

# Detección de eventos delictivos usando Deep Learning: un caso de estudio en el TecNM/TEST

## Crime events detection using Deep Learning: a case study on TecNM/TEST

Ballesteros-Reyes, L.F.<sup>1</sup>, Cleofas-Sánchez, L.<sup>2\*</sup> , Guzmán-Ponce, A.<sup>3</sup> , Yañez-Badillo, H.<sup>4</sup> , Pineda-Briseño, A.<sup>5</sup>   
(\*laura\_cs@test.edu.mx)

Recibido: 2024-11-15 Aceptado: 2024-12-09

### RESUMEN

En 2023, se observó un incremento en la incidencia y prevalencia de actos delictivos en México, con variaciones significativas entre las diferentes entidades federativas. El Estado de México destacó particularmente por registrar un aumento en los niveles de eventos delictivos con respecto al año anterior. Los delitos más comunes incluyeron robos perpetrados mediante distintas modalidades, destacando el uso de armas blancas y de fuego, así como agresiones físicas directas. Por otro lado, la violencia en los entornos escolares en México ha mostrado tendencias alarmantes en varios tipos de agresión. De lo antes mencionado, se destaca la importancia de implementar técnicas avanzadas que permitan la identificación oportuna de hechos delictivos

en entornos escolares a fin de salvaguardar la integridad de los estudiantes. Este estudio explora el uso del modelo YOLOv5s para la detección de arrebatos, armas blancas y de fuego, específicamente en el Tecnológico de Estudios Superiores de Tlanguistenco utilizando imágenes digitales obtenidas a través de cámaras fotográficas. Los resultados obtenidos indican que este modelo logra clasificar incidentes que involucran el uso de armas blancas, armas de fuego y arrebatos, alcanzando un promedio de precisión media del 60 %.

**Palabras clave:** YOLOv5, Reconocimiento de actos violentos, Deep Learning.

<sup>1</sup>Tecnológico Nacional de México/Tecnológico de Estudios Superiores de Tlanguistenco, Carretera Tenango, Santiago-La Marquesa 22, 52650 Santiago Tlapa, México.

<sup>2</sup>Tecnológico Nacional de México/Tecnológico de Estudios Superiores de Tlanguistenco, Carretera Tenango, Santiago-La Marquesa 22, 52650 Santiago Tlapa, México.

<sup>3</sup>Departamento de Lenguajes y Sistemas Informáticos, Instituto de Nuevas Tecnologías de Imagen, Universitat Jaume I, 12071 Castelló de la Plana, España.

<sup>4</sup>Tecnológico Nacional de México/Tecnológico de Estudios Superiores de Tlanguistenco, Carretera Tenango, Santiago-La Marquesa 22, 52650 Santiago Tlapa, México.

<sup>5</sup>Tecnológico Nacional de México/Tecnológico Nacional de México / Instituto Tecnológico de Matamoros, H. Matamoros, Tamps., México.



## ABSTRACT

In 2023, an increase in the incidence and prevalence of criminal acts was observed in Mexico, with significant variations across different federal entities. The State of Mexico particularly stood out for recording a rise in crime levels compared to the previous year. The most common offences included thefts committed through various methods, notably involving the use of bladed weapons and firearms, as well as direct physical assaults. Moreover, violence in school settings in Mexico has shown alarming trends in various types of aggression. From the aforementioned, the importance of implementing advanced techniques for the opportune identification of criminal events in school settings to

safeguard the integrity of students is highlighted. This study explores the potential use of the YOLOv5s model for detecting snatch thefts, bladed weapons, and firearms, specifically at the Technological Institute of Higher Studies in Tianguistenco, using digital images captured by photographic cameras. The results obtained indicate that this model is capable of classifying achieves a classification of incidents involving bladed weapons, firearms, and snatch thefts, reaching an average precision of 60 %.

**Keywords:** YOLOv5, Recognition of violent acts, Deep Learning.

## INTRODUCCIÓN

En 2023, se registró un aumento en las tasas de criminalidad del 83 % de la población global<sup>1</sup>, este fenómeno ha afectado significativamente a nivel internacional. Los reportes de la Interpol<sup>2</sup> destacaron como las principales actividades delictivas el crimen organizado, el tráfico de drogas, la ciberdelincuencia y la explotación sexual.

Lankford [1] argumentó que los tiroteos masivos y otros ataques públicos se originaron por algún tipo de violencia, donde los tiradores atacan a víctimas inocentes en espacios públicos. En Estados Unidos, estos tiradores suelen actuar solos y son autodestructivos, representando solo el 1.25 % de los casos globales. En otros países, los ataques generalmente involucran grupos, como organizaciones terroristas o milicias, que tienen como motivación la lucha y la supervivencia.

En las últimas cuatro décadas, los incidentes con armas en México han aumentado debido a la producción y el tráfico de armas desde Estados Unidos, impulsados por su política de 1791 que favorece el derecho a portar armas [1, 2]. El tráfico de armas hacia México ha incrementado la violencia reflejándose en los años 2011 y 2017, con 24 y 25 homicidios por cada 100,000 ciudadanos [3] (ver Figura 1).

Al respecto, México ocupa el sexto lugar mundial en mortalidad por el uso y portación de armas, con un aumento en las muertes entre los años 2015 y 2022. Las víctimas incluyen principalmente hombres (95.5 %), pero también mujeres (53.5 %) y niños. Los tiroteos ocurren en espacios públicos, calles y hogares, relacionados con el crecimiento de organizaciones delictivas y el tráfico de armas, las cuales cada año, se adquieren 213,000 armas, muchas de ellas provenientes de Estados Unidos [5].

En 2023, México registró 15,082 homicidios, con una tasa de 12 por cada 100,000 habitantes, de los cuales el 71.3 % fueron causados por armas de fuego, seguido por arma blanca (9.1 %) y ahorcamiento, estrangulamiento o sofocación (6.7 %) [4].

<sup>1</sup><https://globalinitiative.net/analysis/ocindex-2023/>

<sup>2</sup><https://www.interpol.int/en/How-we-work/Criminal-intelligence-analysis/Our-analysis-reports>

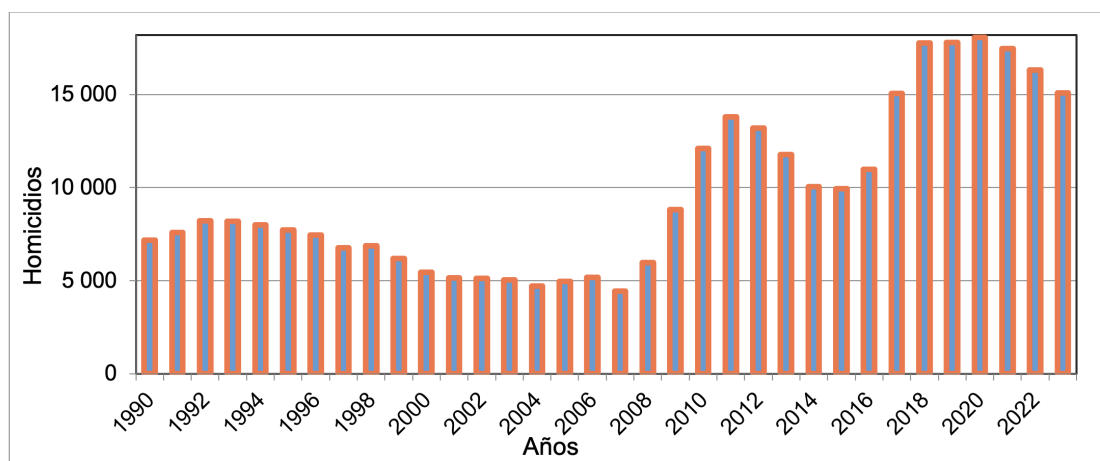


Figura 1: *Homicidios en México, entre los años 1990 y 2023* [4].

El estudio realizado por Weigend [5] analizó que el 26 % de los incidentes violentos con armas de fuego ocurridos en escuelas y oficinas en México entre 2015 y 2022, impactan negativamente en la economía del país. La inseguridad reduce las oportunidades de inversión, empleo y desarrollo local. Asimismo, estos eventos tienen un efecto considerable en la salud mental de los mexicanos, generando un ambiente de miedo y estrés crónico.

Derivado de lo anterior, según la Encuesta Nacional de Victimización y Percepción sobre la Seguridad Pública del Instituto Nacional de Estadística y Geografía (INEGI)<sup>3</sup>, en los últimos años, el Estado de México y el país han experimentado un aumento en los índices de violencia. En 2006, la criminalidad era el segundo problema más grave, pero entre 2012 y 2014 ascendió al primer lugar de las preocupaciones nacionales, afectando tanto a los residentes del Estado de México como al resto del país [6].

De igual modo, en 2023, el INEGI reportó 2,849 homicidios en el Estado de México, con un 85.3 % de víctimas masculinas y un 10.31 % femeninas [5]. En 2024, los homicidios por arma de fuego aumentaron un 5.8 % en comparación con 2022, una tendencia atribuida al tráfico de armas entre Estados Unidos y el norte de México. En ese mismo año, más del 70 % de los homicidios en México fueron cometidos con armas de fuego, reflejando una grave situación de seguridad, exacerbada por la proliferación de armamento ilegal en el país.

En la Tabla 1, se describen las casusas socioeconómicas del incremento de violencia en México y Estado de México entre los años 2006 y 2014. Como puede verse, en el año 2006, el desempleo agravó la inseguridad, mientras que en 2014 se observó un incremento en delitos. En términos económicos, en el 2010 se presentó una situación relativamente más estable.

La criminalidad en México ha afectado diversas áreas, incluido el sector educativo, en donde se requiere atención para prevenir escaladas de violencia como las ocurridas en Estados Unidos, en el cual la violencia escolar involucra armas blancas y otras agresiones físicas [7].

El uso de aprendizaje profundo ha mejorado significativamente en la detección de delitos. Investiga-

<sup>3</sup>[https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2024/ENVIPE/ENVIPE\\_24.pdf](https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2024/ENVIPE/ENVIPE_24.pdf)

Tabla 1: Causas de violencia en el Estado de México y a nivel nacional en México: a.- Desempleo (México), b.- Desempleo (Estado de México), c.- Delincuencia y Crimen (México), d.- Delincuencia y Crimen (Estado de México), e.- Economía (México), f.- Economía (Estado de México), g.- Encuestados (México), h.- Encuestados (Estado de México) [3].

|      | <b>a</b> | <b>b</b> | <b>c</b> | <b>d</b> | <b>e</b> | <b>f</b> | <b>g</b> | <b>h</b> |
|------|----------|----------|----------|----------|----------|----------|----------|----------|
| 2014 | 13.5 %   | 11.3 %   | 25.7 %   | 29.6 %   | 21.3 %   | 13.8 %   | 1519     | 203      |
| 2012 | 12.2 %   | 13.9 %   | 21.7 %   | 18.8 %   | 13.4 %   | 9.9 %    | 778      | 101      |
| 2010 | 20.0 %   | 23.4 %   | 15.4 %   | 22.9 %   | 33.1 %   | 24.9 %   | 1540     | 201      |
| 2008 | 16.7 %   | 15.9 %   | 15.8 %   | 16.9 %   | 16.5 %   | 23.4 %   | 1530     | 201      |
| 2006 | 21.5 %   | 21.8 %   | 14.9 %   | 14.4 %   | 11.3 %   | 9.9 %    | 1533     | 202      |

dores han explorado diversos métodos para optimizar los protocolos de seguridad y reducir actividades ilegales [8, 9, 10]. Mukto et al [11] argumentaron que los algoritmos de aprendizaje profundo se están implementando en sistemas de monitoreo criminal en tiempo real, lo que permite identificar y responder rápidamente a incidentes delictivos.

Has [12] y Sandhya [13] afirmaron que los modelos de aprendizaje profundo para el reconocimiento de objetos son efectivos en la detección de intenciones delictivas y mejoran las operaciones policiales al identificar sospechosos, en tanto que, Jenga et al. [14] y Kumar [15] subrayaron la necesidad de adaptar estas técnicas para entornos educativos. Stalidis et al [16] exploraron cómo las arquitecturas de aprendizaje pueden predecir delitos en contextos educativos, proponiendo un marco para implementar medidas de detección específicas.

Por otro lado, Divya et al [17] desarrollaron un sistema que utiliza VGG-19, conocido por su eficacia en el reconocimiento de imágenes de armas, para ello emplearon redes convolucionales de regiones (RCNN) para delimitar objetos en imágenes relevantes, logrando una precisión de entrenamiento superior al 80 %. Anandhi, et al [18] propusieron un sistema automatizado para detectar actos de violencia en tiempo real usando cámaras de vigilancia mediante algoritmos de aprendizaje profundo.

La implementación de aprendizaje profundo en el contexto mexicano, particularmente en el sector educativo, crea una brecha significativa en las estrategias preventivas que permitan identificar armas portadas por individuos que busquen cometer actos de violencia, como tiroteos en instituciones educativas.

El Estado de México enfrenta desafíos de seguridad en el ámbito educativo, especialmente en áreas de crecimiento como Tianguistenco<sup>4</sup>. El TecNM/TEST se encuentra en el municipio de Tianguistenco (Tlaxcala), en donde se han reportado fosas clandestinas en un sembradío de maíz en el año 2024<sup>5</sup>. Asimismo, en el 2023 se suscitaron 18 feminicidios en un lapso de dos meses en los municipios de Oztolotepec, Almoloya de Juárez, Toluca y Tianguistenco, dichos delitos fueron cometidos con arma de fuego (2), arma blanca (5) y otros (11). Uno de los casos más mencionados fue el homicidio de Ameyali de 17 años quien

<sup>44</sup> [https://www.ieem.org.mx/maxima\\_publicidad/maxima17\\_18/InfGral/PLATAFORMAS-2018/docs/Plataformas\\_Mpa-les/PRI/102\\_Tianguistenco.pdf?utm\\_source=chatgpt.com](https://www.ieem.org.mx/maxima_publicidad/maxima17_18/InfGral/PLATAFORMAS-2018/docs/Plataformas_Mpa-les/PRI/102_Tianguistenco.pdf?utm_source=chatgpt.com)

<sup>5</sup> [https://www.milenio.com/policia/hallan-3-cuerpos-en-fosa-clandestina-de-tianguistenco-edomex?cx\\_testId=32&cx\\_testVariant=cx\\_1&cx\\_artPos=0&cx\\_experienceId=EXZC9VQ5Y2SQ&cx\\_experienceActionId=showRecommendationsI8BLQC9TX2D091#cxrecs\\_s](https://www.milenio.com/policia/hallan-3-cuerpos-en-fosa-clandestina-de-tianguistenco-edomex?cx_testId=32&cx_testVariant=cx_1&cx_artPos=0&cx_experienceId=EXZC9VQ5Y2SQ&cx_experienceActionId=showRecommendationsI8BLQC9TX2D091#cxrecs_s)

vivía en Tianguistenco, Tilapa, a ella se le encontró sin vida en el paraje llamado las ratoneras rumbo al Ajusco<sup>6</sup>.

Esta situación de violencia en Santiago Tianguistenco y su comunidad de Tilapa, Estado de México, ha sido motivo de preocupación en los últimos meses. Entre los hechos recientes, destaca el asesinato de tres comerciantes en septiembre de 2024. Estas personas fueron secuestradas y posteriormente encontradas sin vida en una calle de la localidad. Este suceso impactó profundamente a la comunidad, evidenciando los desafíos en materia de seguridad pública que enfrenta la región<sup>7</sup>.

Por lo antes mencionado, la falta de tecnologías avanzadas para detectar violencia en el Tecnológico de Estudios Superiores de Tianguistenco (TecNM/TEST) es preocupante. Este estudio propone implementar YOLO v5 para el reconocimiento de armas en el TecNM/TEST, buscando mitigar los riesgos de violencia mediante detección temprana y efectiva, y así hacer eficiente los protocolos de seguridad, pudiendo ser adaptable para otras instituciones, fomentando un entorno académico seguro.

## YOU ONLY LOOK ONCE

YOLO es un modelo de redes neuronales convolucionales para la detección de objetos de una sola pasada (*single-shot detection*), esto lo hace ideal para aplicaciones en problemas que requieren respuesta en tiempo real, como sistemas de vigilancia o vehículos autónomos. Funciona transformando el proceso de detección de objetos en una tarea de regresión lineal, en la cual se analizan directamente las imágenes de entrada para determinar las coordenadas de cuadros delimitadores y las probabilidades asociadas a cada clase de objeto detectado [19].

YOLO presenta varias ventajas con respecto a otros modelos de aprendizaje automático, particularmente en tareas de detección de objetos, debido a que se requiere menos recursos computacionales en comparación con modelos como Faster R-CNN o SSD, lo que permite implementar el modelo de YOLO en dispositivos con hardware limitado.

En su funcionamiento YOLO [19] se puede describir en los siguientes pasos:

1. Segmentar una imagen de entrada en una cuadrícula  $S \times S$ .
2. Para cada celda de la cuadrícula, el modelo genera predicciones para  $B$  cuadros delimitadores y asigna a cada uno un valor de confianza. Este valor de confianza refleja la precisión con la que el cuadrado delimitador encierra un objeto.
3. Cada celda de la cuadrícula estima probabilidades condicionales asociadas a  $C$  clases posibles de objetos.
4. Para obtener la predicción final, la red convolucional combina estas probabilidades condicionales con los valores de confianza correspondientes, multiplicándolos para evaluar la clase más probable de objeto presente dentro de cada cuadro delimitador.

YOLOv5 [20] optimiza el modelo YOLO en tres componentes principales:

<sup>6</sup><https://adnoticias.mx/ameyali-alvarado-localizan-sin-vida/>

<sup>7</sup><https://paralelo19.tv/nacional/2024/09/11/tragedia-en-santiago-tianguistenco-tres-comerciantes-asesinados> y <https://www.milenio.com/policia/matan-a-3-personas-en-santiago-tianguistenco-eran-polleros>

- **Backbone:** Procesa las imágenes de entrada *reduciendo* sus dimensiones espaciales y aumentando la profundidad de sus canales para extraer características detalladas. Se beneficia de los módulos *Focus*, que optimiza la captura de información espacial en menos operaciones y CSPDarknet, mejora la eficiencia de la extracción de características con un diseño eficiente que reduce el coste computacional y mejora el flujo de información. Estas mejoras hacen que YOLOv5 sea más rápido y preciso en la detección de objetos [20].
- **Neck:** En este componente se utilizan los módulos PAN (Path Aggregation Network) y FPN (Feature Pyramid Network) para integrar características detectadas en múltiples escalas espaciales, mejorando notablemente la precisión en la localización de objetos de diversos tamaños. FPN propaga características ricas en contexto de niveles superiores a niveles inferiores, mientras que PAN mejora la precisión de localización optimizando la fusión de características de bajo nivel con alta resolución [20].
- **Head:** En este componente se optimiza la detección simultánea de objetos grandes y pequeños, aumentando la precisión del modelo.

El modelo utiliza la métrica de intersección sobre la unión (IoU, por sus siglas en inglés) de la Ec. (1) para seleccionar los cuadros con mayor precisión en la detección de objetos [20].

$$IoU = \frac{\min(W_a, W_b) * \min(H_a, H_b)}{W_a H_a + W_b H_b - \min(W_a, W_b) * \min(H_a, H_b)} \quad (1)$$

Donde de la Ec. (1),  $(W_a, H_a)$  y  $(W_b, H_b)$  representan el alto y ancho de los cuadros.

YOLOv5s, por su tamaño reducido y baja demanda de recursos, es ideal para dispositivos con hardware limitado, como Raspberry Pi y teléfonos móviles [19].

## METODOLOGÍA

En respuesta a los desafíos de seguridad en el TecNM/TEST, se propone una metodología para desarrollar un sistema de detección de armas blancas, armas de fuego y arrebatos con el modelo YOLOv5.

### Conjunto de datos

Ante la falta de conjuntos de datos que involucren armas y arrebatos, se creó un conjunto de datos especializado. Se recopilaron 2400 imágenes de la web mediante el motor de búsqueda de Google<sup>8</sup>, utilizando palabras clave como pistolas de comando, mausser, luger, revólveres, rifles, escopetas, fusiles mosquetones para categorizar armas de fuego y punzocortantes como cuchillos de cocina, navajas, machetes zapapicos, hachas, para armas blancas y comportamientos de arrebatos. Las imágenes se seleccionaron cuidadosamente de sitios públicos, priorizando el formato JPEG, dimensiones de 640x640 y visibilidad clara de las armas. Todo el conjunto recolectado se destinó al entrenamiento del modelo, respetando los derechos de autor y utilizando material que cumple con las normativas de uso y distribución de imágenes públicas.

<sup>8</sup>[https://es.m.wikipedia.org/wiki/Archivo:Colt\\_Official\\_Police\\_38\\_Revolver-NMAH-AHB2015q021271.jpg](https://es.m.wikipedia.org/wiki/Archivo:Colt_Official_Police_38_Revolver-NMAH-AHB2015q021271.jpg)

Para el conjunto de prueba, se recopilieron 400 imágenes autorizadas por autoridades del TecNM/TEST. Se usaron réplicas de juguete para las armas de fuego y utensilios de oficina como armas blancas (cutters). Las imágenes se tomaron bajo iluminación controlada, protegiendo la identidad de los participantes con cubrebocas. Además, las imágenes se mantuvieron consistentes en tamaño y formato con respecto al conjunto de entrenamiento. Todos los participantes otorgaron su consentimiento informado previo a la inclusión en el estudio.



Figura 2: Representación de objetos en un entorno controlado para pruebas de reconocimiento en el TecNM/TEST.

El etiquetado de las imágenes se llevó a cabo manualmente utilizando la herramienta LabelImg, que está desarrollada en Python. Este proceso implicó la selección cuidadosa de las regiones de interés en las imágenes, donde se identificaban claramente armas de fuego, armas blancas y/o acciones de arrebató. Estas regiones seleccionadas constituyen las clases dentro del conjunto de datos.

#### Modelo YoloV5s

La red neuronal convolucional *YOLOV5s* cuenta con 117 capas, 5447688 parámetros, 11.4 GFLOPs que realizan 11.4 mil millones de operaciones de punto flotante para procesar una imagen de entrada, un `batch_size` de 4 imágenes, 100 épocas, una función de activación llamada LeakyReLU(0.1), considera hasta 80 clases de un entrenamiento previo con el conjunto de datos llamado COCO (pesos `yolov5s.pt`), utiliza una profundidad de modelo de 0.33 (`depth_multiple` : 0.33), un número de canales en cada capa convolucional del modelo de 0.50 ( `width_multiple`: 0.50).

#### Métricas de validación

Con el fin de medir la eficiencia del modelo *YOLOv5s* la métrica del promedio de precisión media (mAP, por sus siglas en inglés) es comúnmente adoptada debido a su capacidad para integrar las medidas de precisión y sensibilidad a lo largo de diferentes umbrales de IoU.

Formalmente la mAP se calcula promediando los valores de precisión en varios niveles de sensibilidad, lo que da como resultado un único valor escalar que indica la precisión general de detección, como se describe en la Ec. (2).

$$mAP = \frac{1}{C} \sum_{i=1}^C AP_i \quad (2)$$

Donde  $C$  denota el número de clases y  $AP_i$  (Ec. (3)) representa el área bajo la curva de precisión-sensibilidad para la clase  $i$  [20].

$$AP_i = \sum_i (R_i - R_{i-1}) P_i \quad (3)$$

Donde  $R$  denota la sensibilidad, definida como la proporción de verdaderos positivos identificados en relación con el total de los verdaderos positivos reales (Ec. (4)) y  $P$  denota precisión, que se calcula como la proporción de verdaderos positivos en comparación con el total de predicciones positivas (Ec. (5)).

$$R = \frac{TP}{TP + FN} \quad (4)$$

$$P = \frac{TP}{TP + FP} \quad (5)$$

Donde  $TP$  representa el número de verdaderos positivos,  $FP$  indica el número de falsos positivos y  $FN$  señala el número de falsos negativos.

#### *Escenario Experimental*

En la experimentación se usaron 1,920 imágenes para entrenamiento y 480 para validación, configurando 100 épocas con un tamaño de lote de 4 imágenes. El entrenamiento se realizó en Google Colab, con Python 3.10, 12 GB de RAM y una GPU NVIDIA® T4. La fase de prueba con 400 imágenes se realizó usando un umbral de confianza del 60 % para evaluar la métrica de la mAP. Los experimentos se realizaron en una computadora HP EliteBook 820, procesador Intel Core i5 6ta generación, velocidad (3.0 Ghz), cache 3 MB, memoria (8 GB), unidad de estado sólido 256 GB, pantalla LED 12.5", gráficos HD 520 y video integrado. En un entorno de Windows y Anaconda, instalando en el entorno las librerías como Pytorch usando el toolkit de CUDA.

## RESULTADOS

En la Figura 3, se muestran los resultados de la función de pérdida en el contexto de entrenamiento (train/cls\_loss) y validación (val/cls\_loss), de los 10 (a, b, c, d, e, f, g, h, i, j) modelos construidos con validación cruzada. En estos resultados se observa que el error de entrenamiento tiende a disminuir de forma constante a medida que aumentan las épocas, lo cual está ajustándose adecuadamente a los datos de entrenamiento, mientras que el error de validación también tiene un descenso progresivo indicando que el modelo está generalizando bien, sin embargo, las oscilaciones del error son normales en los datos de validación debido a que son un conjunto más reducido o diverso.

En la Figura 4, el mAP de los resultados de entrenamiento varía entre 83.8 % y 90.4 %, lo que indica que los modelos son capaces de detectar y clasificar todas las clases adecuadamente. Este rango indica un rendimiento bastante sólido, ya que un valor más alto de mAP refleja una mayor precisión en la detección de objetos, mientras que los valores cercanos al 100 % son ideales.

En la Figura 5, el mAP de los resultados de prueba que varía entre 53.16 % y 60.59 %, significa que el modelo logra detectar y clasificar correctamente los objetos en las imágenes con un rendimiento moderado.

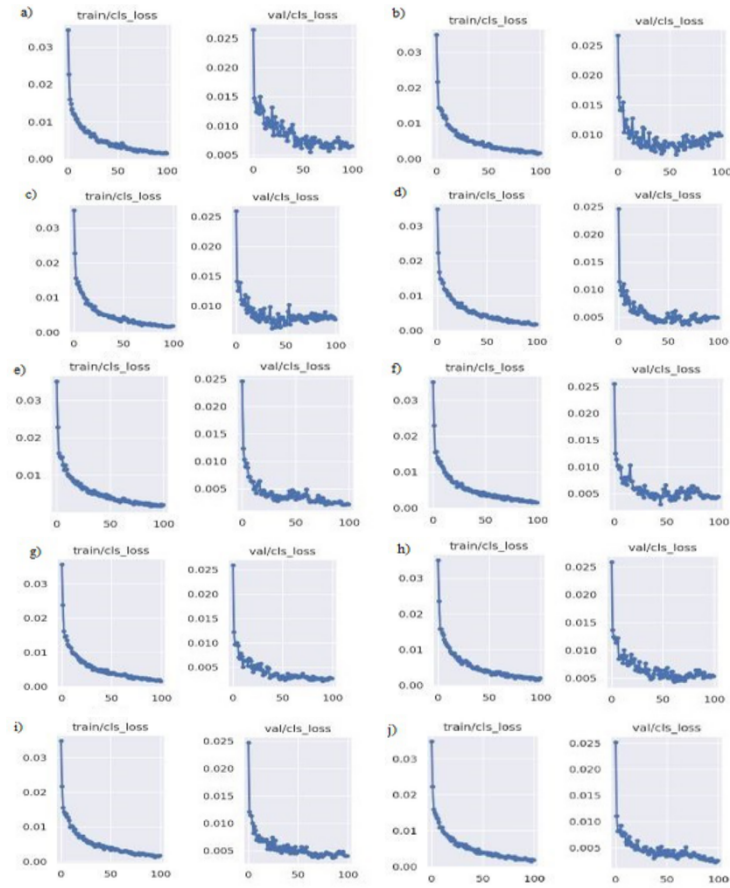


Figura 3: Resultados de la función de pérdida del entrenamiento y la validación, considerando una validación cruzada de 10.

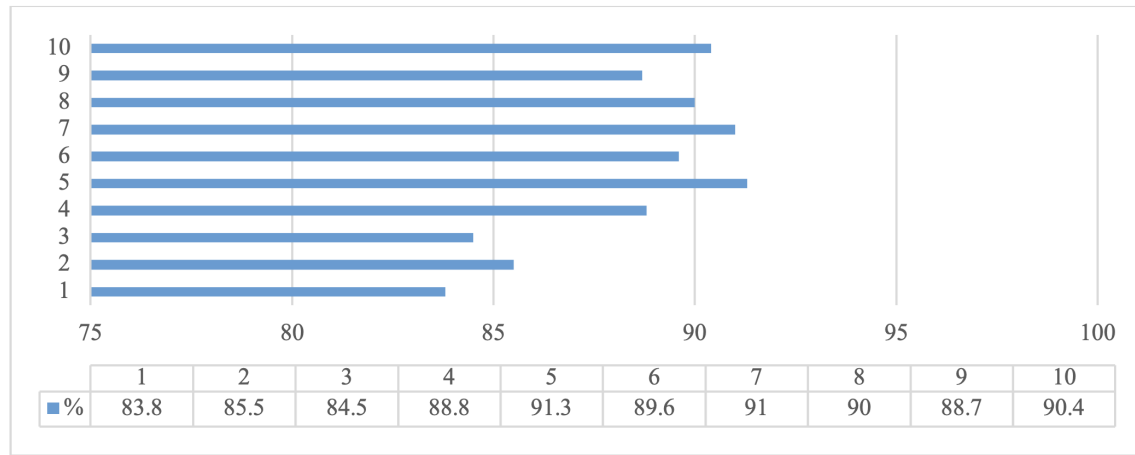


Figura 4: Resultados del mAP en porcentaje de cada entrenamiento, considerando una validación cruzada de 10.

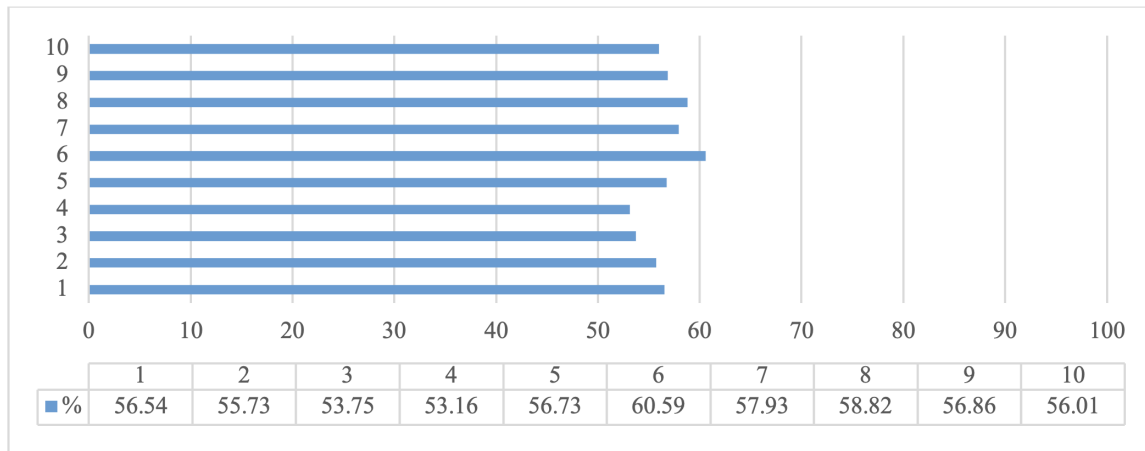


Figura 5: Resultados del mAP en porcentaje de cada prueba, considerando una validación cruzada de 10.

## CONCLUSIONES

En 2023, el Estado de México experimentó un aumento significativo en delitos como robos y asaltos con armas de fuego y blancas, además de agresiones directas, lo que generó preocupación sobre la posible expansión de estos incidentes a entornos escolares. Ante esta situación, este estudio evaluó la implementación del modelo YOLOv5s en el Tecnológico de Estudios Superiores de Tianguistenco. La investigación se centró en la capacidad del modelo para detectar y clasificar eventos relacionados con armas blancas y de fuego, así como actos de arrebató, buscando mejorar los protocolos de seguridad y ofrecer respuestas más eficientes a los incidentes, contribuyendo a un ambiente académico más seguro y favorable para el aprendizaje.

Los resultados de mAP indica que el modelo YOLOv5s muestra un desempeño notable en la detección de armas y acciones de arrebató durante las fases de entrenamiento, con valores entre 83.8 % y 90.4 %. Estos resultados destacan la capacidad del modelo para identificar correctamente los objetos de interés. Sin embargo, en las pruebas realizadas en escenarios no conocidos por el modelo en el entrenamiento, el mAP disminuye a un rango de 53.16 % a 60.59 %, lo que refleja una reducción significativa en la precisión general y la recuperación del modelo bajo condiciones nuevas y variables.

El contraste en el rendimiento sugiere que, aunque el modelo es eficaz en entornos controlados, necesita ajustes para mejorar su consistencia y precisión en escenarios no previstos durante la fase de entrenamiento, donde las diferencias en el contexto y las condiciones presentan desafíos adicionales para la detección efectiva de violencia.

Como trabajo futuro se encuentra la implementación de técnicas de aprendizaje incremental y transferencia de aprendizaje. Esto permitirá al modelo adaptarse mejor a nuevas condiciones y variabilidades que no fueron presentes en la fase de entrenamiento inicial. Además, sería beneficioso incrementar la diversidad y cantidad de datos en las etapas de entrenamiento para cubrir un espectro más amplio de situaciones potenciales, incluyendo diferentes entornos, tipos de armas y comportamientos de arrebató.

## DECLARACIÓN DE CONTRIBUCIÓN DE AUTORES Y COLABORADORES

Ballesteros-Reyes, L. F.: Ejecución de experimentos y redacción. Cleofas-Sánchez, L.: Diseño experimental, redacción y supervisión. Guzmán-Ponce, A.: Revisión y redacción. Yañez-Badillo, H.: Revisión. Pineda-Briseño, A.: Revisión

## REFERENCIAS

- [1] A. Lankford. The importance of analyzing public mass shooters separately from other attackers when estimating the prevalence of their behavior worldwide. *Econ Journal*, vol. 17, no. 1, pp. 40–55, 2020. <https://journaltalk.net/articles/5999/>.
- [2] D. Arboleda Andrade. Tiroteo en Stoneman Douglas: discursos políticos, funcionamiento y materialidades discursivas del suceso. *Revista Científica de Comunicación*, vol. 10, no. 1, pp. 49–68, 2019. DOI: 10.31207/rch.v10i1.177.

- 
- [3] D. Perez Esparza, S. D. Johnson, y P. Gill. Why did Mexico become a violent country? Assessing the role of firearms trafficked from the US. *Security Journal*, vol. 33, pp. 179–209, 2020. DOI: 10.1057/s41284-019-00178-6.
  - [4] E. Weigend Vargas, M. Degli Esposti, y S. Hargarten. Examining firearm-related deaths in Mexico. *Econ Journal*, vol. 11, no. 33, pp. 2–7, 2024. DOI: 10.1186/s40621-024-00519-z.
  - [5] Instituto Nacional de Estadística y Geografía (INEGI). 2023. [Online]. Available: <https://en.www.inegi.org.mx/inegi/acercade.html>. [Consultado el 1 de noviembre de 2024].
  - [6] A. Pulido Gómez y B. J. Almaraz Calderón. Violencia y comportamiento electoral: el caso del Estado de México. *Apuntes Electorales*, vol. 16, no. 56, pp. 9–38, 2017. Recuperado de <https://aelectorales.ieem.org.mx/index.php/ae/article/view/80>.
  - [7] J. E. Goldstick, R. M. Cunningham, y P. M. Carter. Current Causes of Death in Children and Adolescents in the United States. *New England Journal of Medicine*, vol. 386, no. 20, pp. 1955–1956, 2022. DOI: 10.1056/nejmc2201761.
  - [8] F. Dakalbab, M. A. Talib, O. A. Waraga, A. B. Nassif, S. Abbas, y Q. Nasir. Artificial intelligence & crime prediction: A systematic literature review. *Social Sciences & Humanities Open*, vol. 6, no. 1, pp. 1–23, 2022. DOI: 10.1016/j.ssaho.2022.100342.
  - [9] A. M. Al-Madani, V. Mahale, y A. T. Gaikwad. Real-time detection of crime and violence in video surveillance using deep learning. in *First International Conference on Advances in Computer Vision and Artificial Intelligence Technologies (ACVAIT 2022)*, Atlantis Press, 2023, pp. 431–441. DOI: 10.2991/978-94-6463-196-8\_33.
  - [10] P. Sivakumar, V. Jayabalaguru, R. Ramsugumar, y S. Kalaisriram. Real-time crime detection using deep learning algorithm. in *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)*, IEEE, 2021, pp. 1–5. DOI: 10.1109/ICSCAN53069.2021.9526393.
  - [11] M. M. Mukto, M. Hasan, M. M. Al Mahmud, I. Haque, M. A. Ahmed, T. Jabid, M. S. Ali, M. R. A. Rashid, M. M. Islam, y M. Islam. Design of a real-time crime monitoring system using deep learning techniques. *Intelligent Systems with Applications*, vol. 21, pp. 1–17, 2024. DOI: 10.1016/j.iswa.2023.200311.
  - [12] A. O. Hashi, A. A. Abdirahman, M. A. Elmi, y O. E. R. Rodriguez. Deep learning models for crime intention detection using object detection. *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 4, pp. 300–306, 2023. DOI: 10.14569/IJACSA.2023.0140434.
  - [13] S. Sandhya, A. Balasundaram, y A. Shaik. Deep learning-based face detection and identification of criminal suspects. *Computers, Materials & Continua*, vol. 74, no. 2, pp. 2331–2343, 2023. DOI: 10.32604/cmc.2023.032715.

- 
- [14] K. Jenga, C. Catal, y G. Kar. Machine learning in crime prediction. *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 3, pp. 2887–2913, 2023. DOI: 10.1007/s12652-023-04530-y.
- [15] J. R. R. Kumar, D. Dhabliya, y S. S. Dari. A comparative study of machine learning algorithms for image recognition in privacy protection and crime detection. *International Journal of Intelligent Systems and Applications in Engineering*, vol. 11, no. 9s, pp. 482–490, 2023. ISSN: 2147-679921.
- [16] P. Stalidis, T. Semertzidis, y P. Daras. Examining deep learning architectures for crime classification and prediction. *Forecasting*, vol. 3, no. 4, pp. 741–762, 2021. DOI: 10.3390/forecast3040046.
- [17] M. Divya, G. S. Priya, R. Abitha, K. Sirisha, A. Manikanta, y K. Jayanth. Automated crime intention detection using deep learning. *International Research Journal of Modernization in Engineering, Technology, and Science*, vol. 4, no. 6, pp. 5670–5676, 2022. e-ISSN: 2582-5208.
- [18] R. J. Anandhi. Edge Computing-Based Crime Scene Object Detection from Surveillance Video Using Deep Learning Algorithms. in *2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, 2023, pp. 1159–1163. DOI: 10.1109/ICIRCA57980.2023.10220800.
- [19] M. Alruwaili *et al.* Deep Learning-Based YOLO Models for the Detection of People With Disabilities. *IEEE Access*, vol. 12, pp. 2543–2566, 2024. DOI: 10.1109/ACCESS.2023.3347169.
- [20] K. Jun-Hwa, K. Namho, P. Yong Woon, y W. Chee Sun. Object Detection and Classification Based on YOLO-V5 with Improved Maritime Dataset. *Journal of Marine Science and Engineering*, vol. 10, no. 3, pp. 2–14, 2022. DOI: 10.3390/jmse10030377.