



UNIVERSIDAD TÉCNICA ESTATAL DE QUEVEDO

FACULTAD DE POSGRADO

MAESTRÍA EN CIENCIA DE DATOS

Proyecto de investigación previo
a la obtención del Grado
Académico de Magíster en
Ciencia de Datos.

TEMA

**APRENDIZAJE PROFUNDO PARA LA DETECCIÓN DE ARMAS DE FUEGO EN
SISTEMAS DE VIDEOVIGILANCIA**

AUTOR

ING. BRIONES MONTALVO CRISTHIAN DANILO

DIRECTOR

ING. ZAMBRANO VEGA CRISTIAN GABRIEL. PHD

QUEVEDO – ECUADOR

AÑO 2024

RESUMEN

Ecuador enfrenta un desafío creciente de inseguridad, potenciado por el narcotráfico y la proliferación de armas de fuego, lo que exige soluciones inteligentes que combinen estrategias de seguridad pública y tecnológica. En este contexto, los sistemas de videovigilancia basados en circuitos cerrados de televisión (CCTV) se han convertido en herramientas cruciales para el monitoreo y registro de eventos en tiempo real. Sin embargo, para combatir eficazmente la amenaza de las armas de fuego, se requieren sistemas precisos de detección en tiempo real.

Este proyecto aborda una necesidad urgente al explorar técnicas de aprendizaje profundo y visión por computadora para desarrollar un sistema de reconocimiento de armas en sistemas de CCTV, con el objetivo de proponer una herramienta informática que pueda contribuir a la lucha contra la proliferación de armas de fuego y la violencia en la sociedad ecuatoriana. Se analizaron diversas herramientas para la construcción de modelos de visión artificial: YOLOv8, Teachable Machine (EfficientNet) y la metodología Haar-Cascade. Se evaluaron los rendimientos de los modelos de detección generados empleando varias métricas de precisión. Entre los modelos evaluados, se encontró que el modelo basado en YOLOv8 demostró ser el más eficaz en términos de precisión y rendimiento.

Palabras claves: Detección de armas, Aprendizaje Profundo, Visión por Computadora, Modelos de detección, CCTV.

ABSTRACT

Ecuador faces a growing challenge of insecurity, driven by drug trafficking and the proliferation of firearms, necessitating smart solutions that blend public security and technological strategies. In this context, closed-circuit television (CCTV) surveillance systems have become crucial tools for real-time monitoring and event recording. However, effectively combating the threat of firearms requires precise real-time detection systems. Thus, this project addresses this urgent need by exploring deep learning techniques and computer vision to develop a firearms recognition system within CCTV systems, aiming to propose a computational approach to combat the proliferation of firearms and violence in Ecuadorian society. Various tools for constructing artificial vision models were analyzed: YOLOv8, Teachable Machine (EfficientNet), and the Haar-Cascade methodology. The performance of detection models was evaluated using multiple precision metrics. Among the evaluated models, it was found that the YOLOv8-based model proved to be the most effective in terms of precision and performance.

Keywords: Weapon Detection, Deep Learning, Computer Vision, Detection Models, CCTV.

CERTIFICACIÓN

Ing. Cristian Zambrano, PhD., Asesor del Proyecto de Investigación previo a la obtención del Grado Académico de Magíster en Ciencia de Datos.

CERTIFICA:

Que el Ing. **Briones Montalvo Cristhian Danilo**, autor del proyecto de investigación titulado “**Aprendizaje profundo para la detección de armas de fuego en sistemas de videovigilancia**” ha sido revisada en todos sus componentes, la misma que está apta para la presentación ante el tribunal respectivo.

Quevedo, 13 de agosto del 2024

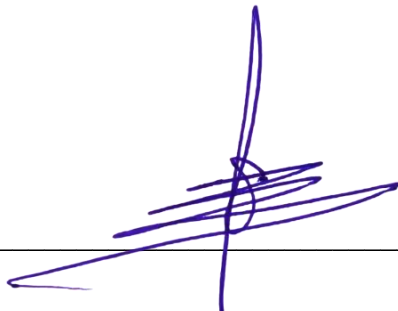
CRISTIAN		Firmado
GABRIEL		digitalmente por
ZAMBRANO		CRISTIAN
VEGA		GABRIEL ZAMBRANO VEGA

Ing. Cristian Zambrano Vega, PhD.

DIRECTOR

AUTORÍA

Yo, Ing. Briones Montalvo Cristhian Danilo declaro que el presente Proyecto de Investigación titulado “**APRENDIZAJE PROFUNDO PARA LA DETECCIÓN DE ARMAS DE FUEGO EN SISTEMAS DE VIDEOVIGILANCIA**”, es de exclusiva autoría y extendiendo los derechos a la Universidad Técnica Estatal de Quevedo, hacer uso del documento como material académico y de consulta.



Ing. Briones Montalvo Cristhian Danilo

AUTOR

DEDICATORIA

A mis padres, hermanos y familia en general, les dedico este logro como un testimonio de su amor incondicional y constante apoyo a lo largo de mi vida. Su presencia ha sido mi mayor fortaleza, guiándome con sabiduría y alentándome en cada paso de este camino hacia el éxito.

A mis profesores, amigos y estudiantes, les expreso mi más sincero agradecimiento por ser fuentes inagotables de inspiración y motivación. Su influencia y camaradería han sido fundamentales en mi desarrollo académico y personal, recordándome constantemente la importancia del aprendizaje, la amistad y el trabajo en equipo.

Finalmente, a todos aquellos que han dejado una huella en mi vida, les dedico este logro con profunda gratitud. Cada encuentro, cada consejo y cada experiencia compartida ha contribuido a mi crecimiento y ha enriquecido mi camino. Agradezco sinceramente su presencia en este viaje, que ha contribuido en gran medida a mi crecimiento personal y profesional.

AGRADECIMIENTO

Este momento representa una profunda satisfacción, donde puedo poner al servicio de mi comunidad los conocimientos y habilidades adquiridos. Mi contribución busca fomentar la construcción de un país más seguro para todos, especialmente para las nuevas generaciones, quienes se ven afectadas de manera insólita en esta realidad que les legamos. Mi aspiración es inspirarlas a convertirse en agentes de cambio y líderes en la edificación de un futuro mejor para las siguientes generaciones.

Expreso mi más sincero agradecimiento a la SENESCYT por el programa "Becas Fortalécete 2022", el cual me ha ofrecido una invaluable oportunidad para continuar con mis estudios de posgrado. Esta oportunidad ha sido fundamental para fortalecer mi formación académica y profesional.

Agradezco de corazón a mi familia, quienes son el pilar de mi vida. El amor incondicional, el apoyo constante y la sabiduría que me han brindado, especialmente mis padres y hermanos, han sido la fuente de mi fuerza y determinación. Les debo todo lo que soy y cada logro alcanzado. También, quiero expresar mi sincero agradecimiento a mis profesores, cuya dedicación y orientación han sido vitales en mi crecimiento académico y personal. Su influencia ha sido invaluable y les estoy profundamente agradecido por su constante apoyo.

Cristhian Danilo Briones Montalvo

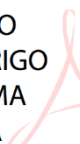
PRÓLOGO

En un mundo marcado por la creciente complejidad de los desafíos en materia de seguridad, la innovación tecnológica se alza como una herramienta fundamental para proteger a nuestras comunidades y preservar la tranquilidad de nuestras sociedades. En este contexto, la detección de armas de fuego en sistemas de videovigilancia emerge como un objetivo crucial en la búsqueda de soluciones efectivas para contrarrestar el crimen y garantizar la seguridad pública.

Esta tesis, titulada "Aprendizaje Profundo para la Detección de Armas de Fuego en Sistemas de Videovigilancia", representa un esfuerzo oportuno por explorar nuevas fronteras en el campo de la visión por computadora y el reconocimiento de objetos. A través del estudio y la aplicación de técnicas de aprendizaje profundo, este trabajo se propone desarrollar un sistema eficiente y preciso capaz de identificar la presencia de armas de fuego en imágenes y videos capturados por cámaras de seguridad.

A lo largo de estas páginas, nos sumergiremos en el fascinante mundo de los algoritmos de detección de objetos, explorando diferentes enfoques y metodologías para abordar esta problemática crítica. Desde la recopilación y etiquetado de datos hasta el diseño y entrenamiento de modelos de aprendizaje automático, cada paso en este viaje de investigación nos acerca más a la creación de una herramienta poderosa para fortalecer la seguridad pública y proteger a nuestras comunidades.

EMILIO
RODRIGO
ZHUMA
MERA



Firmado
digitalmente
por EMILIO
RODRIGO
ZHUMA MERA

Ing. Emilio Zhuma Mera, M. Sc.

Docente titular UTEQ

Índice de contenidos

DEDICATORIA	vi
AGRADECIMIENTO	vii
PRÓLOGO.....	viii
INTRODUCCIÓN	8
CAPÍTULO I.....	11
1. MARCO CONTEXTUAL DE LA INVESTIGACIÓN	11
1.1. Ubicación y contextualización de la problemática	12
1.2. Situación actual de la problemática	13
1.3. Problema de investigación	13
1.3.1. Problemas derivados	14
1.4. Delimitación del problema.....	14
1.5. Objetivos.....	14
1.5.1. Objetivo general.....	14
1.5.2. Objetivos específicos	14
1.6. Justificación	15
CAPÍTULO II	16
2. MARCO TEÓRICO	16
2.1. Fundamentación conceptual	17

2.1.1.	Detección de objetos	17
2.1.2.	Aprendizaje de máquina	17
2.1.3.	Aprendizaje profundo	18
2.1.4.	Aprendizaje supervisado	18
2.1.5.	Inteligencia artificial	18
2.1.6.	OpenCV	19
2.1.7.	Armas de fuego	19
2.1.8.	Cámaras IP	20
2.1.9.	Videovigilancia CCTV	20
2.1.10.	Redes neuronales convolucionales	21
2.1.11.	Algoritmo YOLOv8.....	22
2.1.12.	Algoritmo Tensorflow	22
2.1.13.	Algoritmo Haar Cascade.....	22
2.2.	Fundamentación teórica	23
2.3.	Marco Legal	25
1.3.1.	Código Orgánico Integral Penal.....	26
1.3.2.	Ley Orgánica de Transparencia y Acceso a la Información Pública	26
CAPÍTULO III		28
3.	METODOLOGÍA DE LA INVESTIGACIÓN	28
3.1.	Tipo de investigación	29
3.1.1.	Investigación descriptiva.....	29

3.1.2. Investigación aplicada.....	29
3.2. Métodos de investigación.....	29
3.2.1. Método de Análisis	30
3.2.2. Método de Síntesis	30
3.2.3. Método Deductivo.....	30
3.3. Construcción metodológica del objeto de investigación.....	30
3.3.1. Técnicas de investigación	30
3.3.2. Instrumentos de investigación.....	36
3.4. Elaboración del marco teórico	37
3.5. Recolección de la información.....	39
3.6. Procesamiento y análisis	41
CAPÍTULO IV	45
4. RESULTADOS Y DISCUSIÓN	45
4.1. Creación del conjunto de datos de imágenes de armas de fuego de alta calidad y diversidad para el entrenamiento de los modelos	46
4.1.1. Recopilación de imágenes.....	46
4.1.2. Extracción y etiquetado de imágenes.....	47
4.1.3. Discusión.....	49
4.2. Aplicación de distintas técnicas de reconocimiento de objetos para determinar cuál es la más eficaz en la detección de armas de fuego.	50
4.2.1. Entrenamiento y evaluación de modelos de detección	50

4.3. Establecimiento de un canal de comunicación que permita la notificación de la presencia de armas de fuego a partir de los resultados generados por el modelo.....	59
4.3.1. Proceso de detección y generación de alertas	60
CAPÍTULO V	64
5. CONCLUSIONES Y RECOMENDACIONES	64
BIBLIOGRAFÍA	68

INTRODUCCIÓN

En las últimas décadas, el fenómeno del narcotráfico ha dejado una huella imborrable en la historia del Estado ecuatoriano, transformando el paisaje social y amenazando la seguridad de los ciudadanos. Ecuador, conocido por su belleza natural y diversidad cultural, se ha visto desafiado por una creciente ola de inseguridad y crimen organizado que afecta profundamente al país. La proliferación de armas de fuego, que en su mayoría ingresan ilegalmente al territorio ecuatoriano, ha contribuido significativamente a este escenario de violencia.

El comercio de armas ilícitas y el tráfico de drogas se han entrelazado de manera preocupante, lo que ha generado una constante amenaza a la paz y la estabilidad de la sociedad ecuatoriana (ACUÑA, 2021). La delincuencia se ha vuelto más sofisticada y ha demostrado la capacidad de adaptarse a las medidas de seguridad existentes. Esta situación plantea la necesidad urgente de desarrollar soluciones innovadoras y efectivas para contrarrestar la proliferación de armas de fuego y proteger a la comunidad.

En este contexto, surge la importancia de la visión por computadora como un aliado en la lucha contra la inseguridad. Los sistemas de videovigilancia, a través de los Circuitos Cerrados de Televisión (CCTV), se han convertido en una herramienta fundamental para monitorear y registrar eventos en tiempo real. Sin embargo, la simple presencia de cámaras de seguridad no es suficiente. La capacidad de detectar armas de fuego en sistemas de CCTV en tiempo real se ha convertido en un objetivo crucial para salvaguardar la seguridad pública y prevenir incidentes violentos ya que permitiría una respuesta rápida y efectiva de las fuerzas de seguridad y la disuasión de crímenes.

Dado este escenario, esta tesis aborda el reconocimiento de armas en redes de videovigilancia, con el objetivo de desarrollar un sistema eficiente y preciso que pueda identificar la presencia de armas de fuego en imágenes y videos capturados por cámaras de seguridad. A lo largo de este trabajo de investigación, se exploran las técnicas de aprendizaje profundo, la visión por computadora y el procesamiento de imágenes para abordar esta problemática crítica. En respuesta a este desafío, esta tesis busca ofrecer una contribución significativa al campo de la seguridad pública en Ecuador.

La estructura del presente trabajo está organizada de la siguiente manera:

Capítulo I: Marco Contextual de la Investigación - Este capítulo abarca la ubicación y contextualización de la problemática, la situación actual de la problemática, la formulación del problema de investigación, la delimitación del problema, los objetivos y la justificación del estudio.

Capítulo II: Marco Teórico - Aquí se presenta la fundamentación conceptual, teórica y legal, incluyendo los aportes de diversos autores relacionados con el tema tratado en el proyecto.

Capítulo III: Metodología de la Investigación - Este capítulo describe los métodos y técnicas utilizados en la investigación, los instrumentos empleados, la elaboración del marco teórico, así como la recolección, procesamiento y análisis de la información.

Capítulo IV: Resultados y Discusión - En este capítulo se presentan y discuten los resultados obtenidos en relación con los objetivos específicos planteados.

Capítulo V: Conclusiones y Recomendaciones - Este capítulo finaliza el trabajo con las conclusiones derivadas de los resultados obtenidos y las recomendaciones pertinentes.

Finalmente, se incluyen las Referencias Bibliográficas y los Anexos.

CAPÍTULO I

1. MARCO CONTEXTUAL DE LA INVESTIGACIÓN

1.1. Ubicación y contextualización de la problemática

Las armas de fuego representan un gran problema en todo el mundo. El comercio y posesión de armas impacta en la sociedad de distintas formas, siendo este uno de los principales desafíos en el panorama de la seguridad y se encuentra dentro de los esfuerzos de aplicación de la ley (UNODC, 2020). Este problema se genera debido a que muchas de las armas circulando alrededor de todos los países son adquiridas de manera ilegal, permitiendo el incremento significativo de crímenes y violencia.

La violencia con armas impacta principalmente a regiones que muestran altos niveles de desigualdad en los ingresos, incapacidad para cubrir necesidades básicas y sistemas judiciales vulnerables (Alvarez, 2020). Ecuador no queda fuera de esta situación; en este Estado, la creciente ola de violencia y crimen organizado ha generado preocupación a nivel nacional. La delincuencia se ha vuelto más sofisticada y ha demostrado la capacidad de adaptarse a las medidas de seguridad existentes (Rojas Daniel, 2024).

Muchos de estos hechos han sido registrados en sistemas de videovigilancia. Sin embargo, lamentablemente no se han podido frustrar o atender de forma inmediata debido a que los sistemas de alarma antirrobo y de videovigilancia actuales no cuentan con la suficiente inteligencia para tomar decisiones en función de la situación de peligro por la que estamos pasando. De esta manera, no pueden notificar a las autoridades competentes de una manera rápida y eficiente.

En la región, se observa una carencia de sistemas de vigilancia sofisticados capaces de brindar información crucial ante actos delictivos que involucren armas de fuego. Esta situación convierte a los establecimientos comerciales e instituciones vulnerables a la acción de los delincuentes.

1.2. SITUACIÓN ACTUAL DE LA PROBLEMÁTICA

De acuerdo con el aumento del uso de armas de fuego y de muertes violentas que está propagándose rápidamente en Ecuador, existe un principio rector: el narcotráfico, como resultado de las disputas entre los grupos por las rutas y otras economías criminales (Doherty, 2023)

Las bandas criminales gestionan su propio armamento tanto dentro como fuera de las cárceles. A nivel internacional, Ecuador ha pasado de ser considerado solo un punto de consumo de drogas a convertirse también en un centro de distribución (Pichel, 2021).

Esta problemática se acentúa aún más debido a la abundante información sin explotar en la actualidad, especialmente en forma de videos. La producción de contenido digital ha alcanzado niveles tan extraordinarios que resulta prácticamente imposible supervisarla con medios humanos (Sánchez Valdés, 2016). Por ello, surge la necesidad de emplear técnicas modernas para poder procesar la información de forma masiva.

Aunque existen equipos y tecnologías avanzadas capaces de abordar este desafío, su alto costo limita significativamente su implementación. En este contexto, surge la propuesta de implementar nuevas tecnologías más accesibles y específicas para el procesamiento y análisis de información digital relacionada con actividades delictivas, con el objetivo de mejorar la capacidad de respuesta y control en situaciones de seguridad.

1.3. PROBLEMA DE INVESTIGACIÓN

¿Cómo se podría detectar la presencia de armas de fuego en un sistema de videovigilancia?

1.3.1. Problemas derivados

- ¿Cómo habilitar cámaras para reconocer armas de fuego de manera efectiva?
- ¿Qué técnica de reconocimiento de objetos es la más eficaz en la detección de armas de fuego?
- ¿Cómo establecer comunicación entre los usuarios y el modelo para alertar la presencia de armas de fuego?

1.4. DELIMITACIÓN DEL PROBLEMA

CAMPO: Tecnologías de la Información y Comunicación

ÁREA: Tecnologías de la Información y Comunicación

LÍNEA: Ciencia de Datos

LUGAR: Esmeraldas

TIEMPO: Septiembre del 2023 – Agosto del 2024

1.5. OBJETIVOS

1.5.1. Objetivo General

Implementar un modelo de detección basado en técnicas de aprendizaje profundo para alertar la presencia de armas de fuego en sistemas de videovigilancia.

1.5.2. Objetivos específicos

- Preparar un conjunto de datos con imágenes de diversas armas de fuego para el entrenamiento de modelos de reconocimiento.
- Aplicar distintas técnicas de reconocimiento de objetos para determinar cuál es la más eficaz en la detección de armas de fuego.
- Establecer un canal de comunicación que permita la notificación de la presencia de armas de fuego a partir de los resultados generados por el modelo.

1.6. JUSTIFICACIÓN

La detección de objetos mediante cámaras tiene aplicaciones en diversas áreas, siendo una de las más relevantes el monitoreo en sistemas de vigilancia (Miranda-Ramos et al., 2022). En la actualidad, se han desarrollado sistemas de vigilancia que emplean visión artificial para mejorar la eficiencia de estos sistemas. Este avance representa un progreso significativo, ya que anteriormente la detección de objetos dependía exclusivamente de la atención y capacidad de los operadores de las cámaras.

El desarrollo de este proyecto supone un logro de gran importancia en el ámbito de la seguridad. Se trata de una propuesta innovadora que ofrece una solución real a una problemática que afecta a la sociedad ecuatoriana en términos de seguridad y economía. Los beneficios de esta tecnología se extienden a nivel macroeconómico, impactando inicialmente en las empresas de seguridad. Estas empresas experimentarían mejoras notables en la automatización de sus procesos, la optimización de los tiempos de respuesta y una mayor eficiencia en la detección de objetos.

Sin embargo, el alcance positivo de esta tecnología va más allá del ámbito empresarial, beneficiando a toda la sociedad. La capacidad de detectar de manera temprana la presencia de armas de fuego ofrece una ventaja crucial para la seguridad pública. Esto permite la posibilidad de una notificación rápida a las autoridades pertinentes en casos de amenazas, lo que puede tener un impacto significativo en la prevención de delitos y en la protección de la comunidad en general.

CAPÍTULO II

2. MARCO TEÓRICO

2.1. Fundamentación conceptual

Este marco conceptual establece los cimientos teóricos que respaldarán el reconocimiento de armas de fuego en sistemas de videovigilancia. Se exploran los conceptos y teorías relacionadas con las armas de fuego, circuitos cerrados de televisión, minería de datos, aprendizaje automático, así como las técnicas más utilizadas. Este marco proporcionará la base necesaria para comprender cómo las técnicas avanzadas de ciencia de datos pueden automatizar tareas de detección de objetos en tiempo real.

2.1.1. Detección de objetos

La detección de objetos es una tarea en el campo de la visión por computadora que implica identificar y localizar objetos específicos en imágenes o videos. El objetivo principal de la detección de objetos es encontrar la presencia de objetos dentro de una imagen y, en algunos casos, proporcionar información adicional, como la ubicación y las dimensiones de los objetos detectados (MATLAB & Simulink, 2020).

2.1.2. Aprendizaje de máquina

El aprendizaje de máquina es una rama de la inteligencia artificial que le proporciona a las máquinas la capacidad de aprender y mejorar por medio de un conjunto de experiencias que son datos e información que provienen de interacciones con el mundo real (Mirjalili & Raschka, 2020).

Este trabajo busca por medio de la experiencia, en este caso un conjunto de vídeos de armas de mano conseguir que en un video se detecten pistolas y revólveres usando algoritmos de aprendizaje profundo.

2.1.3. Aprendizaje profundo

El aprendizaje profundo es una subcategoría del aprendizaje de máquina, que permite a las computadoras aprender. Este aprendizaje sigue teorías acerca de cómo funciona el cerebro humano, específicamente las neuronas, buscando fundamentarse en ideas básicas al inicio para después basarse en ellas para abstraer ideas más amplias (Jesús Alfonso López Sotelo, 2023).

2.1.4. Aprendizaje supervisado

Es un conjunto de algoritmos de inteligencia artificial que hace uso de un conjunto de datos, que se conoce como conjunto de entrenamiento, con el cual realiza predicciones (APD, 2019). El conjunto de datos de entrenamiento incluye datos de entrada y valores de respuesta (etiquetas). Todo esto con el fin de obtener un modelo que permita realizar predicciones en conjuntos de datos nuevos.

El aprendizaje profundo contiene dos categorías de algoritmos, los algoritmos de clasificación que dan como respuesta valores categóricos; los algoritmos de regresión que trabajan con predicciones continuas, según los datos de entrada (Ramasubramanian & Moolayil, 2019).

2.1.5. Inteligencia artificial

La inteligencia artificial es una rama de la informática con bases en la lógica y las ciencias cognitivas. Esta disciplina estudia y desarrolla sistemas capaces de realizar tareas que normalmente se atribuyen a la inteligencia humana, como traducir un documento o reconocer a una persona por su rostro, así como sistemas que aprenden a comportarse de manera autónoma, como vehículos sin conductor o robots (Gonçalves & Partyk, 2021).

La inteligencia artificial se utiliza por medio de algoritmos que permiten programar una máquina para simular comportamientos inteligentes, como la detección automática de armas de fuego.

2.1.6. OpenCV

OpenCV (*Open Source Computer Vision Library*) es una biblioteca abierta que proporciona las herramientas y los marcos necesarios para implementar aplicaciones de visión por computadora. Esta librería proporciona una amplia gama de funciones y herramientas que permiten a los desarrolladores y científicos realizar una variedad de tareas relacionadas con el procesamiento de imágenes y el análisis de video (Sunila Gollapudi, 2019).

2.1.7. Armas de fuego

Ante el concepto amplio del término arma de fuego, ya que las armas de fuego no se limitan a los dispositivos con al menos un cañón, sino que el derecho internacional también ha agregado algunas armas extremadamente pesadas, para los fines de este estudio, el estudio se limita a las armas de fuego.

Así, como primera aproximación al concepto de armas, tenemos denominaciones de motores térmicos capaces de expulsar o propulsar un proyectil a causa de la pólvora.

Con trayectorias y fuerzas definidas, el destino es el objetivo preseleccionado o no seleccionado. McNab (2009) señala que las armas de fuego como máquinas con compuestos químicos como propulsores que, al encenderse o quemarse, crean presión de aire, debido a que propulsan un tiro con fuerza y dirección como puede observarse en la Ilustración 1.

Ilustración 1

Armas de fuego



Fuente: Diario La Hora (Fedexpor, 2023)

2.1.8. Cámaras IP

Las cámaras IP, están proyectadas específicamente al uso en una red de área local “LAN” desde un navegador. Hay cámaras en las que se puede observar en tiempo real lo que sucede. Las cámaras IP también se conectan a un punto de red inalámbrica o Ethernet (Silhavy & Silhavy, 2023).

A cambio de las cámaras web que deben conectarse a una computadora a través de un puerto USB o de video, no requiere hardware adicional porque procesa la imagen y el sonido internamente.

2.1.9. Videovigilancia CCTV

Los sistemas de CCTV son muy fáciles de usar ya que se procesan de la misma manera que el video normal. El CCTV se utiliza para la grabación de imágenes mediante cámaras cuando nos encontramos fuera, podemos observar todo el mundo en tiempo real, controlar distintas dependencias y rincones sin movernos y tener una visión unificada desde la comodidad (Armstrong & Norris, 2020), tal como podemos observar en la Ilustración 2.

Ilustración 2

Sistema CCTV



Fuente: LatinPyme (LatinPyme, 2023)

Estas cámaras de circuito cerrado de televisión también actúan como elemento disuasorio contra el robo y el vandalismo, lo que es de especial interés para nosotros en este artículo. Lo mejor de registrar todo lo que sucede es que no tenemos que estar allí para ver las cosas, cada vez que sucede un evento, podemos mirar el registro y ver por nosotros mismos si sucedió o no.

2.1.10. Redes neuronales convolucionales

Las (RNC) o *CNN* en sus siglas en inglés Estos e refieren específicamente a datos con una topología de cuadrícula conocida, los datos de series de tiempo se consideran una cuadrícula unidimensional muestreada en ciertos intervalos de tiempo, y los datos de imagen son lo que consideran una red bidimensional. Píxeles (Liu, 2020). Este tipo de red realiza una operación de convolución sobre la señal de entrada con respecto a la propagación a la salida de la neurona.

2.1.11. Algoritmo YOLOv8

YOLOv8 es una herramienta popular de detección de objetos basada en redes neuronales, que puede ser entrenada con Python y Torch. Es una excelente opción para la detección de armas en tiempo real debido a su alta velocidad de ejecución y precisión en la detección de objetos. A diferencia de otras herramientas similares como RetinaNet,

YOLOv8 ha demostrado ser especialmente eficiente en situaciones de tiempo real y su facilidad de uso lo hace accesible para usuarios con menos experiencia en el campo de las redes neuronales (Singh Sisodia et al., 2023).

2.1.12. Algoritmo Tensorflow

TensorFlow es una plataforma de código abierto de extremo a extremo para el aprendizaje automático. Cuenta con un ecosistema completo y flexible de herramientas, bibliotecas y recursos comunitarios que permite a los investigadores avanzar en el estado del arte en el aprendizaje automático y a los desarrolladores construir y desplegar fácilmente aplicaciones impulsadas por el aprendizaje automático (Shetty et al., 2021).

2.1.13. Algoritmo Haar Cascade

Haar Cascading es el método de aprendizaje automático donde un clasificador se entrena con una gran cantidad de fotos positivas y negativas (Shetty et al., 2021). Este clasificador sigue un procedimiento de aprendizaje automático en el que una operación en cascada se inculca a partir de las fotos para descubrir elementos en fotos adicionales. El entrenamiento se realiza ofreciendo al clasificador imágenes positivas y negativas. Luego, se extraen las características de la imagen. Cada característica es un valor individual, que

se obtiene restando la suma de píxeles en un rectángulo blanco de la suma de píxeles en un rectángulo negro. Esto permite detectar patrones en distintos entornos.

2.2. Fundamentación teórica

El trabajo de Manikandan & Rahamathunnisa (2022) su trabajo se basó en la detección de objetos dañinos en imágenes de cámaras de CCTV para mejorar la seguridad. Utiliza un esquema llamado AODS que se basa en una red neuronal convolucional (*CNN*) para detectar y clasificar objetos peligrosos. Este proceso implicó analizar las características de los objetos y ajustarlas para una detección precisa. Se evaluó el sistema utilizando métricas como precisión y exactitud, y se encontró que el entrenamiento con datos externos mejoró el rendimiento. Además, se sugiere la futura implementación de la clasificación basada en etiquetas.

El artículo presentado por Debnath & Bhowmik (2021) se centra en la detección automática de armas de fuego, ya sea que estén visibles o escondidas en el cuerpo, utilizando sistemas de visión por computadora. Se compara esta tarea con la habilidad humana de detectar armas y su importancia en la seguridad. El artículo revisa la literatura sobre este tema, propone una taxonomía, discute las dificultades y compara los enfoques existentes. Finalmente, el artículo concluye señalando posibles brechas de investigación en campos relacionados con la detección de armas de fuego.

En su trabajo Salazar González et al. (2020), destaca las mejoras recientes en los detectores de objetos, pero señala que la detección de objetos pequeños, como armas en sistemas de CCTV, sigue siendo un desafío. Presenta un nuevo conjunto de datos obtenido de un sistema de CCTV real en una universidad e imágenes generadas de forma sintética. Los autores utilizaron *Faster R-CNN* con *Feature Pyramid Network* y ResNet-50 para

crear un modelo de detección de armas en tiempo casi real. Se realiza un estudio exhaustivo y se destacó el impacto de los conjuntos de datos sintéticos en el entrenamiento. También se señalan las limitaciones actuales de estos sistemas y se pone a disposición de la comunidad de investigación los conjuntos de datos generados.

El artículo presentado por Qasim & Verdu (2023) se centra en la detección automática de anomalías en videos de vigilancia utilizando una red neuronal convolucional profunda (CNN) y una unidad recurrente simple (SRU). El enfoque combina características espaciales y temporales para identificar acciones inusuales en los videos. Se proponen tres modelos diferentes y se evalúan en un conjunto de datos llamado UCF-Crime.

Los resultados muestran que estos modelos logran altos niveles de precisión en la detección de acciones inusuales, con la mayor precisión alcanzada por el modelo ResNet50 + SRU. Estos modelos superan a otros enfoques de aprendizaje profundo comparables.

El estudio de Veranyurt & Sakar (2020) se enfoca en evaluar métodos de aprendizaje profundo para detectar armas de fuego en imágenes. Se comparan tres enfoques: entrenar una red neuronal convolucional (CNN) desde cero, utilizar transferencia de aprendizaje con el modelo VGG-16 y realizar un ajuste fino basado en el mismo modelo. Los resultados se basan en 8300 imágenes que contienen o no armas de fuego. Se encontró que el método de ajuste fino obtiene las precisiones más altas en la detección de armas de fuego en imágenes. Además, se destaca que este enfoque logra un equilibrio en la precisión al identificar imágenes con armas de fuego y aquellas sin ellas.

El estudio presentado por Dextre et al. (2021) aborda el desafío de la detección automatizada de armas en imágenes de vigilancia de video, que es complicado debido a factores como la iluminación, el enfoque y la resolución. La resolución de este problema tendría aplicaciones valiosas en la seguridad ciudadana. En este trabajo de investigación, se entrena un sistema de detección de armas basado en YOLOv5 (*You Only Look Once*) utilizando diferentes fuentes de datos. El sistema logra una precisión del 98.56% en imágenes de vigilancia de video y puede realizar inferencias en tiempo real a una velocidad de 33 fps en el Nvidia's Jetson AGX Xavier, lo cual es un resultado destacable en comparación con otros trabajos de investigación en el estado del arte.

2.3. Marco Legal

Ecuador enfrenta un desafío creciente en materia de seguridad pública debido al narcotráfico y la proliferación de armas de fuego. Para abordar esta problemática, se han implementado sistemas de videovigilancia basados en CCTV como una herramienta crucial para la prevención del crimen y la protección de la sociedad.

En este contexto, el marco legal, en particular el Código Orgánico Integral Penal (COIP) y la Ley Orgánica de Transparencia y Acceso a la Información Pública, desempeñan un papel fundamental. Estas regulaciones establecen delitos, penas y el derecho a acceder a información, lo que es esencial para garantizar el uso adecuado de los sistemas de videovigilancia y, en este proyecto, para respaldar la detección de armas en sistemas de CCTV.

1.3.1. Código Orgánico Integral Penal

El Código Orgánico Integral Penal (COIP) de Ecuador es la pieza legislativa fundamental que regula el sistema penal en el país. Este código establece una amplia variedad de normativas y sanciones en relación con delitos y penas. Algunos de los artículos clave vinculados al COIP son los siguientes (Oficial Suplemento, 2021):

Art. 178: Este artículo, prohíbe acceder, grabar o divulgar datos personales, mensajes, grabaciones de audio y vídeo, y otras comunicaciones privadas de una persona sin su consentimiento o autorización legal. La violación de esta prohibición puede resultar en una pena de prisión de uno a tres años. Sin embargo, existen excepciones, como cuando una persona divulga grabaciones en las que participa o cuando se trata de información pública según lo establecido por la ley. El objetivo del artículo es proteger la privacidad y la confidencialidad de la información personal y las comunicaciones.

Art. 179: Este artículo establece sanciones para quienes comercien o trafiquen con información personal grabada o difundida sin autorización. Es decir, prohíbe la venta o intercambio de información privada de personas que ha sido grabada o divulgada sin su consentimiento. El objetivo de este artículo es proteger la privacidad de las personas y prevenir la explotación comercial de información personal sin permiso. La violación de esta disposición puede dar lugar a sanciones legales de acuerdo con las leyes ecuatorianas.

1.3.2. Ley Orgánica de Transparencia y Acceso a la Información Pública

La Ley Orgánica de Transparencia y Acceso a la Información Pública, conocida como LOTAI, es una legislación fundamental en Ecuador que regula el acceso de los ciudadanos a la información pública. Algunos de los aspectos clave de esta ley son:

Art 4: Este artículo establece los principios fundamentales para el acceso a la información pública en Ecuador. Entre ellos, se destaca que la información pública pertenece a los ciudadanos y debe ser accesible de manera general y gratuita, fomentando la transparencia en la función pública y la participación ciudadana en la toma de decisiones.

La implementación de sistemas de detección de armas en tiempo real a través de cámaras de videovigilancia no vulnera la privacidad de las personas, ya que se rige por las disposiciones legales establecidas en la Constitución Política de la República y en la Ley Orgánica de Transparencia y Acceso a la Información Pública (LOTAI). Estas normativas garantizan el derecho a la información y, al mismo tiempo, protegen la privacidad y los datos personales de los ciudadanos. Los sistemas de detección de armas están diseñados para cumplir con los principios de apertura y publicidad de actuaciones, lo que promueve la transparencia en la gestión pública, la seguridad ciudadana y la prevención de delitos, sin afectar la privacidad de las personas.

CAPÍTULO III

3. METODOLOGÍA DE LA INVESTIGACIÓN

3.1. TIPO DE INVESTIGACIÓN

Con base en los objetivos planteados, los tipos de investigación a emplear son los siguientes:

3.1.1. Investigación descriptiva

La investigación descriptiva se enfoca en describir detalladamente un fenómeno, situación o problema específico. En este estudio, se empleó para proporcionar una descripción detallada de las técnicas de detección de objetos aplicadas a la identificación de armas de fuego, las métricas de evaluación y los resultados obtenidos. Además, se describió el proceso de comunicación entre el modelo y los encargados de seguridad.

3.1.2. Investigación aplicada

La investigación aplicada se centra en la aplicación práctica del conocimiento existente para abordar problemas específicos o satisfacer necesidades concretas. En este estudio, se adopta un enfoque de investigación aplicada para desarrollar y aplicar algoritmos de inteligencia artificial preexistentes en el análisis de imágenes de videovigilancia. Se utiliza el conocimiento y las técnicas disponibles en el campo de la inteligencia artificial para desarrollar soluciones prácticas y efectivas para el reconocimiento de armas de fuego en entornos de videovigilancia.

3.2. MÉTODOS DE INVESTIGACIÓN

En el marco de este proyecto de detección de armas de fuego en sistemas de videovigilancia, se aplican los siguientes métodos.

3.2.1. Método de Análisis

Este método se utiliza para evaluar y comparar diferentes técnicas y algoritmos de detección de armas de fuego. Se descompone cada técnica en sus componentes esenciales, como características, parámetros y métricas de rendimiento, con el fin de comprender su funcionamiento y efectividad en la detección.

3.2.2. Método de Síntesis

La síntesis se emplea para integrar y combinar información proveniente de diversas fuentes, incluyendo la literatura existente sobre detección de armas de fuego y los resultados de las pruebas realizadas. El objetivo es desarrollar un enfoque unificado y efectivo para la detección en tiempo real.

3.2.3. Método Deductivo

Se sigue un enfoque deductivo para comprobar el rendimiento de las técnicas de detección de armas de fuego en el contexto de sistemas de videovigilancia. Se basa en teorías previas y conocimientos existentes para establecer expectativas antes de la realización de pruebas y experimentos.

3.3. CONSTRUCCIÓN METODOLÓGICA DEL OBJETO DE INVESTIGACIÓN

3.3.1. Técnicas de investigación

Para llevar a cabo esta investigación sobre la detección de armas en sistemas de CCTV, se emplearon las siguientes técnicas.

Revisión bibliográfica y documental

Esta técnica implica la recopilación y revisión exhaustiva de la literatura especializada en detección de armas, sistemas de videovigilancia, algoritmos de aprendizaje profundo y métodos relacionados. Esto proporciona una base sólida de conocimientos teóricos para la investigación. En la Tabla 1 se describen los medios utilizados.

Tabla 1

Fuentes bibliográficas

Fuente	Descripción
Scopus	Base de datos sobre ciencia y tecnología que permite la consulta y el acceso a las referencias bibliográficas de 14.000 publicaciones científicas procedentes de 4.000 editoriales distintas.
IEEE Xplore	Base de datos de investigación académica que proporciona acceso al texto completo de artículos y trabajos sobre Ciencias de la Computación, Ingeniería Eléctrica y Electrónica.
Google Scholar	Buscador de Google enfocado y especializado en la búsqueda de contenido y bibliografía científico-académica.

Nota. En esta tabla se describen las fuentes y medios utilizados para la investigación.

Diseño de investigación

El grupo de investigación utilizado en este estudio es de tipo cuasi-experimental. Se definen tres grupos que corresponden a cada una de las técnicas de detección de objetos seleccionadas.

Grupo A: You Only Look Once (YOLO)

Para la implementación de YOLO se ha decidido utilizar la YOLOv8 debido a su rapidez y a su capacidad de procesar imágenes con un gran número de etiquetas. Esto permitirá facilitar la creación de los datos de entrenamiento permitiendo el etiquetado múltiple en las imágenes.

Implementación de YOLO

El algoritmo YOLO emplea redes neuronales convolucionales (CNN) para predecir varias probabilidades de clase y cuadros delimitadores simultáneamente.

En el caso de este algoritmo se variará sus 3 parámetros principales:

- **Batch:** El tamaño del lote (batch size) es un hiper parámetro que define la cantidad de muestras a través de las cuales trabajar antes de actualizar los parámetros internos del modelo.
- **InputSize:** Es el tamaño de entrada de las imágenes. Una resolución de píxeles más grande puede ayudar a su modelo a detectar objetos pequeños. La resolución de píxeles del modelo YOLO debe ser un múltiplo de 32.
- **Epoch:** El número de épocas (epoch) es un hiper parámetro que define el número de veces que el algoritmo de aprendizaje funcionará en todo el conjunto de datos de entrenamiento.

Estas variaciones en los parámetros se llevan a cabo con el objetivo de explorar diversas configuraciones del modelo, prevenir el sobreajuste, ajustar la convergencia del mismo y optimizar el uso de recursos computacionales durante el entrenamiento. Estas adaptaciones permitieron encontrar el equilibrio óptimo entre el rendimiento del modelo

y la eficiencia computacional, contribuyendo así a mejorar la precisión y generalización del modelo en diferentes conjuntos de datos y escenarios de aplicación.

Métricas de Evaluación para YOLO

La métrica de evaluación para YOLO en este caso son principalmente la precisión para medir la exactitud, la matriz de confusión, que permitió analizar el desempeño de los modelos en términos de verdaderos positivos (TP), falsos positivos (FP), verdaderos negativos (TN) y falsos negativos (FN). Además, se utilizó la F1 Confidence Curve, que proporciona una medida equilibrada de la precisión y el recall del modelo a lo largo de diferentes umbrales de decisión, ayudando a evaluar la robustez del modelo en diversos escenarios.

Grupo B: TensorFlow No-Code a través de Teachable Machine

Para el segundo grupo, se ha seleccionado la plataforma Teachable Machine, la cual utiliza TensorFlow No-Code. Esta elección se basa en su facilidad de uso y su capacidad para entrenar modelos de aprendizaje automático sin necesidad de escribir código. Utilizar esta plataforma permitió desarrollar un modelo de detección de armas de manera rápida y eficiente, sin requerir experiencia previa en programación.

Al usar Teachable Machine, cuando se entrena un modelo de reconocimiento de imágenes, se está reentrenando un modelo EfficientNet preentrenado con datos propios. Esto garantiza que los modelos creados no solo son altamente precisos y eficientes en el reconocimiento de objetos, sino que también están específicamente adaptados a las necesidades del proyecto mediante la utilización de datos personalizados.

Implementación de TensorFlow

La plataforma Teachable Machine emplea una interfaz intuitiva que permite entrenar modelos de aprendizaje automático sin necesidad de escribir código. En el caso de este proyecto, se utilizará para entrenar un modelo de detección de armas utilizando el mismo conjunto de imágenes etiquetadas.

Los parámetros principales que se variarán en la configuración del modelo son los siguientes:

- **Clases de armas a detectar:** Se etiquetarán las imágenes con las distintas clases de armas que se desean detectar, como pistolas, cuchillos, etc.
- **Número de épocas:** Se define el número de épocas como el número de veces que el algoritmo de aprendizaje funcionará a través del conjunto de datos de entrenamiento.
- **Tamaño de entrada de las imágenes:** El tamaño de entrada de las imágenes se ajustará para optimizar la detección de armas en diferentes tamaños y resoluciones.

Estas variaciones en los parámetros del modelo se llevarán a cabo con el objetivo de explorar diversas configuraciones del mismo, prevenir el sobreajuste y mejorar la precisión de la detección de armas en diferentes escenarios.

Métricas de Evaluación para TensorFlow No-Code

La métrica de evaluación para el grupo que utiliza Teachable Machine es principalmente la precisión para medir la exactitud del modelo, la matriz de confusión, que permite

analizar el desempeño del modelo en términos de verdaderos positivos (TP), falsos positivos (FP), verdaderos negativos (TN) y falsos negativos (FN).

Grupo C: Haar Cascade

Se optó por este algoritmo debido a su capacidad para realizar detecciones con un alto rendimiento en tiempo real. Haar Cascade aborda el problema de la detección de objetos mediante un enfoque basado en características de Haar y clasificadores en cascada. Este método es eficiente en términos de velocidad de detección, especialmente cuando se compara con métodos más complejos como Faster R-CNN, ya que no requiere generar propuestas de región.

Implementación de Haar Cascade

Haar Cascade utiliza una serie de clasificadores en cascada, cada uno entrenado con un conjunto de características simples de Haar extraídas de imágenes con y sin el objeto de interés. Estas características son usadas para crear un clasificador fuerte a partir de muchos clasificadores débiles, que filtran las regiones de la imagen donde el objeto podría estar presente.

Para optimizar el rendimiento del modelo en términos de precisión y velocidad, se realizaron ajustes en los parámetros del algoritmo, como el número de etapas de la cascada, el tamaño mínimo del objeto y el factor de escala. Estos ajustes permiten que Haar Cascade detecte objetos de diferentes tamaños y en diversas condiciones de iluminación con alta precisión y rapidez.

La experimentación con diferentes configuraciones permitió encontrar la combinación óptima de parámetros que maximizara la precisión de detección de estos algoritmos para esperar una alta eficiencia computacional.

Métricas de Evaluación para Haar Cascade

La matriz de confusión proporciona una visión detallada del rendimiento del modelo, permitiendo identificar posibles áreas de mejora y ajuste. Esta evaluación es fundamental para garantizar que el modelo de Haar Cascade pueda detectar de manera precisa y eficiente los objetos de interés en diferentes escenarios y condiciones.

3.3.2. Instrumentos de investigación

En el desarrollo de este estudio, se emplearon diversos instrumentos y tecnologías especializadas para llevar a cabo la investigación de manera efectiva y rigurosa. A continuación, se presenta una tabla detallada que resume los principales instrumentos utilizados, destacando su función específica y su valiosa contribución al éxito del proyecto:

Tabla 2

Instrumentos de investigación

Instrumento	Descripción
Python	Lenguaje de programación fundamental para el desarrollo de algoritmos y procesamiento de datos, permitiendo la implementación de técnicas de visión por computadora y aprendizaje automático.
OpenCV	Biblioteca de visión por computadora que proporciona herramientas para el procesamiento de imágenes y análisis de video, esenciales para la detección y reconocimiento de armas de fuego en los sistemas de CCTV.

Google Colab	Plataforma en línea que permite ejecutar código Python en un entorno de notebook colaborativo, facilitando el acceso a recursos computacionales y la colaboración en el desarrollo del proyecto entre diferentes investigadores.
PyCharm	Entorno de desarrollo integrado (IDE) que ofrece herramientas avanzadas para la escritura, depuración y ejecución de código Python, mejorando la eficiencia y precisión en el desarrollo de los algoritmos necesarios para la investigación.
Roboflow	Plataforma utilizada para la gestión y etiquetado de conjuntos de datos de imágenes, simplificando el proceso de preparación de datos y generación de anotaciones, cruciales para entrenar modelos de detección de armas de fuego con precisión.

Nota. Los instrumentos mencionados han sido seleccionados por su relevancia y utilidad en el contexto de la investigación.

3.4. ELABORACIÓN DEL MARCO TEÓRICO

En la elaboración del marco teórico para este proyecto, se siguieron los pasos detallados en la Ilustración 3:

Ilustración 3

Diagrama de elaboración del marco teórico



Paso 1: Identificación de conceptos clave

Se comenzó por identificar los conceptos clave relacionados con la detección de armas de fuego en sistemas de videovigilancia, como algoritmos de aprendizaje profundo, sistemas de seguridad, tecnología de cámaras, entre otros.

Paso 2: Búsqueda de fuentes de información

Se llevó a cabo una búsqueda exhaustiva de fuentes de información secundaria en libros, revistas académicas, tesis, ensayos, informes técnicos y páginas web confiables. Se emplearán bases de datos académicas y bibliotecas digitales para acceder a la información pertinente.

Paso 3: Análisis de la literatura

Se procedió a leer y analizar críticamente la literatura encontrada. Se buscarán estudios previos relacionados con la detección de armas de fuego en sistemas de videovigilancia, métodos de aprendizaje profundo, tecnología de cámaras de seguridad y aplicaciones prácticas en la seguridad pública. Se identificaron tendencias, avances tecnológicos y desafíos actuales en el campo.

Paso 4: Organización de conceptos

Los conceptos clave y los hallazgos del análisis se organizaron en categorías temáticas, lo que facilitará la estructuración lógica y coherente del marco teórico.

Paso 5: Identificación de marco conceptual

Se definirá un marco conceptual que presente los conceptos clave y las relaciones entre ellos. Se utilizarán diagramas conceptuales, mapas conceptuales o esquemas para visualizar estas conexiones.

Paso 6: Revisión de marco legal

Si el proyecto involucra aspectos legales, se revisaron y analizaron las legislaciones y regulaciones pertinentes relacionadas con la videovigilancia, la privacidad y la seguridad. Se aseguró de entender las implicaciones legales del proyecto.

Paso 7: Síntesis y redacción

Se sintetizó la información recopilada en un marco teórico coherente. El marco teórico se redactó de manera clara y organizada, relacionando los conceptos clave, teorías y marco legal con la investigación.

Paso 8: Validación y actualización

El marco teórico se compartió con expertos en el campo para obtener retroalimentación. Se asegurará de que el marco teórico sea completo y actualizado de acuerdo con las últimas investigaciones.

3.5. RECOLECCIÓN DE LA INFORMACIÓN

La recolección de información para este proyecto se llevó a cabo creando conjuntos de datos de imágenes de armas de fuego utilizando cámaras de alta resolución. Estos datasets proporcionaron la información primaria necesaria para entrenar y evaluar el sistema de

detección de armas de fuego en sistemas de videovigilancia. Para ello se planteó seguir los siguientes pasos:

Obtención de imágenes: Se extrajeron imágenes para la creación de un dataset propio de armas de fuego capturadas por las cámaras de un sistema de videovigilancia de una unidad educativa, asegurando que las armas estuvieran claramente visibles en el campo de visión de las cámaras para optimizar los algoritmos en diversos escenarios. En este caso particular, se trabajó con dos tipos de armas de fuego: pistolas y carabinas. La razón de seleccionar estos tipos se debió a su disponibilidad; fueron las armas a las que se tuvo acceso. Es importante destacar que se contó con una mayor cantidad de imágenes de pistolas en comparación con las de carabinas. Además, se consideraron imágenes en distintos ángulos, condiciones de iluminación y escenas, utilizando herramientas de clasificación de imágenes y técnicas de aprendizaje automático para garantizar la calidad y diversidad del conjunto de datos.

Construcción del conjunto de datos de entrenamiento: Se construyó un dataset de entrenamiento que incluya una variedad de imágenes de armas de fuego de distintos tipos en diferentes condiciones y escenarios. Esto permitirá que el modelo de detección se entrene de manera efectiva.

Etiquetado de imágenes: Las imágenes en el conjunto de datos se etiquetaron minuciosamente para identificar la presencia de armas de fuego en cada imagen. El etiquetado se realizó con información precisa que indique la ubicación y la clase del objeto detectado.

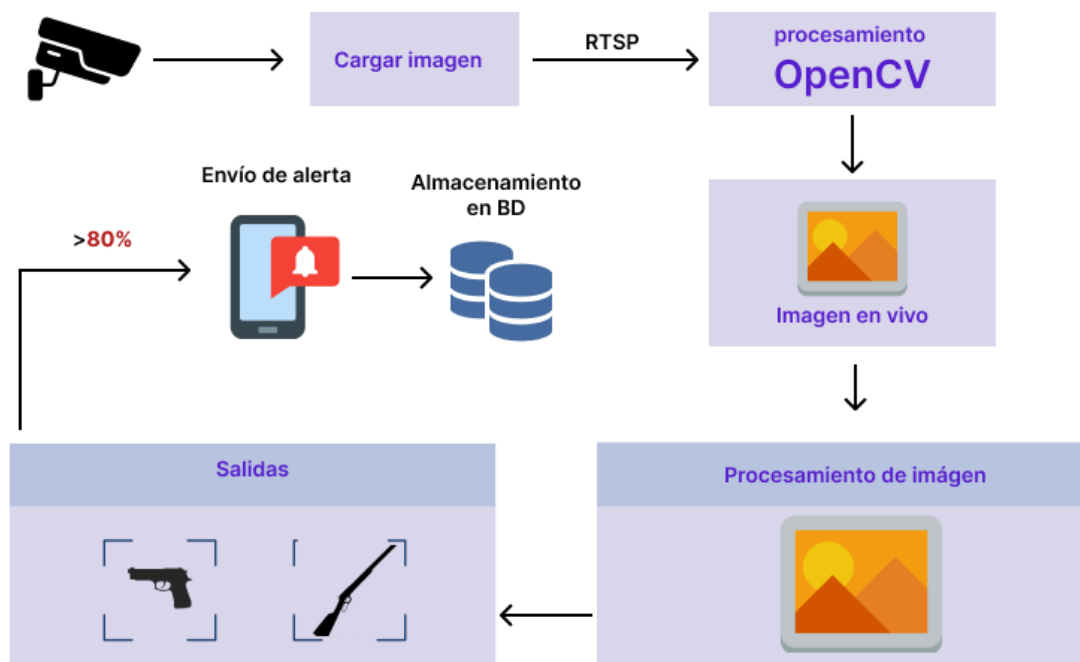
Validación y prueba: Se dividió el conjunto de datos en conjuntos de entrenamiento, validación y prueba para evaluar el rendimiento del modelo. El conjunto de validación se utilizó para ajustar hiper parámetros y el conjunto de prueba para evaluar la precisión del sistema.

3.6. PROCESAMIENTO Y ANÁLISIS

Para el procesamiento de la información, se empleó principalmente el lenguaje de programación Python, reconocido por su versatilidad y la amplia gama de bibliotecas y herramientas disponibles para el procesamiento de datos y visión artificial, convirtiéndolo así en la elección idónea para este proyecto.

Ilustración 4

Arquitectura de alerta de objetos



Nota. La captura de vídeo de la cámara se lleva a cabo utilizando el protocolo RTSP, en el cual se implementa el algoritmo de detección de objetos.

Los mecanismos de procesamiento se desarrollarán mediante software y se dividirán en las siguientes etapas:

Captura de video

Para la captura de secuencias de video en tiempo real de sistemas de videovigilancia, se empleará la biblioteca OpenCV (cv2 en Python). Esta herramienta permite procesar cada cuadro de video de manera individual para su análisis y procesamiento posterior.

Detección de armas de fuego

La detección de objetos, específicamente armas de fuego, se realizará mediante algoritmos avanzados de aprendizaje profundo. Se utilizarán las siguientes bibliotecas especializadas:

- **YOLO (Ultralytics):** Este modelo preentrenado permitirá una detección rápida y precisa de objetos en imágenes y videos.
- **TensorFlow:** Utilizado para la construcción y entrenamiento de modelos de aprendizaje profundo, será fundamental para la detección de objetos en tiempo real.
- **PyTorch:** Esta biblioteca se aplicará en tareas específicas de visión por computadora y redes neuronales, mejorando la precisión y eficiencia del sistema de detección.

Variación de Parámetros del Modelo

Para optimizar el rendimiento del modelo de detección de armas, se llevaron a cabo varios experimentos variando los parámetros clave del modelo. Los parámetros ajustados incluyen el número de épocas, la tasa de aprendizaje y el tamaño del lote.

Generación de alertas

Cuando se detecte la presencia de un arma de fuego, el sistema activará alertas tanto visuales como auditivas. Estas alertas estarán diseñadas para informar de manera inmediata a los usuarios sobre la situación detectada.

Además de las herramientas mencionadas para el procesamiento de video y detección de objetos, se utilizarán las siguientes bibliotecas complementarias:

- **os:** Para gestionar y almacenar imágenes capturadas en el sistema local.
- **pyodbc:** Utilizado para la conexión y manipulación de datos en una base de datos SQL Server, donde se almacenarán los registros y resultados del sistema.
- **telebot:** Implementado para enviar imágenes procesadas a través de un bot en Telegram, facilitando la notificación y visualización remota de incidentes.
- **datetime:** Para registrar la fecha y hora exactas de cada detección, proporcionando contexto temporal a los eventos capturados.

Esta infraestructura tecnológica asegura un funcionamiento integral y eficiente del sistema de vigilancia y detección de armas de fuego, contribuyendo significativamente a la seguridad y protección pública.

Análisis de Datos

El análisis de datos se llevó a cabo utilizando técnicas de procesamiento de imágenes y datos para evaluar el rendimiento del sistema de detección. Los criterios para el análisis de datos incluirán:

- **Precisión de detección:** Se evalúa la capacidad del sistema para detectar armas de fuego con precisión. Se calculará la tasa de verdaderos positivos, falsos positivos, verdaderos negativos y falsos negativos.
- **Tiempo de respuesta:** Se mide el tiempo que el sistema tarda en detectar un arma de fuego desde el momento en que aparece en el cuadro de video.
- **Rendimiento del modelo:** Se evalúa el rendimiento del modelo de detección, incluyendo su capacidad para manejar diferentes condiciones de iluminación, resoluciones de video y ángulos de cámara.

El procesamiento y análisis de datos se llevó a cabo de manera sistemática y se utilizarán herramientas y bibliotecas específicas en Python para cumplir con los objetivos del proyecto.

CAPÍTULO IV

4. RESULTADOS Y DISCUSIÓN

4.1. CREACIÓN DEL CONJUNTO DE DATOS DE IMÁGENES DE ARMAS DE FUEGO DE ALTA CALIDAD Y DIVERSIDAD PARA EL ENTRENAMIENTO DE LOS MODELOS

Para el primer objetivo específico, se desarrolló un conjunto de datos de imágenes de armas de fuego, incluyendo pistolas y carabinas, con el fin de entrenar los modelos de detección. A continuación, se describe el proceso paso a paso:

4.1.1. Recopilación de imágenes

Imágenes Capturadas por Cámaras de Seguridad

- ❖ Se utilizaron cámaras de seguridad de 4MP instaladas en un CCTV de una unidad educativa, capturando imágenes en las entradas principales.
- ❖ Las cámaras lograron registrar con claridad las personas que entraban y salían, proporcionando una base de datos sólida para el análisis.

Simulaciones con Personal de Seguridad

- ❖ Durante las simulaciones, se grabaron personas portando pistolas y carabinas en diversas posiciones y distancias.
- ❖ Las simulaciones incluyeron tomas a diferentes distancias, especificadas a continuación:
 - **Cámara 1:** Capturó imágenes con claridad a distancias entre 1 y 3 metros.
 - **Cámara 2:** Capturó imágenes efectivas a distancias entre 2 y 6 metros.

Tomas en Superficie Plana

- ❖ Se filmaron tomas detalladas de las armas sobre una superficie plana con fondo sólido.
- ❖ La claridad y definición de las imágenes en estas condiciones fueron óptimas para el propósito del análisis.

Estas imágenes de armas siendo manipuladas por personas demostraron ser altamente efectivas para el entrenamiento del modelo, como se muestra en la Ilustración 5.

Ilustración 5

Construcción del conjunto de imágenes



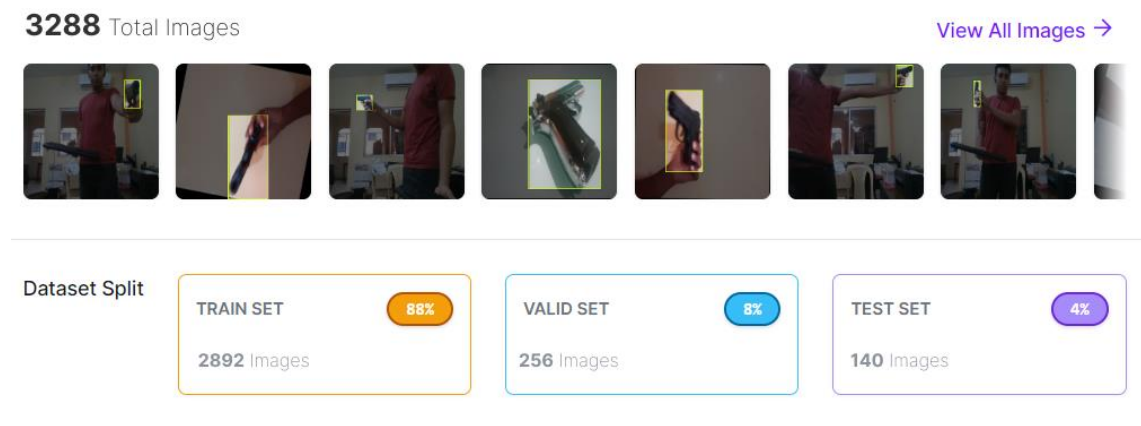
Se repitió el proceso con pistolas, revólveres, escopeta de mano, pistola de juguete y cuchillo como objeto de interés adicional.

4.1.2. Extracción y etiquetado de imágenes

El etiquetado de las imágenes de armas de fuego utilizando la herramienta GroundingDino de Python resultó en un conjunto de datos de alta calidad, con un notable grado de precisión gracias a la integración de inteligencia artificial. La validación de este conjunto de datos a través de la plataforma Roboflow permitió identificar y corregir los errores en el etiquetado inicial, garantizando así la fiabilidad del conjunto de datos para el entrenamiento del modelo. Como resultado, el conjunto de datos refinado contiene imágenes etiquetadas con precisión, proporcionando una base sólida para el desarrollo de modelos de detección de armas altamente efectivos.

Ilustración 6

Conjunto de datos para entrenamiento (Algoritmo YOLOv8)



La Ilustración 6 muestra el proceso de recopilación y etiquetado que dio como resultado un conjunto de datos compuesto por 3288 imágenes de alta calidad y diversidad. De estas, 2892 imágenes se destinaron al entrenamiento de los modelos, 140 imágenes a la prueba y 256 imágenes a la validación. Esta distribución estratégica permite garantizar:

- **Un entrenamiento robusto:** La cantidad considerable de imágenes de entrenamiento (2892) expone a los modelos a una amplia variedad de escenarios y características, fortaleciendo su capacidad para detectar armas de fuego con precisión.
- **Una validación confiable:** Las 256 imágenes de validación, también independientes, corroboran la precisión y confiabilidad de los modelos en diferentes condiciones.
- **Una evaluación precisa:** Las 140 imágenes de prueba, independientes del conjunto de entrenamiento, permiten evaluar el rendimiento real de los modelos en situaciones nuevas.

Finalmente, se exportó el conjunto de datos etiquetado y validado para su aplicación en el modelo de visión por computadora YOLOv8, el cual será entrenado para la detección de armas de fuego en tiempo real.

Este proceso de etiquetado y validación ha permitido obtener un conjunto de datos de alta calidad y precisión, lo cual es fundamental para entrenar este modelo de aprendizaje profundo.

Roboflow no solo facilitó la creación y validación del conjunto de datos, sino que también permitió su exportación en diversos formatos. Esto facilita el entrenamiento de modelos de detección de objetos con diferentes arquitecturas y técnicas, como YOLOv8, Haar Cascade o Tensorflow. De esta manera, se pueden comparar y evaluar diferentes enfoques para la detección de armas de fuego, seleccionando la solución más adecuada para las necesidades específicas del proyecto.

4.1.3. Discusión

La creación del conjunto de datos de imágenes de armas de fuego fue un proceso meticuloso y estratégico. Las cámaras del sistema de CCTV permitieron capturar imágenes realistas en las entradas principales y realizar simulaciones controladas con la colaboración del personal de seguridad. Estas incluyeron personas manipulando pistolas y carabinas, así como tomas con un arma de juguete realista para garantizar la seguridad durante las grabaciones adicionales.

Las imágenes recopiladas fueron etiquetadas con precisión utilizando GroundingDino de Python y validadas en Roboflow para corregir posibles errores. El conjunto final consistió en 3288 imágenes, distribuidas estratégicamente para entrenamiento, validación y prueba.

Esta diversidad aseguró un entrenamiento robusto y una evaluación confiable del modelo de detección.

Se exportó el conjunto de datos en varios formatos para ser compatible con diferentes arquitecturas de modelos como YOLOv8, Haar Cascade y TensorFlow, permitiendo la flexibilidad en el desarrollo y evaluación de los modelos. Este proceso es fundamental para mejorar la seguridad pública mediante tecnologías de visión por computadora en entornos de vigilancia.

4.2. APLICACIÓN DE DISTINTAS TÉCNICAS DE RECONOCIMIENTO DE OBJETOS PARA DETERMINAR CUÁL ES LA MÁS EFICAZ EN LA DETECCIÓN DE ARMAS DE FUEGO.

4.2.1. Entrenamiento y evaluación de modelos de detección

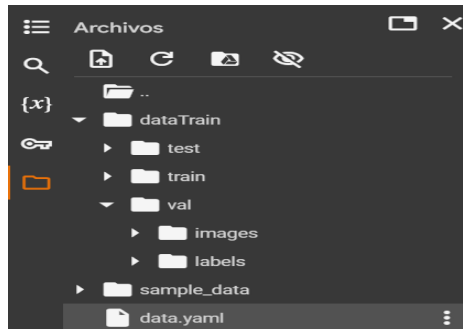
GRUPO A:

La implementación del modelo de detección de armas de fuego se llevó a cabo utilizando el algoritmo YOLOv8 en el entorno de programación Python. Los pasos seguidos durante la implementación incluyeron:

- **Preparación de los datos:** Se utilizó el conjunto de datos compuesto por imágenes de armas de fuego etiquetadas con sus respectivas clases y localizaciones y se optimizaron a través de roboflow para su correcta compatibilidad con YOLO. Este conjunto de datos se utilizó para el entrenamiento del modelo de detección con la estructura que puede apreciarse en la figura:

Ilustración 7

Estructura del conjunto de imágenes de entrenamiento, validación y evaluación



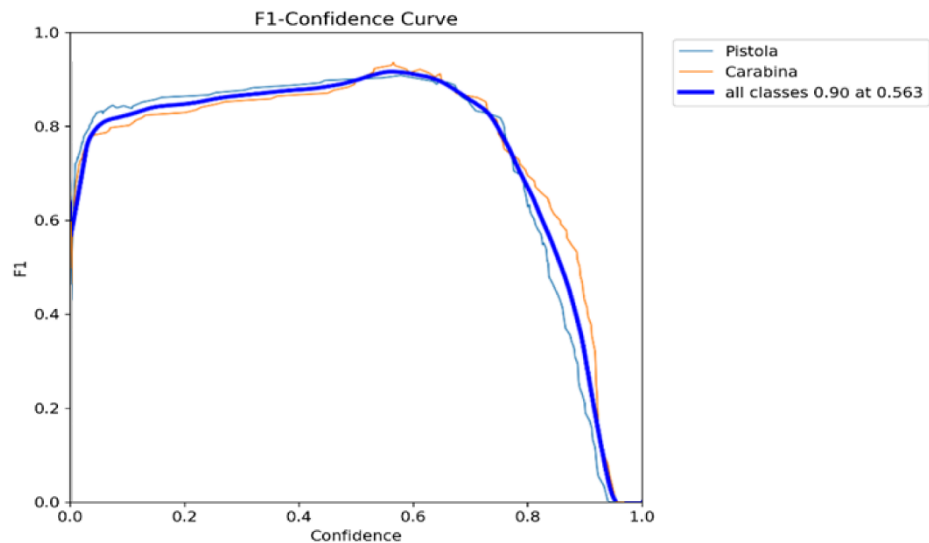
Nota. Esta imagen muestra la distribución estratégica de las imágenes en conjuntos de entrenamiento, validación y prueba. Es crucial para asegurar que el modelo se entrene adecuadamente con datos diversificados y se evalúe de manera efectiva su rendimiento en diferentes condiciones.

- **Entrenamiento del modelo:** Se llevó a cabo el entrenamiento del algoritmo YOLOv8 utilizando el conjunto de datos preparado. Durante este proceso, se variaron los parámetros del modelo, como el tamaño del lote, el tamaño de entrada de las imágenes y el número de épocas, con el fin de encontrar la configuración óptima que maximizara la precisión y el rendimiento del modelo.
- **Evaluación del modelo:** Una vez completado el entrenamiento, se evaluó el rendimiento del modelo utilizando un conjunto de datos de prueba independiente. Se calcularon métricas como la precisión, la sensibilidad y el valor F1 para evaluar la capacidad del modelo para detectar con precisión la presencia de armas de fuego en las imágenes.

La Ilustración 8 muestra una curva de precisión-confianza. La curva de precisión-confianza es un gráfico que muestra la relación entre la precisión de las predicciones de un modelo y la confianza que el modelo tiene en esas predicciones.

Ilustración 8

Curva de confianza – YOLOv8



Nota. Como se observa, la elección del umbral de confianza adecuado depende de la aplicación específica y de los requisitos del usuario. Un umbral alto (por ejemplo, 0,95) minimiza los falsos positivos (instancias que se identifican erróneamente como armas) a costa de perder algunas armas reales. Por otro lado, un umbral bajo (por ejemplo, 0,5) aumenta la probabilidad de detectar todas las armas reales, pero también incrementa el número de falsos positivos.

- **El eje X:** Representa la confianza del modelo en su clasificación. Una confianza de 1.0 significa que el modelo está 100% seguro de que un punto de datos dado pertenece a una clase particular (pistola, carabina, etc.).

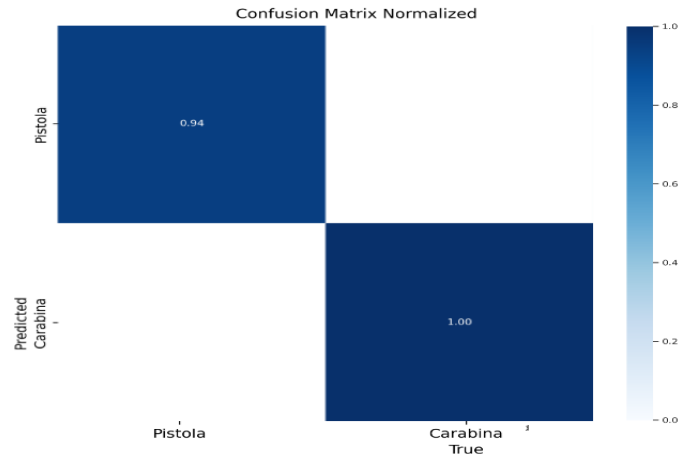
- **El eje Y:** Muestra la precisión, que es la proporción de predicciones positivas que fueron realmente correctas. Una precisión de 1.0 significa que todas las clasificaciones positivas fueron correctas.

La precisión del modelo, medida en 0.884, indica que tiene un 88% de probabilidades de clasificar correctamente una instancia como "Pistola" cuando la confianza del modelo es de al menos 0.884. En términos prácticos, esto significa que si el modelo predice que una instancia es una "Pistola" con una confianza de 0.884 o superior, existe un 88% de probabilidad de que efectivamente sea una "Pistola".

A continuación, se observa la matriz de confusión:

Ilustración 9

Matriz de confusión normalizada



Nota: La matriz de confusión muestra que el modelo es muy bueno para clasificar pistolas. De la 1 instancia que en realidad era una pistola, el modelo la clasificó correctamente como pistola 1 vez y la clasificó incorrectamente como carabina 0 veces. Esto significa que el modelo tiene una **precisión** de 1.0 para pistolas, que es la precisión más alta posible.

El modelo también es bueno para clasificar carabinas. De la 1 instancia que en realidad era una carabina, el modelo la clasificó correctamente como carabina 1 vez y la clasificó incorrectamente como pistola 0 veces. Esto significa que el modelo tiene una precisión de 1.0 para carabinas.

GRUPO B:

La implementación del modelo de detección de armas de fuego se llevó a cabo utilizando el algoritmo Tensorflow - Teachable Machine en el entorno de programación Python. Los pasos seguidos durante la implementación incluyeron:

Se utilizó la plataforma Teachable Machine para entrenar un modelo de reconocimiento de armas de fuego, específicamente pistolas y carabinas. El proceso incluyó las siguientes etapas:

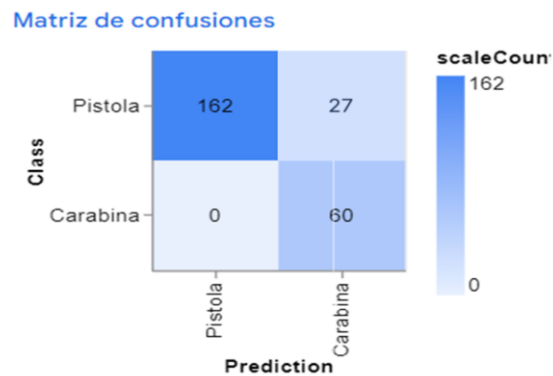
- **Recopilación de Datos:** Se reunieron imágenes de pistolas y carabinas en diversas condiciones de iluminación y ángulos para asegurar la diversidad del conjunto de datos.
- **Entrenamiento del Modelo:** Las imágenes se cargaron en Teachable Machine, clasificándose en dos categorías: pistolas y carabinas. El modelo se entrenó utilizando la interfaz intuitiva de la plataforma.
- **Evaluación del Modelo:** Tras el entrenamiento, se evaluó el modelo con un conjunto de imágenes de prueba no incluidas en el entrenamiento inicial.

El modelo logró una alta precisión en la identificación de pistolas, mostrando una capacidad robusta para reconocer esta categoría. Además, se observó que el modelo sufrió

de sobreajuste en la identificación de carabinas, mostrando menor precisión y una tendencia a reconocer incorrectamente esta categoría en las imágenes de prueba.

Ilustración 10

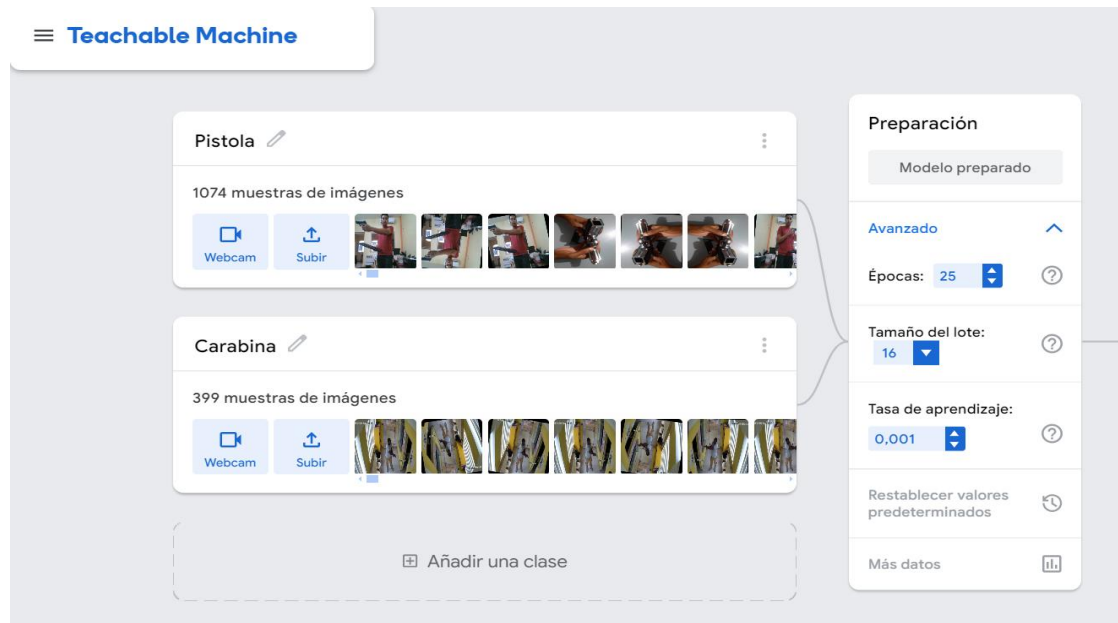
Matriz de confusión – Teachable Machine



Teachable Machine facilitó un entrenamiento rápido y sencillo del modelo de reconocimiento de armas de fuego. Aunque el modelo mostró alta precisión para pistolas, el sobreajuste en carabinas indica la necesidad de ajustar el conjunto de datos y parámetros de entrenamiento para mejorar su rendimiento en esta categoría.

Ilustración 11

Interfaz de Teachable Machine



GRUPO C:

Se utilizó la técnica Haar Cascade para entrenar un modelo de reconocimiento de armas de fuego, específicamente pistolas y carabinas. A diferencia de otros métodos, Haar Cascade requiere imágenes tomadas en el lugar con y sin el objeto de estudio, lo que limitó el tamaño del conjunto de datos utilizado para el entrenamiento.

- **Recopilación de Datos:** Se reunieron imágenes específicas del entorno, tanto con las armas de fuego (pistolas y carabinas) como sin ellas, debido a los requisitos de la técnica Haar Cascade.
- **Entrenamiento del Modelo:** Utilizando la técnica Haar Cascade, se entrenó el modelo con el conjunto de imágenes reducido. El proceso implicó la creación de clasificadores en cascada basados en características de Haar.

- **Evaluación del Modelo:** Se evaluó el rendimiento del modelo con un conjunto de imágenes de prueba que incluían las mismas condiciones del entorno utilizado para el entrenamiento.

El modelo mostró buenos resultados en la identificación de pistolas y carabinas cuando se evaluó en el mismo entorno utilizado para el entrenamiento. La precisión fue alta debido a la similitud de las condiciones.

La aplicación del modelo en entornos que no fueron previamente entrenados mostró deficiencias significativas. La precisión disminuyó notablemente debido a la falta de imágenes de esos nuevos entornos en el conjunto de entrenamiento.

La técnica presenta limitaciones significativas al aplicarse en entornos no entrenados, lo que sugiere la necesidad de recopilar un conjunto de datos más amplio y diverso para mejorar su capacidad de generalización.

Resumen

En este estudio, se evaluaron tres técnicas diferentes para la detección de armas de fuego: YOLOv8, Teachable Machine con Tensorflow, y Haar Cascade. A continuación, se presentan los principales hallazgos de cada método:

YOLOv8

- **Implementación:** Utilización de Python y optimización de datos con Roboflow.
- **Rendimiento:** Alcanzó una precisión del 88%, destacándose por su capacidad para identificar pistolas con alta precisión bajo diferentes umbrales de confianza.

- **Discusión:** Se resalta la importancia de seleccionar el umbral de confianza adecuado para equilibrar la precisión y la detección de falsos positivos.

Teachable Machine (Tensorflow)

- **Implementación:** Entrenamiento intuitivo a través de la plataforma Teachable Machine.
- **Rendimiento:** Mostró alta precisión en la detección de pistolas, pero evidenció sobreajuste en la clasificación de carabinas.
- **Discusión:** Se identificó la necesidad de ajustar los datos y parámetros del modelo para mejorar su rendimiento en categorías menos representadas.

Haar Cascade

- **Implementación:** Utilización de clasificadores en cascada basados en características de Haar.
- **Rendimiento:** Buen desempeño en entornos específicos con alta precisión en la identificación de pistolas y carabinas.
- **Discusión:** Limitaciones significativas al aplicarse en entornos no entrenados, sugiriendo la expansión del conjunto de datos para mejorar la generalización del modelo.

Variación de Parámetros del Modelo YOLOv8

Dado el rendimiento superior de YOLOv8, se procedió a optimizar este modelo ajustando varios parámetros clave, tales como el número de épocas, la tasa de aprendizaje y el tamaño del lote. Los resultados de estos experimentos se resumen en la siguiente tabla.

Tabla 3

Tabla 3 Variación de parámetros YOLOv8

Experimento	Épocas	Tasa de Aprendizaje	Tamaño del Lote	Precisión (%)	Recall (%)	F1-Score (%)
1	50	0.01	32	85	80	82.5
2	100	0.001	32	90	85	87.5
3	100	0.01	64	88	83	85.5
4	150	0.001	64	92	88	90.0
5	150	0.005	32	89	87	88.0

Los resultados de la variación de parámetros del modelo YOLOv8 demuestran que el experimento con 150 épocas, una tasa de aprendizaje de 0.001 y un tamaño del lote de 64 (Experimento 4) mostró el mejor rendimiento general, con una precisión del 92%, un recall del 88% y un F1-Score del 90%.

4.3. ESTABLECIMIENTO DE UN CANAL DE COMUNICACIÓN QUE PERMITA LA NOTIFICACIÓN DE LA PRESENCIA DE ARMAS DE FUEGO A PARTIR DE LOS RESULTADOS GENERADOS POR EL MODELO

Para capturar las imágenes de las cámaras del sistema de videovigilancia NVR, se habilitó el protocolo RTSP. Esto permitió la obtención del video en vivo de las cámaras mediante OpenCV, en este caso en particular se trabajó con una cámara de 4MP para la integración del modelo de detección en tiempo real.

Ilustración 12

Código – Detección de objetos en tiempo real

```
#Realizar detección de objetos

if len(results) > 0:
    for res in results:
        print("Tool detect")
        # Crear imagen etiquetada
        annotated_frames = results[0].plot()
        # Generar una marca de tiempo
        timestamp
datetime.datetime.now().strftime("%Y%m%d %H%M%S")
```

```

        #Se guarda la imagen
        save_image(annotated_frame, timestamp, output_folder,
class_name)

    # Redimensionar el frame para ajustarlo a la pantalla
    frame_redimensionado = cv2.resize(frame, (1280, 720))

    # Mostrar la imagen
    cv2.imshow('Tool detect', annotated_frames)

    t = cv2.waitKey(5)
    if t == 27:
        break

cap.release()

cv2.destroyAllWindows()

```

4.3.1. Proceso de detección y generación de alertas

El sistema opera continuamente utilizando un bucle de lectura de video donde se aplica la detección de objetos mediante un modelo preentrenado. Cuando se detecta la presencia de un arma de fuego con una confianza mínima del 80%, se procede con la generación de alertas automáticas. Este proceso incluye la creación de una imagen etiquetada con el objeto detectado y la inclusión de una marca de tiempo precisa utilizando la biblioteca.

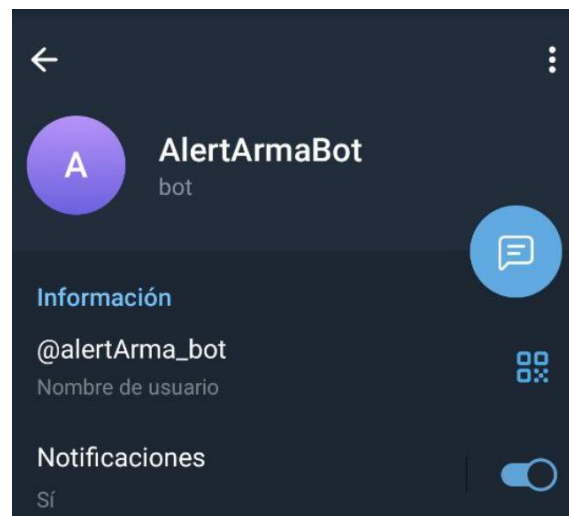
Para garantizar una comunicación efectiva y rápida de las notificaciones, se optó por utilizar la aplicación de mensajería Telegram debido a su versatilidad y robustez en cuanto a servicios disponibles. El primer paso consistió en identificar y obtener el ID del grupo específico al cual se dirigirían las alertas. Para esto, se empleó el bot de Telegram **@Rose**, una herramienta que facilita la extracción automática del ID del grupo objetivo.

Con el ID del grupo obtenido, se procedió a configurar un bot personalizado en Telegram utilizando la librería Telebot en Python. Este bot fue nombrado "**@alertArma_bot**"

(Ilustración 13), diseñado específicamente para integrar y gestionar la función de generación de alertas. La configuración incluyó la programación de comandos y funciones necesarias para enviar imágenes capturadas en tiempo real cuando se detecta la presencia de armas de fuego con una precisión del 80% o superior.

Ilustración 13

Bot de Telegram



Posteriormente, se desarrolló una función personalizada que permite al bot enviar una imagen al grupo de Telegram cuando se detecta la presencia de un arma con una precisión del 80 % o superior.

El bot @alertArma_bot se convirtió en una pieza central para la operación del sistema de vigilancia, permitiendo la entrega instantánea de alertas visuales al grupo de Telegram configurado. Esta solución no solo facilita la supervisión remota y la respuesta inmediata ante incidentes, sino que también optimiza la gestión de emergencias al proporcionar información crucial de manera eficiente y directa a los responsables de la seguridad.

Ilustración 14

Código – Envío de alertas

```
def enviar_imagen_con_mensaje(chat_id, imagen_path, mensaje):  
    with open(imagen_path, "rb") as imagen:  
  
        bot.send_photo(chat_id, imagen, caption=mensaje)
```

La función de generación de alertas se diseñó para ejecutarse en tiempo real, integrándose con el flujo de procesamiento de video y detección de armas. Cuando el sistema identifica la presencia de un arma de fuego en un cuadro de video con la precisión establecida, se activa automáticamente la función de alerta del bot.

Ilustración 15

Chat de telegram –Alarma de detección de objeto



Nota: La imagen ilustra cómo se envían automáticamente las alertas visuales al grupo de chat mediante un bot de Telegram, facilitando la supervisión en tiempo real.

6.1.1. Discusión

El envío de la notificación es automático y se produce inmediatamente después de que el sistema reconoce la presencia de un arma. Sin embargo, persiste un ligero retardo desde el momento en que la cámara captura la imagen y la transmite a través del protocolo RTSP, aproximadamente de un segundo, seguido de un segundo adicional mientras el algoritmo realiza la detección.

También se observa que, en algunas tomas, especialmente cuando el arma está parcialmente cubierta, el reconocimiento no es tan preciso. A pesar de que se estableció un umbral de precisión del 80%, aún se presentan momentos leves de falsos positivos. Si bien podría considerarse trabajar con un umbral de precisión del 85%, existe el riesgo de pasar por alto objetos relevantes, por lo que se optó por mantenerlo en el 80%.

Este análisis llevó a concluir que, si bien el sistema responde de manera automática y casi instantánea a la detección de armas, aún existen algunos aspectos a mejorar, como la reducción del retardo en la transmisión de imágenes y la minimización de los falsos positivos en situaciones donde las armas están parcialmente cubiertas. Estos hallazgos proporcionan una base para futuras mejoras en el sistema de detección y alerta.

CAPÍTULO V

5. CONCLUSIONES Y RECOMENDACIONES

Conclusiones

En base a los resultados obtenidos en esta investigación, a modo de conclusión se puede resaltar lo siguiente:

- Se llevó a cabo la creación de un conjunto de datos de imágenes para el entrenamiento de los modelos de detección. Este proceso implicó la recopilación de imágenes de cámaras de seguridad, así como la grabación controlada de situaciones con armas de fuego, incluyendo pistolas y carabinas, en diferentes distancias y ángulos de visión. Posteriormente, se procedió al etiquetado meticuloso de las imágenes utilizando herramientas especializadas como GroundingDino y Roboflow. El conjunto de datos resultante, compuesto por 3288 imágenes de alta calidad y diversidad, proporciona una base sólida para el entrenamiento de los modelos de detección, garantizando un proceso de aprendizaje robusto y una evaluación precisa de su rendimiento.
- Con el objetivo de determinar la técnica más eficaz en la detección de armas de fuego, se llevaron a cabo pruebas con distintas técnicas de reconocimiento de objetos, incluyendo YOLO, Tensorflow y Haar Cascade. Los resultados obtenidos indican que la técnica YOLO, específicamente la versión YOLOv8, demostró ser la más efectiva en la detección de armas de fuego en términos de precisión y rendimiento. Este análisis se sustenta en métricas como la precisión, la matriz de confusión y F1, así como en la evaluación cualitativa de las detecciones realizadas por cada técnica. Estos hallazgos respaldan la elección de YOLOv8 como la técnica preferida para la detección de armas de fuego en este proyecto.
- Para establecer un canal de comunicación efectivo que permitiera la notificación de la presencia de armas de fuego, se desarrolló un sistema automatizado de alerta

utilizando la aplicación de mensajería Telegram. Este sistema, integrado con el modelo de detección de armas, envía notificaciones en tiempo real a un grupo de Telegram designado cuando se identifica la presencia de un arma con una precisión del 80% o superior. A pesar de algunos desafíos relacionados con la transmisión de imágenes y la precisión de la detección, el sistema de notificación proporciona una solución efectiva para alertar a los usuarios sobre la presencia de armas de fuego en entornos de vigilancia, mejorando así la capacidad de respuesta y seguridad del sistema de vigilancia.

Recomendaciones

A continuación, se presentan algunas sugerencias para ampliar las prestaciones de la aplicación y aumentar la precisión del algoritmo de reconocimiento de armas de fuego.

- Se recomienda incorporar técnicas de aumento de datos durante el entrenamiento del modelo. Aplicar transformaciones como rotaciones, traslaciones y cambios en la escala de las imágenes puede aumentar la diversidad del conjunto de datos. Esto mejorará la capacidad de generalización del modelo y reducirá el riesgo de sobreajuste.
- Se sugiere explorar el uso de varios modelos de detección de objetos optimizados para detectar tipos específicos de armas de fuego. Al combinar los resultados de estos modelos especializados, se puede mejorar la precisión global del sistema y reducir la incidencia de falsos positivos.
- Se recomienda implementar un sistema automatizado de retroalimentación que permita a los usuarios proporcionar comentarios sobre las detecciones realizadas por el modelo. Este proceso será fundamental para identificar y corregir áreas de mejora en la capacidad de detección y clasificación de armas de fuego, facilitando el reentrenamiento automático del modelo para una mejora continua.

BIBLIOGRAFÍA

- ACUÑA, R. J. (2021). EL TRÁFICO ILEGAL DE ARMAS COMO UNA AMENAZA A LA SEGURIDAD INTEGRAL DEL ESTADO. *Revista de La Academia Del Guerra Del Ejército Ecuatoriano*, 14(1), 11. <https://doi.org/10.24133/age.n14.2021.05>
- Alvarez, C. (2020). Lesiones con armas de fuego_ sobrevivir a las balas en Ecuador. *Coyuntura*.
- APD. (2019). ¿Cuáles son los tipos de algoritmos del machine learning? *ADP*.
- Armstrong, G., & Norris, C. (2020). *The_Maximum_Surveillance_Society*. Taylor & Francis.
- Debnath, R., & Bhowmik, M. K. (2021). A comprehensive survey on computer vision based concepts, methodologies, analysis and applications for automatic gun/knife detection. *Journal of Visual Communication and Image Representation*, 78. <https://doi.org/10.1016/j.jvcir.2021.103165>
- Dextre, M., Rosas, O., Lazo, J., & Gutiérrez, J. C. (2021). Gun Detection in Real-Time, using YOLOv5 on Jetson AGX Xavier. *2021 XLVII Latin American Computing Conference (CLEI)*, 1–7. <https://doi.org/10.1109/CLEI53233.2021.9640100>
- Doherty, S. (2023, April 7). *Ecuador facilita acceso a armas en medio de un repunte de violencia*. Insightcrime.Org.
- Fedexpor. (2023, January 24). *La importación legal de armas casi se triplicó en 2022 en medio del aumento de la inseguridad en el país*. Diario La Hora.
- Gonçalves, R. M., & Partyk, A. (2021). *Artificial_intelligence_and_human_rights*. Editorial Dykinson, S.L.
- Jesús Alfonso López Sotelo. (2023). *Deep_Learning_teoría_y_aplicaciones*. Marcombo.
- LatinPyme. (2023, December 26). *Cámaras de vigilancia, la clave para la seguridad en las ciudades*. LatinPyme.

- Liu, H. (2020). Unmanned Driving Systems for Smart Trains. *Elsevier Science*.
- Manikandan, V. P., & Rahamathunnisa, U. (2022). A neural network aided attuned scheme for gun detection in video surveillance images. *Image and Vision Computing*, 120, 104406. <https://doi.org/10.1016/J.IMAVIS.2022.104406>
- MATLAB & Simulink. (2020). *What Is Object Detection? - MATLAB & Simulink*. <https://www.mathworks.com/discovery/object-detection.html>
- McNab, C. (2009). *Armas_de_fuego_Enciclopedia_universal*. Tikal Ediciones.
- Miranda-Ramos, M. M., Hallon, J. O., & Suriaga, J. D. (2022). Sistema de cámaras para la detección de mascarillas con Python y Raspberry PI con comunicación inalámbrica a través del sistema global de comunicación móvil (GSM). *Información Tecnológica*, 33(3), 23–30. <https://doi.org/10.4067/s0718-07642022000300023>
- Mirjalili, V., & Raschka, S. (2020). *Python_machine_learning*. Marcombo.
- Oficial Suplemento, R. (2021). *CÓDIGO ORGÁNICO INTEGRAL PENAL, COIP*. www.lexis.com.ec
- Pichel, M. (2021, November 21). *Cómo Ecuador pasó de ser país de tránsito a un centro de distribución de la droga en América Latina (y qué papel tienen los carteles mexicanos)*. BBC.
- Qasim, M., & Verdu, E. (2023). Video anomaly detection system using deep convolutional and recurrent models. *Results in Engineering*, 18, 101026. <https://doi.org/10.1016/J.RINENG.2023.101026>
- Ramasubramanian, K., & Moolayil, J. (2019). *Applied Supervised Learning with R*. Packt Publishing.
- Rojas Daniel. (2024). *La delincuencia en el Ecuador: Un problema complejo con soluciones complejas*. CADIEC.

- Salazar González, J. L., Zaccaro, C., Álvarez-García, J. A., Soria Morillo, L. M., & Sancho Caparrini, F. (2020). Real-time gun detection in CCTV: An open problem. *Neural Networks*, 132, 297–308. <https://doi.org/10.1016/j.neunet.2020.09.013>
- Sánchez Valdés, V. M. (2016). ¿Son efectivas las cámaras de video vigilancia para reducir los delitos? *URVIO - Revista Latinoamericana de Estudios de Seguridad*, 19, 162. <https://doi.org/10.17141/urvio.19.2016.2406>
- Shetty, A. B., Bhoomika, Deeksha, Rebeiro, J., & Ramyashree. (2021). Facial recognition using Haar cascade and LBP classifiers. *Global Transitions Proceedings*, 2(2), 330–335. <https://doi.org/10.1016/j.gltp.2021.08.044>
- Silhavy, P., & Silhavy, R. (2023). *Networks_and_Systems_in_Cybernetics*. Springer International Publishing.
- Singh Sisodia, D., Garg, L., M., T., & Pachori, R. B. (2023). *Machine_Intelligence_Techniques_for_Data*. Springer Nature Singapore.
- Sunila Gollapudi. (2019). *Learn_Computer_Vision_Using_OpenCV*. Apress.
- UNODC. (2020). *Estudio Mundial sobre el Tráfico de Armas de Fuego 2020*.
- Veranyurt, O., & Sakar, C. O. (2020). Hand-Gun Detection in Images with Transfer Learning-Based Convolutional Neural Networks. *2020 28th Signal Processing and Communications Applications Conference (SIU)*, 1–4. <https://doi.org/10.1109/SIU49456.2020.9302394>

Anexos

Anexo 1: Informe de COMPILATIO

	UNIVERSIDAD TÉCNICA ESTATAL QUEVEDO FACULTAD DE POSGRADO	
---	---	---

PARA: Ing. Byron Oviedo Bayas, PhD

Decano de Posgrado UTEQ

FECHA: Quevedo, 13 de agosto de 2024

Adjunto al presente sírvase encontrar el documento final del proyecto de investigación titulado: **Aprendizaje profundo para la detección de armas de fuego en sistemas de videovigilancia**, elaborado por el Ing. **Briones Montalvo Cristhian Danilo** posgradista de la **Maestría en Ciencia de Datos**. El Proyecto de investigación fue elaborado bajo mi dirección según lo asignado en el contrato PG-1239-2023-C-CIV- SERV-PROF de fecha 27 de noviembre de 2023, el mismo que cumple el informe de la herramienta COMPILATIO, el cual avala los niveles de originalidad, en un **95%** del trabajo investigativo.

 **CERTIFICADO DE ANÁLISIS**
magister

Final - APRENDIZAJE PROFUNDO PARA LA DETECCIÓN DE ARMAS DE FUEGO EN SISTEMAS DE VIDEOVIGILANCIA[1]

5%
Textos sospechosos

5% Similitudes
0% similitudes entre comillas
0% entre las fuentes mencionadas
< 1% Idiomas no reconocidos

Nombre del documento: Final - APRENDIZAJE PROFUNDO PARA LA DETECCIÓN DE ARMAS DE FUEGO EN SISTEMAS DE VIDEOVIGILANCIA[1].pdf ID del documento: 3a8da4f35bf3e961aa4b6a975394e464a8a8d926 Tamaño del documento original: 2,64 MB	Depositante: CRISTIAN GABRIEL ZAMBRANO VEGA Fecha de depósito: 13/8/2024 Tipo de carga: interface fecha de fin de análisis: 13/8/2024	Número de palabras: 13.027 Número de caracteres: 93.100
--	--	--

Atentamente,

Firmado
CRISTIAN GABRIEL digitalmente por
ZAMBRANO VEGA CRISTIAN GABRIEL
ZAMBRANO VEGA

Ing. Cristian Zambrano Vega, PhD.

DIRECTOR DE PROYECTO DE INVESTIGACIÓN

