

La Inteligencia Artificial como apoyo a la gestión de la seguridad ciudadana: un estado del arte

Artificial Intelligence as support for citizen security management: a state of the art

Oscar M. Morales¹, Luis A. Fletscher Bocanegra²,
Juan F. Botero Vega²

¹ Grupo Konecta, Colombia

² Universidad de Antioquia, Colombia

omauricio.morales@udea.edu.co , luis.fletscher@udea.edu.co ,
juanf.botero@udea.edu.co

RESUMEN. Cada vez más, existe un mayor interés en los gobiernos de múltiples ciudades de garantizar la seguridad de sus ciudadanos, siendo la tecnología una de las herramientas clave para enfrentar los diferentes retos que en este campo se plantean. En este trabajo se presenta una aproximación al estado del arte de algoritmos estadísticos y técnicas de inteligencia computacional utilizados en el campo de la gestión de problemas de seguridad ciudadana. Para esto se escogieron dos de las aplicaciones más relevantes en el área, las cuales son: detección de disparos y detección de armas. Para cada una de estas aplicaciones se hizo un cuadro con los trabajos más relevantes junto con un posterior análisis, se formularon las conclusiones y se describió el trabajo futuro. Finalmente, se organizó una lista de recomendaciones con los algoritmos más utilizados y que mejores resultados han arrojado.

ABSTRACT. Increasingly, there is a greater interest in the governments of multiple cities to guarantee the safety of their citizens, technology being one of the main tools to face the different challenges that arise in this field. This paper presents an approach to the state of the art of statistical algorithms and computational intelligence techniques used to manage some citizen security problems. For this, we studied two of the most relevant applications in the area: detection of shots and detection of weapons. For each application, a resume of the main works is presented, followed by an analysis, conclusions, and future work. Finally, it is presented a list of recommendations with the most used algorithms and those with the best results.

PALABRAS CLAVE: Seguridad ciudadana, Inteligencia artificial, Detección de disparos, Detección de armas, Ciudades inteligentes.

KEYWORDS: Citizen security, Artificial intelligence, Gunshot detection, Weapon detection, Smart cities.



1. Introducción

Al hablar de seguridad ciudadana, es posible encontrar un gran abanico de fenómenos, diferentes entre sí, con el denominador común de su potencial para afectar negativamente a la ciudadanía. Dentro de estos fenómenos se encuentran los asaltos a transeúntes, robos de vehículos, desorden público, abuso policial, vandalismo, terrorismo, y porte y uso ilegal de armas. Esta problemática acarrea consecuencias sociales difíciles de solucionar para los gobiernos locales. Por ejemplo, se ha encontrado que ser víctima de un delito está relacionado con la disminución del bienestar colectivo y la calidad de vida (Manjarres & Baca, 2019). Igualmente, el aumento de la inseguridad tiende a disminuir la intención de inversión de origen nacional y extranjero, bajar la valorización de los bienes, causar el desvío de recursos de empresas a gastos de protección, entre otras afectaciones económicas a la población (Soria Romo, 2017).

Si bien existen diferentes posturas respecto a la aproximación que se debe dar al fenómeno de la seguridad ciudadana, hoy en día toma mayor relevancia aquellas aproximaciones centradas en la prevención, monitoreo, detección y acción, ya que muestran un gran potencial para convertirse en factores que salvan vidas, elevan la calidad de vida de los habitantes y mejoran el ambiente económico y social del entorno. Por estas razones, es necesario que la seguridad ciudadana sea tomada en cuenta como un factor determinante en el desarrollo de una región.

Si bien la seguridad ciudadana ha sido un aspecto presente en los planes de gobierno de los territorios, desde hace unos años hacia acá se ha considerado un enfoque de prevención en conjunto con enfoques de acción, siendo la implementación de ciudades inteligentes el marco propicio para estos desarrollos (Su, Li & Fu, 2011; Departamento Nacional de Planeación DNP, 2020). De esta forma, este aumento en la inversión y gasto en bienes tecnológicos tanto por parte de empresas privadas como estatales para la implementación de espacios y ciudades inteligentes, acompañado de la baja progresiva del precio de componentes electrónicos en el mercado se convierte en un incentivo para la investigación y desarrollo de estos paradigmas de ciudad.

Otro elemento que se debe resaltar es que, con el incremento de las capacidades computacionales y la mayor cantidad de datos disponibles, la inteligencia artificial ha tomado un rol protagónico en el abordaje de problemáticas relacionadas con la seguridad ciudadana, por lo que es fundamental conocer las principales técnicas implementadas, así como las soluciones desarrolladas para los diversos fenómenos identificados.

De acuerdo con lo anterior, este trabajo busca generar un estado del arte en los métodos de inteligencia computacional y algoritmos estadísticos que han sido o puedan ser utilizados en proyectos de ciudades inteligentes para medir y mejorar la seguridad ciudadana. Lo anterior implica que se incluyen también trabajos cuyo producto es aplicable en entornos de ciudades inteligentes, aunque su primer objetivo no sea este expresamente. Además, como regla general de un estado del arte, fue esencial que estos trabajos fueran recientes o se mantuvieran vigentes al momento de la búsqueda. De los resultados de esta búsqueda, se extrajeron las herramientas tecnológicas, especialmente de software, que más se han implementado y han arrojado mejores resultados.

De igual forma, se espera que este trabajo pueda servir como una introducción al tema para aquellos interesados y también como punto de referencia sobre las principales aplicaciones, enfoques y herramientas que se ven de forma más amplia en el ámbito de inteligencia computacional y seguridad ciudadana para investigadores que van a iniciar sus proyectos. El resto del documento se encuentra estructurado de la siguiente forma: en la sección 2 se presenta la metodología usada para la búsqueda y clasificación de la información, la sección 3 aborda los estudios más relevantes en cada uno de los fenómenos considerados y las líneas de trabajo futuro identificada y finalmente en la sección 4 se presentan las conclusiones.

1.1. Referentes Teóricos

El marco de referencia para el desarrollo de este proyecto parte del concepto de Ciudad Inteligente, que al igual que muchos otros conceptos tecnológicos recientes, no está completamente consensuado dentro de la



comunidad científica y, por lo tanto, no existen límites reales con relación a sus enfoques de implementación y otros conceptos similares. Este término evoluciona constantemente con el tiempo y con el surgimiento de nuevas tecnologías y propuestas.

Sin embargo, revisando diferentes fuentes, es posible calificar a una ciudad como inteligente si es aquella que hace uso de las tecnologías de la información y la comunicación para medir, analizar e integrar información clave de funcionamiento cotidiano. Esto permite tener una respuesta inteligente a diferentes necesidades de tipo ambiental, industrial, comercial, de tráfico o de seguridad (Su, Li & Fu, 2011).

De (Tabane, Ngwira & Zuva, 2016) y (Zhao et al., 2021), es posible extraer las siguientes definiciones de Ciudad Inteligente, las cuales permiten diversificar el concepto y se encuentran a continuación:

- Aquella ciudad que utiliza las tecnologías de la información y la comunicación para intentar solucionar los retos urbanos que emergen en la actualidad como el crecimiento económico, la creación de trabajos, sostenibilidad ambiental, y resiliencia social.
- Aquella ciudad cuyas inversiones en capital social y humano, en transporte y en las tecnologías de la información y la comunicación incentivan el desarrollo económico sostenible y la mejora de la calidad de vida con un manejo responsable de los recursos naturales y a través de la acción participativa.
- Aquella ciudad que enfrenta los retos de gobernabilidad como urbanización rápida, cambios demográficos y económicos, cambios climáticos y ambientales, entre otros para incrementar el crecimiento económico y el desarrollo social haciendo uso de colaboraciones tecnológicas y de innovación.

En cualquier caso, se puede notar como la gran mayoría de definiciones tienen varios aspectos en común. El más notorio y decisivo de ellos es el uso de las tecnologías de la información y comunicación, junto con el mejoramiento del estilo de vida de los ciudadanos y la economía (Mitchell & Villa, 2013). Otras características como las descritas en (Caragliu, Del Bo & Nijkamp, 2011) pueden ser enunciadas para homogeneizar el concepto de Ciudad Inteligente, por lo que a continuación, se encuentra un resumen de las características propias que tienen generalmente:

1. La utilización de infraestructura de redes para mejorar la eficiencia económica y política y para habilitar el desarrollo social, cultural y urbano. Aquí se incluye el uso de servicios, sensores, cámaras, teléfonos, internet, etc.
2. Un énfasis subyacente en el desarrollo urbano liderado por los negocios. Esto para aumentar el atractivo de la ciudad para la inversión.
3. Un enfoque fuerte en el objetivo de lograr inclusión social de los residentes urbanos a los servicios ofrecidos. Esto hace referencia a la idea de que todas las clases sociales se deben beneficiar de los proyectos.
4. Esfuerzo en el rol tan importante de las industrias de tecnología y de creatividad para el crecimiento urbano a largo plazo.
5. Profunda atención al papel del capital social en el desarrollo urbano que busca aproximar a la comunidad a los beneficios de la ciudad inteligente para incrementar su uso y eficiencia.
6. Sostenibilidad social y ambiental como un componente estratégico. Esto está relacionado a la responsabilidad con el medio ambiente en un mundo donde los recursos son limitados y escasos.

De acuerdo con lo anterior, para este trabajo se entenderá que: “una ciudad inteligente es aquella que emplea estructuras tecnológicas y utiliza herramientas de software para encontrar soluciones a sus diferentes problemáticas a partir de la recolección de información relevante con el objetivo de aumentar el desarrollo económico y social”.

De acuerdo con lo anterior, las ciudades inteligentes requieren recolectar una gran cantidad de información. Particularmente, en el caso de un proyecto enfocado en la seguridad ciudadana, estos datos podrían incluir videos provenientes de cámaras personales, de seguridad o en drones, grabaciones y reportes



de sensores o sistemas de radiofrecuencia. Todo esto debe ser procesado y analizado para obtener patrones de comportamiento, predicciones, mediciones y otra información relevante en términos de seguridad.

1.2. La inteligencia computacional como soporte a las ciudades inteligentes

Como el análisis e interpretación de los datos forman una parte esencial en el éxito de un proyecto de Ciudad Inteligente, es necesario ahondar en los términos de Algoritmos Estadísticos e Inteligencia Computacional que, cómo veremos más adelante, son las herramientas que permiten darle sentido a la información recolectada y reconocer patrones.

La Inteligencia Computacional se puede definir como el estudio del diseño de agentes inteligentes, siendo un agente inteligente algo que interactúa con el ambiente de una forma que se pueda catalogar como inteligente (Poole & Mackworth, 2017). El problema con este tipo de definiciones es que pueden ser poco concretas debido a la naturaleza del término "inteligencia". Aunque algunos autores intentan definir inteligencia como, por ejemplo, la habilidad para alcanzar objetivos, esto se puede quedar corto al momento de describir los diferentes niveles y clases de inteligencia que existen. Esto hace que sea difícil el poder caracterizar qué tipos de procedimiento computacionales podemos denominar como inteligentes y, por lo tanto, no es posible consensuar una definición que no dependa de la inteligencia humana o de un criterio subjetivo (McCarthy, 2007).

La Inteligencia Computacional comprende diferentes conceptos, paradigmas, algoritmos e implementaciones que muestran un comportamiento en ambientes complejos catalogables como inteligentes (Kacprzyk & Pedrycz, 2015) y tiene una gran variedad de aplicaciones actualmente, entre las cuales podemos contar el reconocimiento y detección de patrones en datos de varias dimensiones, predicción de eventos, optimización, sistemas de decisión y sistemas de clasificación. La estrategia se centra en aplicar técnicas y métodos con enfoque heurístico que por lo general encuentran soluciones parcialmente válidas o incompletas, pero con el nivel de exactitud suficiente requerido por la aplicación y con gran eficiencia. Estos algoritmos en general no proveen al investigador de un entendimiento completo o determinístico del fenómeno estudiado ni permiten su descripción con modelos matemáticos. En cambio, ofrecen soluciones de simple implementación, rápida ejecución y eficiente en el empleo de recursos.

Por otro lado, un algoritmo estadístico permite crear un modelo estadístico de los datos recolectados con el fin de comprender el fenómeno y predecir o clasificar datos nuevos (Wandelt, Bux & Leser, 2014). A diferencia de los algoritmos de inteligencia computacional, esto nos permite tener un conocimiento más profundo de la información y de su fuente, su elaboración depende de la capacidad de análisis del investigador y, por lo tanto, pueden representar un reto mayor. Algunos ejemplos de algoritmos estadísticos se encuentran a continuación (Farrelly, 2018):

- Modelos lineales generalizados
- Métodos de series de tiempo
- Agrupamiento K-Means
- Métodos bayesianos
- Modelos de regresión penalizada
- Procesos estocásticos y cadenas de Markov
- Estadística inferencial

2. Metodología

Teniendo en cuenta el tipo de estudio a realizar se diseñó una metodología que estuviera en la capacidad de responder a los siguientes requerimientos:

- Identificación de las métricas más relevantes para monitorear y mejorar la seguridad ciudadana utilizando métodos estadísticos y de inteligencia computacional.



- Categorización y clasificación de los principales algoritmos estadísticos y de inteligencia computacional implementados de forma recurrente y productiva en los últimos años.
- Identificación de las herramientas informáticas más ampliamente utilizadas para la integración de los algoritmos categorizados previamente de acuerdo con la calidad de sus prestaciones

Una vez definidos los requerimientos metodológicos, se adelantó el siguiente proceso:

1. **Búsqueda:** Se realizó una búsqueda de artículos científicos y trabajos académicos publicados en los últimos años que se relacionen con el tema de investigación en dos bases de datos diferentes y en la herramienta de Google Académico. Las bases de datos son IEEE Xplore y Scopus. Ambas bases de datos gozan de prestigio entre la comunidad científica y cuentan con un extenso número de publicaciones. Como palabras claves para orientar la búsqueda se utilizaron las presentadas en la Tabla 1.

Smart city	City security and safety	Intelligence algorithms	Smart algorithms
Gunshot detection	Gun noise detection	Gun Detection	Firearm recognition
Face detection and recognition	Body worn cameras	Surveillance video analysis	Disaster management
Natural disaster prevention	Fire detection	Landslide detection	Landslide detection

Tabla 1. Criterios de búsqueda. Fuente: Elaboración propia.

Posteriormente, se hizo una revisión de las fuentes de los artículos encontrados para ampliar el conjunto de trabajos a revisar. Para preseleccionar un artículo, se tomó en cuenta el título, el resumen y la fecha de publicación. Con estos campos se decidió a criterio de los autores si el trabajo era de relevancia para la investigación.

2. **Evaluación:** Se llevó a cabo una revisión de la relevancia y confiabilidad de los documentos encontrados de acuerdo con la reputación de sus escritores, el lugar de publicación, el año y la institución dónde fueron realizados. Estos criterios se utilizaron como muestra de la calidad y veracidad de las fuentes y para el descarte de documentos poco aptos para el trabajo de investigación. La reputación de los autores se midió con el índice h y el número de citas según Google Académico cuando el perfil existía o Scopus y la IEEE en caso contrario.

3. **Lectura:** Se hizo una lectura detallada de los artículos y trabajos encontrados. Se resaltaron las partes más relevantes de los trabajos a medida que eran leídos.

Las siguientes dos etapas permitieron: “Reunir las métricas más relevantes para monitorear y mejorar la seguridad ciudadana utilizando métodos estadísticos y de inteligencia computacional”.

4. **Extracción:** Se identifican las métricas utilizadas en los artículos consultados que permitan cuantizar diversos aspectos de la seguridad ciudadana y calificar el desempeño de los métodos utilizados.

5. **Valoración:** Se realizó una revisión de cuales métricas resultan más dicentes e informativas, también como cuáles son sus restricciones y campos de aplicación.

De igual forma, para el desarrollo del proyecto se debía cumplir con el siguiente objetivo: “Categorizar de manera ordenada los principales algoritmos estadísticos y de inteligencia computacional implementados de forma recurrente y productiva en los últimos años para trabajar las métricas encontradas anteriormente”. Para ello se adelantaron las siguientes actividades:

6. **Caracterización:** Simultáneamente a la lectura, se fueron extrayendo las características principales de cada trabajo. Ítems como palabras claves, algoritmos, data sets, librerías de código y plataformas utilizadas o mencionadas fueron organizadas en un documento auxiliar. De cada trabajo, se realizó un resumen detallando: descripción, objetivo, tipo de aproximación, lugar de ejecución, entre otros, según correspondiera.

7. **Planeación:** Una vez se tuvieron identificadas las principales características de los artículos y trabajos, se procedió a proponer una estructura general para el informe final. Se realizó una lista con todos los temas



y subtemas que se querían abordar en cada sección del trabajo y se plantearon las conexiones entre estos. De esta manera, se aseguró tener una hoja de ruta que diera coherencia y cohesión al escrito.

8. Escritura: Luego de los pasos anteriores, se inició la escritura del cuerpo del estado del arte. Se sigue la estructura planeada anteriormente para incluir la bibliografía más relevante de acuerdo con los criterios de selección mencionados previamente.

Finalmente, buscando contribuir a los investigadores que trabajan en el campo de las soluciones de seguridad ciudadana en el entorno de ciudades inteligentes, las actividades finales de la metodología se orientaron a: “Recomendar los algoritmos y herramientas informáticas más ampliamente utilizadas para la integración de los algoritmos encontrados anteriormente teniendo en cuenta la calidad de sus prestaciones”.

9. Indagación: Se profundiza la información sobre los algoritmos y herramientas informáticas encontradas durante la investigación. Esto incluye pros, contras, forma de uso, entre otros aspectos que se consideren relevantes.

10. Listado: La información producida durante la indagación referente a los algoritmos y herramientas informáticas más utilizadas y aceptadas se condensa en forma de tabla simplificada a modo de recomendación para futuros trabajos en el tema.

3. Estado del arte

En general, existe una gran cantidad de aplicaciones que pueden ser enfocadas en la seguridad ciudadana e incluyen algoritmos de inteligencia computacional o estadísticos en su implementación. Revisar cada una de ellas no es posible en términos prácticos debido a la cantidad de información que representan y variedad de temas que se incluyen. Por esta razón, se escogieron dos aplicaciones principales de las cuales se realizará la aproximación a su estado del arte. Estas aplicaciones fueron escogidas a criterio de los autores debido a su relevancia en el panorama investigativo en el momento y a su potencial para influir positivamente en la seguridad de los cascos urbanos.

Las aplicaciones seleccionadas son: detección de disparos y detección de armas. Para mostrar el estado del arte de estos tópicos se dedica una sección a cada uno en la cual se hace una introducción al tema, se esquematizan los trabajos más relevantes en una tabla junto con un análisis posterior de trabajos secundarios, se mencionan las conclusiones y se proponen opciones a ser exploradas en el futuro para expandir la investigación en el tema.

3.1. Detección de disparos

Una de las aplicaciones más mencionadas y trabajadas en proyectos de Ciudades Inteligentes enfocados a la seguridad ciudadana es la detección y ubicación de disparos. Esto porque el sonido de los disparos en áreas urbanas puede aumentar la sensación de inseguridad de sus habitantes junto con todas las consecuencias sociales y económicas que esto conlleva. Además, el sonido de los disparos está altamente relacionado con actividades ilícitas y violentas con resultados fatales para la vida humana.

Detectar un disparo y localizar su origen ayuda a que la respuesta de las autoridades y servicios de ayuda sea bastante más rápida y efectiva, lo que se traduce en mejor atención a los posibles heridos en la escena y a la identificación y posible captura de los responsables involucrados. Adicionalmente, estos sistemas de detección proveen información sobre dichos eventos que puede servir para el análisis de las zonas de incidencia y la creación de algoritmos de predicción o reconocimiento de patrones.

Debido a que existen diferentes enfoques en esta área; como la criminalística (Lindström et al., 2014; Głomb et al., 2018) o la clasificación e identificación de armas (Kiktova et al., 2015), es importante definir el tipo de trabajos que se cubren en esta sección. Son de interés aquellos trabajos con desarrollos que permitan identificar sonidos de disparos por armas de fuego en áreas urbanas en contraste a otros tipos de sonidos explosivos y a sonidos de la cotidianidad de una ciudad. Los trabajos no tienen que estar enfocados en o



pertenecer a un proyecto de ciudad inteligente estrictamente para ser considerados.

Gran parte del éxito de los sistemas de reconocimiento de disparos es debido al trabajo realizado previamente para entender la física detrás del disparo de un proyectil de bala. Diversos estudios como (McCoy, s.f.; Fansler, 1998; Jiang, 2003; Freytag, Begault & Peltier, 2006; Maher, 2007), han podido extraer características comunes de los disparos y proveen esclarecimiento sobre las fuentes sonoras involucradas en un disparo de arma de fuego. Estos trabajos permiten conocer la firma sonora de los disparos y han encaminado las investigaciones de detección a enfocarse principalmente en el sonido de explosión proveniente de la boca del arma y no tanto en otros sonidos como la onda de choque del proyectil que no es producido por todos los tipos de armas o el sonido del impacto del martillo que es poco perceptible.

A continuación, se encuentran condensados en una tabla los artículos más relevantes de esta sección. Se describen: el nombre del artículo junto con sus autores y año de publicación, un breve contexto que incluye el objetivo y enfoque con el cual fue desarrollado, los algoritmos utilizados y los resultados obtenidos. Esta tabla contiene los diferentes enfoques más representativos del área de investigación (Tabla 2).

Nombre, autores y año	Contexto	Algoritmos utilizados	Resultados
Analysis of Linear Predictive Coefficients for Gunshot Detection Based on Neural Networks. (Hrabina, 2017) Autor: Martin Hrabina. Año: 2017.	Este trabajo busca identificar sonidos de disparos entre otros tipos de sonidos para ayudar a la protección de sistemas en África central de la guerra ilegal. Su enfoque es evaluar diferentes métodos de extracción de características.	Se utilizaron Coeficientes de autocorrelación (ACC), Coeficientes de codificación lineal predictiva (LPC), Coeficientes cepstrales de predicción lineal (LPDC) y Coeficientes cepstrales de frecuencia de Mel (MFCC) para la extracción de características y se utilizaron redes neuronales multicapa con una capa oculta para la clasificación.	Se midió el éxito de los algoritmos con el porcentaje de falsas alarmas detectadas. En este campo, ACC obtuvo el menor desempeño con un promedio de error de 55 % seguido por MFCC con un 16%. LPC y LPDC obtuvieron los mejores valores de error. Específicamente con 20 coeficientes, se lograron los valores mínimos de 10.67% y 11.62 para LPC y LPDC respectivamente.
Gunshot Classification and Localization System using Artificial Neural Network (ANN). (Galinggus y Guimbal, 2019) Autor: Cherry Mae J. Galinggus, Shervin A. Guimbal. Año: 2019.	Este trabajo se centra en desarrollar un sistema capaz de clasificar disparos y encontrar la localización aproximada de ellos usando una red de 4 neuronas. Esto se realiza con el fin de identificar y localizar disparos en zonas de conflicto bélico armado y con ellos aumentar la ventaja y tiempo de preparación y respuesta de los soldados involucrados. El proyecto utiliza el Eui HMF.	Se utilizó la transformada rápida de Fourier para obtener la información espectral de la señal en el preprocesamiento. Se utilizaron también dos redes neuronales. Perceptrón multicapa que reciben para información la detección del disparo en el sonido de la onda de choque producida por la bala y el sonido de la explosión en el arma. Finalmente, se utilizó una red neuronal Perceptrón multicapa para detectar el ángulo de orientación del evento.	Los resultados se presentan para distancias entre 10 y 100 metros y 24 Ángulos distribuidos omnidireccionalmente. El sistema es capaz de detectar disparos correctamente con un 98% de eficiencia y detectar el ángulo de ubicación con un 90%. Se nota una mayor variación en los resultados cuando se hace un cambio de ángulo que con una variación de distancia en los rangos establecidos.
Gunshot detection using Convolutional Neural Networks. (Bajik, Pritosi y Komar, 2020) Autor: Jelub Bajik, Jit Pritosi, Dusan Komar. Año: 2020.	Este trabajo busca crear un sistema de detección de disparos usando redes neuronales convolucionales sobre representaciones bidimensionales de las grabaciones. No se encuentra enmarcado en un contexto en particular.	Los autores extrajeron el espectrograma, los Coeficientes Cepstrales de Frecuencia de Mel (MFCC) y las matrices de auto-similitud de las grabaciones de entrada para obtener una representación bidimensional que pudiera mostrarse en forma de imágenes. También usaron los modelos diferentes de redes neuronales convolucionales: VGG18 e Inception V3 con pesos preentrenados y sin entreno y ResNet18.	En general, se obtuvo un porcentaje de éxito mayor al 95% entre los diferentes modelos de redes neuronales convolucionales. El valor más alto obtenido fue de 99.14% usando ResNet18 con pesos preentrenados. También se demostró que bajar la frecuencia de muestra de 44 kHz a 8 kHz no afectó significativamente la precisión del sistema para detectar disparos, obteniendo una exactitud de 99.07% con ResNet18.
Low-Cost Gunshot Detection using Deep Learning on the Raspberry Pi. (Morehead et al., 2019) Autor: Alex Morehead, Laurin Ogden, Gabe Mages, Ryan Hostler, Bruce White, George Mohr. Año: 2019.	Este trabajo desarrolla un sistema de detección de disparos usando redes neuronales convolucionales y con especial énfasis en lograr un bajo costo de implementación con el fin de aumentar los niveles de seguridad en ciudades.	Se implementaron dos sistemas diferentes para la detección. El primer sistema pasa la señal de audio como una representación de series de tiempo por una red neuronal convolucional de una dimensión, cuatro capas ocultas y dos capas completamente conectadas. El segundo sistema extrae el espectrograma de la señal de audio y la pasa por una red neuronal convolucional de dos dimensiones e igual arquitectura.	Los resultados obtenidos son bastante buenos con una exactitud del 99.4% en ambos sistemas. En otros medidores la diferencia es poco significativa también. Además, se obtuvo una exactitud de 99.4% al cambiar el sistema para distinguir entre sonidos de disparos y de protesta.
Implementation of developed gunshot Detection algorithm on TM3200C673 processor. (Hrabina y Sigmond, 2016) Autor: Martin Hrabina, Milan Sigmond. Año: 2016.	Este trabajo implementa un sistema de detección de disparos usando una combinación de tres diferentes técnicas de detección conjuntas, con énfasis en la eficiencia computacional con el propósito de cumplir los eventos de guerra ilegal de instantes en África.	Los tres métodos utilizados son: 1) Coeficientes de codificación predictiva lineal LPC para los cuales se usan 8 coeficientes calculados con el algoritmo Levinson-Durbin. 2) Detección de energía en bandas de frecuencia en lo que se utiliza la Transformada rápida de Fourier FFT y 3) Detección de la señal en el tiempo usando la transformada encontrada anteriormente. Datas técnicas se usan en cascada.	Se obtiene un porcentaje de detecciones correctas de 92%, un porcentaje de falsas negativas de 18% y un porcentaje de falsas positivas de 3%. Se puede aprender como la tasa de detecciones erróneas mejoró drásticamente al utilizar las tres técnicas en cascada. Esto a cambio de un decremento en la exactitud del sistema.

Tabla 2. Trabajos más relevantes de detección de disparos. Fuente: Elaboración propia.

Existen diversos trabajos que también utilizan Coeficientes de máxima correlación cruzada, Coeficiente de codificación lineal predictiva y Coeficientes cepstrales de Mel para la extracción de características con buenos resultados (mayores al 95% en exactitud) (Ahmed, Uppal & Muhammad, 2013; Shiekh, Tahir & Uppal, 2017). Estos, además, comparten el hecho de que hacen uso de Máquinas de Soporte Vectorial para la etapa de clasificación.

Otros trabajos enfocados en el análisis de las características temporales de las señales de audio obtienen resultados bajos en comparación. En el caso de (Samireddy, Carletta & Lee, 2017), se utilizó un filtro digital



que computa la correlación cruzada entre una señal de muestra y la señal de entrada para determinar el nivel de similitud entre ellas y se clasifican las señales utilizando un umbral de mínima similitud escogido previamente con base en las observaciones de los autores.

En el caso de (Hrabina & Sigmund, 2018), los autores parten del supuesto de que los sonidos de disparos tienen formas de onda lo suficientemente particulares como para poder ser identificadas entre otros tipos de sonidos. Para esto, evaluaron diferentes características de la señal de audio en el dominio del tiempo tales como: forma de la onda, distancia entre el primer máximo y mínimo, distribución de la distancia entre niveles contiguos de amplitud cero, distancia entre picos de la señal, ángulos y área entre diferentes puntos de la señal y aproximación de la envolvente a la función exponencial. Estas características fueron pasadas por una red neuronal con una capa escondida de diez neuronas. Los resultados obtenidos no sobresalen con respecto a otras técnicas al obtener niveles máximos de exhaustividad de 82.2% y precisión de 69.3%. Los autores concluyen que la detección de disparos es posible usando técnicas de análisis temporal pero que se deben buscar formas de mejorar los resultados. Algunos métodos de mejora pueden ser la búsqueda de otras características en la forma de onda de la señal o la búsqueda de combinación óptima de características que se deben pasar al sistema de clasificación.

Conclusiones respecto a los trabajos analizados

Como es de esperarse, la gran mayoría de trabajos en esta área utilizan el análisis de grabaciones de sonido para la detección, por lo que las mayores diferencias se encuentran en los algoritmos y el hardware utilizados en el proceso. De hecho, existen empresas privadas que ofrecen soluciones de seguridad que pueden articularse con las autoridades e incluyen detección de disparos usando análisis de sonido tales como ShotSpotter (Company, 2020).

Curiosamente, el concepto de Ciudad Inteligente es poco mencionado en los trabajos de detección de disparos, a pesar de que un despliegue de estos sistemas encaja en la definición de Ciudad Inteligente. Este parece ser un caso en el que un desarrollo se une a una aplicación mayor de forma orgánica debido a su naturaleza y a que se vuelve necesario para esta.

Uno de los primeros patrones que se pueden notar en este campo de aplicación es cómo los algoritmos estadísticos y de inteligencia computacional se aplican generalmente en dos etapas distintas del proceso. La primera es en la etapa de preprocesamiento de la señal. En esta, no se limita el tratamiento de la señal a la normalización que se realiza usualmente en análisis de audio, sino que se busca extraer características de la señal que faciliten el trabajo del sistema clasificador, usualmente con métodos matemáticos o estadísticos. Esto puede implicar cambiar el dominio o el tipo de representación de la señal de audio, en muchos casos, a características que no poseen un sentido físico para el entendimiento humano. La segunda es la etapa de clasificación, en ella se exploran diferentes tipos de métodos, pero se nota la predominancia de los algoritmos con redes neuronales.

Otro aspecto que se da a notar revisando el conjunto de trabajos realizados en esta aplicación es la tendencia y mayor éxito de los algoritmos que analizan características relacionadas al espectro de frecuencias de las señales de audio, en comparación con el uso menos frecuente y menor éxito del análisis de características temporales.

En términos de hardware, los desarrollos en esta área generalmente proponen el diseño de dispositivos con la capacidad de grabar, almacenar temporalmente la grabación y realizar el procesamiento y clasificación por sí mismos, por lo que se pueden clasificar como nodos activos. Lo anterior da como resultado un sistema descentralizado en términos del procesamiento, es decir, que el procesamiento se realiza de forma individual por cada nodo final. Esto hace que sea necesario además un sistema adicional de envío de alertas a un ente central para poder tener visualización y registro de las detecciones.



En términos de software, es posible concluir que, con la tecnología y desarrollos actuales, ya es posible detectar sonidos de disparos y distinguirlos de otros sonidos de fondo con una precisión lo suficientemente alta como para construir sistemas confiables incluso cuando existe ruido de fondo. La investigación por venir en este campo no se ve en la obligación de mejorar los márgenes de detección, sino que tendrá el objetivo de hacer esta tecnología más accesible y eficiente.

En este sentido, se sigue que las redes neuronales en diferentes presentaciones sean la opción más frecuente en los trabajos más relevantes, ya que solo requieren una gran carga computacional en la etapa de entrenamiento, la cual se puede realizar en máquinas potentes. Una vez el sistema está entrenado y los pesos óptimos de la red han sido encontrados, se requiere un nivel de procesamiento relativamente bajo que puede ser alcanzado con sistemas de bajo coste.

Sin embargo, existe una tendencia a la búsqueda de otros métodos de clasificación de sencilla implementación que pueden ser potenciados al combinarse varios entre sí o con técnicas de preprocesamiento también de fácil implementación. Entre estas técnicas de preprocesamiento más utilizadas y con mejores resultados se encuentran la extracción de Coeficientes cepstrales de frecuencia de Mel (MFCC) y Coeficientes de codificación lineal predictiva (LPC), el cálculo de la transformada de Fourier y la utilización de espectrogramas. Mientras que en lo referente a técnicas de clasificación se encuentra principalmente la correlación. Otras aproximaciones como el análisis de formas de onda o medición de densidad de energía de la señal no han mostrado resultados exitosos hasta el momento.

Es importante notar que, para la mayoría de las aplicaciones, resulta más conveniente minimizar la tasa de falsas detecciones que maximizar la tasa de detecciones correctas. Esto es debido a que, en los eventos de disparo de armas de fuego, generalmente involucran múltiples disparos, por lo que los sistemas implementados pueden tener varias oportunidades para realizar detecciones correctas y enviar la alerta según corresponda (Pikrakis, Giannakopoulos y Theodoridis, 2008). El sistema no tiene que detectar el 100% de los disparos registrados, pero si debe evitar malinterpretar todos los demás sonidos al máximo posible.

Finalmente, en lo referente a la localización del origen del disparo, diferentes autores han encontrado que la mayor parte de la energía en un disparo, y por lo tanto el sonido detectable producido, se distribuye en la parte delantera del arma hacia donde la bala sale disparada. Esta observación es necesaria al hacer el análisis de qué sensores tendrán más éxito en la detección estando a la misma distancia, pero con diferente orientación, ya que el sonido producido no es omnidireccional. Esto también debe ser considerado en sistemas de localización del disparo.

Trabajo futuro en cuanto a detección de disparos

Es necesario seguir investigando formas de realizar sistemas de detección de disparos de bajo costo, bajo consumo energético, con compactación en el hardware y especialmente que sean capaces de realizar la detección en tiempo real. De igual manera, podría ser enriquecedor evaluar las decisiones en la arquitectura de los sistemas de detección de disparos tal que se puedan explorar otras topologías o maneras de distribuir las tareas, teniendo en mente que estos sistemas deben estar ampliamente distribuidos en las zonas urbanas para lograr su objetivo.

Por ejemplo, en lugar de utilizar los nodos finales como nodos activos para todas las tareas de grabación, almacenamiento, procesamiento y clasificación, los cuales además requieren un sistema de comunicación con un nodo central para el envío de alertas, se podría aprovechar este sistema de comunicación centralizado para enviar las señales recolectadas y así realizar el procesamiento y clasificación en dicho nodo central. Esto podría reducir los costos en los nodos finales, que pasarían a ser nodos pasivos, aunque obliga a tener un nodo central con mayores recursos y un sistema de comunicación con mayor capacidad. A continuación, las Figuras 1 y 2 se muestran los diagramas de un sistema descentralizado con nodos activos y centralizado con nodos pasivos.



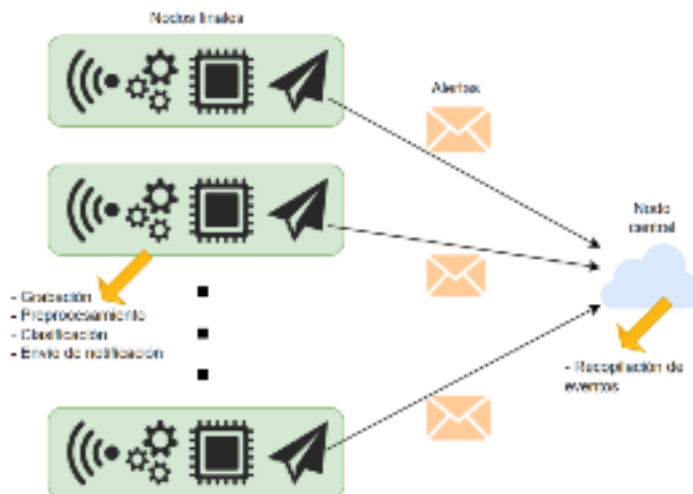


Figura 1. Arquitectura descentralizada con nodos finales activos. Fuente: Elaboración propia.

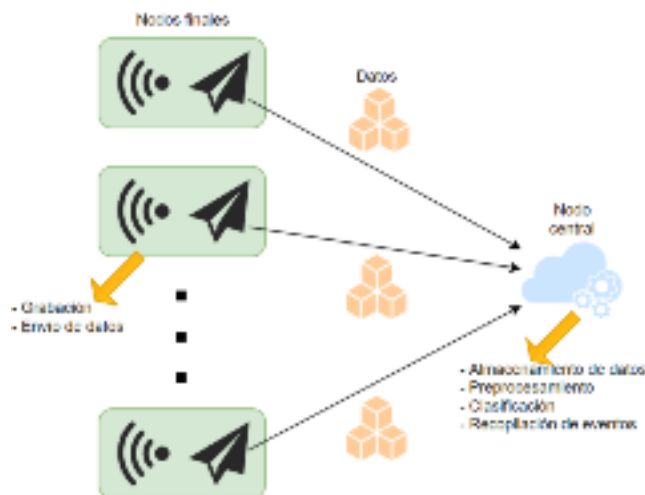


Figura 2. Arquitectura centralizada con nodos finales pasivos. Fuente: Elaboración propia.

Además, es conveniente generar sistemas capaces de generalizar los sonidos producidos por diferentes tipos de armas de fuego, lo que podría lograrse añadiendo estos otros sonidos a los sistemas en la etapa de entrenamiento. Con esto, se puede lograr un mayor alcance en los sistemas de detección y una disminución en la tasa de falsos negativos. Otro enfoque dirigido a aplicaciones urbanas y a modo de recomendación sería utilizar estudios, si existen, que muestren los tipos de armas de fuego más utilizadas en la zona y añadir especial énfasis en los sonidos de disparos producidas por estas en la fase de entrenamiento.

Finalmente, podría ser de gran valor para esta aplicación la creación de sistemas adaptativos que puedan mantener un nivel aceptable de precisión al recalibrarse automáticamente con base en las condiciones auditivas de detecciones exitosas previas y el ruido en el ambiente del lugar de instalación.

3.2. Detección de armas

En menor medida y con menor diversidad de algoritmos utilizados en comparación con la sección anterior, se encuentran los proyectos de detección de armas de fuego en grabaciones de video. Sin embargo, esto no resta ni importancia ni afinidad a esta aplicación en el marco de la gestión de la seguridad ciudadana en ciudades inteligentes. La presencia de armas en espacios públicos y cerrados tiene el potencial de desembocar en lesiones u homicidios y alteraciones al orden público y por lo tanto, se traduce en la afectación de la salud



y vida de las personas y la disminución de la sensación de seguridad en la población, especialmente en un país como Colombia donde el porte de armas no es libremente permitido para civiles (Ministerio de Justicia de Colombia, 2000).

La detección de armas en videos de vigilancia tiene la capacidad de: 1) ayudar en la prevención de siniestros al alertar a los entes de control de la presencia de un arma en cierto espacio y 2) ayudar en el seguimiento e identificación del portador tanto si el arma se alcanzó a disparar o no.

Al realizar una primera inspección, se puede notar como la gran mayoría de los trabajos en el área concuerdan en que los sistemas de videovigilancia que utilizan el factor humano como actor vigilante para la detección de armas son ineficientes debido a la incapacidad humana de revisar diferentes videos simultáneamente y de mantener la concentración de manera continua en uno solo (Smith, 2004). Por esto, se hace necesario recurrir a la tecnología, que ofrece maneras de abordar el problema y brindar soluciones de forma más eficaz.

Es importante notar que el problema de detección de objetos en general ha sido abordado exhaustivamente, produciendo algoritmos bastante capaces de detectar objetos ampliamente visibles en la imagen, incluyendo detección de personas, animales, vehículos y plantas (Salazar González et al., 2020). Sin embargo, la detección de armas es compleja porque éstas son objetos de menor tamaño, son movidas con rapidez, a menudo son ocluidas por otros objetos como las manos y tienen gran cantidad de formas y tamaños. De igual manera, se suman otros inconvenientes propios de la detección de objetos en imágenes tales como el tamaño relativo del objeto en la imagen, el contraste con el fondo y la rotación del objeto (Debnath & Bhowmik, 2020).

Como es de esperarse, las personas que portan y utilizan un arma intentan llevarla oculta tanto tiempo como sea posible. Esto y otros factores producen dos fenómenos que dificultan la detección: la autooclusión y la interoclusión, el segundo con más probabilidad de ocurrencia que el primero. La autooclusión se produce cuando las armas de fuego se solapan entre sí en una imagen, desdibujando su forma, mientras que la interoclusión se produce cuando otros objetos están posicionados por delante del arma de fuego en la imagen, siendo las manos el objeto más común (Verma & Dhillon, 2017).

Son de interés para esta sección los trabajos recientemente publicados que proponen desarrollos para la detección de armas de fuego en imágenes, especialmente si el enfoque se encuentra en extraer dichas imágenes de grabaciones de videos de vigilancia. Se priorizan también aquellos trabajos que incluyan la capacidad de detección en tiempo real. No son de interés trabajos que únicamente se centran en la clasificación de diferentes tipos y referencias de armas de fuego.

Estos trabajos mencionados, tomarían ventaja del vasto número de sistemas de videovigilancia que existen tanto públicos como privados para obtener las grabaciones necesarias para la detección. En términos de la arquitectura de los sistemas de monitoreo videográfico, se da el caso de éstos que tienen sus cámaras conectadas a un sistema central que recibe y almacena las grabaciones obtenidas. Este sistema central sería el encargado de ejecutar los desarrollos para detección de armas o de reunir y entregar las grabaciones a un sistema de detección externo. A diferencia de en la sección anterior, la centralización de los sistemas de videovigilancia permite que la capacidad de procesamiento sea mayor, lo cual, es además necesario tratándose de imágenes.

A continuación, se encuentran condensados en una tabla los artículos más relevantes de esta sección. Todos los artículos a continuación comparten la preocupación por las tasas de violencia armada en diferentes partes del mundo y realizan su investigación con el fin de ayudar a las autoridades a controlar este flagelo, por lo que esto no se agregó a la columna de contexto (Tabla 3).



Nombre, autores y año	Contexto	Algoritmo utilizado	Resultados
Weapon Detection for Security and Video Surveillance Using CNN and YOLOv5. (Ahrari et al., 2021) Autores: Abou Hanan Ahrari, Muhammad Imran, Abdurrahman H. Cahani, Abdulmajed Aluayyil, Omar Alshayb, Hussa Mahmood, Muhammad Alsouq, Mohamed Hachem. Año: 2021.	Este trabajo busca implementar un sistema de detección de armas presentes en imágenes basado en el algoritmo YOLOv5 de forma que se diferencie los fondos uniformes y fondos no uniformes mientras se mantiene la velocidad de detección.	Los autores utilizan Diferenciación Gaussiana para la eliminación del fondo en las imágenes a analizar y YOLOv5 para la clasificación y Faster R-CNN para comparación.	Se obtuvo una precisión de detección en imágenes de 99.5% y 77.9% en una prueba de videos. En comparación, Faster R-CNN obtuvo una precisión menor de 94.2%. Se logró una velocidad de detección de 0.34 segundos en comparación con Faster R-CNN cuya velocidad de detección es de 0.12 segundos.
Application of Deep Learning for Weapons Detection in Surveillance Videos. (Hosaini et al., 2021) Autores: Tarek Magid Elshar, Mohamed, Hamed El-Haq, Mohamed, Mohamed Hoss, Mohamed El-Hadi. Año: 2021.	Este trabajo busca implementar y comparar dos sistemas de detección de armas en imágenes con los algoritmos con las relaciones de la familia de algoritmos YOLO para analizar su desempeño y realizar una comparación en los videos.	Los algoritmos utilizados son YOLOv3 y YOLOv4 que con dos modelos del sistema de código abierto YOLO (You Only Look Once) que consta de una única red neuronal convolucional de 21 capas convolucionales y 2 capas completamente conectadas.	Los autores, en conclusión, como era esperado, que el sistema YOLOv4 fue mejor que el sistema YOLOv3, con valores de precisión de 89% y 89% respectivamente. A pesar de que las ventajas de precisión y robustez del sistema YOLOv4 obtiene una precisión en videos, algunos errores humanos para los resultados de algoritmos que a menudo requieren mayor carga de procesamiento.
Automatic Gun Detection from Images Using Faster R-CNN. (Majidi et al., 2020) Autores: Rana M. Aliq, Jaleel A. Alsharhan, Basim A. Alsharhan, Raed A. Alsharhan, Raafiq D. Alsharhan, Hedi Benhbour. Año: 2020.	Este trabajo desarrolla diferentes sistemas de detección de armas empotrados en videos utilizando diferentes arquitecturas de Redes Neuronales como extractores de características en Faster R-CNN con el fin de comparar su desempeño y contrastar resultados con el sistema YOLOv2. El artículo describe de forma en su totalidad el funcionamiento del algoritmo Faster R-CNN.	Se utilizó el algoritmo Faster R-CNN con su estructura de redes neuronales convolucionales para extracción de características (VGG19) y también con videos empotrados (como: ResNet50, Inception ResNetV2, y MobileNetV2). Adicionalmente, se probó el algoritmo YOLOv2 sin modificaciones. Estas modificaciones fueron predefinidas utilizando el conjunto de datos ImageNet.	Se obtuvo una precisión igual o superior al 70% en todos los escenarios siendo el más sobresaliente Inception ResNetV2 con una precisión del 87%. En comparación, el sistema YOLOv2 obtuvo una precisión del 79%. Los autores muestran capacidad para detectar armas parcialmente ocultas e fondo de la imagen, con poco ruido en el fondo y desde la otra frontal. Sin embargo, mostró propósitos fallidos en el sistema de extracción de los resultados de videos.
Automatic Visual Gun Detection Carried by a Moving Person. (Doroudi y Eshooreh, 2020) Autores: Rajib Doroudi, Mraz Karo Ghannini. Año: 2020.	Este trabajo desarrolla un sistema de detección de armas siendo portadas por personas, utilizando métodos de detección de objetos, sustracción de fondo y Template Matching. La detección de objetos se utiliza para analizar la región donde se encuentra una persona sosteniendo un arma y pasar entonces al algoritmo de Template Matching para detección de arma como tal. Esto con el fin de aumentar la rapidez del sistema. Se comparan los resultados obtenidos con estos diferentes métodos y se contrastan con desarrollos realizados en otros trabajos.	Se utilizó el algoritmo para detección de objetos (MaskRCNN) para detectar personas llevando un arma. Para la misma tarea, también se evaluaron los algoritmos Gaussian Mixture Model (GMM), Modelo Temporal Diferencia (MTD) y VLD. Se aplicó el algoritmo de Template Matching que utiliza Redes Neuronales Convolucionales con la Transformada Generalizada de Hough sobre el área previamente delimitada para encontrar el arma.	El sistema de detección obtuvo una precisión y exactitud de 84.21% y 80.17% respectivamente y toma 1.0 segundos suerando a los otros trabajos relacionados y evaluados en la investigación tanto en desempeño como en rapidez. Los autores más cercanos de precisión, exactitud y tiempo de detección obtenidos por otras pruebas evaluados en el artículo es de 82.87%, 85.34% y 2 segundos, teniendo en cuenta que no todos los valores pertenecen al mismo sistema. El desarrollo es capaz de detectar armas en diferentes escenas y locaciones, pero no lo proceso en armas parcialmente ocultas.

Tabla 3. Trabajos más relevantes de detección de armas. Fuente: Elaboración propia.

Para esta aplicación, existen en realidad muy pocos algoritmos que se estén utilizando y provean resultados aceptables. A pesar de que, anteriormente en la tabla, se pretende mostrar variedad de algoritmos con el fin de realizar una cobertura amplia, la revisión de la literatura permite ver que unos se están usando en mucho mayor medida que otros. Lo que se quiere decir no es que la cantidad de investigación realizada para esta aplicación sea reducida, sino que el conjunto de investigaciones tiende a utilizar una lista reducida de algoritmos para sus implementaciones.

El primer problema que enfrentan los investigadores está relacionado con la búsqueda de bases de datos para entrenar los algoritmos de inteligencia computacional que pretenden utilizar. Esto sucede porque la mayoría de las bases de datos ofrecen imágenes de armas en primer plano que reposan con fondos uniformes o imágenes con mayor definición de personas sosteniendo un arma muy cerca de la cámara y ocupando gran parte de la imagen.

Esto dista bastante de las imágenes que se espera encontrar en aplicaciones reales donde estas deben ser extraídas de videos de grabación afectados por el ruido y tomados por cámaras ubicadas en lugares altos. Un porcentaje considerable de investigadores deciden utilizar estas bases de datos de todas formas para la implementación de sistemas de detección. Otros investigadores, como en (Lim et al., 2019; Warsi et al., 2019), decidieron construir sus propios conjuntos de imágenes. Para esto, tomaron grandes cantidades de fotografías o videograbaciones de personas utilizando distintos tipos de armas de fuego y logran simular lo que obtendrían de un sistema de videovigilancia. Por otro lado, otros investigadores, como en (Gaidon et al., 2016; Salazar González et al., 2020), decidieron utilizar simulaciones y videojuegos para generar estos bancos de imágenes. Este método de generación de conjuntos de datos tiene sus contras con respecto a la extrapolación a aplicaciones reales, pero presenta grandes pros en cuestión de costos y accesibilidad.

Algunos trabajos como (Olmos, Tabik & Herrera, 2018; Jain, Aishwarya & Garg, 2020; Jain et al., 2020) obtienen una precisión mayor al 80% en la detección de armas. Sin embargo, el conjunto de datos que se utiliza para entrenamiento y pruebas se construye a partir de imágenes con armas de fuego totalmente descubiertas, en primer plano y sobre un fondo uniforme, que, para efectos de esta investigación, se denominan imágenes sin contexto. Su contraparte se denomina imágenes con contexto y se refiere a imágenes cuyas características son o se asemejan a las que se extrajeren de una grabación de seguridad y en las cuales el arma está siendo

empuñada. A continuación, en la Figura 3, se puede observar un ejemplo comparativo.



Figura 3. A la izquierda: Imagen de arma de fuego sin contexto. A la derecha: Imagen de arma con contexto. Fuente: Tomado de (Desk, 2018; BEN BOTKIN, The Statesman Journal, 2019).

En términos de algoritmos, se puede notar como la gran mayoría de los trabajos utilizan Redes Neuronales Convolucionales (CNN) en sus implementaciones. Este algoritmo se utiliza en diferentes etapas del sistema, en compañía con otros algoritmos y con diferentes arquitecturas de tal manera que cerca del 80% de los trabajos revisados los utilizan y cerca del 95% los mencionan en algún punto del escrito. Entre los trabajos mencionados anteriormente resaltan 2 algoritmos, los cuales son descritos a continuación.

1. **Faster R-CNN**: Es la versión más avanzada de R-CNN hasta el momento, es el sucesor directo de Fast R-CNN y fue propuesto en (Ren et al., 2015). Este algoritmo utiliza una red neuronal convolucional en la etapa de extracción de características para obtener un mapa convolucional de características que será sometido a una red neuronal independiente encargada de encontrar segmentos de la imagen con mayor probabilidad de detectar el objeto a encontrar, llamadas Regiones de Interés. Estas regiones de interés pasan por un adaptador de tamaño y resolución para finalmente pasar por un clasificador basados en una máquina de soporte vectorial (SVM) con función Softmax para decisión y un regresor lineal para encontrar el segmento de la imagen con el objeto. Este algoritmo es utilizado por ejemplo en (Verma & Dhillon, 2017; Alaqil et al., 2020; Jain et al., 2020).

2. **YOLO**: Sus siglas significan You Only Look Once (en español: Solo miras una vez) debido a su capacidad de detección después de solo un paso de la imagen por la red neuronal convolucional, a diferencia de otros algoritmos como R-CNN que necesitan múltiples pasos. Este algoritmo utiliza una cuadrícula cuyas dimensiones son definidas a criterio del desarrollador para segmentar la imagen en un conjunto de recuadros. La imagen pasa por una red neuronal convolucional que producirá un vector por cada cuadro, el cual indica la probabilidad de la presencia del objeto buscado y las coordenadas y dimensiones del rectángulo que contiene al objeto en caso de estar presente. Estos datos son posteriormente analizados por un clasificador para determinar la presencia del objeto en la imagen. Hasta el momento, existen 5 versiones de este algoritmo, cada una presentando mejoras con respecto a su predecesora. Este algoritmo es utilizado en sus diferentes versiones en (Afandi & Isa, 2021; Ashraf et al., 2021; Dextre et al., 2021; Narejo et al., 2021).

En menor medida, se exploran otras arquitecturas de redes neuronales convolucionales como en (Lim et al., 2019) que usa M2Det que es piramidal y también otro tipo de algoritmos como el clasificador tipo Haar Cascade en (Jain, Aishwarya & Garg, 2020; Salazar González et al., 2020). Este último se observa en mayor medida en aplicaciones de detección de rostros, pero ha mostrado resultados competitivos en los trabajos citados anteriormente. Sin embargo, se debe tener en cuenta que las bases de datos utilizadas en ellos son en su mayoría compuestas por fotos de armas sin contexto.

Para cerrar esta sección, es apropiado mencionar que para ambos algoritmos Faster R-CNN y YOLO se han logrado velocidades de detección lo suficientemente cortas para permitir el análisis de videos en tiempo



real. Sin embargo, es necesario tener en cuenta que, al momento de utilizarlos en aplicaciones reales, con los inconvenientes técnicos que esto acarrea, es necesario hacer ajustes que aumentan la complejidad y por lo tanto el tiempo de detección. Por otro lado, la carga computacional de entrenamiento es tal que algunos desarrollos como (Alaqil et al., 2020) optan por realizarlos en servicios en la nube como Microsoft Azure o Amazon Web Services.

Conclusiones respecto a los trabajos analizados

A diferencia de otras aplicaciones de detección de objetos en imágenes, la detección de armas de fuego en grabaciones de vigilancia enfrenta el reto del tamaño del objeto a detectar. El hecho de que las armas de fuego tengan un tamaño tan pequeño en la imagen, si lo comparamos con los otros objetos a su alrededor, afecta el desempeño de los algoritmos que se utilizan y exigen medidas adicionales. Por esta razón, una cantidad considerable de artículos se centran en la detección de armas que estén siendo empuñadas por personas. Esto, además de ofrecer un enfoque más cercano a las aplicaciones reales, soluciona un gran factor de los problemas causados por la interoclusión con las manos de los portadores convirtiéndolos en información en favor de la detección.

Las aproximaciones con mayor potencial en esta aplicación hasta el momento son aquellas que realizan la detección de las armas en el contexto. Los trabajos revisados que se enfocan en la detección de armas sin contexto ofrecen un aporte en la investigación, pero son más apropiados para otro tipo de aplicaciones o como ejercicios académicos. Estos verían su desempeño afectado al momento de enfrentar imágenes sacadas de videos de vigilancia.

Diversos desarrollos utilizan bases de datos cuyas imágenes tienen buena resolución. A menudo, estas fotos son sacadas de producciones audiovisuales profesionales o son tomadas específicamente para el proyecto y cuentan con una definición superior a la cual tienen generalmente los sistemas de vigilancia con videos. Por ejemplo, las imágenes utilizadas en (Debnath & Bhowmik, 2020) tienen una resolución de 1920x1080 y fueron tomadas por una Nikon D5100. Esto genera una preocupación de la existencia de un sesgo en algunos desarrollos que podrían ver decaído su desempeño en aplicaciones reales. Sin embargo, la evolución en las mejoras de las cámaras de vigilancia en los últimos años podría hacer alcanzables estos estándares de calidad de imagen.

De lo anterior se concluye la importancia de las decisiones que se toman al momento de escoger o construir el conjunto de datos, no solamente para obtener buenos resultados al probar los desarrollos con este mismo sino también con los datos que producirán las aplicaciones cuando deban ser implementadas.

Otra gran limitación se encuentra en la capacidad para la detección de armas en tiempo real que requiere un análisis cuadro a cuadro lo suficientemente eficiente para no quedar rezagado. Esto obliga a los algoritmos a ser altamente eficientes. Después de la revisión en la literatura disponible, es posible concluir que el algoritmo por excelencia escogido para la detección de armas en imágenes son las Redes Neuronales Convolucionales. La inmensa mayoría de artículos publicados en esta área usan de una forma u otra este algoritmo, que, además, es bastante popular en la detección de objetos en general. Una de sus principales desventajas radica en el tiempo y recursos necesarios para realizar el entrenamiento. No obstante, como este procedimiento generalmente se realiza una sola vez, es posible realizarlo en un dispositivo externo con mayor capacidad para posteriormente pasar los valores entrenados al sistema de detección que pudieran tener especificaciones técnicas más modestas.

Las redes neuronales convolucionales generalmente se utilizan en las primeras etapas de los sistemas de detección para los procesos de extracción de características o para la determinación preliminar de la probabilidad de la presencia del objeto en diferentes segmentos de la imagen.

Un ejemplo de lo anterior es la familia de algoritmos R-CNN, que ha pasado de utilizar las redes



neuronales convolucionales para la extracción de características en subsegmentos de imágenes a extraer mapas convolucionales de características de una imagen completa para su posterior clasificación. Este método de detección presenta un menor número de falsos positivos y falsos negativos y puede alcanzar niveles de precisión mayores al 80%, sin embargo, su exhaustividad, es decir, la habilidad para detectar diferentes instancias del objeto en la imagen, son menores.

Por otro lado, las aproximaciones YOLO pueden detectar con mejor precisión diferentes objetos en una misma imagen. La resolución en la detección del algoritmo YOLO para diferentes objetos depende del tamaño escogido para los rectángulos definidos por la cuadrícula trazada en la imagen. De esta manera, entre menor sea el espacio ocupado por el objeto a detectar, menor debe ser el área de los rectángulos. Esto implica que se deba dividir la imagen en una mayor cantidad de rectángulos, lo que incrementa el tamaño de la matriz de resultados en la salida de la red neuronal convolucional y, por lo tanto, aumenta la carga computacional del algoritmo.

A pesar de que las diferentes aplicaciones mencionadas en esta sección están enfocadas en detección de armas de fuego en grabaciones de video, las métricas para determinar el éxito del desarrollo se basan en las tasas de detección del arma en las imágenes que representarían los cuadros de las grabaciones. Esto significa que se mide la precisión de la detección contando cada imagen por separado, lo cual ignora que en aplicaciones en funcionamiento, las imágenes vendrían en un conjunto serializado, que cada cuadro compartiría una alta similitud con el anterior y el siguiente, y que su velocidad de concurrencia es alta. Estas características permiten vislumbrar el hecho de que al ser captada un arma por una cámara de vigilancia aparecerá no solo en un cuadro si no en varios, por lo que el sistema implementado tendría varias oportunidades para realizar una detección positiva. Este hecho aumenta las posibilidades de realizar una detección positiva y por lo tanto aumentaría el desempeño general del sistema.

Finalmente, es necesario comparar las velocidades de detección que han obtenido los algoritmos mencionados durante esta sección. El mejor hasta el momento es YOLOv5 que ha obtenido tiempo de detección de 0.01 segundos permitiendo el análisis de 100 cuadros por segundo, lo cual es una velocidad mayor a la necesaria para procesar videograbaciones que corren por lo general de 30 a 60 cuadros por segundo. Faster R-CNN obtiene una velocidad mínima de 0.17 segundos permitiendo el análisis de cerca de 5 cuadros por segundo, lo cual es insuficiente para equiparar la velocidad a la que corren los videos de vigilancia, pero aun así podría ser implementado. Finalmente, otras aproximaciones como Haar Cascade y métodos de Ventana Deslizantes obtienen velocidades mayores a 0.2 segundos y quedan rezagados frente a otros algoritmos.

Trabajo futuro en cuanto a detección de armas

En primer lugar, son muy necesarias la construcción y libre difusión de bases de datos que sean adecuadas para esta aplicación. Esta tarea es difícil por las implicaciones logísticas y la alta cantidad de imágenes que necesita un conjunto de datos para considerarse amplio, pero encaminaría la investigación a su verdadero objetivo y proporcionaría los insumos para evaluar los desarrollos realizados con mayor verosimilitud.

Existe un objetivo común entre los investigadores de mejora constante en los algoritmos a utilizar, como se pueden evidenciar en los avances logrados con las diferentes versiones del algoritmo YOLO y la familia de algoritmos R-CNN, lo cual es necesario para el avance del área de investigación. Por esta razón, se debe prestar cuidado en futuras investigaciones en optimizar los algoritmos y reducir su tiempo de detección manteniendo o mejorando su precisión tal que se puedan utilizar cómodamente en grabaciones en tiempo real. Mientras tanto, podría significar un aporte relevante el realizar una evaluación del éxito de detección de los algoritmos actuales sobre grabaciones en tiempo real, de forma que se evalúen solo un porcentaje de los cuadros generados según la velocidad de detección lo permite. Este experimento podría tener éxito basándose en la hipótesis de que no es necesario analizar todos y cada uno de los cuadros en una grabación para detectar la presencia de un arma, ya que se esperaría que el arma esté presente durante varios cuadros. Se esperaría



una tasa de éxito menor a la que obtendría un sistema que pueda analizar todos los cuadros generados, pero una tasa mayor a la de un sistema que analiza un conjunto de imágenes no relacionadas entre sí y que basa sus métricas de desempeño en el éxito de detección de cada imagen independientemente.

Seguidamente, resultaría enriquecedor para el área de investigación la exploración de diferentes arquitecturas de redes neuronales convolucionales y diferentes tipos de clasificadores en los algoritmos que se están usando actualmente para diversificar los métodos utilizados y también para encontrar configuraciones más precisas y eficientes o descartarlas.

4. Conclusiones

Este estado del arte permitió observar que para las diferentes problemáticas de seguridad ciudadana estudiadas existen diversos abordajes desde las técnicas de inteligencia artificial, siendo por lo tanto un campo de gran interés para los investigadores hoy en día. Adicionalmente, se evidenció que independientemente de la aplicación, siempre que se requiere la detección de algún objeto, sean rostros, armas, o fuego, el algoritmo más utilizado e investigado son las Redes Neuronales Convolucionales en sus diferentes variantes.

Desde el punto de vista tecnológico, se evidencia una tendencia a utilizar sistemas con grandes capacidades de procesamiento que permitan suplir las demandas de las técnicas de inteligencia artificial utilizadas. Sin embargo, dado el panorama que abre el Internet de las Cosas y los sistemas 5G, es necesario enfocar esfuerzos al desarrollo de soluciones acordes con las particularidades de la computación en el borde y las limitadas capacidades de procesamiento de estos sensores.

Finalmente, es importante mencionar que, si bien existen diferentes abordajes desde el punto de vista de la inteligencia artificial para hacer frente a las problemáticas de seguridad ciudadana, llama la atención que dentro de la revisión realizada no fue posible encontrar trabajos que utilizaran métodos estadísticos o de inteligencia computacional para determinar métricas útiles para el monitoreo de los niveles de percepción de la seguridad ciudadana.

Cómo citar este artículo / How to cite this paper

Morales, O. M.; Fletscher Bocanegra, L. A.; Botero Vega, J. F. (2023). La Inteligencia Artificial como apoyo a la gestión de la seguridad ciudadana: un estado del arte. *Revista de Pensamiento Estratégico y Seguridad CISDE*, 8(2), 55-72. (www.cisdejournal.com)

Referencias

- Afandi, W.E.I.B.W.N.; Isa, N.M. (2021). Object Detection: Harmful Weapons Detection using YOLOv4. In 2021 IEEE Symposium on Wireless Technology Applications (ISWTA) (pp. 63-70). doi:10.1109/ISWTA52208.2021.9587423.
- Ahmed, T.; Uppal, M.; Muhammad, A. (2013). Improving efficiency and reliability of gunshot detection systems. In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 513-517). doi:10.1109/ICASSP.2013.6637700.
- Alaqil, R. M.; ... (2020). Automatic Gun Detection From Images Using Faster R-CNN. In 2020 First International Conference of Smart Systems and Emerging Technologies (SMARTTECH) (pp. 149-154). doi:10.1109/SMARTTECH49988.2020.00045.
- Ashraf, A. H.; ... (2021) Weapons Detection for Security and Video Surveillance Using CNN and YOLO-V5s. (https://www.researchgate.net/profile/Muhammad-Khan-730/publication/354871087_Weapons_Detection_for_Security_and_Video_Surveillance_Using_CNN_and_YOLO-V5s/links/6151fef3d2ebba7be7522142/Weapons-Detection-for-Security-and-Video-Surveillance-Using-CNNand-YOLO-V5s.pdf).
- Bajzik, J.; Prinosil, J.; Koniar, D. (2020). Gunshot Detection Using Convolutional Neural Networks. In 2020 24th International Conference Electronics (pp. 1-5). doi:10.1109/IEEECONF49502.2020.9141621.
- BEN BOTKIN, The Statesman Journal (2019). Oregon weighs multiple new restrictions on firearms. (<https://kval.com/news/local/in-oregon-it-isnt-a-question-of-if-firearms-legislation-will-pass-its-which-ones>).
- Caragliu, A.; Del Bo, C.; Nijkamp, P. (2011). Smart Cities in Europe. *Journal of Urban Technology*, 18(2), 65-82.

Morales, O. M.; Fletscher Bocanegra, L. A.; Botero Vega, J. F. (2023). La Inteligencia Artificial como apoyo a la gestión de la seguridad ciudadana: un estado del arte. *Revista de Pensamiento Estratégico y Seguridad CISDE*, 8(2), 55-72.



- doi:10.1080/10630732.2011.601117.
- Company (2020). ShotSpotter. (<https://www.shotspotter.com/company/>).
- Debnath, R.; Bhowmik, M. K. (2020). Automatic Visual Gun Detection Carried by A Moving Person. In 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS) (pp. 208-213). doi:10.1109/ICIIS51140.2020.9342681.
- Desk, T. (2018). CCTV VIDEO: Robber pulls out gun; man in cowboy hat tackles him to the ground. The Indian Express, 26 abril. (<https://indianexpress.com/article/trending/viral-videos-trending/armed-robber-steps-shop-man-cowboy-hat-tackles-him-mexico-5152558/>).
- Dextre, M.; ... (2021). Gun Detection in Real-Time, using YOLOv5 on Jetson AGX Xavier. In 2021 XLVII Latin American Computing Conference (CLEI) (pp. 1-7). doi:10.1109/CLEI53233.2021.9640100.
- Farrelly, C. M. (2018). Key algorithms and statistical models for aspiring data scientists, KDnuggets. (<https://www.kdnuggets.com/2018/04/key-algorithms-statistical-models-aspiring-data-scientists.html>).
- Departamento Nacional de Planeación DNP (2020). Documento de Lineamientos de Política de Ciudades Inteligentes (Borrador). (<https://www.dnp.gov.co/DNPN/Paginas/Lineamientos-de-politica-de-ciudades-inteligentes.aspx>).
- Fansler, K. S. (1998). Description of muzzle blast by modified ideal scaling models. Shock and Vibration Digest, 5(1), 1-12. doi:10.1155/1998/640253.
- Freytag, J. C.; Begault, D. R.; Peltier, C. A. (2006). The acoustics of gunfire. In INTER-NOISE and NOISE-CON Congress and Conference Proceedings. Institute of Noise Control Engineering (pp. 1165-1174).
- Gaidon, A.; ... (2016). Virtual worlds as proxy for multi-object tracking analysis. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4340-4349). (https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Gaidon_Virtual_Worlds_as_CVPR_2016_paper.html).
- Galangque, C. M. J.; Guirnaldo, S. A. (2019). Gunshot Classification and Localization System using Artificial Neural Network (ANN). In 2019 12th International Conference on Information Communication Technology and System (ICTS) (pp. 1-5). doi:10.1109/ICTS.2019.8850937.
- Glomb, P.; ... (2018). Application of hyperspectral imaging and machine learning methods for the detection of gunshot residue patterns. Forensic science international, 290, 227-237. doi:10.1016/j.forsciint.2018.06.040.
- Hashmi, T. S. S.; ... (2021). Application of Deep Learning for Weapons Detection in Surveillance Videos. In 2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2) (pp. 1-6). doi:10.1109/ICoDT252288.2021.9441523.
- Hrabina, M.; Sigmund, M. (2016). Implementation of developed gunshot detection algorithm on TMS320C6713 processor. In 2016 SAI Computing Conference (SAI) (pp. 902-905). doi:10.1109/SAI.2016.7556087.
- Hrabina, M.; Sigmund, M. (2018). Gunshot recognition using low level features in the time domain. In 2018 28th International Conference Radioelektronika (RADIOELEKTRONIKA) (pp. 1-5). doi:10.1109/RADIOELEK.2018.8376372.
- Jain, A.; Aishwarya; Garg, G. (2020). Gun Detection with Model and Type Recognition using Haar Cascade Classifier. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 419-423). doi:10.1109/ICSSIT48917.2020.9214211.
- Jain, H.; ... (2020). Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications. In 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 193-198). doi:10.1109/ICESC48915.2020.9155832.
- Jiang, Z. (2003). Wave dynamic processes induced by a supersonic projectile discharging from a shock tube. Physics of fluids, 15(6), 1665-1675. doi:10.1063/1.1566752.
- Kacprzyk, J.; Pedrycz, W. (eds.) (2015). Springer Handbook of Computational Intelligence. Springer, Berlin, Heidelberg. doi:10.1007/978-3-662-43505-2.
- Kiktova, E.; ... (2015). Gun type recognition from gunshot audio recordings. In 3rd International Workshop on Biometrics and Forensics (IWBF 2015) (pp. 1-6). doi:10.1109/IWBF.2015.7110240.
- Lim, J.; ... (2019). Gun Detection in Surveillance Videos using Deep Neural Networks. In 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (pp. 1998-2002).
- Lindström, A.-C.; ... (2014). Detection of gunshot residues (GSR) on a self-inflicted gunshot wound. Pathology, 46(3), 260-263. doi:10.1097/PAT.0000000000000083.
- McCarthy, J. (2007). What is artificial intelligence?. Stanford University.
- McCoy, R. L. (s.f.). Modern exterior ballistics: the launch and flight dynamics of symmetric projectiles, 1999. Schiffer Publishing Ltd [Preprint]. Schiffer Publishing Ltd.
- Maher, R. C. (2007). Acoustical Characterization of Gunshots. In 2007 IEEE Workshop on Signal Processing Applications for Public Security and Forensics (pp. 1-5). doi:10.1109/IEEECONF12259.2007.4218954.
- Manjarres, W.; Baca, W. (2019). Victimization por crimen, percepción de seguridad y satisfacción con la vida en Colombia. Revista de Economía Institucional, 21(41), 133-160. doi:10.18601/01245996.v21n41.06.
- Ministerio de Justicia de Colombia (2000). Código Penal Colombiano.
- Mitchell, S.; Villa, N.; Stewart-Weeks, M.; Lange, A. (2013). The Internet of Everything for Cities. Connecting People, Process, Data, and Things to Improve the 'Livability' of Cities and Communities. San Jose Cisco. 2013.
- Morehead, A.; ... (2019). Low Cost Gunshot Detection using Deep Learning on the Raspberry Pi. In 2019 IEEE International Conference on Big Data (Big Data) (pp. 3038-3044). doi:10.1109/BigData47090.2019.9006456.
- Narejo, S.; ... (2021). Weapon Detection Using YOLO V3 for Smart Surveillance System. Mathematical Problems in Engineering, 2021. doi:10.1155/2021/9975700.
- Olmos, R.; Tabik, S.; Herrera, F. (2018). Automatic handgun detection alarm in videos using deep learning. Neurocomputing, 275, 66-72. doi:10.1016/j.neucom.2017.05.012.



- Pikrakis, A.; Giannakopoulos, T.; Theodoridis, S. (2008). Gunshot detection in audio streams from movies by means of dynamic programming and Bayesian networks. In 2008 IEEE International Conference on Acoustics, Speech and Signal Processing (pp. 21-24). doi:10.1109/ICASSP.2008.4517536.
- Poole, D.; Mackworth, A. (2017). *Artificial Intelligence: Foundations of Computational Agents*. Cambridge University Press.
- Ren, S.; ... (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28. (<https://proceedings.neurips.cc/paper/2015/hash/14bfa6bb14875e45bba028a21ed38046-Abstract.html>).
- Salazar González, J. L.; ... (2020). Real-time gun detection in CCTV: An open problem. *Neural networks: the official journal of the International Neural Network Society*, 132, 297-308. doi:10.1016/j.neunet.2020.09.013.
- Samireddy, S. R.; Carletta, J.; Lee, K.-S. (2017). An embeddable algorithm for gunshot detection. In 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS) (pp. 68-71). doi:10.1109/MWSCAS.2017.8052862.
- Shiekh, A. A.; Tahir, M.; Uppal, M. (2017). Accurate gunshot detection in urban environments using blind deconvolution. In 2017 International Multi-topic Conference (INMIC) (pp. 1-4). doi:10.1109/INMIC.2017.8289452.
- Smith, G. J. D. (2004). Behind the Screens: Examining Constructions of Deviance and Informal Practices among CCTV Control Room Operators in the UK. *Schweizerische Monatsschrift für Zahnheilkunde = Revue mensuelle suisse d'odonto-stomatologie / SSO*, 2(2/3). doi:10.24908/ss.v2i2/3.3384.
- Soria Romo, R. (2017). El impacto de la inseguridad pública en la competitividad empresarial. Análisis comparativo de las entidades federativas en México. *Economía y Sociedad*, XXI(36), 19-41. (<https://www.redalyc.org/articulo.oa?id=51052064002>).
- Su, K.; Li, J.; Fu, H. (2011). Smart city and the applications. In 2011 International Conference on Electronics, Communications and Control (ICECC) (pp. 1028-1031). doi:10.1109/ICECC.2011.6066743.
- Tabane, E.; Ngwira, S. M.; Zuva, T. (2016). Survey of smart city initiatives towards urbanization. In 2016 International Conference on Advances in Computing and Communication Engineering (ICACCE) (pp. 437-440). doi:10.1109/ICACCE.2016.8073788.
- Verma, G. K.; Dhillon, A. (2017). A Handheld Gun Detection using Faster R-CNN Deep Learning. In *Proceedings of the 7th International Conference on Computer and Communication Technology*. New York, NY, USA: Association for Computing Machinery (ICCT-2017) (pp. 84-88). doi:10.1145/3154979.3154988.
- Wandelt, S.; Bux, M.; Leser, U. (2014). Trends in Genome Compression. *Current bioinformatics*, 9(3). doi:10.2174/1574893609666140516010143.
- Warsi, A.; ... (2019). Gun Detection System Using Yolov3. In 2019 IEEE International Conference on Smart Instrumentation, Measurement and Application (ICSIMA) (pp. 1-4). doi:10.1109/ICSIMA47653.2019.9057329.
- Zhao, F.; ... (2021). Smart city research: A holistic and state-of-the-art literature review. *Cities*, 119, 103406. doi:10.1016/j.cities.2021.103406.

