoU. La exactitud, la cobertura y el puntaje F1 se obtienen a partir de la cantidad de aciertos verdaderos (TP), errores de tipo I (FP) y errores de tipo II (FN). La IoU mide qué tan bien se superponen los cuadros de predicción con los cuadros reales, y usualmente se usa para decidir si una detección es un acierto verdadero, ver figura 14.

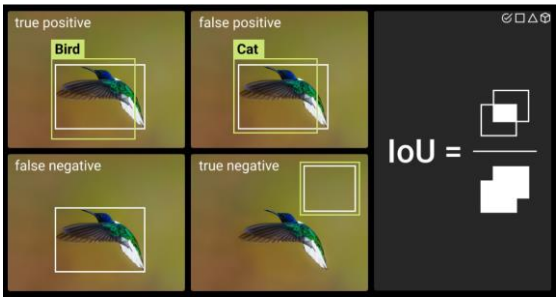


Figura 14. Matriz de confusión en detección de objetos.

2.2.18. Roboflow para etiquetado de datos

Roboflow ofrece herramientas para la anotación de datos y la exportación de conjuntos de datos en diversos formatos. La plataforma presenta dos servicios principales para facilitar la recolección de datos destinada a los modelos YOLO de Ultralytics: Universe y Collect.

- **Roboflow universe**

Roboflow Universe es un repositorio en línea de conjuntos de datos de visión. Puede exportar conjuntos de datos para usarlos con modelos Ultralytics [28].

- **Roboflow collect**

Roboflow Collect es un proyecto de código abierto que permite la recopilación automática de imágenes a través de una cámara web en dispositivos periféricos. Puede utilizar texto o indicaciones de imagen para especificar los datos que se van a recopilar, lo que ayuda a capturar solo las imágenes necesarias para el modelo como se observa en la figura 15 [28].

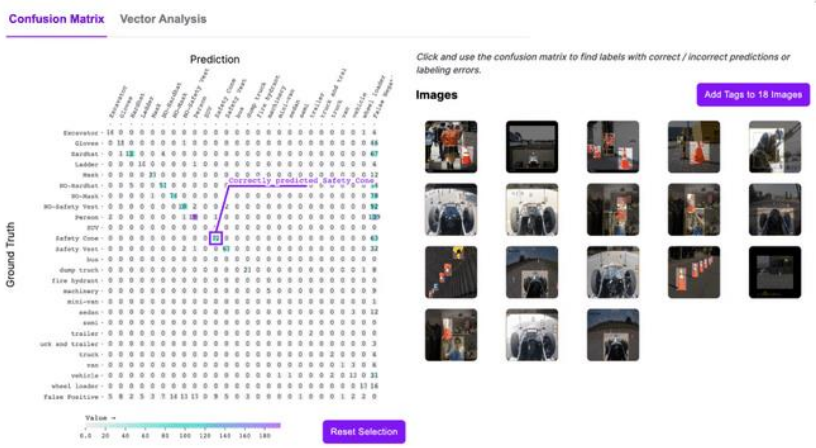


Figura 15. Etiquetado de imágenes mediante Roboflow [28].

CAPÍTULO III

3. METODOLOGÍA

3.1. Tipo de investigación

El presente proyecto de investigación se trata de una metodología de tipo cuantitativa y experimental, puesto que lo que buscamos es evaluar la eficacia del sistema para detectar eventos sospechosos en escenarios simulados, el índice de la delincuencia mediante el uso de la tecnología moderna, visión artificial, detección de objetos y de poses en un individuo, con el respectivo uso de componentes previamente estudiados e indagados, como cámaras de seguridad IP que funcionan en tiempo real. Se considera cuantitativa debido a que se medirá y analizará datos numéricos relacionados con la efectividad del sistema de visión artificial en la detección de poses y accesorios, a través de sistemas de seguridad convencionales y la debida comparativa con la propuesta que engloba este proyecto. Por otro lado, es de tipo experimental debido a que se manipulan variables de tipo dependiente e independiente que permiten comprobar la eficiencia del proyecto de investigación.

3.2. Técnicas de recolección de datos

- **Observación**

Mediante el análisis constante de secuencias de imágenes, el sistema obtiene rasgos externos y corporales directamente de la estructura de píxeles de los videos empleando arquitecturas de redes neuronales convolucionales.

- **Experimentación de escenarios**

Método en el que se reproducen escenarios de atracos, hurtos, entradas dudosas con elementos que ocultan el rostro como cascos, pasamontañas, gorras, etc., comportamientos de pánico o tensión laboral como manos en la cabeza, manos arriba, para medir la solidez técnica de la inteligencia artificial comparándola con el sistema tradicional convencional.

- **Control de pruebas**

Cuadro de análisis donde se compararon los resultados de aciertos y errores de detección, producidos al momento por el sistema automático como alerta inmediata por Telegram contra la intuición y comprensión de contexto de un monitorista de seguridad convencional.

3.3. Población de estudio y tamaño de muestra

La población está conformada por los datos de eficiencia en la detección de alertas en comparación al sistema implementado con visión artificial y el método tradicional. El experimento se llevó a cabo en la sucursal comercial Global Cell ubicada en la ciudad de Ambato, una importadora caracterizada por integrar equipos, accesorios y repuestos de calidad para sus distribuidores. El sujeto de estudio fue una persona de 28 años que cumplía el rol de sospechoso y delincuente para ejecutar las distintas pruebas y simulaciones dentro del entorno de la sucursal. Además, otras personas miembros de Global Cell de 19, 22 y 29 años que controlaba la caja registradora y atención al cliente. Entre los escenarios realizados se consideraron, compras del día común, se han presentado casos emulados de asaltos, robos, que afectan al día a día, dando como resultado un entorno de riesgo en un 80% a ser afectado.

3.4. Operacionalización de variables

Las variables del proyecto se muestran en la Tabla 5. Se evalúa el impacto sobre la variable dependiente como es la efectividad del sistema. El sistema se armoniza para la detección de accesorios, postura y la prevención de atracos dentro del área establecida.

Tabla 5. Operacionalización de variables.

Nombre variable	Tipo de variable	Descripción operacional	Instrumento de medición	Indicador de variable
Tipo de sistema	Independiente	Es la forma en la que se lleva a cabo el monitoreo, para la prevención y detección de comportamientos sospechosos.	Observación directa	Nominal
Efectividad	Dependiente	Capacidad del sistema para detectar o prevenir atracos, evaluando la eficiencia del sistema.	Observación directa	Porcentual (%)

3.5. Procedimiento/Desarrollo

El desarrollo de este proyecto se ha dividido en tres fases principales, cada una con sus respectivas actividades las cuales se ejecutan de forma rigurosa y secuencial para alcanzar los objetivos propuestos con relación a un enfoque técnico y progresivo mediante el empleo de redes neuronales avanzadas. En la Figura 16 se presenta el diagrama de flujo que muestra las distintas fases que engloban el proyecto de investigación.

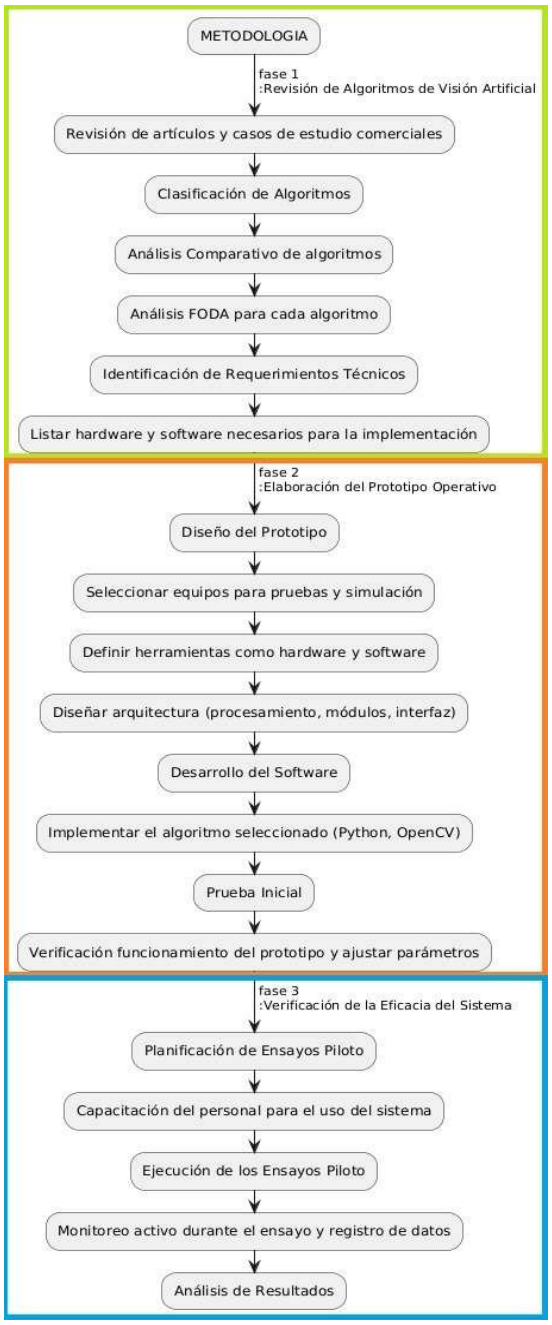


Figura 16. Diagrama de flujo – actividades a realizar por cada fase del proyecto.

Fuente: Autor.

3.5.1. Artículos y casos de estudio

En esta fase se realizaron varias indagaciones en diferentes artículos, páginas sobre las diferentes técnicas y algoritmos usados en visión artificial incluyendo la detección de postura, con la finalidad de seleccionar la que mejor se adapte a nuestras necesidades para desarrollar el análisis de comportamientos sospechosos mediante una técnica de postura y la prevención de atracos armados. El caso de estudio comercial que nos dio el enfoque necesario, aclarando varias dudas, fue el proyecto del señor Lovato Sebastián [9]. Se basa en un sistema que utiliza modelos de yolov5 para implementar un sistema de vigilancia

inteligente mediante el uso de visión artificial y reconocimiento de lenguaje natural que incorpora mecanismos de atención para enfocarse en las partes más relevantes dentro del entorno del local comercial el chaval [9]. Durante varias pruebas de campo, su eficiencia total fue del 75%, logrando así dar por hecho que el sistema se adapta de acuerdo con los mecanismos programados, además de ser efectivo a diferencia que los mecanismos tradicionales.

3.5.2. Fase 1: Clasificación de algoritmos

A continuación, se detallan los algoritmos mejor empleados y sus características para el desarrollo e implementación de proyectos con visión artificial, como se muestra en la tabla 6.

Tabla 6. Comparativa entre algoritmos de visión artificial modernos.

Algoritmo	Análisis Comparativo	Análisis FODA	Requerimientos Técnicos (HW)	Requerimientos Instalación (SW)
Red Neuronal Convolucional (CNN)	-Base de las arquitecturas de aprendizaje profundo. -Procesa la imagen por partes (capas) y es extremadamente costosa en cómputo.	F: Máxima precisión y versatilidad. O: Ideal para investigación médica o satelital. D: Lentitud en tiempo real. A: Costo de hardware prohibitivo.	GPU: NVIDIA A100 - RTX 3090 (24GB VRAM). RAM: 32GB+. CPU: Intel i9 - AMD Threadripper.	Python 3.x, TensorFlow-PyTorch, CUDA Toolkit, cuDNN, Drivers NVIDIA específicos.
Haar Cascade Trainer	-Detecta objetos mediante patrones. -No usa aprendizaje profundo complejo; se basa en rasgos de luz-sombra (Haar). Muy rápido, pero poco preciso.	F: Consumo de recursos casi nulo. O: Dispositivos IoT muy limitados. D: Falsos positivos altos, no tolera rotaciones. A: Obsolescencia técnica.	GPU: No requiere (corre en CPU). RAM: 4GB - 8GB. CPU: Intel i3 o procesadores ARM (Raspberry Pi).	OpenCV, Python 2.7-3.x, Compilador C++, XML para el clasificador.
You Only Look Once (YOLO)	-Usa una base de CNN. -Mira la imagen una sola vez (regresión única), permitiendo velocidad de tiempo real con alta precisión.	F: Velocidad en tiempo real y alta precisión. O: Mejor balance recursos-eficacia. D: Requiere instalación de librerías. A: Dependencia de drivers CUDA.	GPU: NVIDIA RTX 3060 - 4060 (8GB VRAM min). RAM: 16GB. CPU: Intel i5 - i7 de 10ma Gen+.	Python 3.10+, Ultralytics (YOLOv8-v10), OpenCV, PyTorch, Drivers NVIDIA (versión estable).

Yolo al ser un algoritmo no tan estricto en cuanto a ejecución y procesamiento de un modelo, el consumo de recursos es regular y se puede optimizar con limitantes dentro del script, nos ayuda a reducir el uso de la GPU, RAM y memoria, además de gestionar los famosos cuellos de botella a diferencia del algoritmo CNN, es óptimo en cuanto a requerimientos para el desarrollo de nuestro proyecto. Entre los algoritmos indagados y propuestos para el proyecto se ha realizado un análisis minucioso y a profundidad con métricas relevantes que destacan

por su gran escalabilidad y adaptabilidad en sus puntos característicos y primordiales como se observa en la tabla 6.

3.5.3. Clasificación de algoritmos de postura

Mientras que la detección de objetos identifica su ubicación en una imagen mediante la delimitación con un recuadro, la estimación de la postura supera esto al pronosticar las ubicaciones precisas de los puntos clave en el objeto.

Entre los algoritmos de estimación de pose indagados y propuestos para el proyecto se ha realizado un análisis minucioso y a profundidad con métricas relevantes que destacan por su gran escalabilidad y adaptabilidad en sus puntos característicos y primordiales como se observa en la tabla 7.

Tabla 7. Comparativa entre algoritmos de postura modernos.

Algoritmo	Análisis Comparativo	Análisis FODA	Requerimientos Técnicos (HW)	Requerimientos Instalación (SW)
OpenPose	- El estándar de oro en puntos clave. -- Es "Bottom-Up" (detecta partes y luego las une). Muy pesado.	F: Altísima precisión (24-135 puntos). O: Análisis biomecánico clínico. D: Extremadamente lento sin GPUs potentes. A: Licencia comercial costosa.	GPU: NVIDIA RTX 3080-4090 (12GB+ VRAM). RAM: 32GB. CPU: Multicore de alto rendimiento.	Caffe, OpenCV, CUDA, cuDNN, Python o C++. Requiere compilación manual compleja.
PoseNet	- Basado en TensorFlow. - Es el más ligero, pero menos preciso. Ideal para web.	F: Accesibilidad total (navegador). O: Apps web sencillas y juegos livianos. D: Mayor consumo frente a oclusiones. A: Precisión insuficiente para seguridad.	GPU: No requiere (usa aceleración WebGL). RAM: 4GB - 8GB. CPU: Básica (Laptop estándar).	TensorFlow.js o PoseNet Python wrapper. Muy fácil de instalar vía npm o pip.
BlazePose (MediaPipe)	- Usa CNN (EfficientNet). - Proporciona 33 puntos en 3D. Muy estable en móviles.	F: Balance perfecto velocidad-puntos (33). O: Fitness apps y Realidad Aumentada. D: Sufre con poses muy extremas-acostadas. A: Muy dependiente del ecosistema Google.	GPU: Integrada o móvil (Snapdragon/Apple A-series). RAM: 8GB. CPU: i5 de 8va Gen en adelante.	MediaPipe Library, Python 3.8+, OpenCV. Instalación simple: pip install mediapipe.
YOLOPose V11	- Basado en arquitectura YOLO. - Detecta la persona y sus puntos en un solo paso ("Top-Down").	F: El más rápido en video de alta resolución. O: Vigilancia y tracking masivo. D: Requiere dataset específico de pose. A: Curva de aprendizaje en la v11.	GPU: NVIDIA RTX 3060 (6GB+ VRAM). RAM: 16GB. CPU: i7 de 10ma Gen.	Ultralytics Framework, Python 3.10+, PyTorch, Drivers NVIDIA actualizados.

Tras una evaluación exhaustiva entre los algoritmos convencionales y de postura, se opta por implementar YOLOv11 en nuestro proyecto, tanto para detección de accesorios y YoloPose para la detección de postura debido a su eficacia y sus requerimientos técnicos. Además, es la más adecuada para trabajar con nuestro computador, ya que cumple todos los requisitos tanto en software y hardware. Posteriormente, se lleva a cabo una exhaustiva selección de imágenes para crear una base de datos, un paso sumamente importante para el

entrenamiento del algoritmo elegido. Se generaron siete conjuntos de datos específicos, entre ellos se encuentran la detección de gorras, gafas, mascarillas, mochilas, capuchas, cascos de motocicleta y pasamontañas, en cuanto a la estimación de postura, usamos una lógica convencional usando los keypoints que nos proporciona YoloPose entre una comparación a la altura de los hombros hacia arriba para los dos casos de estimación de pose.

3.5.4. Fase 2: Diseño del prototipo

3.5.4.1 Definición de equipos y herramientas a utilizar

Se obtuvo una idea simple pero completa de todo lo que conlleva el esquema de conexión electrónico, puesto que al utilizar un computador se evita generar en lo más mínimo los picos de botella y no provocar ralentización o congelamiento al momento de enviar las alertas, además de ser eficaz y lo más preciso, cumpliendo los objetivos del proyecto. Se procede a enlistar los equipos y herramientas que se utilizaron en el desarrollo e implementación del sistema.

3.5.4.2 Equipos utilizados para la implementación

Se divide en cuatro módulos tales como, periféricos, enlace, etapa de control y etapa de alimentación, son consideradas el complemento de cada una, puesto que si una tiende a fallar nuestro sistema no será eficiente.

3.5.4.2.1 Periféricos

- **Cámara de seguridad MI home 360 1080p**

Para el desarrollo de este proyecto en general, la clave es el tipo de cámaras a utilizar, por lo que se usaron dos cámaras de seguridad IP modernas, específicamente unas MI Home 360 a 1080p de la compañía de Xiaomi como se puede observar en la figura 17.

Figura 17. Cámara de seguridad MI home 360 1080p [29].

Este innovador producto se encuentra entre una de las mejores en el mercado por sus características como se observa en la tabla 8. Además, es accesible en cuanto a calidad precio, ofrece una tecnología incorporada como lo es una nube de Xiaomi Home, donde se almacenan todos los datos cómo por ejemplo grabaciones del día a día, haciendo parte fundamental en cuestión de seguridad en la compañía de Global Cell. Fue lanzada específicamente para ambientes interiores de bajo contraste de luz o visibilidad ya que incorpora una apertura focal de f/2,1 y una resolución de 1920 x 1080p, capaz de detectar

rostros, accesorios, y artículos complicados de observar. A diferencia de las cámaras convencionales o que utilizan DVR puesto que por el costo logran alcanzar una grabación de 720p dando como resultado una baja calidad de video e imagen. La estructura de las cámaras se diseñó en plástico blanco fuerte y tiene un estilo sencillo y pequeño, Se pueden ubicar en una mesa, o atornillarla al techo, ya que incluye una base y tornillos para montarla por ser tan liviana, su peso es alrededor de 228 gramos aproximadamente.

Tabla 8. Especificaciones técnicas de la cámara Mi 360° [29].

ESPECIFICACIONES TÉCNICAS	
Modelo	MJSXJ10CM
Apertura	f/2,1
Resolución	1920 x 1080
Memoria ampliable	Tarjeta Micro SD (hasta 32 GB)
Temperatura de funcionamiento	De -10 a 40 °C
Conectividad inalámbrica	Wi-Fi IEEE 802.11b/g/n, 2,4 GHz
Peso neto	254 g
Entrada de alimentación	5 V/3 A
Ángulo de visión	110°
Dimensiones	108 × 75 × 75 mm
Compatibilidad	Android 4.4, iOS 9.0 o versiones posteriores

- **Teléfono móvil**

Otro equipo importante es un dispositivo móvil, ya sea Android o IOS ya que cumple la función de la recepción de mensajes de alerta, en este caso se utilizó un iPhone 8 plus con características básicas de 64gb de almacenamiento e IOS 16.7.11, el cual es óptimo para nuestros requerimientos, ver figura 18.

Figura 18. IPhone 8 plus para recepción de alertas.

Fuente: Autor

- **Pantalla externa**

Si bien es cierto, necesitamos una pantalla de mejor tamaño para obtener una visualización más detallada y solida al momento del monitoreo, es por ello por lo que se implementó una pantalla adicional para una mejor comprensión. El modelo de monitor es de la marca Dell LCD de la serie E176FPf de 17 pulgadas y sus características son las más adecuadas para lograr nuestro propósito, ver figura 19.

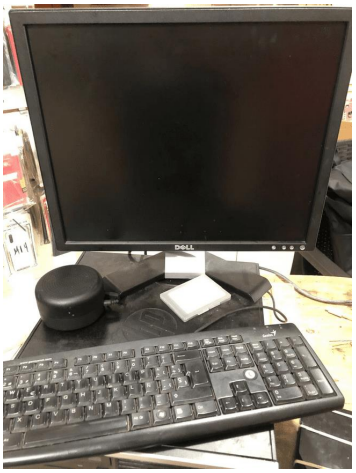


Figura 19. Monitor Dell LCD de la serie E176FPf

Fuente: Autor

3.5.4.2.2 Enlace

Se produce mediante un router ONT (Terminal de Red Óptica) de la empresa CNT, siendo la Corporación Nacional de Telecomunicaciones de Ecuador.

Específicamente, corresponde al fabricante ZTE, modelo ZXHN F6610M compatible con tecnologías GPON/XPON y conectividad inalámbrica Wi-Fi 6, sus características son las más adecuadas para lograr el objetivo del proyecto, ver figura 20.

Figura 20. Router ONT, terminal de red óptica.

Fuente: Autor

3.5.4.2.3 Etapa de control

Se desarrolló el entorno de virtualización e implementación del código, por ello se utilizó una máquina que cumple los requerimientos técnicos, al ser una laptop de marca Dell Core I5 de octava generación con un núcleo central de 2.50 GHz, con 16 GB de RAM y 500 GB de disco sólido, un equipo para el procesamiento y desarrollo de la interfaz, ver figura 21.

Figura 21. Computador maraca Dell Core I5.

Fuente: Autor

3.5.4.2.4 Etapa de alimentación.

Para las dos cámaras IP se utilizaron fuentes de características un tanto globales como otorga la misma empresa de Xiaomi, utiliza una entrada de 5 voltios y 3 Amperios cada una.
El router utiliza una fuente de 12 voltios a 2.5 amperios, el computador se alimenta con una fuente de 19.5 voltios a 6.5 amperios, para finalizar el equipo móvil utiliza un cargador convencional de tipo lightning de 5 voltios y 2.5 amperios.

3.5.4.3 Herramientas de software

Bajo los estudios realizados, dirigidos directamente en la implementación de un entorno virtual para no afectar el software de nuestro computador, puesto que, al instalar varias librerías estas pueden terminar interrumpiendo el funcionamiento del equipo. Como eje principal tenemos el uso del software base que se detalla a continuación.

3.5.4.3.1 Visual studio code

En su última versión 1.114.0, constituye el editor de texto predilecto para la mayoría de los programadores en la actualidad. Visual Studio Code es una herramienta ligera, ágil y sumamente adaptable para un entorno a desarrollar. A continuación, se detallan las librerías utilizadas:

- **cv2:** Permite la activación de cámaras, la lectura de archivos de imagen, el ajuste de dimensiones de fotografías, la delineación de recuadros de detección y la visualización de contenido de video en ventanas.
- **Requests:** Facilita la transmisión de resultados de detección a una interfaz de programación de aplicaciones (API).
- **Threading:** Previene la interrupción del programa.
- **Time:** Posibilita el cálculo de la tasa de fotogramas por segundo (FPS) del video.
- **numpy as np:** Es frecuente para la manipulación de píxeles, la aplicación de filtros matemáticos o la transformación de coordenadas.
- **ultralytics import YOLO:** Carga los modelos preentrenados.

3.5.4.3.2 Roboflow para el etiquetado de imágenes

Diseñada para optimizar el ciclo de vida completo de los proyectos que incluyen Visión Artificial. Las bibliotecas previamente mencionadas, constituyen las herramientas para la construcción del proyecto, Roboflow ayuda a representar la fuente que suministra los materiales preprocesados y estructurados.
La utilización de esta plataforma proporciona beneficios esenciales, particularmente en entornos de colaboración o ante la escasez de datos.

- **Etiquetado**

Para que una inteligencia artificial logre identificar un accesorio, es necesario demarcar miles de instancias del objeto en imágenes. Roboflow dispone de una interfaz web que facilita las siguientes acciones:

- Trazar manualmente cajas delimitadoras bounding boxes o polígonos.
- Emplear "Smart Click", una funcionalidad que utiliza inteligencia artificial para la selección automática de objetos con un solo clic

- **Aumento de datos**

Constituye uno de sus principales atributos. En el caso de disponer de cien fotografías, Roboflow puede generar quinientas adicionales de manera automatizada mediante la aplicación de diversos efectos:

- Rotar las imágenes, modificar la luminosidad, incorporar ruido o aplicar zoom. Este proceso incrementa significativamente la robustez del modelo, permitiéndole identificar objetos incluso en condiciones de iluminación adversas o desde ángulos inusuales.

- **Google colab**

Se trata, fundamentalmente, de un servicio en la nube gratuito, basado en cuadernos de Jupyter, que posibilita la escritura y ejecución de código Python directamente desde el navegador web. Muy utilizado en el entrenamiento de redes neuronales, Colab funciona como un préstamo de una computadora de alto rendimiento por parte de Google, sin costo alguno.
El entrenamiento de inteligencia artificial demanda una considerable capacidad computacional. Si se intentara realizar en una computadora portátil convencional, el proceso podría extenderse por varios días o incluso sobrecargar el procesador.

3.5.5. Diseño de la arquitectura electrónica

La arquitectura de este proyecto se puede apreciar en la figura 22, presenta una estructura electrónica global, que garantiza un prototipo eficaz y seguro para la empresa, mismo que fue diseñado para usarse dentro de un entorno controlado como lo es la sucursal de Global Cell. Además, el diseño es asíncrono, es decir, mientras el hilo principal muestra el video, un hilo secundario (threading) gestiona las comunicaciones con Telegram, lo que evita que el sistema se ralentice cuando hay congestión de red en el lugar de la instalación.

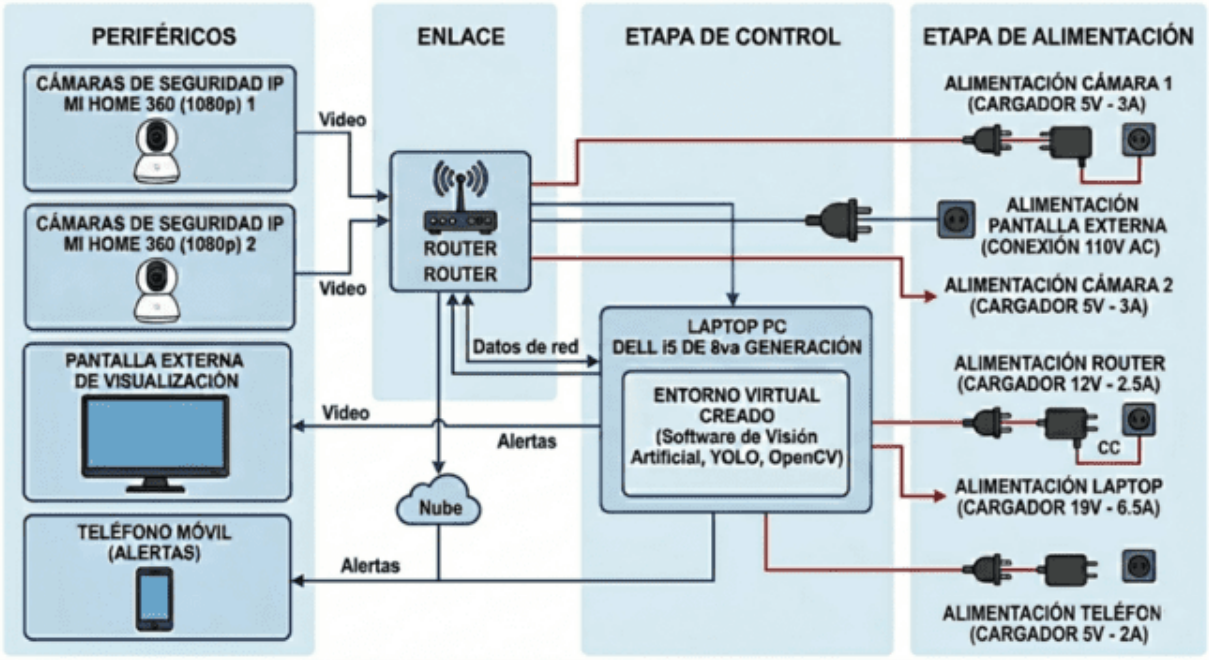


Figura 22. Arquitectura general del sistema electrónico distribuido.

Fuente: Autor

3.5.6. Desarrollo del software

El software desarrollado, incorpora una arquitectura de monitoreo híbrido en tiempo real. Este sistema emplea modelos de aprendizaje profundo para la identificación de objetos y la determinación de posturas, los cuales se integran a través de una lógica de decisión para la emisión de alertas. A continuación, se detalla el proceso del desarrollo del software, desde el punto más básico, hasta el más complejo sin eludir ningún paso relevante.

3.5.6.1 Creación del entorno virtual

Para establecer el entorno virtual en el sistema operativo Windows, es necesario ejecutar el comando 'python -m venv sis_seguridad_ws' en la interfaz de línea de comandos de Cmd o powerShell, sis_seguridad_ws se trata del nombre de la carpeta que llevará nuestro entorno virtual. Esta acción resultará en la creación de una estructura de directorios aislada, la cual contendrá su propio ejecutable de Python y un directorio de bibliotecas independiente. Una vez completada su creación, resulta importante activar dicho entorno empleando el script localizado en '.\sis_seguridad_ws\Scripts\activate'. La correcta activación se evidenciará por la aparición del nombre '(sis_seguridad_ws)' al inicio de la línea de comandos y la

carpeta oficial en Windows como se aprecia en la figura 23. Este procedimiento asegura que todas las dependencias instaladas a través de 'pip' permanezcan encapsuladas dentro de este contenedor virtual, previniendo así conflictos con otros proyectos, conflictos con el sistema operativo de Windows y garantizando un proceso de desarrollo ordenado y supervisado para el sistema de seguridad en cuestión.

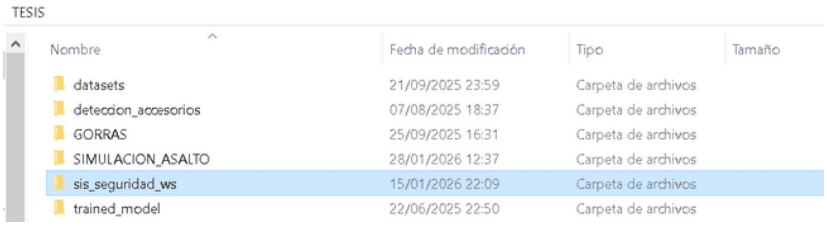


Figura 23. Creación del entrono virtual.

Fuente: Autor

3.5.7. Etiquetado y entrenamiento de la red

Se detalla a continuación el proceso del etiquetado y entrenamiento de la red neuronal.

3.5.7.1 Dataset y procesamiento mediante Roboflow

- **Número de imágenes y fuente**

El sistema se basa en datos recolectados para reconocer siete tipos de accesorios como, gorras, gafas, mascarillas, mochilas, capuchas, cascos de moto y pasamontañas dando un número total de 3537 imágenes. En cuanto al número de videos, se utilizaron aproximadamente 60 videos puesto que, al realizar el entrenamiento de accesorios de manera individual fue necesario probar la eficiencia de detección de cada uno de manera independiente. Con respecto a la partición del Dataset, el entrenamiento se realizó con 3186 muestras, con 210 imágenes para validación, y 141 para conjuntos de prueba, dando un número de 3537 imágenes totales mediante la plataforma de Roboflow, ver figura 24.

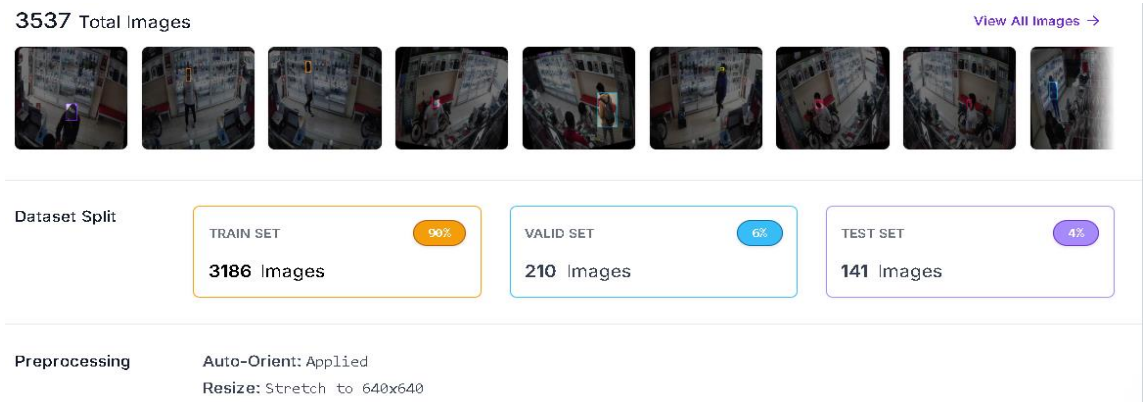


Figura 24. Número de imágenes y partición del Dataset en Roboflow.

Fuente: Autor

3.5.7.2 Proceso de etiquetado en muestras recolectadas

- **Cargar los datos**

Dentro de la plataforma simplemente se cargan las muestras en imágenes, se espera el tiempo en que tarda en subirse a la plataforma, ver figura 25.

Figura 25. Cargar las muestras en imágenes [28].

- **Etiquetar imágenes**

Se etiquetan las imágenes al respectivo tipo de accesorio asignado y a la clase. Esto se realiza para las 7 clases que son gorras, gafas, mascarillas, mochilas, capuchas, cascos de moto y pasamontañas, ver figura 26.

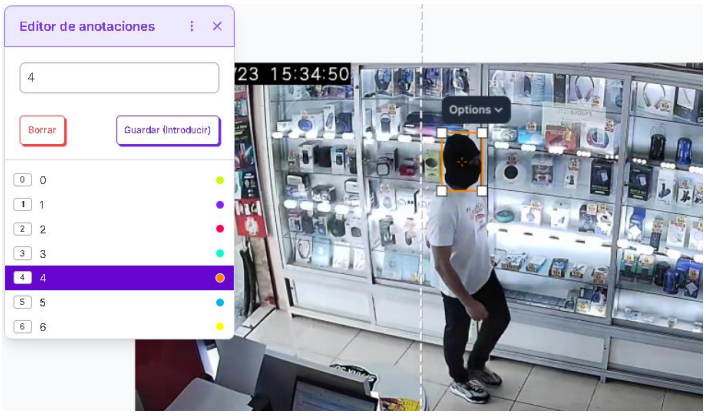


Figura 26. Proceso de etiquetado de accesorios, ejemplo pasamontañas.

Fuente: Autor

- **Entrenamiento del conjunto de datos**

Posterior al etiquetado de todo el conjunto de datos, se debe escoger la división del conjunto de muestras, esto implica, el número de muestras para entrenamiento, para validación y para prueba. Para este caso, se optó por un 90% para entrenamiento, 6% validación y 4% prueba, ver figura 27.

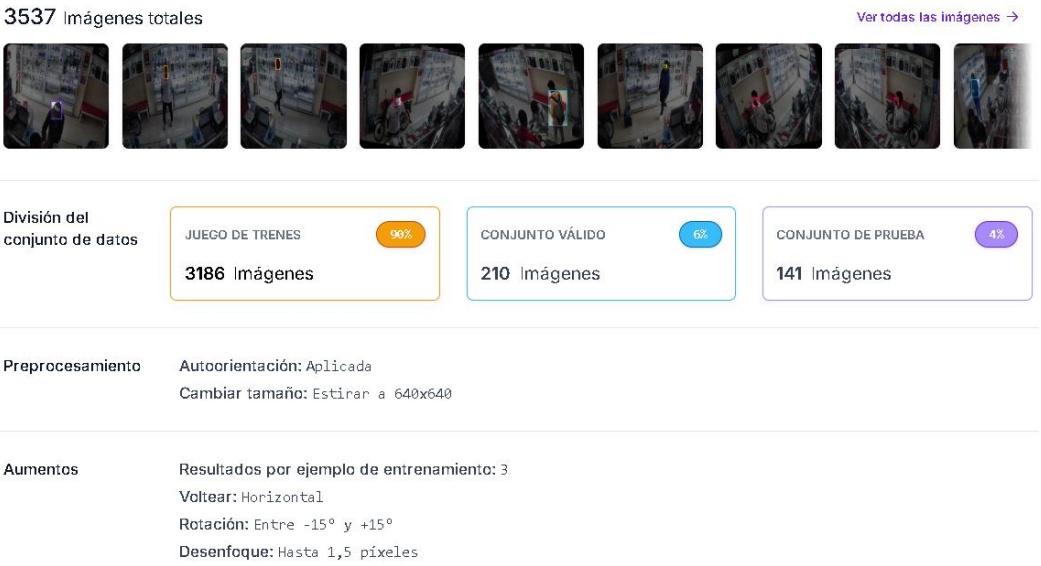


Figura 27. División del conjunto de datos en Roboflow.

Fuente: Autor

- **Descarga del dataset**

Una vez finalice el entrenamiento basta en descargar el conjunto de datos o simplemente hacer uso del link que nos proporciona Roboflow, es más práctico y optimiza muchos recursos en varias plataformas como Google Colab o Visual Studio Codec, ver figura 28.

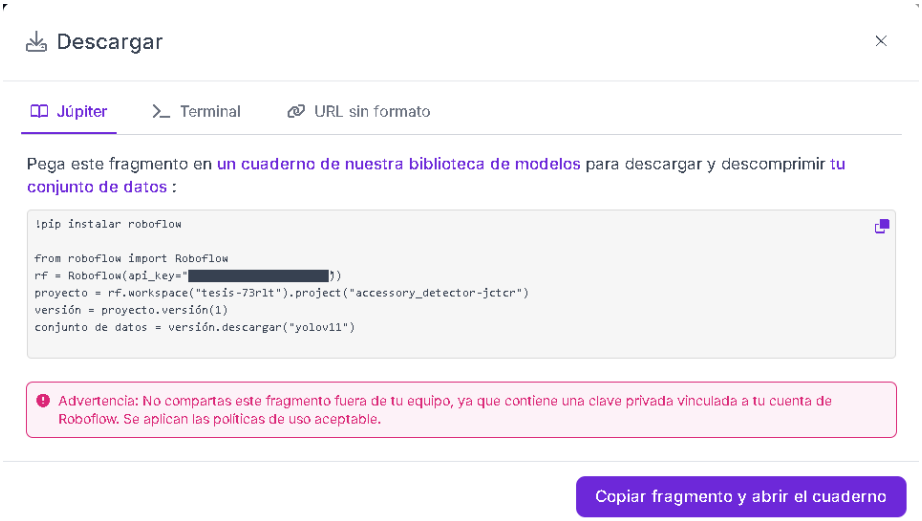


Figura 28. Dataset proporcionado por Roboflow.

Fuente: Autor

3.5.8. Parámetros del modelo y configuración

Para que el sistema funcione de manera eficiente, se han establecido pautas y ajustes técnicos específicos. Cabe recalcar que el parámetro depende de la variable asignada dentro del software, así mismo se utilizaron Dataset previamente entrenados como por ejemplo la detección de personas.

3.5.8.1 Parámetros ajustables para optimización y rendimiento de GPU

En este apartado se integraron los parámetros esenciales y ajustables para mejor rendimiento del computador, ya que se ralentiza al integrar tres modelos cargados, entre ellos tenemos el ajuste para el tamaño de la imagen, el umbral de confianza de detección de accesorios, el umbral de confianza de detección de personas, la mínima cantidad de accesorios a detectarse, el ROI, se basa en que cada vez que se detecta a un individuo, se crea un recorte dinámico sobre el cual se ejecuta el detector de accesorios, los tipos de clases que se ejecutan son siete y puede ser escalable, es posible incorporar nuevas clases de accesorios mediante el reentrenamiento del modelo 'm_acc', además se añaden los colores que toman los cuadros por alerta y emergencia.

En la tabla 9 se muestra las configuraciones que ajustan el balance entre la efectividad del sistema, como resultado tenemos la figura 29, en donde se aprecia el resultado de la ventana de optimización, además, en la tabla 10 se puede observar la configuración del Dataset generalizado y las condiciones de las cámaras IP (Xiaomi MI 360°).

Tabla 9. Parámetros técnicos para optimizar recursos de CPU y GPU.

Variable	Valor	Beneficio
INFERENCE_IMG_SIZE	416	**Optimización de recursos:** Achicar la imagen a 416 píxeles hace que el modelo procese la información más rápido que el tamaño normal (640 píxeles), lo cual es perfecto para dispositivos con poca capacidad.
CONF_ACC	0.5	**Filtro de precisión:** Se usa un nivel de confianza del 50% para los accesorios, lo que evita errores sin dejar de detectar objetos pequeños.
CONF_PER	0.6	**Estabilidad de rastreo:** Un nivel de confianza mayor para las personas asegura que el sistema solo siga objetivos claramente visibles, lo que evita gastar energía en sombras o interferencias.
SKIP_N	2	**Rendimiento de CPU/GPU:** El sistema solo hace un procesamiento intensivo cada dos imágenes. En teoría, esto duplica la velocidad de procesamiento (FPS) sin perder fluidez visual.
ROI_PAD	50	**Precisión de contexto:** Se añade un borde de 50 píxeles al área de la persona. Esto permite que el modelo de accesorios vea "un poco más allá" del cuerpo, detectando mochilas o prendas que sobresalgan.

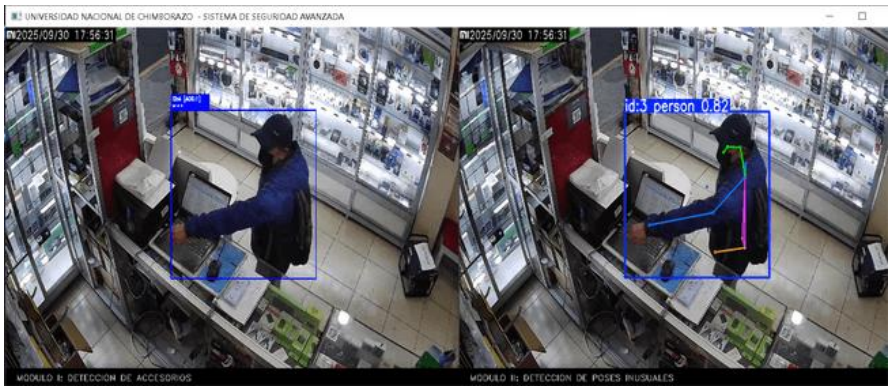


Figura 29. Optimización para mejorar el rendimiento de GPU y CPU.

Fuente: Autor

Tabla 10. Configuración del Dataset y condiciones de cámara

Categoría	Especificación
Número de Clases	7 (Gafas, Gorra, Mascarilla, Casco, Pasamontañas, Mochila, Capucha)
Proceso de Etiquetado	Anotación manual en plataforma Roboflow
División del Dataset	Entrenamiento 90%, Validación 6%, Pruebas 4%
Fuente de Datos	Capturas propias en entorno de simulación (Global Cell)
Resolución de Captura	1080p (Full HD)
Tipo de Cámara	Cámaras IP (Xiaomi MI 360°)
Ángulo de Inclinación	~30° respecto al plano horizontal
Umbral de Confianza	Variable según módulo (0.5 para accesorios, 0.6 para personas/pose)

3.5.9. Configuración del entorno y carga de modelos

Durante esta fase preliminar, se establecen los parámetros de seguridad fundamentales, las credenciales necesarias para la comunicación con la interfaz de programación de aplicaciones (API) de Telegram, se procede a la carga de los tres modelos de visión artificial indispensables para el análisis multiproceso como se observa en la tabla 11.

Tabla 11. Propósito de modelos entrenados para el funcionamiento del software.

Variable	Modelo	Propósito Técnico	Beneficio
m_per	**yolo11n.pt**	**Identificación de Personas:** Un modelo fundamental que ha sido instruido para reconocer seres humanos.	**Agilidad:** Siendo la variante nano, su velocidad es altísima, logrando detectar elementos en fracciones de segundo.
m_acc	**MODEL_ACC_PATH (Pre-entrenado)**	**Clasificación de Accesorios:** Un modelo específico (que fue entrenado) para las siete categorías de prendas.	**Especialización:** Al emplear un modelo a medida (besty11s.pt), el programa identifica elementos que un modelo común no detectaría, como gorros o capuchas.
m_pose	**yolo11n-pose.pt**	**Estimación de Pose:** Un modelo diseñado con una estructura de puntos de referencia importantes.	**Análisis Biométrico:** No solo reconoce la presencia de un individuo, sino que también localiza 17 puntos clave en el cuerpo (como los ojos, hombros y muñecas), lo que hace posible determinar sus ángulos y posturas.
persist=True	**Parámetro en .track()**	**Rastreo de Objetos:** Asigna una identificación exclusiva a cada individuo que se encuentra.	**Continuidad:** Previene que el sistema pierda el rastro de la identidad de cada persona. Si un individuo cambia de posición, el sistema sigue relacionando su información de elementos y avisos con su identificador exclusivo.

- **YOLO11n:** El modelo se emplea para el seguimiento de personas mediante un cuadro ROI, como se observa en la figura 30.

Figura 30. Detección automática de una persona en un cuadro ROI.

Fuente: Autor

- **Besty11s.pt:** Constituye un modelo entrenado, personalizado y diseñado para la detección de accesorios, tales como gafas, gorras, mascarillas, como se puede apreciar en la figura 31.

Figura 31. Detección de un accesorio (gorra) dentro del cuadro ROI.

Fuente: Autor

- **YOLO11n-pose:** Modelo especializado, cuya función principal es la extracción de puntos clave del cuerpo humano para la lectura de poses como se puede observar en la figura 32.

Figura 32. Detección de pose (manos arriba) dentro del cuadro ROI.

Fuente: Autor

3.5.10. Parte I: Detección y seguimiento de accesorios

Este apartado se basa en la identificación de individuos y al análisis de si estos portan elementos que puedan encubrir su identidad. El procedimiento emplea una técnica de Región de Interés (ROI) dinámica, tras la detección de una persona, el sistema delimita el área de su cuerpo para aplicar el segundo modelo de detección, optimizando de este modo los recursos computacionales.

- **Persistencia**

Se emplea un registro histórico por identificador ID, para retener información sobre los complementos utilizados por una persona a lo largo del tiempo, lo cual previene falsos negativos derivados de oclusiones transitorias como se aprecia la figura 33.

Figura 33. Registro histórico por identificador ID.

Fuente: Autor

- **Lógica de alerta**

En caso de que un individuo exceda el límite de MIN_ACC (tres accesorios), se le clasifica como sospechoso cambiando el color del cuadro ROI, como se observa en la figura 34.

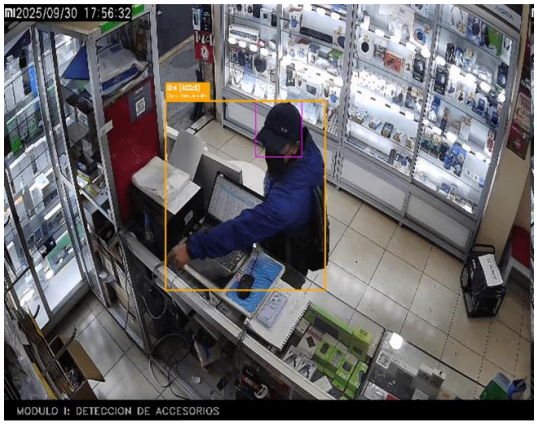


Figura 34. Registro de sospechoso en color naranja.

Fuente: Autor

3.5.10.1 Algoritmo de identificación de accesorios

Este apartado opera mediante un procedimiento de detección en cascada. Su propósito es identificar individuos y determinar si estos poseen elementos que dificulten apreciar su identidad (tales como gafas, mascarillas o capuchas), empleando una lógica de acumulación de pruebas para cada sujeto rastreado.

- **Seguimiento de personas (Tracking) y región de interés (ROI)**

La fase inicial emplea el modelo yolo11n.pt con el fin de localizar a los individuos dentro del fotograma. Una vez que una persona es detectada y se le asigna un identificador único (tid), el software procede a generar una Región de Interés (ROI).

- **Optimización mediante relleno (Padding)**

Se aplica un margen adicional (ROI_PAD = 50) alrededor del recuadro delimitador de la persona. Esto garantiza que, incluso ante un movimiento rápido del sujeto, el modelo de accesorios disponga de suficiente entorno visual para detectar elementos situados en los márgenes (como mochilas o capuchas) como se observa en la figura 35.

Figura 35. Margen adicional para la detección de mochilas.

Fuente: Autor

- **Detección especializada de accesorios (inferencia en cascada)**

Una vez que se obtiene la imagen recortada (ROI), esta ingresa en el segundo modelo neuronal previamente entrenada (besty11s.pt). Este modelo ha sido preparado específicamente para la clasificación de siete categorías de accesorios sospechosos.

- **Precisión localizada**

Al llevar a cabo la inferencia sobre el recorte del objetivo y no sobre la imagen completa de la cámara, la precisión se incrementa significativamente, dado que el modelo se concentra en los pequeños detalles de la vestimenta como el tipo de accesorios.

- **Memoria histórica y lógica de alerta SOS**

Para evitar la pérdida de un accesorio de la vista si el individuo cambia de orientación, el software implementa un diccionario de persistencia (historial) para no perder de vista el objetivo.

- **Acumulación de evidencias**

Cada vez que se detecta un accesorio nuevo, este se incorpora a un conjunto (set) vinculado al identificador del individuo, permitiendo así realizar la suma de los tres accesorios para dar como posible sospechoso.

- **Umbral de peligrosidad**

Sólo cuando el número de accesorios distintos de la variable (n_acc) es igual o superior a la variable (MIN_ACC) (tres elementos), el cuadro de seguimiento cambia de azul a naranja, lo que indica un nivel crítico de sospecha como se observa en la figura 36.

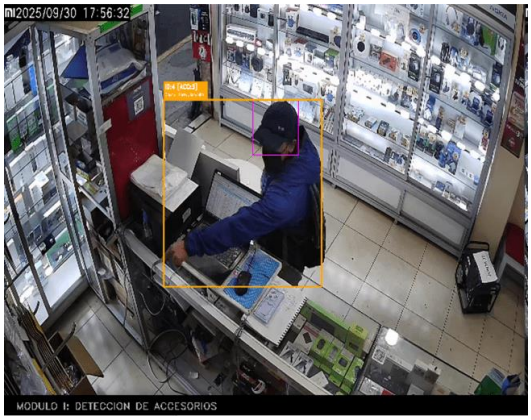


Figura 36. Nivel crítico de sospecha en color naranja.

Fuente: Autor

- **Integración del mosaico de monitoreo**

Finalmente, los resultados de ambos sistemas (Prendas y Poses) se unifican en un mosaico dinámico. Se utiliza una barra de estado inferior para etiquetar cada sección, facilitando la interpretación del operador de seguridad como se puede observar en la figura 37.

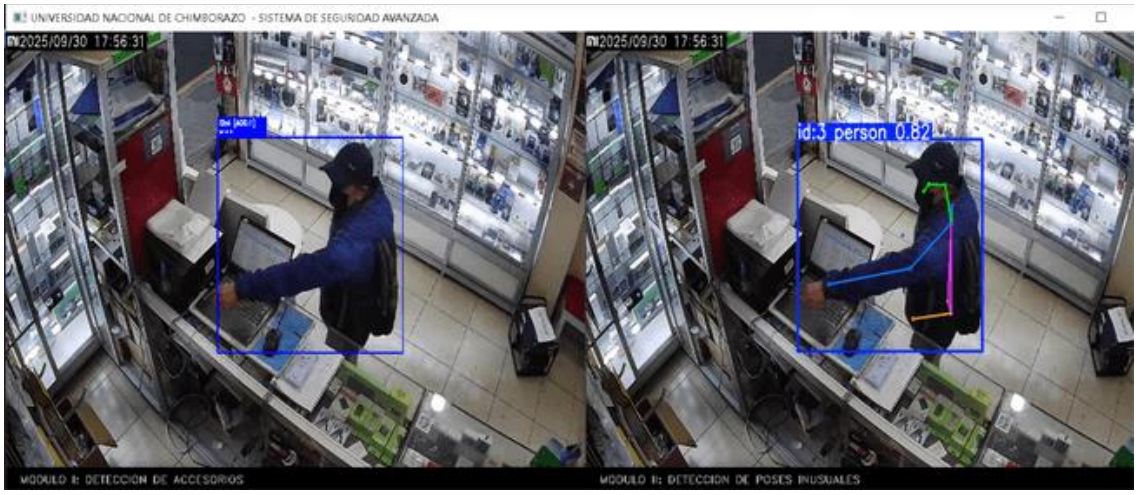


Figura 37. Mosaico dinámico unificado.

Fuente: Autor

3.5.11. Parte II: Análisis biomecánico de posturas de alerta

Esta sección representa una de las fases más avanzadas del proyecto. No se limita a la detección de objetos, sino que procede al análisis de la geometría corporal mediante la segmentación de puntos clave (keypoints). El programa calcula la distancia relativa entre las muñecas (lw_y, rw_y) y la parte superior de la cabeza (top_head_y), este sistema emplea una detección por umbral de plano vertical, optimizando la velocidad de respuesta en tiempo real.

- **Manos a la cabeza**

  Se identifica esta postura si las manos se encuentran próximas a la cabeza (un gesto comúnmente asociado con rendición, agresión o protección) como se observa en la figura 38.

Figura 38. Integración de poses, manos sobre la cabeza.

Fuente: Autor

- **Manos arriba**

  Se clasifica de esta manera si las manos superan la altura de la cabeza como un gesto de prevención como se observa en la figura 39.

Figura 39. Integración de poses, manos arriba.

Fuente: Autor

3.5.11.1 Cinemática corporal para detección de amenazas

El sistema utiliza el modelo YOLO11-pose, el cual identifica 17 puntos clave en el cuerpo humano. Para la lógica de seguridad, el algoritmo se centra en la relación espacial entre la región craneal y las extremidades superiores (muñecas).

- **Definición del marco de referencia (Cabeza)**

En lugar de usar un punto fijo, el software calcula dinámicamente el tamaño y posición de la cabeza para adaptarse a la distancia de la cámara. El algoritmo identifica los 5 puntos faciales (ojos, nariz, oídos) para establecer una métrica de referencia denominada head_h. Esta variable es fundamental, ya que actúa como una unidad de medida dinámica, es decir, si el sujeto se aleja, head_h disminuye; si se acerca, aumenta. Esto permite que el sistema sea invariante a la distancia, ver figura 40.

Figura 40. Extracción de las métricas de referencia (ojos, nariz, oídos).

Fuente: Autor

- **Punto superior (top_head_y)**

Se obtiene mediante el valor mínimo de la coordenada Y de los puntos faciales, representando la coronilla.

- **Altura de la cabeza (h_head)**

Se calcula como la diferencia entre el punto más bajo y el más alto de la cara, añadiendo un factor de corrección (padding) de 10 píxeles.

- **Extracción de coordenadas de las muñecas**

Se monitorizan los puntos de las muñecas izquierda (ID 9) y derecha (ID 10). Para asegurar la fiabilidad del sistema, se aplica un filtro de confianza.

- **Umbral de confianza (c > 0.5)**

Sólo si el modelo está más del 50% seguro de la ubicación de la muñeca, se procede al cálculo. Esto evita falsas alarmas por oclusiones o ruido en la imagen.

- **Lógica algorítmica de clasificación**

El sistema evalúa la distancia relativa $dist = Y_{wrist} - Y_{top}$ (muñecas) para clasificar la postura en tres estados posibles:

- **A. Postura de manos arriba (Amenaza/Rendición total)**

Se activa si ambas muñecas están físicamente por encima del límite superior de la cabeza, es decir en coordenadas de imagen, un valor de Y menor indica mayor altura. Si la distancia es negativa y supera la altura de la cabeza, se clasifica como una postura de rendición o amenaza alta como se muestra en la figura 41.

$$Y_{wrist} < (Y_{top} - head_h) \quad (1)$$

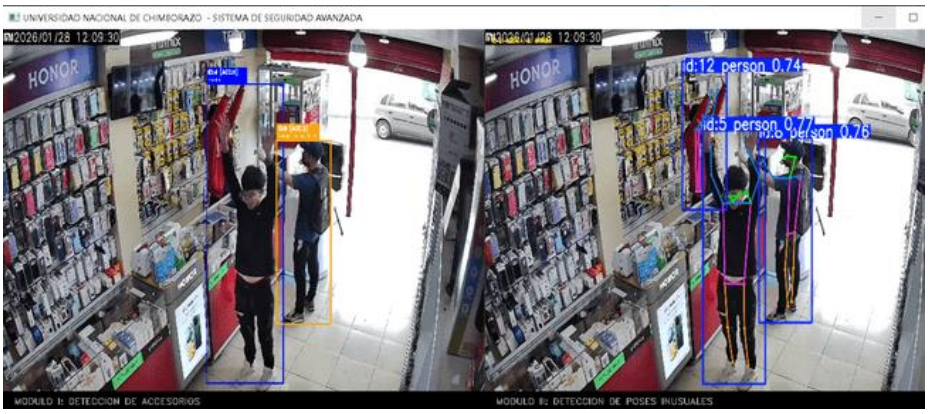


Figura 41. Integración de pose, manos arriba.

Fuente: Autor

- **B. Postura de manos a la cabeza (Rendición/Protección)**

El proceso de detección de manos a la cabeza se basa en comparar la ubicación vertical de las muñecas (lw_y , rw_y) respecto al límite superior de la cabeza top_{head_y}. Se calcula la diferencia de altura ($dist$) y se aplica un valor absoluto. Si el resultado es menor al 90% de la altura de la cabeza ($head_h \times 0.9$), el sistema interpreta que las manos han entrado en el mismo plano horizontal que el rostro como se observa en la figura 42.

$$|Y_{wrist} - Y_{top}| < (head_h \times 0.9) \quad (2)$$

Esta condición detecta cuando la persona tiene las manos pegadas a las sienes o nuca, gesto común en situaciones de asalto.

Figura 42. Integración de pose, manos sobre la cabeza.

Fuente: Autor

- **Gestión de alertas y persistencia visual**

Para evitar el parpadeo de alertas de falsos positivos momentáneos, el sistema almacena el estado en ($pose_{alerts\_history}$). Esto permite que la alerta se dibuje de forma estable sobre ($\text{annotated}_{\text{frame2}}$) y se envíe a Telegram mediante un hilo secundario, garantizando que la evidencia quede registrada sin detener el procesamiento de video.

3.5.12. Parte III: Notificaciones de alerta en tiempo real

Para la operatividad del sistema en un entorno práctico, se ha implementado un proceso de ejecución independiente mediante un token ID que despacha alertas automáticas a un bot de Telegram como se observa la figura 43 y 44. Esta funcionalidad permite que el personal de seguridad o el administrador reciban la evidencia visual, como la captura de pantalla y la identificación del individuo sospechoso sin interrupción del flujo de video en tiempo real.

```
8   #CONFIGURACIÓN INTEGRADA TOKEN ID PARA ALERTAS DIRECTAS A TELEGRAM
9   TELEGRAM_TOKEN = "7568383590:AAHUOQM_ndxUHJjpnFYHck93XQx4Qb08Vlo"
10  TELEGRAM_CHAT_ID = "6139265200"
```

Figura 43. Token ID de Telegram para el envío de Alertas.

Fuente: Autor

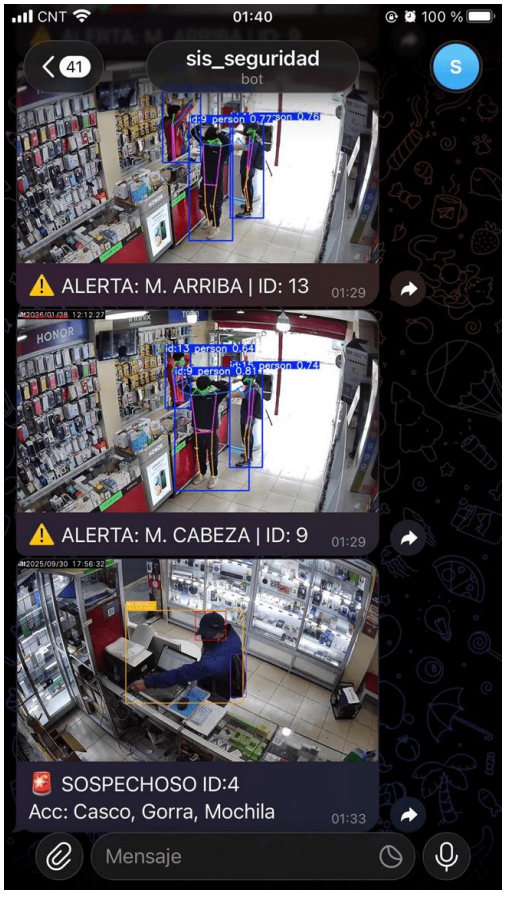


Figura 44. Recepción de imágenes y alertas en Telegram.

Fuente: Autor

3.5.12.1 Renderizado y mosaico de monitorización final

Finalmente, el programa unifica ambos flujos de análisis en una única interfaz visual. Emplea técnicas de manipulación de matrices con NumPy para apilar los fotogramas de video, lo que posibilita la visualización simultánea de la detección de accesorios y el análisis de posturas, ver figura 45.

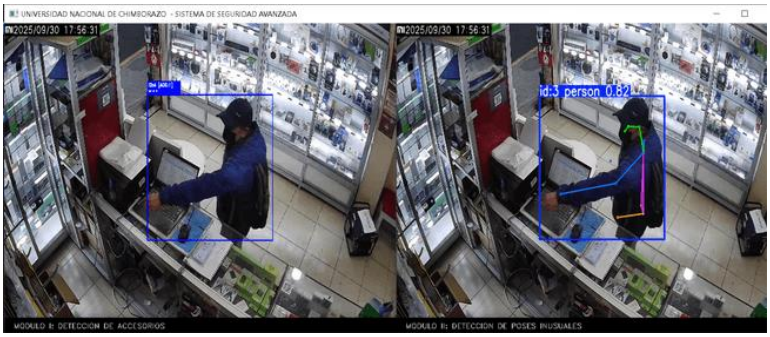


Figura 45. Resultado de la unificación de los multiprocesos en una sola ventana.

Fuente: Autor

Se eligió analizar los umbrales en el eje Y porque es computacionalmente eficiente. Al usar videos de alta resolución y varios modelos YOLO al mismo tiempo para accesorios y posturas, este método ayuda a mantener una velocidad de fotogramas constante. Así, podemos monitorear en tiempo real sin perder exactitud al identificar posturas comunes de agresión, el sistema de seguridad cuenta con diferentes etapas como se muestra en el diagrama de flujo de la Figura 41, y se detalla la respectiva fase a continuación.

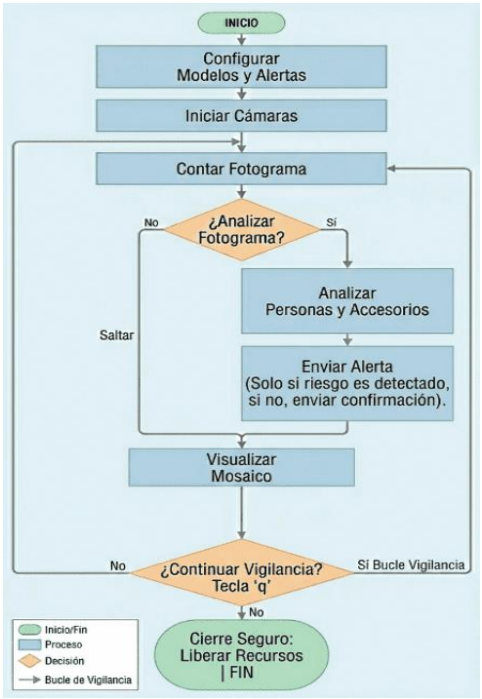


Figura 46. Diagrama de flujo del sistema de seguridad avanzado.

Fuente: Autor.

- **Configurar modelos y alertas**

Primero, el programa carga en la memoria RAM los modelos del sistema los archivos YOLOv11s, YoloPose y besty11s.pt y conecta con el servidor de Telegram.

- **Iniciar cámaras**

El sistema enciende el módulo de video de las cámaras de seguridad. Este paso activa la visualización del programa, de modo que las imágenes puedan capturarse y tratarse al instante.

- **Contar fotogramas y decisión de análisis**

Aquí ocurre la optimización. El sistema cuenta los fotogramas que recibe.

  - Si es una imagen de salto, la descarta para el análisis detallado, ahorrando recursos de la computadora.
  - Si es una imagen de análisis, el sistema continúa con la parte de visión artificial. Esto impide que el programa se ralentice o se bloquee.

- **Analizar personas y accesorios**

Cuando el sistema opta por analizar, emplea los modelos cargados para localizar personas y elementos específicos como mochilas, cascos o pasamontañas y evaluar la posición del cuerpo.

- **Enviar alerta**

Este punto funciona como un filtro inteligente. Solo cuando el sistema detecta una combinación peligrosa, por ejemplo, 3 accesorios sospechosos o una postura amenazante como manos arriba o manos sobre la cabeza, envía un mensaje de alerta a Telegram.

- **Visualizar mosaico**

El sistema une todos los videos procesados en una sola pantalla para que el usuario pueda ver toda la información en un solo lugar.

- **Continuar vigilancia**

El programa mantiene el monitoreo activo todo el tiempo hasta que el usuario aprieta la tecla 'q'. Así se puede vigilar sin pausas durante el tiempo que se necesite.

- **Cierre seguro y liberar recursos**

Cuando termina, el sistema hace un cierre ordenado. Esto es importante para liberar la memoria RAM y la GPU, evitando problemas con el sistema operativo y asegurar que las cámaras se desconecten bien.

3.5.13. Fase 3: Verificación de la eficacia del sistema

En esta última etapa se realizaron tareas que aseguran el cumplimiento de la fase de pruebas del prototipo implementado. Estas tareas se explican a continuación:

3.5.13.1 Planificación de ensayos piloto

Abarca planificación, preparación del personal, implementación, recolección de datos medibles y el análisis comparativo posterior sobre la eficacia de ambos sistemas de seguridad usando un método cualitativo y cuantitativo.

La planificación se enfocó en crear un conjunto de 80 pruebas aleatorias controladas, organizadas en dos grupos separados de 40 ensayos cada uno, 40 para el sistema con inteligencia artificial (IA) y 40 para el método tradicional (TD) manejado por un monitorista.

- **Propósito**

Medir la precisión general del sistema frente a señales de peligro reales para alertas positivas y casos normales de control como alertas negativas.

- **Criterio de evaluación**

Se estableció que el resultado de cada prueba se juzgaría mediante una regla de coincidencia simple. La eficiencia total del Sistema (εf) valdrá 1 únicamente cuando el acierto del sistema evaluado (A) sea idéntico a la clasificación correcta del caso (C). Se expresa a continuación:

$$\varepsilon f = \begin{cases} 1 \; si \; C \; = A \\ 0 \; si \; C \neq A \end{cases} \tag{3}$$

3.5.13.2 Capacitación del personal para uso del sistema

Antes de realizar las pruebas, se organizó una capacitación técnica para el personal en general, quienes iban a someterse ante las pruebas y simulaciones necesarias.

- **Sistema con IA:**

Se realizó una enseñanza técnica sobre el sistema automático de alertas. Los usuarios internos de Global Cell aprendieron a reconocer las advertencias iniciales que envía automáticamente el sistema con inteligencia artificial.

- **Método tradicional**

Repaso de las normas actuales sobre seguridad física, asegurando que los usuarios siguieran los métodos habituales de detección visual de riesgos y el uso de medios de comunicación tradicionales sin ayuda de tecnología externa.

3.5.13.3 Ejecución de los ensayos piloto

La prueba se realizó de manera secuencial, registrando el comportamiento real de cada grupo experimental de forma independiente:

- **Sistema con IA**

Se llevaron a cabo 40 pruebas simuladas en total. Entre ellas se encontraban 27 situaciones de peligro real (Alertas Positivas) y 13 situaciones de prueba creadas para medir falsas alarmas (Alertas Negativas). El sistema con Inteligencia Artificial funcionó muy bien al detectar correctamente 26 de las 27 amenazas verdaderas. No obstante, reaccionó de manera sensible a algunas situaciones de prueba, el sistema de visión artificial activó una alerta de sospecha, lo que significa que baja de manera notoria la eficiencia al dar falsas alarmas.

- **Método tradicional**

No obstante, el sistema tradicional examinó las mismas situaciones, pero su juicio se vio afectado por errores humanos. A diferencia de la inteligencia artificial, el monitor fue muy asertivo ignorando eventos no importantes, pero ineficiente detectando peligros reales, alertando solo en 24 de los 40 casos verdaderos. El sistema tradicional falló gravemente por descuido, cansancio, obstrucciones visuales y cosas que lo distrajeron en la mayoría de las alertas importantes donde debía haber actuado, pero no lo hizo.

3.5.13.4 Monitoreo y procesamiento de datos

Por consiguiente, se plasman las respectivas pruebas realizadas con respecto a los sistemas convencionales y el sistema implementado con visión artificial, se demuestra la destreza y eficiencia del sistema con respecto a condiciones específicas tales como la distancia de las cámaras frente a los usuarios, la iluminación mediante la distancia y enfoque de cada cámara, el número de repeticiones por simulación. Estos datos se los aprecia en la tabla del anexo C. La información se organizó cuidadosamente de manera cualitativa, puesto que los eventos que se asignaron fueron rigurosamente analizados, fueron ejecutados en una comparativa independiente para los dos sistemas, se puede observar el anexo D, donde se aprecian los múltiples escenarios organizados. Además, para el análisis de manera eficaz y obtener resultados eficientes, se organizó de manera binaria en una nueva tabla, se ubicaron los datos en un apartado como matriz consolidada de simulaciones y eficiencia del sistema, uniendo las respuestas de ambos sistemas, como se aprecia en el anexo E.

Una vez recolectado los datos y la información necesaria se obtuvo el margen de error y eficiencia de los dos sistemas. Con esto se realizó la respectiva comparación en cuanto al margen de error y la eficiencia final con las correspondientes pruebas estadísticas aplicando Chi cuadrado.

CAPÍTULO IV

4. RESULTADOS Y DISCUSIÓN

Para saber qué tan bien funcionó el modelo, se revisó una tabla denominada matriz de confusión. La tabla 12 muestra qué tan bien el modelo clasificó correctamente cada uno de los siete tipos de accesorios. La figura 47 y la figura 48 representan, respectivamente, los valores absolutos y la normalización de las predicciones del sistema frente a las etiquetas reales del conjunto de validación.

La diagonal principal de la matriz evidencia una alta tasa de aciertos, lo que significa que identificó correctamente la mayoría de las clases como gafas, gorra, mascarilla, casco, pasamontañas, mochila y capucha. Los valores situados fuera de esta diagonal, que representan los errores de clasificación, son muy pequeños. Esto indica que la red neuronal está bien entrenada. Al convertir la tabla a porcentajes, se confirmó que el modelo detectó clases importantes para la seguridad, como el casco y la capucha, con una precisión superior al 98%. Esto demuestra que el sistema es efectivo ante situaciones peligrosas.

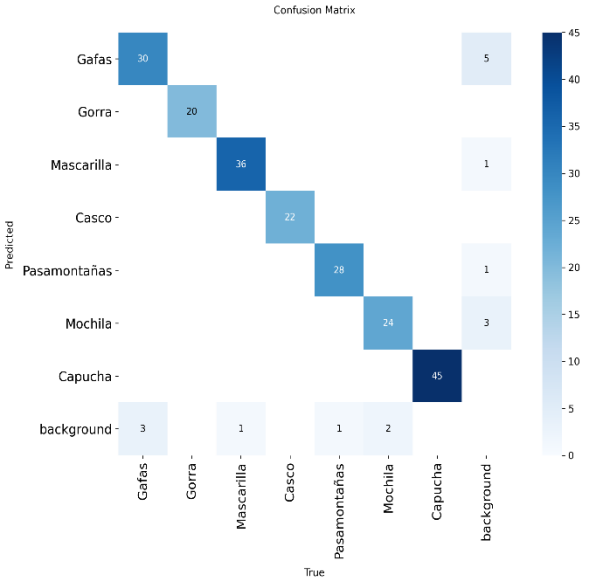


Figura 47. Matriz de confusión.

Fuente: Autor

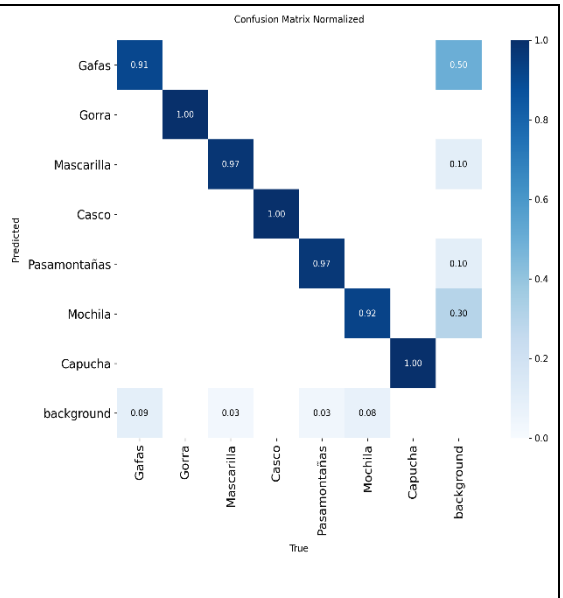


Figura 48. Matriz de confusión normalizada.

Fuente: Autor

4.1. Métricas de rendimiento del modelo

Para complementar el análisis, se plasman los indicadores cuantitativos de desempeño extraídos directamente del proceso de validación (val) del modelo, como se observa en la tabla 12.

Tabla 12. Métricas de rendimiento del modelo.

Clase	Precisión (P)	Recall (R)	mAP50	mAP50-95
Gafas	0.960	0.909	0.974	0.522
Gorra	0.981	1.000	0.995	0.862
Mascarilla	0.941	0.946	0.934	0.631
Casco	0.981	1.000	0.995	0.944
Pasamontañas	0.950	0.966	0.966	0.770
Mochila	0.882	0.867	0.850	0.696
Capucha	0.993	1.000	0.995	0.747
Promedio Global	**0.956**	**0.955**	**0.958**	**0.739**

El modelo alcanzó una precisión promedio (mAP50) de 0.958, lo que demuestra una fiabilidad superior al 95% en la detección de accesorios en entornos controlados. La paridad entre la Precisión (0.956) y el Recall (0.955) indica un punto de equilibrio óptimo en el entrenamiento donde el sistema logra detectar la gran mayoría de objetos, evitando los falsos negativos. Estos hallazgos técnicos demuestran que la red neuronal implementada es adecuada para usarla en sistemas de seguridad que necesitan estar activos todo el tiempo.

- **Evaluación operativa y latencia**

Debido a la naturaleza del sistema de seguridad implementado, la velocidad de respuesta es regular. La evaluación operativa realizada durante los 40 escenarios simulados arrojó los siguientes indicadores de desempeño en tiempo real, ver tabla 13.

Tabla 13. Métricas evaluadas en el sistema.

Métrica	Valor	Método de Medición
FPS (Inferencia)	14 FPS	Promedio en bucle de ejecución (Visual Studio Code).
Latencia de Telegram	1.2 s	Cronometraje del envío asíncrono vía API.
Falsos Positivos (FP)	6	Registro en bitácora de pruebas (Alerta errónea).
Falsos Negativos (FN)	4	Registro en bitácora de pruebas (Omisión de evento).

Las métricas evaluadas muestran que el sistema es mejor que el sistema tradicional. Aunque el sistema a veces identifica erróneamente acciones como ajustarse la gorra o estirarse como amenazas, es mejor que sea medianamente sensible en lugares donde la seguridad es lo primero, para no pasar por alto escenarios reales. La comunicación rápida por Telegram, que tarda 1.2 segundos, permite que el personal de seguridad o las autoridades actúen de inmediato.

4.2. Hipótesis
4.3. Prueba de hipótesis

Sabiendo que la Eficiencia Final (εf) es la variable principal que se está estudiando. Esta variable es de tipo binario, lo que significa que sólo puede ser 1 (si es correcta) o 0 (si es incorrecta). Las pruebas de normalidad sólo se aplican a variables cuantitativas continuas por ejemplo el tiempo en segundos o un porcentaje exacto. Una variable que sólo tiene dos valores, 0 y 1, nunca podrá ser normal. Por lo tanto, no hay razón para realizar una prueba de normalidad en este caso.
Dado que se va a comparar la eficiencia entre dos grupos independientes (Sistema IA vs Sistema Tradicional), la prueba no paramétrica estándar y correcta por excelencia es la Prueba de Chi-cuadrado de Independencia (x^2) o, en su defecto, el test exacto de Fisher. Siendo la hipótesis planteada la siguiente:
$\boldsymbol{H_o}$**:** No existe diferencia significativa en la tasa de aciertos entre el sistema de seguridad basado en visión artificial y el sistema tradicional en entornos controlados.

$\boldsymbol{H_i}$**:** Existe diferencia significativa en la tasa de aciertos entre el sistema de seguridad basado en visión artificial y el sistema tradicional en entornos controlados, siendo el sistema propuesto superior en la detección de eventos.

Al analizar el desempeño operativo en 80 pruebas de la empresa Global Cell, se logró observar que el sistema con inteligencia artificial acertó el 75% de las veces. Esto significa que, de 40 alertas verdaderas, procesó correctamente 30 y falló en 10. Por otro lado, el sistema tradicional, que usa supervisión de usuarios, solo acertó el 40% de las veces, con 16 aciertos y 24 errores como se observa en la tabla 14.

Tabla 14. Resultado porcentual obtenido en ambos sistemas.

Tabla cruzada TIPO_SISTEMA*Eficiencia_Final					
			Eficiencia_Final		Total
			FALLO	**ACIERTO**	
TIPO_SISTEMA	**IA**	Recuento	10	30	40
		% dentro de TIPO_SISTEMA	**25,0%**	**75,0%**	**100,0%**
	TRADICIONAL	Recuento	24	16	40
		% dentro de TIPO_SISTEMA	**60,0%**	**40,0%**	**100,0%**
Total		Recuento	34	46	80
		% dentro de TIPO_SISTEMA	42,5%	57,5%	100,0%

Para ver si la diferencia en la efectividad entre ambos sistemas es estadísticamente significativa, se aplicó la prueba no paramétrica de Chi-cuadrado de Independencia (χ^2) en una tabla de contingencia de 2x2.
Las muestras de las pruebas fueron satisfactorias, dado que el 0.0% de las casillas tuvieron un recuento esperado menor que 5, siendo el recuento mínimo esperado de 17.00. El resultado de la prueba Chi-cuadrado fue de 10.026, con 1 grado de libertad como se observa en la tabla 15. Dado que el valor de significación asintótica bilateral es sustancialmente

menor al nivel de alfa establecido en la investigación ($p = 0.002 < \propto = 0.05$), se puede decir con seguridad que se rechaza la hipótesis nula y se acepta la hipótesis alternativa.

Tabla 15. Resultados obtenidos mediante la prueba de Chi-Cuadrado.

Pruebas de chi-cuadrado					
	Valor	**gl**	**Significación asintótica (bilateral)**	**Significación exacta (bilateral)**	**Significación exacta (unilateral)**
Chi-cuadrado de Pearson	10,026[a]	1	,002		
Corrección de continuidad[b]	8,645	1	,003		
Razón de verosimilitud	10,269	1	,001		
Prueba exacta de Fisher				,003	,002
Asociación lineal por lineal	9,900	1	,002		
N de casos válidos	80				
a. 0 casillas (0,0%) han esperado un recuento menor que 5. El recuento mínimo esperado es 17,00.					
b. Sólo se ha calculado para una tabla 2x2					

Hay una diferencia considerable y medible en qué tan bien los dos sistemas detectan comportamientos sospechosos y previenen asaltos. El sistema de visión artificial mejora de forma clara y comprobable la seguridad de la empresa. Además, los resultados de las pruebas estadísticas como la prueba exacta de Fisher o la Corrección de Continuidad, fueron de .003, lo cual es mucho menor que 0.05 como se observa en la tabla 15, lo que confirma totalmente la eficiencia final del sistema con IA.

4.4. Discusión

Los resultados del análisis estadístico no paramétrico, usando la Prueba de Chi-cuadrado de Pearson, muestra la existencia de diferencias estadísticamente significativas en las detecciones ($p = 0.002$), El sistema basado en inteligencia artificial logró una precisión del 75.0%, lo que representa una mejora significativa en comparación con el método tradicional, que solo alcanzó un 40.0% de éxito, como se muestra en la Figura 49. Estos resultados apoyan la idea inicial de que la automatización con modelos YOLO mejora mucho la detección de comportamientos sospechosos, haciendo la vigilancia más fiable, exacta y sin el cansancio que afecta al personal con el método tradicional. Esta gran mejora técnica ayuda a lograr el objetivo principal del proyecto: mitigar delitos en la empresa de telecomunicaciones Global Cell. Al hacer que las respuestas sean más rápidas con alertas automáticas e inmediatas a Telegram, se disminuye la debilidad del lugar ante delitos, lo que protege los bienes de la empresa y mejora la seguridad y el bienestar de los empleados y clientes.

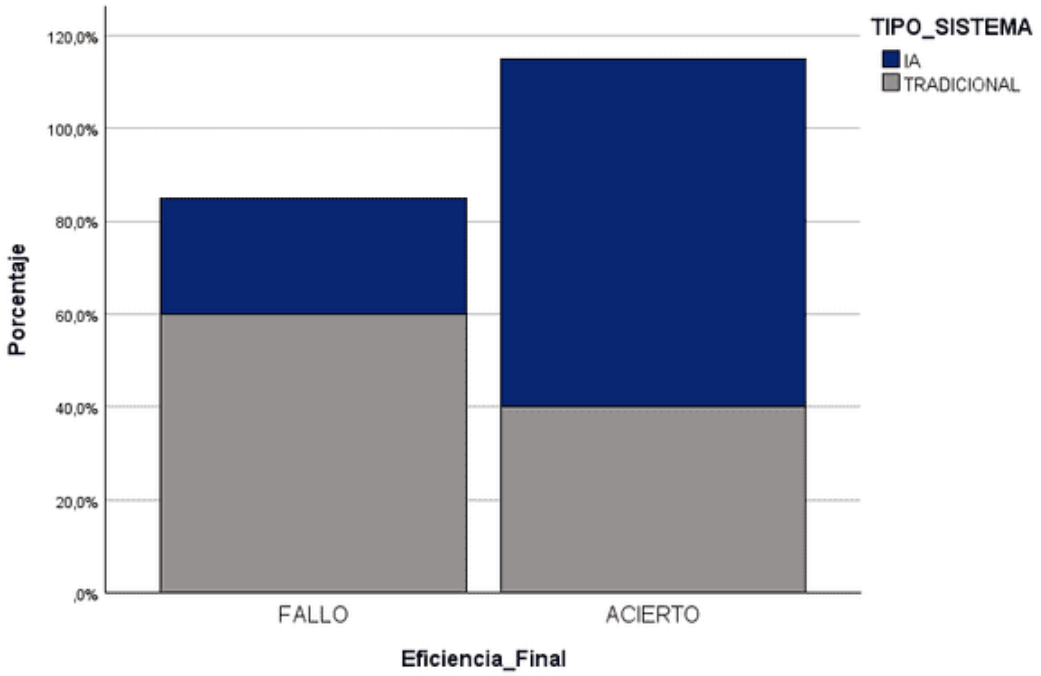


Figura 49. Eficiencia obtenida de manera porcentual de los dos sistemas.

El sistema que se implementó es mejor que el método tradicional (40.0%) porque puede procesar varios datos a la vez. Los métodos de vigilancia tradicional dependen de que una persona esté atenta al monitor, pero esta atención disminuye con el tiempo por el cansancio visual, oclusiones etc. En cambio, el sistema con visión artificial modelo YOLOv11 garantiza un monitoreo constante. A diferencia de los estudios de Villa Yuquilema, 2024 [10], que tuvieron problemas para la detección de armas con poca luz, el sistema implementado resuelve esto usando un modelo propio (besty11s.pt) y guardando imágenes hasta tener suficientes pruebas antes de alertar. Esto disminuye en gran cantidad las alarmas falsas, que son comunes en sistemas más simples.

CAPÍTULO V

5. CONCLUSIONES Y RECOMENDACIONES

5.1. CONCLUSIONES

- A partir del análisis de las técnicas y algoritmos de visión artificial, se determinó que los modelos de aprendizaje profundo, específicamente la arquitectura YOLOv11, son una solución técnica altamente eficaz para la detección automatizada de eventos sospechosos en escenarios simulados. Se determinó que los modelos de aprendizaje profundo como YOLOv11, son la mejor opción para un entorno comercial. Esto se debe a que pueden procesar varios escenarios al mismo tiempo como accesorios y la estimación de pose sin afectar la atención laboral de la sucursal, asegurando que todo sea automático, rápido y exacto. La investigación teórica y técnica confirman que esta tecnología es una herramienta efectiva para mejorar los sistemas de seguridad tradicionales, ofreciendo una solución sólida para prevenir delitos de manera anticipada.

- Se logró diseñar y construir un modelo funcional de un sistema de visión por computadora que funciona mediante cámaras de seguridad en la sucursal de Global Cell, siendo esta una empresa de telecomunicaciones. Este sistema, que usa varios procesos al mismo tiempo, puede vigilar automáticamente en tiempo real, detectando con precisión accesorios sospechosos y cambios drásticos en la postura del cuerpo humano como manos arriba o en la cabeza. Además, puede enviar alertas al instante a Telegram. Esto prueba que se lo puede implementar y puede ser utilizado en entornos comerciales importantes.

- Tras la evaluación del sistema y estudiar los resultados, se encontró una gran diferencia en el valor de p de 0.002 según la prueba exhaustiva de Chi-cuadrado de Pearson. Esto confirmó que usar visión artificial hace que las operaciones sean mucho más eficientes, logrando un 75.0% de éxito, a diferencia del 40.0% que se obtuvo con el método tradicional. Estos resultados validan la hipótesis de investigación, demostrando que el sistema impacta positivamente en la calidad de vida de las personas. Se observa que la vigilancia es mucho más confiable y que las amenazas se detectan con precisión, lo que disminuye el índice de atracos frente a errores o falta de atención humana. Por lo tanto, el sistema propuesto es una herramienta tecnológica eficaz, fuerte y anticipada para la seguridad completa de la empresa.

5.2. RECOMENDACIONES

- Se sugiere que el sistema se combine con otras formas de control y datos adicionales, además del aumento del Dataset para mejorar la precisión en la detección de accesorios, además, incluir programas que identifiquen armas haciendo uso de una red neuronal existente, o sistemas que detecten si una persona se queda mucho tiempo en un lugar para notar comportamientos no habituales, buscando crear una plataforma de seguridad más fuerte y adaptable. Así, se tendrían más funciones para encontrar peligros, se reducirían los errores en las alertas y se lograría una mejor exactitud, confianza y alcance en la prevención de atracos en varias sucursales.

- Se recomienda que el sistema de visión por computadora se desplace a tener un diseño de software flexible y adaptable con la integración de un Jetson nano, permitiendo que configuraciones importantes como el nivel mínimo de accesorios o el tiempo de espera entre avisos, se puedan ejecutar fácilmente a través de una interfaz más rápida, sin tocar el código principal. Esto haría que el sistema sea compatible con diferentes tipos de cámaras de seguridad modernas y de mejores características, aumentaría el costo de implementación, pero facilitaría la reutilización y mejora rápida del programa con los avances tecnológicos o nuevas redes neuronales integradas.

BIBLIOGRAFÍA

[1] F. P. para el D. Crime-Lab, "Informe de caracterización del crimen organizado en ecuador," Jan. 2025.

[2] S. P. Bravo López, F. X. Noroña Moreno, V. L. Meneses Segura, A. O. Loor Saldarriaga, and J. R. Flores Tunja, "Impacto del Crimen Organizado y la Inseguridad en el Desempeño Económico del Ecuador: Una Perspectiva Costo-Beneficio," *Ciencia Latina Revista Científica Multidisciplinar*, vol. 9, no. 5, pp. 8109–8121, Oct. 2025, doi: 10.37811/cl_rcm.v9i5.20149.

[3] Beatriz Borella, "Introducción a la visión artificial - procesos y aplicaciones," Madrid, Sep. 2022.

[4] J. Miño, "41 Revista de Ciencias de Seguridad y Defensa," Latacunga, 2016.

[5] R. Khanam and M. Hussain, "YOLOv11: An Overview of the Key Architectural Enhancements," Oct. 2024, [Online]. Available: http://arxiv.org/abs/2410.17725

[6] Impulse, 2018, "Breve historia de la videovigilancia," Impulsecctv.com.

[7] DTIC, 2019, "¿Cómo evolucionó el circuito cerrado de televisión?," Dtic.com.

[8] D. Gutiérrez, "UNIVERSIDAD POLITÉCNICA SALESIANA SEDE CUENCA," Mar. 2021.

[9] S. Lovato, "Sistema de vigilancia inteligente implementado en el local comercial el chaval," Quito, Sep. 2024.

[10] V. Yuquilema Jimmy Andrés and M. Cuzco Silva Edgar Giovanny Riobamba, "UNIVERSIDAD NACIONAL DE CHIMBORAZO FACULTAD DE INGENIERÍA INGENIERÍA EN TELECOMUNICACIONES," Riobamba, Aug. 2024.

[11] I. DE Un Sistema De Control De Postura Y Velocidad En El Patinaje De Velocidad Basado En Procesamiento De Imágenes, G. González Deysi Tamara Muñoz Muñoz Darío Javier Tutor, and L. Fabian Rentería Bustamante Riobamba, "UNIVERSIDAD NACIONAL DE CHIMBORAZO FACULTAD DE INGENIERÍA CARRERA DE ELECTRÓNICA Y TELECOMUNICACIONES.", Riobamba, 2023.

[12] D. Orlando Romero Mogrovejo, "Desarrollo de un sistema de detección de armas de fuego cortas en el monitoreo de videos de cámaras de seguridad," Cuenca, 2018.

[13] J. Carlos *et al.*, "Control de Seguridad y Acceso Vehicular, Mediante Visión Artificial. Security control and vehicular access, through artificial vision," Jun. 2020. [Online]. Available: http://tesis.pucp.edu.pe/repositorio/bitstream/han

[14] M. Guerri, 2021, "Las secuelas psicológicas de sufrir un robo o una intrusión en casa," PsicoActiva.

[15] La Hora, 2019, "La economía y el entorno para hacer negocios son los principales damnificados de la creciente inseguridad," La Hora Periódico Digital.

[16] M. M. Chowdhury, R. S. Ayon, and M. S. Hossain, "An investigation of machine learning algorithms and data augmentation techniques for diabetes diagnosis using class imbalanced BRFSS dataset," *Healthcare Analytics*, vol. 5, Jun. 2024, doi: 10.1016/j.health.2023.100297.

[17] **CLAC LACAC, "Curso de Identificación de Personas con Comportamientos Sospechosos," Oct. 2020.**

[18] **D. Vejar, "SISTEMAS DE SEGURIDAD ELECTRÓNICOS e-Learning," May 2023. [Online]. Available: www.libercap.com**

[19] **NOVASEP, 2026, "¿QUÉ ES UN SISTEMA DE SEGURIDAD INTEGRAL Y EN QUÉ NOS BENEFICIA?," ToGrow.**

[20] **I. Culjak, D. Abram, T. Pribanic, H. Dzapo, and M. Cifrek, 2024, "A brief introduction to OpenCV."**

[21] **J. J. Sanabria and J. F. Archila, "Agosto de 2012," *Scientia et Technica Año XVII*, vol. 51.**

[22] **G. Santillán, I. Danilo, C. Sánchez, and V. Manuel, "Ciencia y Biodiversidad LA VISIÓN ARTIFICIAL Y LOS CAMPOS DE APLICACIÓN ARTIFICIAL VISION AND FIELDS OF APPLICATION", doi: 10.32645/26028131.76.**

[23] **F. Lubinus Badillo, C. A. Rueda Hernández, B. Marconi Narváez, and Y. E. Arias Trillos, "Redes neuronales convolucionales: un modelo de Deep Learning en imágenes diagnósticas. Revisión de tema," *Revista colombiana de radiología*, vol. 32, no. 3, pp. 5591–5599, Sep. 2021, doi: 10.53903/01212095.161.**

[24] **L. Orisel, "Herramienta para la generación de clasificadores haar," La Habana, 2013.**

[25] **Abirami Vina, 2019, "Cómo utilizar Ultralytics YOLO11 para la estimación de la pose," Ultralytics.**

[26] **A. Mustofa, " YoloV8-Pose-Keypoint-Classification." Accessed: Mar. 15, 2026. [Online]. Available: https://github.com/Alimustoofaa/YoloV8-Pose-Keypoint-Classification?tab=readme-ov-file**

[27] **N. Karyna, 2023, "Detección de objetos: Métricas clave para el rendimiento de la visión artificial," Label Your Data.**

[28] **G. Jocher, "Roboflow," Ultralytics . Accessed: Mar. 15, 2026. [Online]. Available: https://docs.ultralytics.com/es/integrations/roboflow/#how-can-i-integrate-and-deploy-yolo26-models-with-roboflow**

[29] **Xiaomi Corporación, 2024, "Mi 360° Camera (1080p)," mi.com.**

ANEXOS

Anexo A. Materiales usados en la implementación del sistema

Para este proyecto, se eligieron materiales basados en la idea de tecnología inclusiva de bajo costo. Esto significa que los materiales son bastante económicos, haciendo el proyecto accesible para cualquier usuario que lo requiera. Los materiales escogidos incluyen:

- Dos cámaras IP de Xiaomi
- Router
- Computador
- Equipo móvil
- Pantalla Externa

Anexo B. Imágenes de instalación de las cámaras de seguridad Xiaomi.

Anexo C. Escenarios y campos con condiciones específicas.

N°	Escenario (Suceso)	Accesorio/Postura	Distancia	Iluminación	Repeticiones	Resultado IA	Resultado Tradicional
1	Asalto móvil	3 accesorios	2m	500 lux	2	Verdadero	Verdadero
2	Intento caja	3 accesorios	2m	500 lux	2	Verdadero	Falso
3	Estrés cajero	Manos cabeza	2m	500 lux	2	Verdadero	Falso
4	Merodeo sospechoso	3 accesorios	4m	300 lux	2	Verdadero	Falso
5	Ingreso sospechoso	3 accesorios	4m	300 lux	2	Verdadero	Falso
6	Repartidor	2 accesorios	4m	300 lux	2	Falso	Falso
7	Robo cuchillo	2 acc. + Pose	2m	500 lux	2	Verdadero	Falso
8	Amenaza arma	Manos cabeza	2m	500 lux	2	Verdadero	Falso
9	Estrés grupal	Manos cabeza	2m	500 lux	2	Verdadero	Falso
10	Ingreso coordinado	3 accesorios	4m	200 lux	2	Verdadero	Falso
11	Intimidación	2 acc. + Pose	2m	500 lux	2	Verdadero	Falso
12	Robo c/arma	1 acc. + Pose	2m	500 lux	2	Verdadero	Falso
13	Robo rápido	Pose (Manos)	2m	500 lux	2	Verdadero	Falso
14	Alcance producto	Manos arriba	4m	300 lux	2	Falso	Verdadero
15	Compra normal	1 acc.	4m	300 lux	2	Falso	Falso
16	Robo clientes	Manos cabeza	4m	300 lux	2	Verdadero	Falso
17	Asalto violento	3 acc. + Pose	2m	500 lux	2	Verdadero	Verdadero
18	Hurto mercancía	2 accesorios	2m	500 lux	2	Falso	Falso
19	Ajuste gorra	Manos cabeza	4m	300 lux	2	Falso	Verdadero
20	Conteo dinero	2 accesorios	4m	300 lux	2	Falso	Falso
21	Sospechoso	3 accesorios	4m	300 lux	2	Verdadero	Falso
22	Amenaza arma	1 acc. + Pose	2m	500 lux	2	Verdadero	Falso
23	Emergencia méd.	Manos cabeza	4m	300 lux	2	Verdadero	Falso
24	Robo vitrina	3 accesorios	2m	500 lux	2	Verdadero	Falso
25	Asalto grupal	3 accesorios	2m	500 lux	2	Verdadero	Falso
26	Estiramiento	Manos arriba	2m	500 lux	2	Falso	Verdadero
27	Ocultamiento	Pose	4m	300 lux	2	Verdadero	Falso
28	Ingreso violento	3 accesorios	4m	300 lux	2	Verdadero	Verdadero
29	Compra normal	1 acc.	2m	500 lux	2	Falso	Falso
30	Forcejeo	Manos cabeza	4m	300 lux	2	Verdadero	Falso
31	Merodeo	3 accesorios	4m	300 lux	2	Verdadero	Falso
32	Acomodo cabello	Manos cabeza	2m	500 lux	2	Falso	Verdadero
33	Robo c/cuchillo	Pose	2m	500 lux	2	Verdadero	Falso
34	Robo a cliente	3 acc. + Pose	4m	300 lux	2	Verdadero	Falso
35	Sudoración	Manos cabeza	2m	500 lux	2	Falso	Verdadero
36	Insulto/Amenaza	3 accesorios	2m	500 lux	2	Verdadero	Falso
37	Arma oculta	Pose	2m	500 lux	2	Verdadero	Falso
38	Juego niño	Manos arriba	4m	300 lux	2	Falso	Verdadero
39	Asalto directo	3 accesorios	2m	500 lux	2	Verdadero	Verdadero
40	Arrinconamiento	Pose	4m	300 lux	2	Verdadero	Falso

Anexo D. Comparativa cualitativa entre los sistemas convencionales y el sistema con visión artificial.

N°	Asignación crítica del escenario	Posible Suceso (Asignación Crítica del escenario)	Respuesta Sistema IA	Respuesta Sistema Tradicional (Monitorista)	Respuesta de los sistemas	Resultado Sistema IA	Resultado Sistema Tradicional	N°
1	Alerta Positiva	**1**	Sujeto ingresa con casco, mochila y mascarilla (3 accesorios), se sustrae un móvil.	**VERDADERO** (Envío automático de alerta a Telegram).	**VERDADERO** (Alertó con señas al resto del personal).	AMBOS	1	1

2	Alerta Positiva	**1**	Sujeto ingresa con Gorra, Mochila y Mascarilla caminando hacia la caja.	**VERDADERO** Envío alerta a Telegram.	**FALSO** No distingue la mascarilla por la ropa de color negra.	SISTEMA IA	1	**0**
3	Alerta Negativa	**0**	Personal de caja levanta las manos sobre la cabeza (síntoma de estrés).	**VERDADERO** (Multiproceso detecta manos arriba).	**FALSO** (no alerta, situación normal).	TRADICIONAL	**1**	**0**
4	Alerta Positiva	**1**	Intruso con Capucha, Gafas y Mochila merodea adentro, observando las esquinas del local.	**VERDADERO** recuadro cambia a naranja de sospecha.	**FALSO** Monitorista distraído atendiendo.	SISTEMA IA	**1**	**0**
5	Alerta Positiva	**1**	Sujeto con casco, pasamontaña y mochila entra durante el día.	**VERDADERO** sospechoso detectado.	**FALSO** (persona normal, motorizado).	SISTEMA IA	**1**	**0**
6	Alerta Negativa	**0**	Repartidor de delivery entra con Casco y Mochila a dejar un pedido.	**FALSO** Ningún proceso	**FALSO** Conoce al repartidor por la rutina diaria.	AMBOS	**0**	**0**
7	Alerta Positiva	**1**	Asaltante entra con gafas y gorra (Solo 2 accesorios), luego saca un cuchillo.	**VERDADERO** (Umbral menor a 3, pero detección de manos arriba).	**FALSO** (Horario de almuerzo, una hora fuera).	SISTEMA IA	**1**	**0**
8	Alerta Positiva	**1**	Cliente en la caja es señalado con un bat	**VERDADERO** Multiproceso activa "Manos arriba".	**FALSO**	SISTEMA IA	**1**	**0**
9	Alerta Positiva	**1**	Personal de caja pone sus manos en la cabeza en momento inusual frente a un grupo de clientes y un asaltante.	**VERDADERO** (Multiproceso, Alerta manos en la cabeza).	**FALSO** Hora de almuerzo rotativo	SISTEMA IA	**1**	**0**
10	Alerta Positiva	**1**	Día de lluvia, dos sujetos con Capucha, Mascarilla y Gafas entran coordinados.	**VERDADERO** Multisistema, accesorios detectados	**FALSO** Tarde de frío, no hay peligro.	SISTEMA IA	**1**	**0**
11	Alerta Positiva	**1**	Sujeto con capucha y mascarilla espera afuera por más de 15 minutos, posterior entra e intimida al personal con arma.	**VERDADERO** (Solo 2 accesorios, multiproceso, alerta manos arriba .	**FALSO** Manos arriba todos, no se mueve	SISTEMA IA	**1**	**0**
12	Alerta Positiva	**1**	Sujeto en una esquina saca un	**VERDADERO** Registra accesorio	**FALSO**	SISTEMA IA	**1**	**0**

			cuchillo de su Mochila.	Mochila. Al levantar el arma, todo personal manos en la cabeza.	Punto ciego lateral del monitor.			
13	Alerta Positiva	**1**	Robo rápido: Sujeto entra, saca arma y obliga a poner manos arriba.	**VERDADERO** IA detecta pose al instante.	**FALSO** todos mantenían las manos arriba, situación de riesgo.	SISTEMA IA	**1**	**0**
14	Alerta Negativa	**0**	Cliente alto estira ambos brazos para alcanzar un producto del estante superior.	**VERDADERO** Dispara falso positivo "manos arriba".	**FALSO** Observa que está agarrando un producto físico del estante. Criterio normal.	SISTEMA TRADICIONAL	**1**	**0**
15	Alerta Negativa	**0**	Dos sujetos, uno sin accesorios y otro con Gorra y mochila entra a la sucursal, compra normal.	**FALSO** Umbral menor a 3, detección por individual.	**VERDADERO** Guardia sospecha por política.	SISTEMA IA	**0**	**1**
16	Alerta Positiva	**1**	Encapuchado obliga a tres clientes a ponerse de rodillas con manos en la cabeza.	**VERDADERO** Alerta de manos en la cabeza	**FALSO** monitorista se encuentra en su hora de almuerzo	SISTEMA IA	**1**	**0**
17	Alerta Positiva	**1**	Sujeto con, pasamontaña, capucha y mochila entra corriendo, saca un destornillador.	**VERDADERO** (Genera oclusión por color de prendas negras, alerta inmediata por manos arriba).	**VERDADERO** (Acción obvia, silbido de auxilio).	AMBOS SISTEMAS	**1**	**1**
18	Alerta Positiva	**1**	Sujeto con Gorra y abrigo con Capucha se guarda mercancía.	**FALSO** 2 accesorios, no detecta capucha.	**FALSO** No se dio cuenta.	NINGUNO	**0**	**0**
19	Alerta Negativa	**0**	Cliente se ajusta la gorra y gafas (Pose de manos cerca de cara).	**VERDADERO** (Falso positivo de pose, manos en la cabeza).	**FALSO** (Ninguna sospecha).	SISTEMA TRADICIONAL	**1**	**0**
20	Alerta Negativa	**0**	Cliente con Gorra y Gafas entra apurado, cuenta fajo de billetes y se retira.	**FALSO** Solo 2 accesorios	**FALSO** Ninguna sospecha	AMBOS	**0**	**0**
21	Alerta Positiva	**1**	Intruso con gorra, mochila y gafas mirando detenidamente las cámaras.	**VERDADERO** (Criterio de 3 accesorios, alerta de	**FALSO** Oclusión por clientes	SISTEMA IA	**1**	**0**

				posible sospechoso).				
22	Alerta Positiva	**1**	Sospechoso con Capucha mete la mano en su chaqueta saca un arma y amenaza al cajero.	**VERDADERO** Detecta Capucha, multiproceso Cajero reacciona subiendo las manos a la cabeza.	**FALSO** El monitorista salió del trabajo a las 8pm.	SISTEMA IA	**1**	**0**
23	Alerta Negativa	**0**	Persona cae al suelo con las manos en la cabeza (Emergencia médica).	**VERDADERO** (Alerta manos en la cabeza).	**FALSO** (Entiende la situación).	TRADICIONAL	**1**	**0**
24	Alerta Positiva	**1**	Intruso con Gorra, Mascarilla y Mochila fuerza la cerradura de la vitrina.	**VERDADERO** umbral de 3 prendas. Alerta a Telegram	**FALSO** No ve el detalle por rapidez.	SISTEMA IA	**1**	**0**
25	Alerta Positiva	**1**	Varios sujetos con mascarillas entran en grupo y sacan armas.	**VERDADERO** (se cuenta por ID individual, activa detección de manos arriba).	**FALSO** (Todo el personal no puede moverse).	SISTEMA IA	**1**	**0**
26	Alerta Negativa	**0**	Cajero sufre de extremo cansancio, bosteza profundamente y eleva los brazos para estirarse.	**VERDADERO** condición de manos arriba de forma errónea. Falso positivo.	**FALSO** Comprende perfectamente la acción corporal de pereza o cansancio del empleado.	SISTEMA TRADICIONAL	**1**	**0**
27	Alerta Negativa	**0**	Sujeto oculta rostro con pasamontañas y capucha mientras sujeto en la caja sube las manos desestresándose.	**VERDADERO** (Accesorios no detectados, pero activa estimación de pose).	**FALSO** (Entiende todo el contexto, no situación de riesgo).	TRADICIONAL	**1**	**0**
28	Alerta Positiva	**1**	Delincuente con Casco, Pasamontañas y mochila entra corriendo directo a las cajas.	**VERDADERO** umbral de accesorios se activa.	**VERDADERO** El ingreso corriendo activa el instinto del monitorista quien observa la pantalla al instante.	AMBOS	**1**	**1**
29	Alerta Negativa	**0**	Cliente habitual con mochila entra a la tienda y realiza una compra de un móvil.	**FALSO** (Ningún multiproceso).	**FALSO** (Comportamiento normal).	AMBOS	**0**	**0**

30	Alerta Positiva	**1**	Sujeto atrapa por la espalda a un usuario y lo obliga a colocar las manos sobre su nuca.	**VERDADERO** Detección de pose manos a la cabeza	**FALSO** La acción se dio detrás de la columna de concreto	SISTEMA IA	1	**0**
31	Alerta Positiva	**1**	Sospechoso con capucha, Gafas y Mascarilla da vuelta sin compras.	**VERDADERO** 3 accesorios detectados en el tiempo.	**FALSO** Fue un periodo de tiempo corto	SISTEMA IA	**1**	**0**
32	Alerta Negativa	**0**	Cliente entra acalorado, se quita la Gorra y se acomoda el cabello con ambas manos en la cabeza.	**VERDADERO** Alerta de manos en la cabeza. Falso positivo.	**FALSO** Visualiza el movimiento natural de refrescarse y acomodarse el cabello.	SISTEMA TRADICIONAL	**1**	**0**
33	Alerta Positiva	**1**	Sujeto saca un cuchillo de cocina y amedrenta de forma directa al cajero.	**VERDADERO** Manos arriba, envío a Telegram.	**FALSO** Monitorista salió a dejar un paquete en Servientrega	SISTEMA IA	**1**	**0**
34	Alerta Positiva	**1**	Dos encapuchados con mascarilla y mochila acorralan a cliente en zona de entrada.	**VERDADERO** Ambos objetivos son clasificados con sospecha por poseer 3 prendas.	**FALSO** Hora pico, oclusión de clientes, todos atendiendo	SISTEMA IA	**1**	**0**
35	Alerta Negativa	**0**	Cliente en la fila de espera se limpia fuertemente el sudor por calor colocándose las manos sobre la frente.	**VERDADERO** Manos en la cabeza**.** Falso positivo.	**FALSO** Sabe que el local no tiene aire acondicionado	SISTEMA TRADICIONAL	**1**	**0**
36	Alerta Positiva	**1**	Sujeto con Casco, Mascarilla y Gafas insulta verbalmente y amenaza al personal de caja.	**VERDADERO** Combinación de 3 accesorios, activa alerta	**FALSO** No escucha la conversación por la distancia	SISTEMA IA	**1**	**0**
37	Alerta Positiva	**1**	Amedrantamiento cauteloso, el sujeto muestra un arma de fuego oculta en la cintura al cajero.	**VERDADERO** Alerta por manos a la cabeza de manera sigilosa.	**FALSO** monitorista no se da cuenta del arma en la cintura.	SISTEMA IA	**1**	**0**
38	Alerta Negativa	**0**	Un niño pequeño juega con su madre y levanta las manos gritando por alegría.	**VERDADERO** Alerta por brazos arriba.	**FALSO** No hay sospecha	SISTEMA TRADICIONAL	**1**	**0**

39	Alerta Positiva	**1**	Sujeto con Gorra, Gafas y Mochila asalta al cajero, metiendo la mano en la caja fuerte.	**VERDADERO** Alerta por accesorios activada, no hay pose.	**VERDADERO** Entiende el asunto, salta a forcejear.	AMBOS	**1**	**1**
40	Alerta Positiva	**1**	Cliente es arrinconado contra una esquina muerta del local por un delincuente alto y encapuchado.	**VERDADERO** Estimación de pose detecta los brazos arriba de la víctima a pesar de la oclusión.	**FALSO** Al estar cerca de la entrada, no le da importancia por presuntos amigos.	SISTEMA IA	**1**	**0**

Anexo E. Matriz consolidada de simulaciones y eficiencia del sistema

N°	**Tipo de Sistema**	**Asignación Crítica (C)**	**Eficiencia / Aciertos (A)**	**Eficiencia Final (Ef)**	**Estado Operativo**
1	IA	**1**	1	**1**	Acierto (Alerta real procesada)
2	IA	**1**	1	**1**	Acierto (Alerta real procesada)
3	IA	**0**	1	**0**	Fallo (Falso Positivo)
4	IA	**1**	1	**1**	Acierto (Alerta real procesada)
5	IA	**1**	1	**1**	Acierto (Alerta real procesada)
6	IA	**0**	0	**1**	Acierto (Alerta real procesada)
7	IA	**1**	1	**1**	Acierto (Alerta real procesada)
8	IA	**1**	1	**1**	Acierto (Alerta real procesada)
9	IA	**1**	1	**1**	Acierto (Alerta real procesada)
10	IA	**1**	1	**1**	Acierto (Alerta real procesada)
11	IA	**1**	1	**1**	Acierto (Alerta real procesada)
12	IA	**1**	1	**1**	Acierto (Alerta real procesada)
13	IA	**1**	1	**1**	Acierto (Alerta real procesada)
14	IA	**0**	1	**0**	Fallo (Falso Positivo)
15	IA	**0**	0	**1**	Acierto (Alerta real procesada)
16	IA	**1**	1	**1**	Acierto (Alerta real procesada)
17	IA	**1**	1	**1**	Acierto (Alerta real procesada)
18	IA	**1**	0	**0**	Fallo (Falso Positivo)
19	IA	**0**	1	**0**	Fallo (Falso Positivo)
20	IA	**0**	0	**1**	Acierto (Alerta real procesada)
21	IA	**1**	1	**1**	Acierto (Alerta real procesada)
22	IA	**1**	1	**1**	Acierto (Alerta real procesada)

23	IA	**0**	1	**0**	Fallo (Falso Positivo)
24	IA	**1**	1	**1**	Acierto (Alerta real procesada)
25	IA	**1**	1	**1**	Acierto (Alerta real procesada)
26	IA	**0**	1	**0**	Fallo (Falso Positivo)
27	IA	**0**	1	**0**	Fallo (Falso Positivo)
28	IA	**1**	1	**1**	Acierto (Alerta real procesada)
29	IA	**0**	0	**1**	Acierto (Alerta real procesada)
30	IA	**1**	1	**1**	Acierto (Alerta real procesada)
31	IA	**1**	1	**1**	Acierto (Alerta real procesada)
32	IA	**0**	1	**0**	Fallo (Falso Positivo)
33	IA	**1**	1	**1**	Acierto (Alerta real procesada)
34	IA	**1**	1	**1**	Acierto (Alerta real procesada)
35	IA	**0**	1	**0**	Fallo (Falso Positivo)
36	IA	**1**	1	**1**	Acierto (Alerta real procesada)
37	IA	**1**	1	**1**	Acierto (Alerta real procesada)
38	IA	**0**	1	**0**	Fallo (Falso Positivo)
39	IA	**1**	1	**1**	Acierto (Alerta real procesada)
40	IA	**1**	1	**1**	Acierto (Alerta real procesada)
41	TD	**1**	1	**1**	Acierto (Alerta real procesada)
42	TD	**1**	0	**0**	Fallo (Falso Positivo)
43	TD	**0**	0	**1**	Acierto (Alerta real procesada)
44	TD	**1**	0	**0**	Fallo (Falso Positivo)
45	TD	**1**	0	**0**	Fallo (Falso Positivo)
46	TD	**0**	0	**1**	Acierto (Alerta real procesada)
47	TD	**1**	0	**0**	Fallo (Falso Positivo)
48	TD	**1**	0	**0**	Fallo (Falso Positivo)
49	TD	**1**	0	**0**	Fallo (Falso Positivo)
50	TD	**1**	0	**0**	Fallo (Falso Positivo)
51	TD	**1**	0	**0**	Fallo (Falso Positivo)
52	TD	**1**	0	**0**	Fallo (Falso Positivo)
53	TD	**1**	0	**0**	Fallo (Falso Positivo)
54	TD	**0**	0	**1**	Acierto (Alerta real procesada)
55	TD	**0**	1	**0**	Fallo (Falso Positivo)
56	TD	**1**	0	**0**	Fallo (Falso Positivo)
57	TD	**1**	1	**1**	Acierto (Alerta real procesada)
58	TD	**1**	0	**0**	Fallo (Falso Positivo)
59	TD	**0**	0	**1**	Acierto (Alerta real procesada)
60	TD	**0**	0	**1**	Acierto (Alerta real procesada)
61	TD	**1**	0	**0**	Fallo (Falso Positivo)
62	TD	**1**	0	**0**	Fallo (Falso Positivo)
63	TD	**0**	0	**1**	Acierto (Alerta real procesada)
64	TD	**1**	0	**0**	Fallo (Falso Positivo)
65	TD	**1**	0	**0**	Fallo (Falso Positivo)

66	TD	**0**	0	**1**	Acierto (Alerta real procesada)
67	TD	**0**	0	**1**	Acierto (Alerta real procesada)
68	TD	**1**	1	**1**	Acierto (Alerta real procesada)
69	TD	**0**	0	**1**	Acierto (Alerta real procesada)
70	TD	**1**	0	**0**	Fallo (Falso Positivo)
71	TD	**1**	0	**0**	Fallo (Falso Positivo)
72	TD	**0**	0	**1**	Acierto (Alerta real procesada)
73	TD	**1**	0	**0**	Fallo (Falso Positivo)
74	TD	**1**	0	**0**	Fallo (Falso Positivo)
75	TD	**0**	0	**1**	Acierto (Alerta real procesada)
76	TD	**1**	0	**0**	Fallo (Falso Positivo)
77	TD	**1**	0	**0**	Fallo (Falso Positivo)
78	TD	**0**	0	**1**	Acierto (Alerta real procesada)
79	TD	**1**	1	**1**	Acierto (Alerta real procesada)
80	TD	**1**	0	**0**	Fallo (Falso Positivo)

Anexo F. Script implementado en el proyecto, realizado en Visual Studio Codec.