



**Aproximación al estado del arte de algoritmos estadísticos y técnicas de inteligencia computacional para la gestión de problemas de seguridad ciudadana**

Óscar Mauricio Morales Mejía

Informe de Trabajo de grado para optar al título de Ingeniero de Telecomunicaciones

Tutor

Dr. Juan Felipe Botero Vega, Ph.D. en Ingeniería Telemática

Universidad de Antioquia  
Facultad de Ingeniería, Departamento de Ingeniería Electrónica y de Telecomunicaciones  
Ingeniería de Telecomunicaciones  
Medellín  
2022

<b>Cita</b>	(Morales Mejía, 2022)
<b>Referencia</b>	Morales Mejía, M. (2022). <i>Aproximación al estado del arte de algoritmos estadísticos y técnicas de inteligencia computacional para la gestión de problemas de seguridad ciudadana</i> [Trabajo de grado]. Universidad de Antioquia, Medellín.
<b>Estilo APA 7 (2020)</b>	



Grupo de Investigación de Telecomunicaciones Aplicadas.



Centro de Documentación de Ingeniería - CENDOI

**Repositorio Institucional:** <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - [www.udea.edu.co](http://www.udea.edu.co)

**Rector:** John Jairo Arboleda Céspedes.

**Decano/Director:** Jesús Francisco Vargas, decano de Ingeniería.

**Jefe departamento:** Augusto Enrique Salazar Jiménez.

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

# Índice

<b>Índice</b>	<b>3</b>
<b>1. Resumen</b>	<b>5</b>
<b>2. Introducción</b>	<b>5</b>
<b>3. Objetivos</b>	<b>7</b>
3.1. General	7
3.2. Específicos	8
<b>4. Marco Teórico</b>	<b>8</b>
<b>5. Metodología</b>	<b>12</b>
5.1. Pasos transversales a los objetivos específicos.	12
5.2. Objetivo 1	13
5.3. Objetivo 2	14
5.4. Objetivo 3	14
5.5. Cronograma de Actividades	15
<b>6. Resultados y análisis</b>	<b>15</b>
6.1. Estado del arte	15
6.1.1. Detección de disparos	16
6.1.1.1. Introducción	16
6.1.1.2. Trabajos más relevantes	17
6.1.1.3. Conclusiones	21
6.1.1.4. Trabajo futuro	23
6.1.2. Detección de armas	25
6.1.2.1. Introducción	25
6.1.2.2. Trabajos más relevantes	27
6.1.2.3. Conclusiones	32
6.1.2.4. Trabajo futuro	35
6.1.3. Detección de rostros y reconocimiento facial	36
6.1.3.1. Introducción	36

6.1.3.2. Trabajos más relevantes	38
6.1.3.3. Conclusiones	42
6.1.3.4. Trabajo futuro	44
6.1.4. Prevención de desastres.	45
6.1.3.1. Introducción	45
6.1.3.2. Trabajos más relevantes	46
Detección de incendios	49
Detección de deslizamientos de tierra	50
Predicción de inundaciones	51
6.1.3.3. Conclusiones	52
6.1.3.4. Trabajo futuro	54
6.2. Lista de métodos recomendados	55
<b>7. Otras conclusiones</b>	<b>57</b>
<b>8. Referencias Bibliográficas</b>	<b>58</b>

## 1. Resumen

Cada vez más, existe un mayor interés en los gobiernos de múltiples ciudades de garantizar la seguridad de sus ciudadanos. Uno de los ámbitos utilizados para enfrentar el problema de la seguridad es el de la tecnología. En este trabajo se realizó 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 4 de las aplicaciones más relevantes en el área las cuales son: detección de disparos, detección de armas, detección y reconocimiento de rostros y prevención de desastres. Esta última a su vez se subdividió en detección de incendios, prevención de deslizamiento de tierra y predicción de inundaciones. 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. Además, se organizó una lista de recomendaciones con los algoritmos más utilizados y que mejores resultados han arrojado.

## 2. Introducción

Uno de los principales motivos de preocupación entre la ciudadanía colombiana tiende a centrarse en la seguridad de los cascos urbanos. De acuerdo con la Encuesta de Convivencia y Seguridad Ciudadana realizada por el Departamento Administrativo Nacional de Estadística (DANE), el 15% y 10% de los colombianos mayores de 15 años, durante el 2018 y el 2019, respectivamente, manifestaron haber sido víctimas de al menos un delito. De igual forma, para el año 2018, el 7.2% de los encuestados reportaron haber sido víctimas de hurto personal en al menos una ocasión y el 8.5% de los propietarios de vehículos declararon haber sido objetivos de hurto de su vehículo o partes de él, lo que lleva a un índice de percepción de inseguridad del 39% (DANE, 2021).

Cuando se estudia esta preocupación, se puede encontrar un gran abanico de fenómenos, diferentes entre sí, con el denominador común de su potencial para afectar negativamente a la ciudadanía. Podemos encontrar, por ejemplo, asaltos a transeúntes, robos a 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 y 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).

Los planes de mejoramiento con respecto a este tema deberían ser cubiertos desde aproximaciones de prevención, monitoreo, detección y acción ya que estos muestran 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.

Además de la seguridad ciudadana relativa al cometimiento de delitos e infracciones, que son provocados por el factor humano, también se debe hablar de seguridad frente a desastres naturales y factores naturales.

Si bien la seguridad ciudadana ha sido un aspecto presente en los planes de gobierno de los principales entes del país, solo desde hace pocos años se ha considerado un enfoque de prevención en conjunto con enfoques de acción, uno de estos enfoques siendo la implementación de ciudades inteligentes (Su, Li y Fu, 2011; Departamento Nacional de Planeación DNP, 2020). 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.

Todo lo anteriormente mencionado conduce a la introducción de aplicaciones de seguridad colectiva en proyectos de ciudades inteligentes, lo cual se traduce en múltiples retos en infraestructura y organización. Entre otros muchos elementos, como el manejo de datos y la participación ciudadana, el modelo de ciudades inteligentes depende altamente de los métodos de análisis y clasificación de los datos recolectados. Estos son vitales para dar sentido a la información obtenida, reconocer patrones de comportamiento y realizar predicciones según corresponda.

Sin embargo, para incursionar en la realización de estos proyectos, es necesario y recomendable tener un recuento de los avances hechos hasta el momento por otros investigadores. A pesar de ser un tema relativamente reciente, es posible encontrar

diversidad de proyectos con resultados prometedores abiertos a la comunidad. Esto provee un panorama sobre las bases y enfoques que se han tomado en el área y deja entrever el trabajo que queda por realizar.

Este trabajo busca generar un estado del arte en los métodos de inteligencia computacional y algoritmos estadísticos que hayan sido o puedan ser utilizados en proyectos de ciudades inteligentes para medir y mejorar la seguridad ciudadana de cascos urbanos. 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.

Además de los objetivos trazados para este trabajo, se espera que éste 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 vayan a iniciar sus proyectos. De igual manera, se espera que este trabajo sea un incentivo para la investigación de este campo.

Este trabajo de investigación está pensado para pertenecer al proyecto “Administración inteligente de problemas de seguridad ciudadana a través de modelos y herramientas generadas a partir de plataformas para territorios inteligentes apoyadas por estrategias de participación ciudadana en la ciudad de Medellín” formulado por la Universidad de Antioquia y la Universidad Nacional de Colombia para la zona de eje cafetero y, más específicamente, para ser aplicado a los municipios del Valle de Aburrá.

## **3. Objetivos**

### **3.1. General**

Construir un estado del arte de algoritmos estadísticos y técnicas de inteligencia computacional utilizadas para la gestión inteligente de problemas de seguridad ciudadana.

## 3.2. Específicos

- Reunir las métricas más relevantes para monitorear y mejorar la seguridad ciudadana utilizando métodos estadísticos y de inteligencia computacional.
- Categorizar de manera ordenada los principales algoritmos estadísticos y de inteligencia computacional implementados de forma recurrente y productiva en los últimos años.
- 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.

## 4. Marco Teórico

Se podría decir que el concepto de Ciudad Inteligente, 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 y Fu, 2011) que es el enfoque de este trabajo.

De (Tabane, Ngwira y 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.

### *Definiciones varias de ciudad inteligente:*

- 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 y Villa, sin fecha). Otras características como las descritas en (Caragliu, Del Bo y 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.

#### *Características generales de una Ciudad Inteligente:*

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.

Lo discutido anteriormente, permite consensuar una definición de Ciudad Inteligente que sea relevante para el trabajo y le sirva de guía, la cual se encuentra a continuación.

*Definición del autor de Ciudad Inteligente:*

*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.*

Las ciudades inteligentes requieren recolectar una gran cantidad de información. En el caso de un proyecto enfocado en la seguridad ciudadana, estos datos podrían incluir vídeos 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.

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 y 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).

En (Russell y Norvig, 1995), se toman diversas definiciones de Inteligencia Computacional de diferentes autores y se realiza una clasificación para cualquier sistema que se considere tiene Inteligencia Computacional, las cuales se muestran en la tabla 1.

*Tabla 1. Diagrama de clasificación para sistemas inteligentes.*

Sistemas que piensan como humanos	Sistemas que piensan racionalmente
Sistemas que actúan como humanos	Sistemas que actúan racionalmente

La Inteligencia Computacional comprende diferentes conceptos, paradigmas, algoritmos e implementaciones que muestran un comportamiento en ambientes complejos catalogables como inteligentes (Kacprzyk y 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 y 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).

*Lista de ejemplos de algoritmos estadísticos:*

- 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.

## 5. Metodología

Para el desarrollo del proyecto se propuso un esquema de trabajo que facilitara su ejecución y brindara lineamientos del paso a paso a seguir. Estos pasos son transversales a los tres objetivos específicos enunciados anteriormente o particulares a cada uno. Estos se enumeran, sin estar relacionados estrictamente con su orden de ejecución, a continuación.

### 5.1. Pasos transversales a los objetivos específicos.

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.

*Listado de algunas palabras claves utilizadas para la búsqueda:*

- 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.
- Flood prediction.
- Landslide detection.

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 del autor 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. Esto sirve como muestra de la 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.
4. Revisión: Se realizó una revisión constante por parte del asesor de los avances hechos en el trabajo de investigación para incluir las correcciones pertinentes. Esto se hizo en reuniones virtuales cada tres o dos semanas. En estas reuniones acudieron como invitados los profesores Luis Alejandro Fletscher Bocanegra (Universidad de Antioquia) y Freddy Bolaños Martínez (Universidad Nacional).
5. Presentación final: Finalmente, se prepara la sustentación final del trabajo de investigación y se elaboran los medios necesarios para la misma.

## **5.2. Objetivo 1**

*“Reunir las métricas más relevantes para monitorear y mejorar la seguridad ciudadana utilizando métodos estadísticos y de inteligencia computacional”.*

6. 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.
7. Valoración: Se realizó una revisión de cuales métricas resultan más dicientes e informativas, también como cuáles son sus restricciones y campos de aplicación.

### **5.3. Objetivo 2**

*“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”.*

8. 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.
9. 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 de coherencia y cohesión al escrito.
10. 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 a criterio del autor.

### **5.4. Objetivo 3**

*“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”.*

11. 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.

12. 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.

## 5.5. Cronograma de Actividades

Semana →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Tarea ↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	■	■														
2		■														
3			■	■	■											
4											■	■	■	■	■	■
5																■
6					■											
7						■										
8							■	■								
9									■							
10										■	■	■	■			
11				■	■											
12														■		

## 6. Resultados y análisis

### 6.1. 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 4 aplicaciones principales de las cuales se realizará la aproximación a su estado del arte. Estas aplicaciones fueron escogidas a criterio del autor 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. Estas son: detección de disparos, detección de armas, detección y reconocimiento de rostros y prevención de desastres. 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 una 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.

### **6.1.1. Detección de disparos**

#### **6.1.1.1. Introducción**

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, sin fecha; Fansler, 1998; Jiang, 2003; Freytag, Begault y 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 en la imagen 1, se muestran las partes de un arma para referencia.



*Imagen 1. Partes de un arma. Tomada de (2 - Armas de fuego y municiones, 2013).*

#### 6.1.1.2. Trabajos más relevantes

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. Trabajos más relevantes de detección de disparos:*

Nombre, autores y año	Contexto	Algoritmos utilizados	Resultados
Nombre: Analysis of Linear Predictive Coefficients for Gunshot Detection	Este trabajo busca poder identificar sonidos de disparos entre otro tipos de sonidos para ayudar a la protección de elefantes en África	Se utilizaron Coeficientes de autocorrelación (ACC), Coeficientes de codificación lineal predictiva (LPC), Coeficientes	Se midió el éxito de los algoritmos con el porcentaje de falsas alarmas detectadas. En este campo, ACC obtuvo el menor desempeño con un

<p>Based on Neural Networks. (Hrabina, 2017)</p> <p>Autor: Martin Hrabina.</p> <p>Año: 2017.</p>	<p>central de la cacería ilegal. Su enfoque es evaluar diferentes métodos de extracción de características.</p>	<p>cepstrales de predicción lineales (LPCC) y Coeficientes cepstrales de frecuencia de Mel (MFCC) para la extracción de características y se utilizaron redes neuronales multicapa con una capa escondida para la clasificación.</p>	<p>promedio de error de 59 % seguido por MFCC con un 15%. LPC y LPCC obtuvieron los menos valores de error. Específicamente con 20 coeficientes, se lograron los valores mínimos de 10.67% y 11.62 para LPC y LPCC respectivamente.</p>
<p>Nombre: Gunshot Classification and Localization System using Artificial Neural Network (ANN). (Galangque y Guirnaldo, 2019)</p> <p>Autores: Cherry Mae J Galangque. Sherwin A. Guirnaldo.</p> <p>Año: 2019.</p>	<p>Este trabajo se centra en desarrollar un sistema capaz de clasificar disparos y encontrar la localización aproximada de ellos usando una red de 4 micrófonos. 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 fusil M16.</p>	<p>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.</p>	<p>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 99% 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.</p>
<p>Gunshot detection using Convolutional Neural Networks. (Bajzik, Prinosil y Koniar, 2020)</p>	<p>Este trabajo busca crear un sistema de detección de disparos usando redes neuronales convolucionales sobre representaciones bidimensionales de las grabaciones. No se encuentra enmarcado</p>	<p>Los autores extrajeron el espectrograma, los Coeficientes Cepstrales de Frecuencia de Mel (MFCC) y la matrices de autosemejanza de las grabaciones de entrada para obtener una representación</p>	<p>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</p>

<p>Autores: Jekub Bajzik. Jiri Prinosil. Dusan Koniar.</p> <p>Año: 2020.</p>	<p>en un contexto en particular.</p>	<p>bidimensional que pudiera mostrarse en forma de imágenes. También usaron tres modelos diferentes de redes neuronales convolucionales: VGG16 e Inception V3 con pesos preentrenados y sin entrenar y ResNet18.</p>	<p>reentrenados. También se demostró que bajar la frecuencia de muestreo de 44 kHz a 8 kHz no afecta significativamente la precisión del sistema para detectar disparos, obteniendo una exactitud de 99.07% con ResNet18.</p>
<p>Low-Cost Gunshot Detection using Deep Learning on the Raspberry Pi. (Morehead <i>et al.</i>, 2019)</p> <p>Autores: Alex Morehead. Lauren Ogden. Gabe Magee. Ryan Hosler. Bruce White. George Mohler.</p> <p>Año: 2019.</p>	<p>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.</p>	<p>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 escondidas 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.</p>	<p>Los resultados obtenidos son bastantes 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 98.4% al cambiar el sistema para distinguir entre sonidos de disparos y de pirotecnia.</p>
<p>Implementation of developed gunshot detection algorithm on TMS320C6713 processor. (Hrabina y Sigmund, 2016)</p> <p>Autores:</p>	<p>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 disminuir los eventos de cacería</p>	<p>Las tres técnicas utilizadas son: 1) Coeficientes de codificación predictiva lineal LPC para los cuales se usan 8 coeficientes calculados con el algoritmo Levinson-Dubrin, 2) Detección de energía en bandas de</p>	<p>Se obtiene un porcentaje de detecciones correctas de 82%, un porcentaje de falsos negativos de 18% y un porcentaje de falsos positivos de 3%. Se puede apreciar como la tasa de detecciones erróneas mejoró drásticamente al</p>

<p>Martin Hrabina. Milan Sigmund.</p> <p>Año: 2016.</p>	<p>ilegal de elefantes en África.</p>	<p>frecuencia en la que se utiliza la Transformada rápida de Fourier FFT y 3) Correlación de la señal en el espectro usando la transformada encontrada anteriormente. Estas técnicas se usan en cascada.</p>	<p>utilizar las tres técnicas en cascada. Esto a cambio de un decremento en la exactitud del sistema.</p>
---	---------------------------------------	--	---

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 y Muhammad, 2013; Shiekh, Tahir y 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 y 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 y 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.

### **6.1.1.3. Conclusiones**

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 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 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.

#### ***6.1.1.4. Trabajo futuro***

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 en las imágenes 2 y 3 se muestran los diagramas de un sistema descentralizado con nodos activos y centralizado con nodos pasivos.

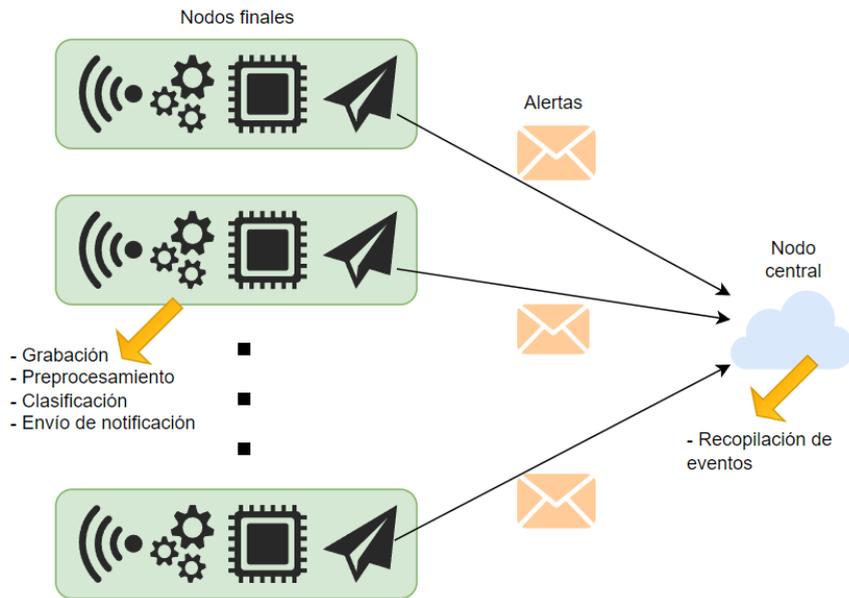


Imagen 2. Arquitectura descentralizada con nodos finales activos.

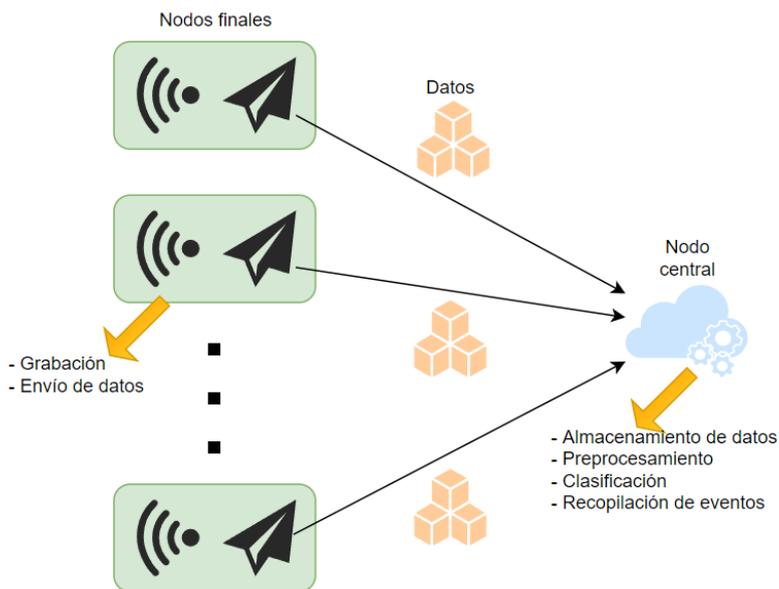


Imagen 3. Arquitectura centralizada con nodos finales pasivos.

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.

## **6.1.2. Detección de armas**

### ***6.1.2.1. Introducción***

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 tienen 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 incapacidades humanas 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 y 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 y 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.

### 6.1.2.2. Trabajos más relevantes

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. 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. Trabajos más relevantes de detección de armas:*

Nombre, autores y año	Contexto	Algoritmos utilizados	Resultados
<p>Nombre: Weapon Detection for Security and Video Surveillance Using CNN and YOLOv5. (Ashraf <i>et al.</i>, 2021)</p> <p>Autores: Abdul Hanan Ashraf. Muhammad Imran. Abdulrahman M. Qahtani. Abdulmajeed Alsufyani. Omar Almutiry. Awais Mahmood. Muhammad Attique. Mohamed Habib.</p> <p>Año: 2021.</p>	<p>Este trabajo busca implementar un sistema de detección de armas presentes en imágenes basado en el algoritmo YOLO de forma que se disminuyan los falsos positivos y falsos negativos mientras se mantiene la velocidad de detección.</p>	<p>Los autores utilizan Difuminación Gaussiana para la eliminación del fondo en las imágenes a analizar y YOLOv5 para la clasificación y Faster R-CNN para comparación.</p>	<p>Se obtuvo una precisión de detección en imágenes del 99.5% y 77.9% en una prueba de video. En comparación, Faster R-CNN obtuvo una precisión menor del 84.2% . Se logra una velocidad de detección de 0.01 segundos en comparación con Faster R-CNN cuya velocidad de detección es de 0.19 segundos.</p>
<p>Nombre:</p>	<p>Este trabajo busca</p>	<p>Los algoritmos</p>	<p>Los autores</p>

<p>Application of Deep Learning for Weapons Detection in Surveillance Videos. (Hashmi <i>et al.</i>, 2021)</p> <p>Autores: Tufail Sajjad Shah Hashmi. Nazeef Ul Haq. Muhammad Moazam Fraz. Muhammad Shahzad.</p> <p>Año: 2021.</p>	<p>implementar y comparar dos sistemas de detección de armas en imágenes con dos versiones de la familia de algoritmos YOLO para confirmar su desempeño y realizar una comparación entre ambos.</p>	<p>utilizados son YOLOv3 y YOLOv4 que son dos versiones del sistema de código abierto YOLO (You Only Look Once) el cual consta de una única red neuronal convolucional de 24 capas convolucionales y 2 capas completamente conectadas.</p>	<p>encontraron, como era esperado, que el sistema YOLOv4 tiene mejor desempeño que YOLOv3, con valores de precisión de 85% y 84% respectivamente. A pesar de que los porcentajes de precisión y exhaustividad son relativamente similares en ambos, siguen siendo superiores para la versión 4 del algoritmo que además requiere menor carga de procesamiento.</p>
<p>Nombre: Automatic Gun Detection From Images Using Faster R-CNN. (Alaqil <i>et al.</i>, 2020)</p> <p>Autores: Rana M. Alaqil. Jaida A. Alsuhaibani. Batool A. Alhumaidi. Raghad A. Alnasser. Rahaf D. Alotaibi. Hafida Benhidour.</p> <p>Año: 2020</p>	<p>Este trabajo desarrolla diferentes sistemas de detección de armas empuñadas 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 muy completa el funcionamiento del algoritmo Faster R-CNN</p>	<p>Se utilizó el algoritmo Faster R-CNN con su arquitectura de redes neuronales convolucionales para extracción de características VGG16 y también con otras arquitecturas tales como: ResNet50, Inception ResNetv2, y MobileNetv2. Adicionalmente, se prueba el algoritmo YOLOv2 sin modificaciones. Estas arquitecturas fueron preentrenadas utilizando el conjunto de datos ImageNet.</p>	<p>Se obtuvo una precisión igual o superior al 70% en todas las arquitecturas siendo la más sobresaliente Inception-ResNetv2 con una precisión del 81%. En comparación, el sistema YOLOv2 obtuvo una precisión del 76%. Los sistemas mostraron capacidad para detectar armas parcialmente ocluidas o fuera de la imagen, con poco contraste con el fondo y desde la vista frontal. Sin embargo, mostró pequeñas falencias en el sistema de ubicación de los cuadros delimitadores.</p>
<p>Nombre: Automatic Visual Gun</p>	<p>Este trabajo desarrolla un sistema de detección de armas</p>	<p>Se utilizó el algoritmo para detección de objetos Illumination</p>	<p>El sistema de detección obtuvo una precisión y</p>

<p>Detection Carried by a Moving Person. (Debnath y Bhowmik, 2020)</p> <p>Autores: Rajib Debnath. Mrinal Kanti Bhowmik.</p> <p>Año: 2020</p>	<p>siendo portadas por personas, utilizando métodos de detección de objetos, sustracción del fondo y Template Matching. La detección de objetos se utiliza para delimitar la región donde se encuentra una persona sosteniendo un arma y poder entregarla al algoritmo de Template Matching para detección del 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.</p>	<p>Sensitive Background Subtraction (ISBS) para detectar personas llevando un arma. Para la misma tarea, también se evaluaron los algoritmos Gaussian Mixture Model (GMM), Multiple Temporal Difference (MTD) y ViBe. 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.</p>	<p>exhaustividad de 94.21% y 95.37% respectivamente y toma 1.5 segundos superando a los otros trabajos relacionados y evaluados en la investigación tanto en desempeño como en rapidez. Los valores más cercanos de precisión, exhaustividad y tiempo de detección obtenidos por otros trabajos evaluados en el artículo son de 92.07%, 93.34% y 2 segundos, teniendo en cuenta que no todos los valores pertenecen al mismo sistema. El desarrollo es capaz de detectar armas en diferentes escalas y rotaciones, pero no fue probado en armas parcialmente ocultas.</p>
--	---	---	---

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 tienden 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 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 a 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 bancos de imágenes. Para esto, tomaron grandes cantidades de fotografías o videgrabaciones 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 y Herrera, 2018; Jain, Aishwarya y 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 como 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 imagen 4, se puede observar un ejemplo comparativo.



*Imagen 4. A la izquierda: Imagen de arma de fuego sin contexto. A la derecha: Imagen de arma con contexto. Tomadas 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.

#### *Algoritmos más utilizados para detección de armas:*

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 y 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 y 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 y 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.

### **6.1.2.3. Conclusiones**

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.

Bastantes 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 y 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 podrían hacer alcanzables estos estándares de calidad de imagen.

De todo 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 aún así podría ser implementado. Finalmente, otras aproximaciones como Haar Cascade y métodos de Ventana Deslizantes obtienen velocidades mayor a 0.2 segundos y quedan rezagados frente a otros algoritmos.

#### **6.1.2.4. Trabajo futuro**

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 una 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.

### **6.1.3. Detección de rostros y reconocimiento facial**

#### **6.1.3.1. Introducción**

El rostro es una de las cosas que más identifica a los seres humanos, constantemente está expuesto, está directamente relacionado a nuestra identidad y es generalmente el principal foco de atención al ver a una persona. Entre otras técnicas de biometría como escaneo del iris o toma de huellas dactilares, la detección de rostros es considerada como la más natural y más utilizada debido a su relativo bajo coste y bajos niveles de intrusividad (Chihaoui *et al.*, 2016). Para el desarrollo de esta sección, es conveniente definir diferentes conceptos relacionados con este tema, basados en (Almadan y Rattani, 2021), con el fin de proveer claridad cuando sean mencionados más adelante.

#### *Definiciones importantes para esta sección:*

- Detección de rostros: Se define como la acción de encontrar y delimitar un rostro humano en un registro fotográfico o videográfico.
- Reconocimiento facial o de rostros: Se define como la acción de corroborar si un rostro pertenece a una persona de la cual ya se tiene un registro fotográfico con el fin de establecer si son la misma persona.
- Identificación de rostros: Se define como la acción de asignar una identidad a la personas cuyo rostro se encuentre en el registro fotográfico y videográfico.

Las técnicas de detección y reconocimiento facial se encuentran en este momento en un estado en el cual cuentan con tal grado de confiabilidad entre la sociedad que se utilizan para diversas aplicaciones presentes en la vida cotidiana tales como: desbloqueo de dispositivos móviles, confirmación de identidad en procesos oficiales, agrupamiento por personas en softwares de galerías de fotos, aplicación de filtros fotográficos, entre otros. Todo esto es una demostración de cómo tanto la detección como el reconocimiento facial son completamente posibles en el momento bajo condiciones suficientes.

En el contexto de programas de mejoramiento de la seguridad en ciudades inteligentes, la detección y reconocimiento de rostros encuentran múltiples aplicaciones, las cuales se listan a continuación, algunas de ellas extraídas de (Praveen y Dakala, 2020).

#### *Aplicaciones de la detección y reconocimiento de rostros:*

- Búsqueda de personas reportadas como desaparecidas.

- Seguimiento y perfilamiento de sospechosos.
- Búsqueda de personas fugitivas de las autoridades.
- Detección de transeúntes infractores.
- Detección de personas en áreas no autorizadas.

Esta sección se centra en la detección y reconocimiento de rostros en videos captados por cámaras de seguridad en espacios públicos, abiertos y cerrados. Además, los trabajos que se desarrollan o fueron publicados en el contexto de las aplicaciones de seguridad y ciudades inteligentes reciben prioridad para pertenecer a esta sección. Esto se hace con el fin de que no se incluyeran otros trabajos de detección y reconocimiento facial cuyas aplicaciones no pueden ser incluidas o extrapoladas al contexto de este trabajo de investigación.

Adicionalmente, se incluyeron y se hizo especial énfasis en los trabajos que utilizaron videos extraídos de cámaras corporales (BWC por sus siglas en inglés: Body-Worn Cameras).

Las cámaras corporales son aquellas que las personas llevan puestas en la ropa o accesorios adicionales como cascos o lentes. Dentro del contexto de seguridad en ciudades inteligentes, estas cámaras son principalmente utilizadas por los cuerpos policiales para la identificación de sospechosos y sus grabaciones podrían ser presentadas para el esclarecimiento de hechos o como pruebas en una corte. De esta forma, se tendrían varias tareas a realizar tales como: la detección y extracción de rostros para la posterior identificación de los individuos y la necesidad de detectar y ocultar los rostros de las otras personas incluidas en las grabaciones en caso de que estos deban convertirse en documentos públicos. En estas tareas, se reconoce la necesidad de automatización al tenerse en consideración la longitud que podrían tener las grabaciones y la cantidad de fuentes de video que tendrían en caso de aplicarse ampliamente entre los miembros de la fuerza pública, lo que da como resultado cantidades muy grandes de datos que representarían un trabajo tedioso y costoso si se quiere tener equipo humano para su tratamiento. Además, es necesario tener desarrollos que se enfoquen específicamente en este tipo de videos debido a los retos técnicos que podrían presentar y que no existen en otras aplicaciones. Estos retos son extraídos de (Brown y Fan, 2016) y se listan a continuación.

#### *Retos a enfrentar en la detección y reconocimiento de rostros con cámaras corporales:*

- Difuminación de la imagen por la captura durante el movimiento corporal.

- Oclusión con el cuerpo del portador o con el ambiente.
- Distorsión en la imagen producida por los lentes cuando hay cercanía de los objetos a los bordes del cuadro o cercanía con la cámara,.
- Condiciones de iluminación pobres o cambiantes.
- Variedad en la presentación de los rostros debido a expresiones faciales, presencia de vello facial, color de piel o el peinado.

### **6.1.3.2. Trabajos más relevantes**

El trabajo encontrado en (Almadan y Rattani, 2021) se considera como relevante para esta aplicación. Sin embargo, no es posible agregarlo a la tabla mostrada más adelante debido a su extensión y alcance. Por esta razón, se decidió que la descripción de este trabajo fuese realizada fuera de la tabla con el fin de abarcarlo de la manera más apropiada. Este artículo desarrolla diferentes sistemas de reconocimiento facial con el propósito de comparar su desempeño. El trabajo fue realizado con el fin de encontrar un método de reconocimiento facial que pueda ser implementado directamente en dispositivos con cámaras corporales, es decir, que el procesamiento se realice en el nodo final y no en un nodo central, por lo que se realiza un enfoque especial en evaluar algoritmos que requieran una menor carga computacional. Por esta razón se dividen los algoritmos en dos conjuntos llamados livianos y pesados que identifican algoritmos con menos y mayor carga computacional en su implementación respectivamente. La ejecución de los algoritmos se realiza usando los celulares iPhone 6, X y XR. Para el conjunto de algoritmos livianos se evaluaron los métodos LightCNN, MobileNet y EfficientNet, mientras que para el conjunto de algoritmos pesados se evaluaron ResNet-50 y VGG-16, los cuales sirven de punto de referencia para el primer grupo. Además, los algoritmos fueron probados usando imágenes capturadas en dos condiciones de luz diferentes. Es decir, cada algoritmo se entrenó dos veces: una vez utilizando únicamente un conjunto de imágenes con condiciones de luz propias de una oficina con luz artificial tenue y otra vez utilizando únicamente un conjunto de imágenes con condiciones de luz correspondientes al exterior en un día soleado. Luego, cada algoritmo fue probado para cada conjunto de condiciones de luz. En general, el artículo evalúa diferentes métodos de reconocimiento facial en tres aspectos: precisión, tiempo de ejecución y resistencia a los cambios en las condiciones de luz de las imágenes. Los resultados obtenidos se encuentran en la tabla a continuación:

Tabla 4. Resultado presentado en (Almadan y Rattani, 2021):

Algoritmo	Tiempo de ejecución promedio [ms]	Condiciones de luz (Entrenamiento - Prueba)	Precisión [%]
LightCNN-9	327	Oficina - Oficina	93.35
		Día - Día	78.70
		Día - Oficina	22.24
		Oficina - Día	22.41
LightCNN-29	820	Oficina - Oficina	<b>94.48</b>
		Día - Día	78.07
		Día - Oficina	22.82
		Oficina - Día	22.93
MobileNet-V2	700	Oficina - Oficina	85.50
		Día - Día	72.43
		Día - Oficina	25.18
		Oficina - Día	22.35
EfficientNet-B0	540	Oficina - Oficina	80.95
		Día - Día	68.95
		Día - Oficina	27.23
		Oficina - Día	27.92
VGG-16	3942	Oficina - Oficina	81.65
		Día - Día	69.49
		Día - Oficina	31.43
		Oficina - Día	27.22
ResNet-50	1387	Oficina - Oficina	96.33
		Día - Día	<b>99.36</b>
		Día - Oficina	84.25
		Oficina - Día	94.45

De la tabla anterior es posible observar varias cosas. En general, el mayor valor de precisión obtenido es logrado por ResNet-50 en condiciones de luz de día tanto para entrenamiento como para prueba. Además, es el único algoritmo cuyos resultados no son significativamente susceptibles a los cambios de iluminación. Por otro lado, entre los algoritmos considerados livianos, el mayor valor de precisión es obtenido por LightCNN-29 en condiciones de luz de oficina para entrenamiento y prueba.

Cabe resaltar que LightCNN-9, cuyos resultados son equiparables a los obtenidos por LightCNN-29, obtiene el menor tiempo de ejecución, y además se toma menos de la mitad del tiempo que toma LightCNN-29 y menos de un cuarto del tiempo de toma ResNet-50. Finalmente, se puede observar que VGG-16 ofrece resultados poco satisfactorios, con valores de precisión menores incluso que los algoritmos livianos en casi todos los escenarios, mostrando susceptibilidad a las condiciones de luz presentes en las imágenes y al mismo tiempo siendo el método que demanda un mayor tiempo de ejecución

A continuación, se encuentran condensados en una tabla los demás 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 5. Trabajos más relevantes de detección y reconocimiento de rostros:*

<b>Nombre, autores y año</b>	<b>Contexto</b>	<b>Algoritmos utilizados</b>	<b>Resultados</b>
Nombre: Enhanced Face Detection Using Body Part Detections for Wearable Cameras. (Brown y Fan, 2016)  Autores: Lisa M. Brown.	Este trabajo realizó la construcción de un conjunto de videos grabados con cámaras corporales con fines de uso para la investigación y de aporte a este campo de investigación. En estos videos, se clasifican los rostros por su dificultad de detección (frontal,	Los autores implementaron dos sistemas para detección facial utilizando Aggregate Channel Features ACF en uno y Faster R-CNN en el otro y además probaron el algoritmo de Viola y Jones (Haar Cascade) (Viola y Jones, 2001) para tomarlo como	Los resultados obtenidos muestran que el sistema con mayor éxito por gran diferencia es Faster R-CNN con una precisión de detección para rostros frontales, de perfil, y con clasificación difícil de 96%, 85% y 30% respectivamente, seguido por ACF que

<p>Quanfu Fan. Año: 2016.</p>	<p>perfil, difíciles). También se implementaron diferentes sistemas de detección facial con enfoque adicional en utilizar la detección de otras partes del cuerpo de forma complementaria con el fin de mejorar la precisión del sistema en general.</p>	<p>referencia. Adicionalmente se utilizó Faster R-CNN para la detección de partes del cuerpo cercanas al rostro con el fin de mejorar la detección del rostro del individuo.</p>	<p>obtuvo valores de 49%, 30%, y 3% para una tasa similar de falsos positivos. Con la adición de la detección corporal, se incrementó la precisión casi un 2% en ambos sistemas.</p>
<p>Nombre: Person Identification by Smart Cameras. (Rothkrantz, 2017)  Autor: Leon Rothkrantz.  Año: 2017.</p>	<p>Este artículo propone un sistema de reconocimiento facial en un sistema de cámaras por circuito cerrado con aplicaciones en ciudades inteligentes para la detección de terroristas. Los autores utilizaron la base de datos FERET que contiene en su mayoría imágenes de rostros frontales.</p>	<p>Los autores utilizaron una reimplementación del algoritmo Haar Cascade propuesto por Viola y Jones (Viola y Jones, 2001) para la detección de rostros, el algoritmo PCA para la posterior extracción de características y máquinas de soporte vectorial usando diferentes versiones de los kernels de Gauss y Laplace para el reconocimiento.</p>	<p>Se obtuvo una tasa de detecciones correctas de 93% y de falsos positivos de 19% al realizar el análisis con 19 características PCA. Se encontró que a medida que se utilizan más características, disminuye la tasa de falsos positivos, pero también disminuye en menor medida la tasa de detecciones correctas. Las versiones más recientes de los kernels (desde Gauss 5.0 y Laplace 3.0) obtuvieron resultados similarmente altos, siendo los mejores Gauss 7.0 y Laplace 6.0.</p>
<p>BWCFace: Open-set Face Recognition using Body-Worn Cameras. (Almadan, Krishnan y</p>	<p>Este artículo describe la creación de una base de datos llamada BWCFace que consta de imágenes con rostros tomadas con cámaras corporales y la implementación de</p>	<p>Los autores utilizan redes neuronales convolucionales para la extracción de características con arquitectura ResNet-50 y SE-ResNet-50 y</p>	<p>En promedio, se obtuvieron valores de precisión de 98.24% y 96.55% para ResNet-50 y SE-ResNet-50 respectivamente. Aunque los valores</p>

<p>Rattani, 2020)</p> <p>Autores: Ail Almadan. Anoop Krisnan. Ajita Rattani.</p> <p>Año: 2020.</p>	<p>un sistema de detección de rostros, el cual se entrena y evalúa con esta base de datos. Esto con el fin de implementarlo en cuerpos policiales para la detección de sospechosos y esclarecimiento de hechos.</p>	<p>funciones de pérdida como Softmax, ArcFace, CosFace, SphereFace y AdaCos para identificación.</p>	<p>son cercanos, la mayor métrica la obtuvo ResNet-50 usando SoftMax siendo su precisión de 98.99%. SoftMax fue la función de pérdida con mejores métricas alcanzando un promedio de precisión de 97.86%.</p>
--	---	--	---

Existen otros trabajos en el área que también utilizan los clasificadores de tipo Haar Cascade tales como (Othman y Aydin, 2018). En este trabajo desarrollan un sistema de detección de rostros utilizando una RaspBerry Pi, OpenCV y Python para su implementación, que además envía notificaciones vía Telegram. Al sistema se le adiciona una etapa en la que se realiza la correlación de las zonas donde se espera que estén ubicadas diferentes partes del rostro con un modelo de rostro predefinido por los autores, con el fin de disminuir los falsos negativos lo máximo posible. La precisión de este sistema se encuentra alrededor del 86% y se observa una disminución de las detecciones erróneas en comparación con el sistema original. Otros ejemplos de trabajos que utilizan este clasificador son (Mantoro, Ayu y Suhendi, 2018; Hoque *et al.*, 2020; Safiullina, Gabdullin y Anikin, 2021).

En (Qezavati, Majidi y Manzuri, 2019), los autores utilizan también clasificadores de tipo Haar Cascade en combinación con Locally Binary Patterns Histograms (LBPH) para extracción de características y una máquina de soporte vectorial para clasificación. Una particularidad de este proyecto es que se centra en detección de rostros incluyendo aquellos que se encuentren parcialmente ocultos con el fin de poderse implementar en regiones del oriente medio donde las personas portan bufandas o hijabs en la cabeza.

### 6.1.3.3. Conclusiones

Como se puede deducir observando los algoritmos utilizados en los trabajos expuestos en esta sección y como es recomendado en (Praveen y Dakala, 2020), al igual que en la sección anterior, una de las principales herramientas a utilizar o incluir en desarrollos de detección de rostros en imágenes son las redes neuronales convolucionales. Estas vienen acompañadas de o incluidas en otros componentes y forman los sistemas más precisos,

robustos, y eficientes hasta la fecha tanto en detección como en reconocimiento facial. Entre los más destacados se encuentran Faster R-CNN y ResNet-50.

Sin embargo, en esta aplicación logran amplio protagonismo los clasificadores tipo Haar Cascade (Viola y Jones, 2001). A pesar de ser un algoritmo desarrollado hace más de dos décadas, el algoritmo propuesto por Viola and Jones para detección de objetos aún sigue siendo relevante en el ámbito investigativo. Si bien es cierto que el algoritmo original ha quedado rezagado en términos de rendimiento frente a desarrollos posteriores, se puede encontrar como referente en la gran mayoría de trabajos realizados en detección y reconocimiento facial, además de ser base para otras implementaciones. Entre las modificaciones que se le han hecho a este algoritmo se encuentra la adición de sistemas previos de extracción de características y remapeo de la imagen y complementos a la etapa de clasificación con filtros de correlación o máquinas de soporte vectorial.

Curiosamente, aparte de (Aung, Bobkov y Tun, 2021), no se encontraron otros trabajos que utilizaran la familia de algoritmos YOLO para esta aplicación, a pesar de su éxito como método para la detección de objetos en general. Esto se atribuye a que las limitaciones de este método con respecto al tamaño y cantidad de los objetos a detectar no lo hacen apropiado para esta aplicación, teniendo en cuenta que en cámaras de vigilancia en espacios públicos, se espera tener una gran cantidad de rostros que ocupen poco espacio en el cuadro. El trabajo citado anteriormente obtiene una precisión del 93% pero su base de datos consiste en imágenes con una o un número muy reducido de personas situadas en primer plano con suficiente iluminación.

Continuando con la comparación con la sección anterior, también se pueden evidenciar los retos que enfrentan los investigadores, algunos propios del análisis de imágenes y otros más específicos. (Brown y Fan, 2016) ejemplifica cómo un mismo desarrollo de detección de rostros puede obtener valores aceptables para este tipo de aplicaciones utilizando un conjunto de datos genérico conformado por imágenes más cercanas al tipo retrato y obtener valores bastante más bajos para un conjunto de datos conformado por imágenes con un encuadre menos planeado y capturadas usando cámaras corporales, pasando de una precisión de alrededor del 90% a una de alrededor del 35%. En (Almadan y Rattani, 2021), también se ejemplifica cómo los cambios en las condiciones lumínicas pueden afectar el desempeño de un sistema. Esto es especialmente importante ya que se espera que los

sistemas de detección y reconocimiento facial ubicados en zonas urbanas sean capaces de cumplir su función de forma satisfactoria tanto de día como de noche.

Finalmente, es necesario mencionar como la cantidad y variedad de bases de datos que consistan de videos de vigilancia de espacios públicos, que se presten para detección de rostros y estén disponibles de forma abierta para la investigación académica son bastante reducidas. Más aún, si se requiere que estas cámaras sean corporales. Por esta razón, diversos trabajos como (Brown y Fan, 2016; Almadan, Krishnan y Rattani, 2020) se toman la tarea de desarrollar sus propios conjuntos de datos. Esto, les permite estar en completo control de la forma y contenido de las videograbaciones tal que se ajusten a las necesidades del proyecto.

#### ***6.1.3.4. Trabajo futuro***

Debido a los resultados adquiridos, podría ser enriquecedor para esta área de investigación seguir impulsando la detección de rostros apoyada en métodos complementarios como detección de partes del cuerpo. Incluir este tipo de sistemas podría ayudar a mejorar los niveles de precisión en la detección de rostros, especialmente en casos de difícil detección como los casos en los que la persona se encuentra de perfil o muy lejos de la cámara. Sin embargo, se debe evaluar si la mejora en los niveles de detección son suficientes para justificar la carga de procesamiento extra que esto conlleva. Un ejemplo de esto se puede encontrar en la imagen 4.



*Imagen 4. Ejemplo de detección de rostros con detección complementaria. Extraído de (Brown y Fan, 2016).*

Por otro lado, puede ser necesario extender la detección y reconocimiento de rostros a aquellos que porten mascarillas, debido a la situación de salud pública desatada por el Covid-19, como se hace en (Bharathi *et al.*, 2021). Esto daría mayor alcance a los sistemas de detección y se podría extender a personas que lleven lentes o gafas oscuras, bufandas o gorros. Esto implica que los algoritmos deberán ser más precisos ya que trabajarían solo con una porción de la información que aporta el rostro.

#### **6.1.4. Prevención de desastres.**

##### **6.1.3.1. Introducción**

La seguridad ciudadana no solo debe ocuparse de las amenazas producidas directamente por las personas, como en las aplicaciones revisadas anteriormente, sino que también debe tener en cuenta las amenazas que vienen del entorno, como por ejemplo de la naturaleza. Por esa razón, esta sección se dedica a explorar el estado del arte de los algoritmos estadísticos y técnicas de inteligencia computacional utilizadas para la prevención y el manejo de desastres naturales. Son de especial interés los trabajos que se desarrollen en tres frentes: detección de incendios, predicción de deslizamiento de tierra, y predicción de inundaciones.

La detección temprana de incendios es un problema que debe concernir a las autoridades por diferentes razones, especialmente en ciudades como Medellín que está rodeada por montañas y cuenta con diferentes elevaciones de terreno con vegetación dentro de su área urbana. Además de las pérdidas humanas y materiales que puede causar un incendio, entre las razones aludidas anteriormente se encuentra la pérdida de la vegetación y la fauna que existe en la zona afectada por el incendio forestal. Si bien existen ecosistemas que tienen la capacidad de sobreponerse después de un incendio forestal, existe otra cantidad de ellos que no lo hacen y sufren posteriores problemas como la erosión, especialmente si el terreno no es propenso a incendios constantes (De las Heras *et al.*, 2013). Adicionalmente, se ha demostrado que los incendios contribuyen a una mala calidad del aire debido a la liberación de material particulado (Chacón Rivera, 2015), lo cual es también un tema de relevancia en ciudades como Medellín (Bedoya y Martínez, 2009).

Por otro lado, cuando existen deslizamientos o derrumbes de tierra en territorios habitados o con infraestructura, se genera un peligro para la población y una serie de pérdidas materiales. Si bien lo recomendable e ideal es no construir ni habitar zonas que tengan riesgo de deslizamiento o inestabilidad geológica, existen casos donde esto es inevitable debido a la ocupación invasiva de poblaciones marginadas o por la importancia de la zona para el acceso y comunicación terrestres de municipios y ciudades, como aquellas ubicadas en zonas con morfología montañosa. En Medellín, este tema ha demostrado ser problemático debido a que se encuentra rodeada de cordilleras, donde seguidamente se presentan derrumbes de tierra que afectan a las personas que habitan en la periferia y además obstaculizan las vías de acceso a la ciudad (Serna Quintana, 2011).

De igual manera, la predicción de inundaciones es una aplicación que puede considerarse de interés para la seguridad ciudadana. Si bien predecir o anticipar una inundación puede no proveer la información o los medios para detenerla o evitarla, permite generar alertas tempranas para la preparación de la zona y la disminución de pérdidas humanas y materiales en áreas propensas a las inundaciones como zonas aledañas a ríos o lagos o zonas con determinados patrones de lluvia.

Se debe tener en cuenta que los fenómenos naturales son difíciles de predecir debido a su naturaleza aleatoria y a la gran cantidad de variables que intervienen, muchas de ellas imposibles de medir. Por esto, es de esperarse que muchos de los desarrollos relacionados

a inundaciones y deslizamientos de tierra no obtengan métricas de resultados como se ha visto en otras secciones.

### 6.1.3.2. Trabajos más relevantes

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 las tres diferentes aplicaciones que contiene esta sección.

*Tabla 6. Trabajos más relevantes de prevención de desastres:*

Nombre, autores y año	Contexto	Algoritmos utilizados	Resultados
<p>Nombre: Efficient Deep CNN-Based Fire Detection and Localization in Video Surveillance Applications. (Muhammad <i>et al.</i>, 2019)</p> <p>Autores: Khan Muhammad. Jamil Ahmad. Zhihan Lv. Paolo Bellavista. Po Yang. Sung Wook Baik.</p> <p>Año: 2018</p>	<p>Este trabajo propone un sistema que requiera poca carga computacional para la detección y segmentación de fuego en videos de vigilancia tanto en entornos domésticos, como urbanos y forestales utilizando una RaspBerry Pi.</p>	<p>Se evaluaron dos arquitecturas de redes neuronales convolucionales: la primera es una versión modificada de SqueezeNet y la segunda es AlexNet. La modificación a SqueezeNet se realiza para disminuir su complejidad disminuyendo la cantidad de neuronas utilizadas en las diferentes capas.</p>	<p>Se realizaron las pruebas con dos conjuntos de datos diferentes: con la primera se obtuvieron valores de precisión de 94.5% y 94.4% para SqueezeNet modificado y AlexNet respectivamente mientras que con la segunda se obtienen valores de 86% y 82%. La diferencia en la base de datos radica en el tamaño del fuego que cada una registra.</p>
<p>Nombre: Forest Fire Detection Using Combined Architecture of Separable Convolution and Image Processing.</p>	<p>Este artículo propone un sistema de detección de incendios forestales para los bosques del</p>	<p>Los autores utilizan dos técnicas en paralelo: la primera es un módulo de procesamiento digital de la imagen que emplea Multichannel Binary Thresholding y Filtros</p>	<p>El modelo combinado es capaz de detectar si existe fuego o no es una imagen con una precisión del 93.6% y</p>

<p>(Dutta y Ghosh, 2021)</p> <p>Autores: Sreejata Dutta. Soham Ghosh.</p> <p>Año: 2021</p>	<p>norte de Arizona utilizando imágenes aéreas con el fin de dar alerta temprana a las autoridades y evitar la expansión del fuego. Se hace uso de la base de datos FLAMES (Fire Luminosity Airborne-based Machine learning Evaluation).</p>	<p>HSV para obtener una imagen binaria que encuentre la región de la imagen donde están presentes las llamas y la segunda es una red neuronal convolucional de dos capas con función de activación ReLu que realiza la clasificación con base en los datos de entrenamiento.</p>	<p>obteniendo una tasa de falsos negativos inferior al 2%.</p>
<p>Comparison of Different Machine Learning Models for Landslide Susceptibility Mapping. (Yi <i>et al.</i>, 2019)</p> <p>Autores: Yaning Yi. Zhijie Zhang. Wanchang Zhang. Chi Xu.</p> <p>Año: 2019.</p>	<p>Este artículo se realiza con el propósito de comparar diferentes modelos de Machine Learning para el mapeo de la susceptibilidad de un terreno situado en una provincia en China propensa a presentar deslizamientos basados en los datos históricos.</p>	<p>Se implementaron tres métodos: Regresión Logística, el clasificador estadístico Naïve Bayes y una Máquina de Soporte Vectorial junto con el kernel de tipo RBF.</p>	<p>La mejor tasa de éxito en la predicción la obtuvo Regresión Lineal con un valor de 86%, seguido de Naïve Bayes con 85.69, y en último lugar está la Máquina de Soporte Vectorial con un valor de 81.84%.</p>
<p>Nombre: New Method of Artificial Intelligence for Disaster Information Floods use Distributed Wireless Sensors. (Wahyono <i>et al.</i>, 2019)</p> <p>Autores: Irawan Dwi Wahyono. Khoirudin Asfani. Aripriharta.</p>	<p>En este artículo, se estudia la implementación de un sistema que puede predecir la probabilidad de que haya una inundación en zonas urbanas de Indonesia debido a ríos cercanos utilizando una red de sensores instalados en</p>	<p>Se utilizó una combinación de KNN y Naive Bayes para esto. Los datos son pasados en paralelo por KNN que separa los datos en tres grupos que representan "Inundación", "Precaución" y "No inundación" y por Naïve Bayes para calcular la precisión de predicción del río.</p>	<p>Los autores consiguen una precisión de 86.7% al probar únicamente KNN, una precisión de 68,4% al probar únicamente Naive Bayes, y una precisión de 93,4% al usar la combinación de los dos algoritmos, como es propuesto</p>

Irham Fadlika. Gwo Jia Jong. Año: 2019.	ellos que midan el nivel del río e información pluvial.		inicialmente.
Nombre: 3 Hours Ahead of Time Flood Water Level Prediction Using NNARX Structure: Case Study Pahang. (Rohaimi, Ruslan y Adnan, 2016)  Autores: Noor Ashikin Rohaimi. Fazlina Ahmat Ruslan. Ramli Adnan.  Año: 2016.	El artículo propone un modelo de predicción del nivel del agua con anticipación de 3 horas en una determinada área de Malasia cercana a diferentes ríos con el fin de poder evacuar a sus habitantes antes de una inundación.	El método propuesto es Neural Network Autoregressive with Exogenous Input (NNARX), el cual consiste en una red neuronal artificial recurrente y dinámica que predecirá los valores siguientes del nivel del agua con base en el histórico reciente de mediciones.	Los autores comparan las gráficas de nivel de agua versus tiempo de los datos reales y del modelo simulado y encuentra una similitud del 73.35%, siendo la raíz del error cuadrático medio entre las señales de 0.1087 metros.

### Detección de incendios

En general, la forma de realizar detección de incendios que se encuentra en la literatura revisada consta del análisis de imágenes, normalmente tomadas desde el aire. Esto trae consigo un problema relacionado con el humo producido por incendio, que se posiciona por encima del fuego y disminuye la visibilidad. Por esto desarrollos como (Lee *et al.*, 2017; Naimeng, Wanjun y Xiaoyu, 2021) han optado por incluirlo como un factor a detectar de forma complementaria. Por otro lado, diversos estudios muestran que el entrenamiento de las redes neuronales convolucionales debe ser diferente según la topología del lugar donde se quiera realizar la detección de los incendios (Ghosh y Dutta, 2020). Esto significa que un sistema que obtenga un desempeño positivo y haya sido entrenado para cierta área, podría verse afectado al enfrentarse a información de un área diferente.

Otros trabajos realizados para la detección de incendios también utilizan redes neuronales convolucionales en sus implementaciones. En (Yaloveha *et al.*, 2019), se realizó la comparación entre un sistema que utiliza Redes Neuronales Convolucionales y otro utilizando Redes Neuronales de Cápsula. Esta comparación mostró un mejor resultado para

el primer método con una precisión máxima de 99.2% frente al segundo método con una precisión también aceptable de 95.2%. Otras arquitecturas de redes neuronales convolucionales tales como: AlexNet, GoogLeNet, VGG13, GoogLeNet modificado y VGG13 modificado son evaluadas en (Lee *et al.*, 2017). En este caso, a excepción de VGG13, que obtiene una precisión de 86.2%, con todas las arquitecturas se obtienen valores de precisión en la detección mayores a 94%. El mayor valor de precisión obtenido es 99% usando GoogLeNet. Sin embargo, todas las arquitecturas mencionadas toman entre 7 y 12 segundos para completar el análisis de una imagen. Más ejemplos de la utilización de redes neuronales convolucionales para la detección de fuego pueden ser encontradas en (Zhang, Wang y Liu, 2019) y en (Shamsoshoara *et al.*, 2021) (en el cual además se construye la base de datos FLAME).

#### Detección de deslizamientos de tierra

Dentro del análisis para la prevención de deslizamientos de tierra se pudieron identificar dos aproximaciones. La primera aproximación es bastante popular y consta de realizar mapas del terreno que indiquen la probabilidad que tiene cada zona de sufrir un deslizamiento de tierra, llamada susceptibilidad. Para esto, utilizan diferentes fuentes de información como los registros históricos de los deslizamientos habidos en la zona como en (Yi *et al.*, 2019) o características intrínsecas de la zona. Por ejemplo, (Pradhan y Lee, 2010) presenta un análisis de la susceptibilidad de deslizamientos de tierras en una zona de Malasia. Para esto, se parametrizaron 11 características del terreno tales como: ángulo de inclinación, curvatura, altitud, distancia a carreteras, distancia a ríos, tipo de suelo, entre otros. Se implementó una red neuronal basada en el Perceptrón Multicapa la cual fue entrenada usando el algoritmo Back Propagation. De las 11 características mencionadas anteriormente, se dejaron solo 7, con las cuales se obtuvieron los mejores resultados siendo la precisión del sistema del 94%. Otros ejemplos de trabajos que implementan sistemas para producir mapas de susceptibilidad de un terreno se pueden encontrar en (Utomo, Hu y Hsiung, 2020) que utiliza Redes Neuronales tipo Long Short-Term Memory (LSTM) o en (Bai *et al.*, 2008; Xu *et al.*, 2019), los cuales utilizan regresión logística con resultados similares.

La segunda aproximación tiene menos popularidad se basa en la obtención de mediciones continuas de diferentes variables propias del terreno tales como la humedad, la temperatura o la presión, con las cuales se intenta encontrar una relación entre estas variables y los deslizamientos que permita anticiparlos. Este es el caso por ejemplo de (Lee *et al.*, 2020) que utiliza Regresión de Soporte Vectorial (una variante para predicción de las Máquinas de

Soporte Vectorial) y de (Munasinghe y Karunanayake, 2021) que utiliza Eliminación Recursiva de Características (RFE).

De acuerdo con la información presentada en (Santos *et al.*, 2020), es posible crear una lista de los algoritmos más usados en esta aproximación entre 2015 y 2020, la cual es mostrada a continuación.

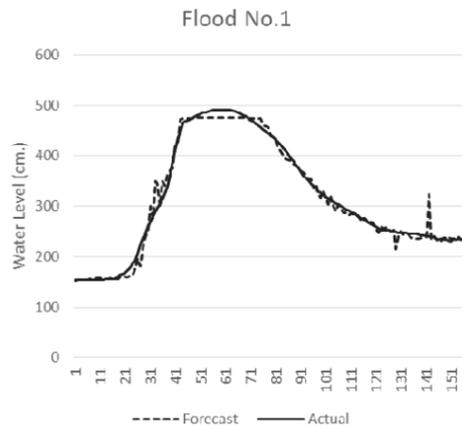
#### *Lista de algoritmos más utilizados para la predicción de deslizamientos de tierra:*

1. Máquinas de soporte vectorial (24%).
2. Redes neuronales artificiales (16%).
3. Random Forest (16%).
4. Regresión logística (12%).
5. Otros (32%).

#### **Predicción de inundaciones**

Además de en (Rohaimi, Ruslan y Adnan, 2016), mostrado anteriormente en la tabla 6, otros trabajos como (Adnan, Samad y Ruslan, 2016; Lukman, Ruslan y Adnan, 2016; Ruslan *et al.*, 2016; Ruslan, Samad y Adnan, 2016, 2018) utilizan el algoritmo NNARX para la predicción del comportamiento del nivel del agua en diferentes cuerpos acuíferos cercanos a una población. Cada artículo trabaja un caso de estudio diferente dependiendo de la estructura hidrológica del terreno pero conservando la esencia del método de predicción. NNARX consiste en una red neuronal artificial basada en el modelo del Perceptrón Multicapa con un regresor en la etapa de entrada de datos y se utiliza principalmente para la predicción de series de tiempo no lineales. Es importante mencionar, que los trabajos aludidos fueron elaborados con la colaboración recurrente del mismo grupo de investigadores.

El Perceptrón Multicapa es nuevamente utilizado también para la predicción del nivel del agua en: (Nikhil Binoy *et al.*, 2019) donde se obtiene una precisión máxima de 78% y se muestra como el aumento excesivo en el número de capas oculta y neuronas es contraproducente en términos de desempeño; y en (Phitakwinai, Auephanwiriyaikul y Theera-Umpon, 2016), donde se obtiene una precisión en el nivel del agua de 6.8 cms. A continuación se encuentra una gráfica ejemplo de este último trabajo donde se muestra el comportamiento del nivel de agua real versus el predicho por el sistema.



*Imagen 5. Gráfico del nivel de agua predicho versus el real en función del tiempo. Tomada de (Phitakwinai, Auephanwiryakul y Theera-Umpon, 2016).*

En (Nunavath y Goodwin, 2019), se puede encontrar una amplia revisión sistemática de la literatura en el campo del manejo de desastres usando técnicas de inteligencia artificial. En este, se lista una cantidad considerable de trabajos realizados junto con su objetivo, conjunto de datos utilizados y los métodos de inteligencia artificial o aprendizaje de máquina que proponen. La revisión concluye con una lista de los algoritmos más utilizados, estos se encuentran enunciados a continuación.

*Lista de algoritmos más utilizados para prevención de desastres:*

- Máquina de soporte vectorial.
- Naive Bayes.
- Random Forest.
- Redes neuronales convolucionales.
- Redes neuronales artificiales.
- K-Nearest Neighbors (KNN).
- Regresión logística.

**6.1.3.3. Conclusiones**

Es curioso observar que los trabajos que realizan análisis de imágenes para la detección de incendios no mencionan o no enfatizan en problemas comunes en el análisis de imágenes como difuminación de la imagen producida por movimiento, distorsión del lente o cambios en la iluminación, lo cual contrasta fuertemente con las aplicaciones vistas en secciones anteriores que realizaban análisis de imágenes y se puede atribuir a las características

propias del fuego. El fuego no tiene una forma definida, por lo que sigue siendo reconocible en imágenes borrosas o distorsionadas y además produce su propia iluminación, por lo que es detectable incluso en ambientes oscuros. El principal reto en varios desarrollos se encuentra en la oclusión producida por los objetos en la imagen y por el humo. Sin embargo, este último puede convertirse en una herramienta para mejorar la precisión de los sistemas de detección, aceptando la hipótesis de que la presencia de humo está relacionada a la presencia de fuego. En la imagen a continuación pueden verse fenómenos de difuminación en la imagen y oclusión con humo.



*Imagen 6. Imágen satelital de un incendio. Tomada de (Priya y Vani, 2019).*

Como era de esperarse, la técnica más utilizada para detección de fuego es principalmente las redes neuronales convolucionales. Esto tiene sentido considerando que ha sido empleado en otras aplicaciones mencionadas en este trabajo que también requieren análisis de imágenes. Estos métodos han mostrado obtener niveles de desempeño muy altos en términos de precisión en la detección a pesar de no ser aptos hasta ahora para un análisis en vivo cuadro a cuadro. Sin embargo, esto no ha sido impedimento para que sean implementados múltiples veces, lo cual se puede atribuir a las necesidades propias de esta aplicación. Las redes neuronales convolucionales pueden alcanzar un ritmo de detección lo suficientemente alto como para realizar detecciones tan cercanas temporalmente una de la otra que el fuego no cambia significativamente su tamaño entre un análisis y otro.

Por otro lado, las aplicaciones que requieren tomar los datos de una red de sensores, se encuentran con una dificultad en muchos casos debido a que es necesario tener un historial de datos recolectados de antemano por estos sensores durante un periodo suficientemente

largo de tiempo tal que sea posible conformar el conjunto de datos de entrenamiento. En muchos casos las redes de sensores no han sido instaladas y los investigadores tienen que empezar por la instalación y recolección de datos, o las redes de sensores instaladas tienen falencias que dificultan el desarrollo. Esto es especialmente importante en el tipo de aplicaciones relativas al nivel de ríos, estabilidad de una montaña, actividad sísmica, entre otros, ya que generalmente dos entornos del mismo tipo no exhiben necesariamente el mismo comportamiento y los algoritmos deben ser entrenados con la información del lugar en donde se realizará la implementación. Este es el caso de (Wahyono *et al.*, 2019), donde los sensores están ubicados de forma no homogénea y no estratégica y por lo tanto no recolectan la versión óptima de la información.

Por lo discutido anteriormente, se puede observar como muchos de los artículos que recogen información de cierta área, no solo mencionan el área con bastante exactitud sino que también ponen énfasis en las condiciones ambientales generales de ella.

En el campo de la detección de la probabilidad de deslizamiento en un área o terreno, los métodos más utilizados son Regresión Logística, Máquinas de Soporte Vectorial y Redes Neuronales (Yi *et al.*, 2019). La aproximación más implementada, en lugar de ser aquella en la que se predice si habrá o no un deslizamiento de tierra basado en información actual del terreno, es aquella que asigna una probabilidad de deslizamiento a cada parte del terreno creando así un mapa de susceptibilidad basados en los datos históricos y características intrínsecas del lugar.

Esta aproximación es muy útil para determinar las zonas sobre las cuales es más o menos recomendable construir infraestructuras, por lo que se puede aplicar en la planeación de carreteras, puentes o nuevas zonas urbanas. Sin embargo, no es de mucha utilidad para predecir posibles deslizamientos en zonas sobre las cuales ya existe infraestructura o donde ya hay personas habitando.

La principal aproximación en los métodos de predicción de inundaciones consiste en crear una serie de tiempo con los valores del nivel del agua en el cuerpo acuífero a observar y realizar predicciones sobre este para determinar si representa un peligro o no en las siguientes horas. Con este fin, el método más utilizado es el Perceptrón Multicapa utilizando el algoritmo Back Propagation para su entrenamiento. Existen también unas cuantas excepciones de trabajos que utilizan una aproximación diferente que consiste en analizar los datos sobre el nivel del agua más reciente para generar un resultado binario que indica si

hay riesgo de inundación o no. Estos trabajos utilizan principalmente KNN y Máquinas de soporte para la clasificación de los datos.

#### **6.1.3.4. Trabajo futuro**

Con respecto a la detección de incendios, aún se deben desarrollar más sistemas que puedan alertar sobre un incendio aún cuando se esté captando únicamente la presencia de humo sin fuego aparente. Esto permitiría poder detectar fuegos de manera aún más temprana.

También, es necesario ahondar en la investigación sobre la predicción de deslizamientos de tierra en zonas que ya se encuentren habitadas con el fin de dar alerta temprana de evacuación. Para esto, podría ser útil tener en cuentas diferentes variables del terreno tales como: humedad, temperatura, actividad sísmica, presión subterránea o vibración. Estas variables podrían ser recolectadas y analizadas en tiempo real utilizando una red de sensores distribuida sobre las zonas de riesgo de deslizamiento.

De igual manera, para la predicción de inundaciones, podría ser conveniente tener en cuenta y evaluar otras variables de los cuerpos de agua que puedan estar relacionadas con los cambios en su nivel tales como: caudal, velocidad del cauce, temperatura, turbidez o dirección de la corriente. Nuevamente, estas variables podrían ser recolectadas y analizadas en tiempo real utilizando una red de sensores distribuida sobre el cuerpo de agua que tenga riesgo de provocar inundaciones.

## **6.2. Lista de métodos recomendados**

En la tabla a continuación, se encuentran condensados todos las técnicas que se mencionaron en secciones anteriores los cuales, a criterio del autor y basados en los resultados obtenidos en los artículos expuestos, se recomiendan para las distintas aplicaciones mencionadas. En dicha tabla, se expone el nombre de la técnica junto con la o las aplicaciones en la que pueden ser implementados y unas breves observaciones en su respecto.

### *Convención de las aplicaciones:*

- DD: Detección de disparos.
- DA: Detección de armas.

- DRR: Detección y reconocimiento de rostros
- DI: Detección de incendios.
- DDT: Detección de deslizamientos de tierra.
- PI: Predicción de inundaciones.

*Tabla 7. Algoritmos recomendados*

Nombre	Aplicación	Observaciones
AlexNet CNN	- DI	Exhibe altos niveles de precisión para la detección de fuego en imágenes. Es de fácil implementación. Baja velocidad de detección que imposibilitan el análisis de vídeos cuadro a cuadro.
Coeficientes cepstrales de Mel	- DD	Realiza una transformación del espacio vectorial de la información para acentuar diferencias entre clases diferentes. Tiene mayor complejidad de cómputo comparado con otras técnicas de preprocesamiento. Baja robustez al ruido.
Coeficientes de correlación	- DD	Requiere baja complejidad de cómputo. Obtiene menores niveles de precisión en comparación con otras técnicas, aunque aceptables.
Faster R-CNN	- DA - DRR	Los valores de precisión obtenidos en la detección de objetos son altos. Presenta una alta velocidad de ejecución. Alto coste computacional para la etapa de entrenamiento. Entrega valores de exhaustividad menores comparados con otros métodos. Necesita múltiples análisis para detectar varios objetos en una imagen.
GoogLeNet	- DI	Exhibe altos niveles de precisión para la detección de fuego en imágenes. Es de fácil implementación. Baja velocidad de detección que imposibilitan el análisis de vídeos cuadro a cuadro.
Haar Cascade	- DRR	Requiere una baja complejidad de cómputo en la implementación. Debe ser modificado o complementado con otros métodos para mantener un buen desempeño. Requiere alto coste computacional durante el entrenamiento.

LightCNN	- DRR	Menor tiempo de ejecución que ResNet y VGG para ser implementado en dispositivos con menor capacidad. Menor precisión que ResNet pero resultados aceptables. Susceptibilidad a los cambios de iluminación en las imágenes.
Máquina de Soporte Vectorial	- DDT	Requiere una complejidad computacional baja. Entrega clasificaciones binarias. Debe evaluarse cuál es el kernel más adecuado para la aplicación.
Perceptrón multicapa	- DD - PI - DDT	Tiene relativa facilidad de entrenamiento e implementación. No necesita grandes cantidades de capas o neuronas para arrojar resultados aceptables. Tiene tendencia al sobreentrenamiento.
Regresión Logística	- DDT	Su implementación e interpretación es simple. Puede clasificar en múltiples clases. Presenta bajo desempeño en aplicaciones con conjuntos de datos no linealmente separables.
ResNet	- DD	Mejor mapeo de los datos de entrada a la información de salida comparada con otros tipos de redes neuronales. Baja variabilidad en el desempeño bajo diferentes condiciones de luz en las imágenes. Mayor éxito de predicción con menor tamaño de la red comparada con otros tipos de redes neuronales.
Transformada de Fourier y espectrograma	- DD	Presenta utilidad para análisis espectral. Ofrece una baja complejidad de cómputo. Es la técnica de preprocesamiento más utilizada en trabajos de análisis de audio para DD.
YOLO	- DA - DRR	Tiene una muy alta velocidad de ejecución. De las más altas encontradas en análisis de imágenes. Mayor precisión para DA en comparación con otras técnicas. Detección de múltiples objetos en una imagen un solo análisis. Presenta un aumento de la complejidad al diseñarse para detectar objetos más pequeños.

## 7. Otras conclusiones

Para cada una de las aplicaciones tratadas anteriormente, se ha destinado una sección de conclusiones que permite mostrar el análisis e interpretación de la información expuesta. Sin embargo, se utiliza esta sección para enunciar algunas conclusiones adicionales y transversales a todo el trabajo de investigación.

- Durante la revisión de la literatura, no se encontraron 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. Por esta razón, las métricas mencionadas durante este trabajo están orientadas a medir el desempeño de los algoritmos utilizados.
- En general, no existen tendencias a utilizar determinadas plataformas para la implementación de los diferentes sistemas y algoritmos revisados. Para el entrenamiento de los algoritmos que lo necesitan, los investigadores optan por utilizar computadores, de preferencia con capacidad de procesamiento alta tal que se pueda disminuir el tiempo que toma este proceso.
- El único patrón encontrado y la única excepción a lo mencionado anteriormente es el uso de los ordenadores tipo RaspBerry Pi. Una RaspBerry PI es un dispositivo multipropósito, de bajo coste, compacto, de gran capacidad dado su tamaño y compatible con una basta cantidad de dispositivos periféricos. Además cuenta con una cantidad de usuarios muy alta que permiten que hay abundante documentación disponible en línea. Por estas razones, las RaspBerry Pi son ampliamente utilizadas en ambientes investigativos.
- En la gran mayoría de casos, las investigaciones revisadas no llegan hasta una etapa de producción en masa o una implementación amplia de los desarrollos, por lo que no es posible estudiar sus características o desempeño lo suficiente como para realizar recomendaciones en términos de hardware.
- 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.

## 8. Referencias Bibliográficas

- 2 - *Armas de fuego y municiones* (2013). Disponible en: <https://tpwd.texas.gov/education/hunter-education/educacion-para-cazadores/capitulo-2> (Consultado: el 13 de abril de 2022).
- Adnan, R., Samad, A.M. y Ruslan, F.A. (2016) "A 3-hours river water level flood prediction model using NNARX with improves modelling strategy", en *2016 6th International Conference on System Engineering and Technology (ICSET)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 23–27. doi:10.1109/ICSEngT.2016.7849617.
- Afandi, W.E.I.B.W.N. y Isa, N.M. (2021) "Object Detection: Harmful Weapons Detection using YOLOv4", en *2021 IEEE Symposium on Wireless Technology Applications (ISWTA)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 63–70. doi:10.1109/ISWTA52208.2021.9587423.
- Ahmed, T., Uppal, M. y Muhammad, A. (2013) "Improving efficiency and reliability of gunshot detection systems", en *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 513–517. doi:10.1109/ICASSP.2013.6637700.
- Alaqil, R.M. *et al.* (2020) "Automatic Gun Detection From Images Using Faster R-CNN", en *2020 First International Conference of Smart Systems and Emerging Technologies (SMARTTECH)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 149–154. doi:10.1109/SMART-TECH49988.2020.00045.
- Almadan, A., Krishnan, A. y Rattani, A. (2020) "BWCFace: Open-set Face Recognition using Body-worn Camera", en *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 1036–1043. doi:10.1109/ICMLA51294.2020.00168.
- Almadan, A. y Rattani, A. (2021) "Towards On-Device Face Recognition in Body-worn Cameras", en *2021 IEEE International Workshop on Biometrics and Forensics (IWBF)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 1–6. doi:10.1109/IWBF50991.2021.9465079.
- Ashraf, A.H. *et al.* (2021) *Weapons Detection for Security and Video Surveillance Using CNN and YOLO-V5s*. [researchgate.net](http://researchgate.net). Disponible en: [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-CNN-and-YOLO-V5s.pdf](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-CNN-and-YOLO-V5s.pdf) (Consultado: el 4 de abril de 2022).
- Aung, H., Bobkov, A.V. y Tun, N.L. (2021) "Face Detection in Real Time Live Video Using Yolo Algorithm Based on Vgg16 Convolutional Neural Network", en *2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 697–702. doi:10.1109/ICIEAM51226.2021.9446291.
- Bai, S.B. *et al.* (2008) "Prediction of Landslide Susceptibility Using logistic Regression: A Case Study in Bailongjiang River Basin, China", en *2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 647–651. doi:10.1109/FSKD.2008.524.
- Bajzik, J., Prinosil, J. y Koniar, D. (2020) "Gunshot Detection Using Convolutional Neural Networks", en *2020 24th International Conference Electronics*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 1–5.

doi:10.1109/IEEECONF49502.2020.9141621.

Bedoya, J. y Martínez, E. (2009) “CALIDAD DEL AIRE EN EL VALLE DE ABURRÁ ANTIOQUIA -COLOMBIA”, *Dynamic medicine: DM*, 76(158), pp. 7–15. Disponible en: [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S0012-73532009000200001](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0012-73532009000200001) (Consultado: el 13 de abril de 2022).

BEN BOTKIN, The Statesman Journal (2019) *Oregon weighs multiple new restrictions on firearms*. Disponible en: <https://kval.com/news/local/in-oregon-it-isnt-a-question-of-if-firearms-legislation-will-pass-its-which-ones> (Consultado: el 13 de abril de 2022).

Bharathi, S. *et al.* (2021) “An Automatic Real-Time Face Mask Detection using CNN”, en *2021 Smart Technologies, Communication and Robotics (STCR)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 1–5. doi:10.1109/STCR51658.2021.9589008.

Brown, L.M. y Fan, Q. (2016) “Enhanced face detection using body part detections for wearable cameras”, en *2016 23rd International Conference on Pattern Recognition (ICPR)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 715–720. doi:10.1109/ICPR.2016.7899719.

Caragliu, A., Del Bo, C. y Nijkamp, P. (2011) “Smart Cities in Europe”, *Journal of Urban Technology*, 18(2), pp. 65–82. doi:10.1080/10630732.2011.601117.

Chacón Rivera, L.M. (2015) “Efecto de los Incendios forestales sobre la calidad del aire en dos ciudades colombianas”, *Departamento de Ingeniería Química y Ambiental* [Preprint]. Disponible en: <https://repositorio.unal.edu.co/handle/unal/57067>.

Chihaoui, M. *et al.* (2016) “A Survey of 2D Face Recognition Techniques”, *Computers*, 5(4), p. 21. doi:10.3390/computers5040021.

Company (2020) *ShotSpotter*. Disponible en: <https://www.shotspotter.com/company/> (Consultado: el 13 de abril de 2022).

DANE (2021) “Encuesta de Convivencia y Seguridad Ciudadana (ECSC) 2019”, *Boletín Técnico* [Preprint].

Debnath, R. y Bhowmik, M.K. (2020) “Automatic Visual Gun Detection Carried by A Moving Person”, en *2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 208–213. doi:10.1109/ICIIS51140.2020.9342681.

De las Heras, J. *et al.* (2013) “Incendios forestales”, *Conservar Aprovechando*, p. 45. Disponible en: [https://www.researchgate.net/profile/Enrique-Doblas-Miranda/publication/261410861\\_Consevar\\_Aprovechando\\_Como\\_integrar\\_el\\_cambio\\_global\\_en\\_la\\_gestion\\_de\\_los\\_montes\\_espanoles/links/00b7d5343a8bdb45ab000000/Conservar-Aprovechando-Como-integrar-el-cambio-global-en-la-gestion-de-los-montes-espanoles.pdf#page=45](https://www.researchgate.net/profile/Enrique-Doblas-Miranda/publication/261410861_Consevar_Aprovechando_Como_integrar_el_cambio_global_en_la_gestion_de_los_montes_espanoles/links/00b7d5343a8bdb45ab000000/Conservar-Aprovechando-Como-integrar-el-cambio-global-en-la-gestion-de-los-montes-espanoles.pdf#page=45).

Departamento Nacional de Planeación DNP (2020) “Documento de Lineamientos de Política de Ciudades Inteligentes (Borrador)”. Disponible en: <https://www.dnp.gov.co/DNPN/Paginas/Lineamientos-de-politica-de-ciudades-inteligentes.aspx>.

Desk, T. (2018) “CCTV VIDEO: Robber pulls out gun; man in cowboy hat tackles him to the

ground”, *The Indian Express*, 26 abril. Disponible en: <https://indianexpress.com/article/trending/viral-videos-trending/armed-robber-steps-shop-man-cowboy-hat-tackles-him-mexico-5152558/> (Consultado: el 13 de abril de 2022).

Dextre, M. *et al.* (2021) “Gun Detection in Real-Time, using YOLOv5 on Jetson AGX Xavier”, en *2021 XLVII Latin American Computing Conference (CLEI)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 1–7. doi:10.1109/CLEI53233.2021.9640100.

Dutta, S. y Ghosh, S. (2021) “Forest Fire Detection Using Combined Architecture of Separable Convolution and Image Processing”, en *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 36–41. doi:10.1109/CAIDA51941.2021.9425170.

Fansler, K.S. (1998) “Description of muzzle blast by modified ideal scaling models”, *Shock and Vibration Digest*, 5(1), pp. 1–12. doi:10.1155/1998/640253.

Farrelly, C.M. (2018) *Key algorithms and statistical models for aspiring data scientists, KDnuggets*. Disponible en: <https://www.kdnuggets.com/2018/04/key-algorithms-statistical-models-aspiring-data-scientists.html> (Consultado: el 31 de enero de 2022).

Freytag, J.C., Begault, D.R. y Peltier, C.A. (2006) “The acoustics of gunfire”, en *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*. Institute of Noise Control Engineering, pp. 1165–1174.

Gaidon, A. *et al.* (2016) “Virtual worlds as proxy for multi-object tracking analysis”, en *Proceedings of the IEEE conference on computer vision and pattern recognition*. [cv-foundation.org](https://cv-foundation.org), pp. 4340–4349. Disponible en: [https://www.cv-foundation.org/openaccess/content\\_cvpr\\_2016/html/Gaidon\\_Virtual\\_Worlds\\_as\\_CVPR\\_2016\\_paper.html](https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Gaidon_Virtual_Worlds_as_CVPR_2016_paper.html).

Galangque, C.M.J. y Guirnaldo, S.A. (2019) “Gunshot Classification and Localization System using Artificial Neural Network (ANN)”, en *2019 12th International Conference on Information Communication Technology and System (ICTS)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 1–5. doi:10.1109/ICTS.2019.8850937.

Ghosh, S. y Dutta, S. (2020) “A comprehensive forecasting, risk modelling and optimization framework for electric grid hardening and wildfire prevention in the US”, *Nuclear engineering and design/fusion: an international journal devoted to the thermal, mechanical, materials, structural, and design problems of fusion energy*, 10(3), pp. 80–89. Disponible en: <https://www.academia.edu/download/65231926/10.5923.j.ijee.20201003.02.pdf>.

Głomb, P. *et al.* (2018) “Application of hyperspectral imaging and machine learning methods for the detection of gunshot residue patterns”, *Forensic science international*, 290, pp. 227–237. doi:10.1016/j.forsciint.2018.06.040.

Hashmi, T.S.S. *et al.* (2021) “Application of Deep Learning for Weapons Detection in Surveillance Videos”, en *2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 1–6. doi:10.1109/ICoDT252288.2021.9441523.

Hoque, M.A. *et al.* (2020) “Autonomous Face Detection System from Real-time Video

- Streaming for Ensuring the Intelligence Security System”, en *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*.  
ieeexplore.ieee.org, pp. 261–265. doi:10.1109/ICACCS48705.2020.9074260.
- Hrabina, M. (2017) “Analysis of linear predictive coefficients for gunshot detection based on neural networks”, en *2017 IEEE 26th International Symposium on Industrial Electronics (ISIE)*. ieeexplore.ieee.org, pp. 1961–1965. doi:10.1109/ISIE.2017.8001552.
- Hrabina, M. y Sigmund, M. (2016) “Implementation of developed gunshot detection algorithm on TMS320C6713 processor”, en *2016 SAI Computing Conference (SAI)*.  
ieeexplore.ieee.org, pp. 902–905. doi:10.1109/SAI.2016.7556087.
- Hrabina, M. y Sigmund, M. (2018) “Gunshot recognition using low level features in the time domain”, en *2018 28th International Conference Radioelektronika (RADIOELEKTRONIKA)*.  
ieeexplore.ieee.org, pp. 1–5. doi:10.1109/RADIOELEK.2018.8376372.
- Jain, A., Aishwarya y Garg, G. (2020) “Gun Detection with Model and Type Recognition using Haar Cascade Classifier”, en *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*. ieeexplore.ieee.org, pp. 419–423.  
doi:10.1109/ICSSIT48917.2020.9214211.
- Jain, H. *et al.* (2020) “Weapon Detection using Artificial Intelligence and Deep Learning for Security Applications”, en *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*. ieeexplore.ieee.org, 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), pp. 1665–1675. doi:10.1063/1.1566752.
- Kacprzyk, J. y Pedrycz, W. (eds.) (2015) *Springer Handbook of Computational Intelligence*. Springer, Berlin, Heidelberg. doi:10.1007/978-3-662-43505-2.
- Kiktova, E. *et al.* (2015) “Gun type recognition from gunshot audio recordings”, en *3rd International Workshop on Biometrics and Forensics (IWBF 2015)*, pp. 1–6.  
doi:10.1109/IWBF.2015.7110240.
- Lee, W. *et al.* (2017) “Deep neural networks for wild fire detection with unmanned aerial vehicle”, en *2017 IEEE International Conference on Consumer Electronics (ICCE)*.  
ieeexplore.ieee.org, pp. 252–253. doi:10.1109/ICCE.2017.7889305.
- Lee, Z.-J. *et al.* (2020) “Rainfall Forecasting of Landslides Using Support Vector Regression”, en *2020 3rd IEEE International Conference on Knowledge Innovation and Invention (ICKII)*.  
ieeexplore.ieee.org, pp. 1–3. doi:10.1109/ICKII50300.2020.9318930.
- Lim, J. *et al.* (2019) “Gun Detection in Surveillance Videos using Deep Neural Networks”, en *2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*. ieeexplore.ieee.org, pp. 1998–2002.  
doi:10.1109/APSIPAASC47483.2019.9023182.
- Lindström, A.-C. *et al.* (2014) “Detection of gunshot residues (GSR) on a self-inflicted gunshot wound”, *Pathology*, 46(3), pp. 260–263. doi:10.1097/PAT.000000000000083.
- Lukman, Q.A., Ruslan, F.A. y Adnan, R. (2016) “5 Hours ahead of time flood water level

prediction modelling using NNARX technique: Case study terengganu”, en *2016 7th IEEE Control and System Graduate Research Colloquium (ICSGRC)*. ieeexplore.ieee.org, pp. 104–108. doi:10.1109/ICSGRC.2016.7813310.

Maher, R.C. (2007) “Acoustical Characterization of Gunshots”, en *2007 IEEE Workshop on Signal Processing Applications for Public Security and Forensics*. ieeexplore.ieee.org, pp. 1–5. doi:10.1109/IEEECONF12259.2007.4218954.

Manjarres, W. y Baca, W. (2019) “Victimización por crimen, percepción de seguridad y satisfacción con la vida en Colombia”, *Revista de Economía Institucional*, 21(41), pp. 133–160. doi:10.18601/01245996.v21n41.06.

Mantoro, T., Ayu, M.A. y Suhendi (2018) “Multi-Faces Recognition Process Using Haar Cascades and Eigenface Methods”, en *2018 6th International Conference on Multimedia Computing and Systems (ICMCS)*. ieeexplore.ieee.org, pp. 1–5. doi:10.1109/ICMCS.2018.8525935.

McCarthy, J. (2007) *What is artificial intelligence?* Stanford University.

McCoy, R.L. (sin fecha) “Modern exterior ballistics: the launch and flight dynamics of symmetric projectiles, 1999”, *Schiffer Publishing Ltd* [Preprint]. Schiffer Publishing Ltd.

Ministerio de Justicia de Colombia (2000) “Código Penal Colombiano”.

Mitchell, S. y Villa, N. (sin fecha) “The Internet of Everything for Cities”.

Morehead, A. *et al.* (2019) “Low Cost Gunshot Detection using Deep Learning on the Raspberry Pi”, en *2019 IEEE International Conference on Big Data (Big Data)*. ieeexplore.ieee.org, pp. 3038–3044. doi:10.1109/BigData47090.2019.9006456.

Muhammad, K. *et al.* (2019) “Efficient Deep CNN-Based Fire Detection and Localization in Video Surveillance Applications”, *IEEE transactions on systems, man, and cybernetics*, 49(7), pp. 1419–1434. doi:10.1109/TSMC.2018.2830099.

Munasinghe, K. y Karunanayake, P. (2021) “Recursive Feature Elimination for Machine Learning-based Landslide Prediction Models”, en *2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC)*. ieeexplore.ieee.org, pp. 126–129. doi:10.1109/ICAIIIC51459.2021.9415232.

Naimeng, C., Wanjun, Y. y Xiaoyu, W. (2021) “Smoke detection for early forest fire in aerial photography based on GMM background and wavelet energy”, en *2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA)*. ieeexplore.ieee.org, pp. 763–765. doi:10.1109/ICPECA51329.2021.9362647.

Narejo, S. *et al.* (2021) “Weapon Detection Using YOLO V3 for Smart Surveillance System”, *Mathematical Problems in Engineering*, 2021. doi:10.1155/2021/9975700.

Nikhil Binoy, C. *et al.* (2019) “Flood Prediction Using Flow and Depth Measurement with Artificial Neural Network in Canals”, en *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)*. ieeexplore.ieee.org, pp. 798–801. doi:10.1109/ICCMC.2019.8819807.

Nunavath, V. y Goodwin, M. (2019) “The Use of Artificial Intelligence in Disaster

- Management - A Systematic Literature Review”, en *2019 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 1–8. doi:10.1109/ICT-DM47966.2019.9032935.
- Olmos, R., Tabik, S. y Herrera, F. (2018) “Automatic handgun detection alarm in videos using deep learning”, *Neurocomputing*, 275, pp. 66–72. doi:10.1016/j.neucom.2017.05.012.
- Othman, N.A. y Aydin, I. (2018) “A face recognition method in the Internet of Things for security applications in smart homes and cities”, en *2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 20–24. doi:10.1109/SGCF.2018.8408934.
- Phitakwinai, S., Auephanwiriyaikul, S. y Theera-Umpon, N. (2016) “Multilayer perceptron with Cuckoo search in water level prediction for flood forecasting”, en *2016 International Joint Conference on Neural Networks (IJCNN)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 519–524. doi:10.1109/IJCNN.2016.7727243.
- Pikrakis, A., Giannakopoulos, T. y Theodoridis, S. (2008) “Gunshot detection in audio streams from movies by means of dynamic programming and Bayesian networks”, en *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 21–24. doi:10.1109/ICASSP.2008.4517536.
- Poole, D. y Mackworth, A. (2017) *Artificial Intelligence: Foundations of Computational Agents*. Cambridge University Press.
- Pradhan, B. y Lee, S. (2010) “Landslide susceptibility assessment and factor effect analysis: backpropagation artificial neural networks and their comparison with frequency ratio and bivariate logistic regression modelling”, *Environmental Modelling & Software*, 25(6), pp. 747–759. doi:10.1016/j.envsoft.2009.10.016.
- Praveen, G.B. y Dakala, J. (2020) “Face Recognition: Challenges and Issues in Smart City/Environments”, en *2020 International Conference on COMMunication Systems NETWORKS (COMSNETS)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 791–793. doi:10.1109/COMSNETS48256.2020.9027290.
- Priya, R.S. y Vani, K. (2019) “Deep Learning Based Forest Fire Classification and Detection in Satellite Images”, en *2019 11th International Conference on Advanced Computing (ICoAC)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 61–65. doi:10.1109/ICoAC48765.2019.246817.
- Qezavati, H., Majidi, B. y Manzuri, M.T. (2019) “Partially Covered Face Detection in Presence of Headscarf for Surveillance Applications”, en *2019 4th International Conference on Pattern Recognition and Image Analysis (IPRIA)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 195–199. doi:10.1109/PRIA.2019.8786004.
- Ren, S. *et al.* (2015) “Faster R-CNN: Towards real-time object detection with region proposal networks”, *Advances in neural information processing systems*, 28. Disponible en: <https://proceedings.neurips.cc/paper/2015/hash/14bfa6bb14875e45bba028a21ed38046-Abst.html> (Consultado: el 5 de abril de 2022).
- Rohaimi, N.A., Ruslan, F.A. y Adnan, R. (2016) “3 Hours ahead of time flood water level prediction using NNARX structure: Case study pahang”, en *2016 7th IEEE Control and System Graduate Research Colloquium (ICSGRC)*. [ieeexplore.ieee.org](https://ieeexplore.ieee.org), pp. 98–103.

doi:10.1109/ICSGRC.2016.7813309.

Rothkrantz, L. (2017) "Person identification by smart cameras", en *2017 Smart City Symposium Prague (SCSP)*. ieeexplore.ieee.org, pp. 1–6. doi:10.1109/SCSP.2017.7973347.

Ruslan, F.A. *et al.* (2016) "7 hours flood prediction modelling using NNARX structure: Case study Terengganu", en *2016 IEEE 12th International Colloquium on Signal Processing Its Applications (CSPA)*. ieeexplore.ieee.org, pp. 263–268. doi:10.1109/CSPA.2016.7515843.

Ruslan, F.A., Samad, A.M. y Adnan, R. (2016) "3 hours flood water level prediction using NNARX structure: Case study Kuala Lumpur", en *2016 6th International Conference on System Engineering and Technology (ICSET)*. ieeexplore.ieee.org, pp. 53–56. doi:10.1109/ICSEngT.2016.7849622.

Ruslan, F.A., Samad, A.M. y Adnan, R. (2018) "4 Hours NNARX flood prediction model using 'traingd' and 'trainoss' training function: A comparative study", en *2018 IEEE 14th International Colloquium on Signal Processing Its Applications (CSPA)*. ieeexplore.ieee.org, pp. 77–81. doi:10.1109/CSPA.2018.8368689.

Russell, S.J. y Norvig, P. (1995) *Artificial Intelligence: A Modern Approach*. Editado por M. Pompili y S. Chavez.

Safiullina, L.K., Gabdullin, A.S. y Anikin, I.V. (2021) "Face Recognition in Biometric Systems Using Haar Cascade Classification", en *2021 Dynamics of Systems, Mechanisms and Machines (Dynamics)*. ieeexplore.ieee.org, pp. 1–5. doi:10.1109/Dynamics52735.2021.9653460.

Salazar González, J.L. *et al.* (2020) "Real-time gun detection in CCTV: An open problem", *Neural networks: the official journal of the International Neural Network Society*, 132, pp. 297–308. doi:10.1016/j.neunet.2020.09.013.

Samireddy, S.R., Carletta, J. y Lee, K.-S. (2017) "An embeddable algorithm for gunshot detection", en *2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS)*. ieeexplore.ieee.org, pp. 68–71. doi:10.1109/MWSCAS.2017.8052862.

Santos, A.S. *et al.* (2020) "Brazilian natural disasters integrated into cyber-physical systems: computational challenges for landslides and floods in urban ecosystems", en *2020 IEEE International Smart Cities Conference (ISC2)*. ieeexplore.ieee.org, pp. 1–8. doi:10.1109/ISC251055.2020.9239011.

Serna Quintana, C.A. (2011) "La naturaleza social de los desastres asociados a inundaciones y deslizamientos en Medellín (1930-1990)", *Historia crítica*, (43), pp. 198–223. doi:10.7440/histcrit43.2011.11.

Shamsoshoara, A. *et al.* (2021) "Aerial imagery pile burn detection using deep learning: The FLAME dataset", *Computer Networks*, 193, p. 108001. doi:10.1016/j.comnet.2021.108001.

Shiekh, A.A., Tahir, M. y Uppal, M. (2017) "Accurate gunshot detection in urban environments using blind deconvolution", en *2017 International Multi-topic Conference (INMIC)*. ieeexplore.ieee.org, 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 fur 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), pp. 19–41. Disponible en: <https://www.redalyc.org/articulo.oa?id=51052064002>.

Su, K., Li, J. y Fu, H. (2011) "Smart city and the applications", en *2011 International Conference on Electronics, Communications and Control (ICECC)*, pp. 1028–1031. doi:10.1109/ICECC.2011.6066743.

Tabane, E., Ngwira, S.M. y Zuva, T. (2016) "Survey of smart city initiatives towards urbanization", en *2016 International Conference on Advances in Computing and Communication Engineering (ICACCE)*, pp. 437–440. doi:10.1109/ICACCE.2016.8073788.

Utomo, D., Hu, L.-C. y Hsiung, P.-A. (2020) "Deep Neural Network-Based Data Reconstruction for Landslide Detection", en *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 3119–3122. doi:10.1109/IGARSS39084.2020.9323124.

Verma, G.K. y Dhillon, A. (2017) "A Handheld Gun Detection using Faster R-CNN Deep Learning", en *Proceedings of the 7th International Conference on Computer and Communication Technology*. New York, NY, USA: Association for Computing Machinery (ICCCT-2017), pp. 84–88. doi:10.1145/3154979.3154988.

Viola, P. y Jones, M. (2001) "Rapid object detection using a boosted cascade of simple features", en *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. I–I. doi:10.1109/CVPR.2001.990517.

Wahyono, I.D. *et al.* (2019) "New Method of Artificial Intelligence for Disaster Information Floods use Distributed Wireless Sensors", en *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 1–5. doi:10.1109/ISEMANTIC.2019.8884315.

Wandelt, S., Bux, M. y Leser, U. (2014) "Trends in Genome Compression", *Current bioinformatics*, 9(3). doi:10.2174/1574893609666140516010143.

Warsi, A. *et al.* (2019) "Gun Detection System Using Yolov3", en *2019 IEEE International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 1–4. doi:10.1109/ICSIMA47653.2019.9057329.

Xu, C. *et al.* (2019) "Landslide Susceptibility Mapping Using Logistic Regression Model Based On Information Value for the Region Along China-Thailand Railway from Saraburi To Sikhio, Thailand", en *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 9650–9653. doi:10.1109/IGARSS.2019.8900041.

Yaloveha, V. *et al.* (2019) "Fire Hazard Research of Forest Areas based on the use of Convolutional and Capsule Neural Networks", en *2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON)*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 828–832.

doi:10.1109/UKRCON.2019.8879867.

Yi, Y. *et al.* (2019) “Comparison of Different Machine Learning Models For Landslide Susceptibility Mapping”, en *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*. [ieeexplore.ieee.org](http://ieeexplore.ieee.org), pp. 9318–9321.

doi:10.1109/IGARSS.2019.8898208.

Zhang, G., Wang, M. y Liu, K. (2019) “Forest Fire Susceptibility Modeling Using a Convolutional Neural Network for Yunnan Province of China”, *International Journal of Disaster Risk Science*, 10(3), pp. 386–403. doi:10.1007/s13753-019-00233-1.

Zhao, F. *et al.* (2021) “Smart city research: A holistic and state-of-the-art literature review”, *Cities*, 119, p. 103406. doi:10.1016/j.cities.2021.103406.