



Instituto Politécnico Nacional
Unidad Profesional Interdisciplinaria de Ingeniería
Campus Tlaxcala



Reporte Técnico

"Algoritmo de detección de patrones de movimiento para la identificación de allanamiento de morada usando visión computacional"

No. TT2025-1_IA-002

Presenta(n):

Valeria Jahzeel Castañón Hernández, Ingeniería en Inteligencia Artificial

Asesores:

Dr. Lauro Reyes Cocoletzi

Dra. María del Rocío Ochoa Montiel

Tlaxcala, Tlaxcala, a **05 de Diciembre de 2024**

"La Técnica al Servicio de la Patria"

AGRADECIMIENTOS

Agradezco a todas las personas que han estado conmigo y me han apoyado a lo largo de mi carrera, principalmente a mi mamá que hizo mucho esfuerzo para que yo pudiera seguir estudiando.

También agradezco a todos mis profesores, que nos permitieron aprender todos los temas, por complejos que fueran y sobre todo, que supieron explicar de manera sencilla para que cada uno de los alumnos pudiéramos entender lo mejor posible.

Agradezco a mis asesores por apoyarme en este proceso y por ayudarme a aclarar dudas, darme ideas y guiarme en el desarrollo de este proyecto.

RESUMEN

El objetivo de este proyecto es desarrollar un algoritmo de detección de comportamientos sospechosos en videos de vigilancia de casa habitación, enfocándose en el merodeo y los intentos de allanamiento. Este algoritmo busca diferenciar entre comportamientos sospechosos y cotidianos, proporcionando una herramienta que podría mejorar la seguridad en el hogar.

Se recopilaron videos del UCF Crime Dataset y YouTube, mejorando su calidad mediante ecualización de histograma. Utilizando el flujo óptico, se analizaron movimientos en los cuadros para identificar posiciones y trayectorias clave. El algoritmo utiliza redes neuronales convolucionales (CNN) para la identificar personas y manos, también utiliza un perceptron multicapa (MLP) clasificar comportamientos en merodeo (permanencia mayor a 15 segundos) y forzar accesos (movimientos sospechosos cerca de puertas o ventanas).

El algoritmo busca diferenciar entre actividades cotidianas y potencialmente delictivas, ofreciendo una posible solución la seguridad en el hogar. La evaluación del algoritmo incluirá métricas como precisión, F1-score y tasa de falsos positivos, con el objetivo de validar su efectividad frente a enfoques similares.

Términos/Palabras Clave

- Aprendizaje máquina
- Cámaras de seguridad
- Patrón de movimiento
- Redes neuronales artificiales
- Visión computacional

ÍNDICE GENERAL

1 Introducción	1
1.1 Planteamiento del Problema	1
1.1.1 Definición del Problema	1
1.1.2 Objetivos	2
1.1.3 Justificación	3
1.2 Hipótesis	3
1.3 Aportación Científica y/o Tecnológica	4
1.4 Organización del Documento	4
2 Marco Teórico	5
2.1 Estado del Arte	5
2.2 Allanamiento de morada en México	6
2.2.1 Definición	6
2.2.2 Factores que contribuyen al allanamiento	7
2.2.3 Estadísticas	7
2.2.4 Impacto del allanamiento de morada en las personas	8
2.3 Patrones de movimiento	8
2.3.1 Definición	8
2.3.2 Movimientos asociados a comportamientos sospechosos	9
2.3.3 Movimientos clave para detectar un allanamiento	10
2.4 Visión computacional	11
2.4.1 Conceptos generales	11
2.4.2 Aprendizaje automático para la detección de comportamientos sospechosos	13
3 Metodología	16
3.1 Creación del conjunto de datos	16
3.2 Detección de objetos usando YOLO	17
3.3 Análisis y clasificación de movimientos	17
3.4 Evaluación del algoritmo	18

CAPÍTULO 1

INTRODUCCIÓN

En la actualidad, la seguridad en el hogar enfrenta desafíos constantes debido al incremento en la percepción de inseguridad y la incidencia de allanamientos de morada, particularmente en zonas urbanas. Aunque los sistemas de videovigilancia han avanzado tecnológicamente, aún dependen en gran medida de la supervisión humana, lo que limita su eficacia. La necesidad de herramientas automáticas y proactivas que alerten sobre comportamientos sospechosos se hace evidente.

Diversos estudios han explorado enfoques basados en visión computacional para abordar esta problemática. Por ejemplo, se han desarrollado algoritmos para detectar actos específicos, como escalar bardas o permanecer de manera inusual en un área [17][13][18]. Sin embargo, estos métodos suelen enfrentar limitaciones en términos de precisión y aplicabilidad en entornos diversos.

Este proyecto propone el desarrollo de un algoritmo que, mediante el análisis de patrones de movimiento en videos, identifique actividades sospechosas como el merodeo prolongado o intentos de forzar accesos. A través del uso de redes neuronales y aprendizaje automático, se busca proporcionar una herramienta innovadora que en un futuro pueda ayudar a mejorar la seguridad doméstica al ofrecer alertas tempranas y minimizar la supervisión constante.

El documento está estructurado de la siguiente manera: la primera sección corresponde al planteamiento del problema, seguido del marco teórico y la metodología. Se espera que este proyecto no solo contribuya a la mejora de los sistemas de seguridad actuales, sino que también siente las bases para su aplicación en otros contextos.

1.1 Planteamiento del Problema

1.1.1 Definición del Problema

Según la revista digital El Economista, en 2023 la mayoría de los delitos en México registraron una baja significativa a nivel nacional; sin embargo,

1.1. PLANTEAMIENTO DEL PROBLEMA

la gran mayoría de ellos siguen quedando impunes [7]. De acuerdo con el Observatorio Nacional Ciudadano, el delito de robo a casa habitación disminuyó alrededor del 16.20% entre 2022 y 2023 [20]. No obstante, encuestas del INEGI revelan que, en México, de cada 100 delitos, solo 6.4 se denuncian, y de cada 100 delitos denunciados, solo 14 se resuelven [4]. Además, la incidencia de allanamientos tiende a aumentar durante las vacaciones, cuando las familias regresan a sus hogares y descubren puertas o ventanas rotas, o que les faltan pertenencias. Esto evidencia que, aunque haya una disminución en la tasa de robos, la percepción de inseguridad sigue siendo alta debido a la baja tasa de resolución de casos y al aumento de allanamientos en situaciones específicas. En respuesta a esta problemática, diversos estudios han explorado el uso de sistemas basados en visión por computadora para detectar comportamientos sospechosos en video. Un primer enfoque, presentado en [17], propone un sistema para identificar el acto de escalar bardas. Por otro lado, [13] aborda la detección de robos relacionados con vehículos estacionados cerca de las viviendas. Finalmente, [18] propone un sistema para la detección de merodeo en sistemas de vigilancia de casa-habitación.

Pese a los avances logrados, estos métodos presentan limitaciones, como la falta de precisión en escenarios complejos, la dependencia de configuraciones específicas del entorno y la ausencia de resultados cuantitativos sólidos. Por ello, este proyecto propone el desarrollo de un algoritmo que, mediante el uso de técnicas de aprendizaje automático, redes neuronales, análisis de datos y visión computacional, sea capaz de identificar patrones de movimiento asociados con intentos de allanamiento. Este enfoque buscará diferenciar movimientos sospechosos de actividades cotidianas, como la llegada de un repartidor o visitas familiares, ofreciendo una solución más precisa y proactiva para mejorar la seguridad en el hogar.

1.1.2 Objetivos

Objetivo General

Desarrollar un algoritmo para detectar y clasificar patrones asociados a comportamientos sospechosos en videos de vigilancia a partir del uso de visión computacional.

Objetivos Específicos

1. Generar un conjunto de datos mediante la recolección de videos de diferentes fuentes.
2. Preprocesar los videos recolectados mediante la aplicación de redimensionamientos, recortes o ajustes de iluminación.
3. Diseñar e implementar un algoritmo para la detección y clasificación

de las siguientes conductas: forzar accesos (puertas y/o ventanas) y merodeo.

4. Evaluar la precisión y efectividad del algoritmo mediante la comparación con trabajos similares y el uso de diversas métricas de evaluación.

1.1.3 Justificación

En la actualidad, estamos viviendo una época de rápidos cambios tecnológicos, donde la inteligencia artificial (IA) está cobrando un papel crucial en la vida cotidiana, en ámbitos como el estudio y el trabajo. Según un estudio realizado por Microsoft [8], alrededor del 56% de la generación Z (18-24 años) y el 43% de los millennials (25-44 años) ya utilizan y experimentan con la IA. Este creciente interés ha llevado a que aproximadamente el 50% de las empresas [29], integren la IA en el desarrollo de sus productos y servicios, en este contexto, el presente proyecto adquiere relevancia.

A pesar de los avances tecnológicos en los sistemas de vigilancia doméstica, los delitos de allanamiento siguen ocurriendo con frecuencia [20]. Si bien muchas viviendas están equipadas con cámaras de seguridad, estas generalmente solo registran los eventos sin detectar ni analizar activamente lo que sucede. La detección de actividades sospechosas aún depende en gran medida de que los propietarios o terceros noten algo inusual, revisen las grabaciones, o reaccionen a las alarmas, lo que a menudo ocurre cuando el delito ya ha sido cometido, esta respuesta tardía limita la eficacia de los sistemas de vigilancia actuales, dejando a los hogares vulnerables.

El desarrollo de un algoritmo capaz de detectar conductas sospechosas relacionadas con intentos de allanamiento, a través del análisis de posturas y movimientos, a futuro puede ayudar a mejorar la seguridad al proporcionar una detección proactiva y reducir la necesidad de supervisión constante. Además, este algoritmo puede integrarse a otros ámbitos, adaptándose para detectar más tipos de comportamientos anómalos en diversos escenarios, lo que lo convierte en una herramienta versátil; su integración en diferentes sistemas de seguridad relacionados a esta problemática podría realizarse con modificaciones mínimas, lo que incrementa su potencial de uso en la prevención de delitos y otros contextos de seguridad.

1.2 Hipótesis

Un algoritmo de detección y seguimiento secuencial de videos de vigilancia puede ayudar a la detección de actividades anómalas dados en allanamientos a casa habitación aproximando la actividad observada a como lo haría un humano.

1.3 Aportación Científica y/o Tecnológica

Dentro de los productos esperados están:

1. Algoritmo capaz de detectar patrones de movimiento relacionados a intentos de allanamiento de morada
2. Conjunto de datos con videos de personas tratando de forzar accesos por medio de puertas y/o ventanas y merodeando
3. Reporte técnico que muestre la metodología del algoritmo y que compare los trabajos relacionados con el trabajo propuesto que explique el funcionamiento del algoritmo, así como el análisis de los resultados obtenidos.

En la Figura 1 se presenta un esquema de los módulos que componen el algoritmo.

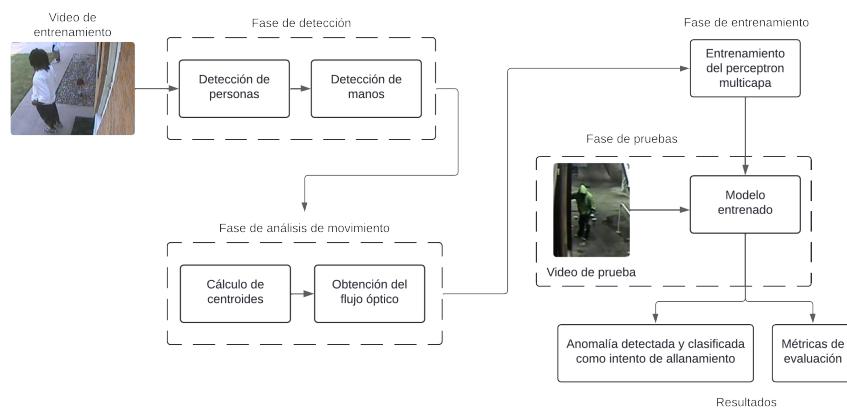


Figura 1: Módulos del algoritmo

1.4 Organización del Documento

El documento está organizado de la siguiente manera, en el capítulo 1 se incluye la introducción y el estado del arte. El marco teórico se incluye en el capítulo 2. La metodología del proyecto se describe en el capítulo 3

CAPÍTULO 2

MARCO TEÓRICO

2.1 Estado del Arte

Entre la literatura revisada, se identificaron los siguientes artículos que están relacionados con la problemática abordada, principalmente se relacionan con la detección de comportamientos sospechosos y de intrusos en diferentes contextos. En [26], se utiliza una combinación de compresión adaptativa de video y redes 3D convolucionales (3D-CNN) para extraer características espaciotemporales. El proceso comienza con la segmentación de los videos en cuadros clave, aplicando compresión selectiva para preservar los objetos de interés y reducir la complejidad computacional. Posteriormente, las características son procesadas por un modelo basado en CNN 3D para clasificar actividades anómalas.

Por otra parte en [32] se emplea un enfoque que integra la extracción de características espacio-temporales y el uso de clasificadores como Random Forest y SVM. El sistema detecta y rastrea personas mediante filtros de partículas y segmentación basada en ViBe, utilizando el tiempo de permanencia y los cambios de ángulo de movimiento para clasificar el comportamiento como merodeo o normal.

Finalmente, [15] propone un sistema basado en módulos embebidos, optimizando algoritmos para detección de intrusos, incendios, caídas y merodeo. Se utilizan sensores y cámaras Kinect, con procesamiento en tiempo real mediante Raspberry Pi, aplicando técnicas de optimización como la reducción de la complejidad de los cálculos y el uso de filtros adaptativos para mejorar la eficiencia del sistema.

A continuación, en la Tabla 1 se muestran las ventajas y limitaciones de los artículos mencionados.

2.2. ALLANAMIENTO DE MORADA EN MÉXICO

Trabajo	Modelo	Ventajas	Limitaciones
Anomalous Human Activity Recognition(2020)	Compresión adaptativa + 3D-CNN	Alta precisión en detección de anomalías, optimización de datos relevantes	Complejidad computacional elevada, requiere conjuntos de datos extensivos
Loitering Detection Using Spatial-Temporal Information (2023)	Ánálisis espacio-temporal + Random Forest/SVM	Eficiencia en detección de merodeo, uso de múltiples características	Limitado a ciertos escenarios, precisión dependiente del entorno
Implementation of real-time intelligent video surveillance system (2021)	Módulo embedido + Algoritmos de detección	Bajo consumo de energía, costo reducido, procesamiento en tiempo real	Limitaciones en potencia computacional, dependencia de hardware específico

Tabla 1: Comparativa de trabajos relacionados

2.2 Allanamiento de morada en México

2.2.1 Definición

En México, la protección de la propiedad privada y la seguridad en el hogar están contempladas en diversas disposiciones legales, aunque con diferencia en los tipos de delitos involucrados. El Código Penal Federal establece en su artículo 285 que se impondrán de un mes a dos años de prisión, además de una multa, a quien, sin motivo justificado, orden de autoridad competente ni permiso de la persona autorizada, ingrese de manera furtiva, con engaño o violencia, a una vivienda o sus dependencias. Este delito, conocido como allanamiento de morada, busca proteger la inviolabilidad del hogar y la privacidad de sus habitantes [14].

Aunque el allanamiento de morada se enfoca en la intrusión no autorizada, el Código Penal también sanciona el delito de robo, que suele involucrar la entrada a una propiedad para sustraer bienes sin consentimiento del propietario. Según el artículo 367, el robo es cometido por quien se apodera de un bien ajeno sin derecho y sin el consentimiento de la persona que puede disponer legalmente de él. En muchos casos, tanto el allanamiento como el robo implican una violación al espacio privado, aunque el robo añade la intención de apropiación de un bien ajeno [1].

Por otro lado, el concepto de merodeo, definido por la Real Academia Española como “vagar por las inmediaciones de algún lugar, en general

con malos fines”, puede ser un factor de alerta, aunque no está tipificado específicamente como delito en México. No obstante, en ciertos casos las autoridades pueden realizar arrestos preventivos bajo la presunción de que una persona que merodea podría cometer un delito, amparados en el artículo 2 de la Ley Nacional del Registro de Detenciones en México [11].

2.2.2 Factores que contribuyen al allanamiento

Hay diversos factores que contribuyen a la ocurrencia de allanamientos de morada, entre ellos, destaca la eficiencia del sistema judicial. Según el Instituto Nacional de Estadística y Geografía (INEGI) [4], en México, de cada 100 delitos, solo 6.4 son denunciados y, de estos, únicamente 14 llegan a resolverse. Este nivel de impunidad puede incentivar a los delincuentes a actuar sin temor a consecuencias legales. Existen también factores de atractividad para el allanamiento, como la facilidad de acceso a la vivienda, su desocupación durante el día y el tiempo de intervención policial en la zona. De manera similar, la policía del Reino Unido [23] ha identificado elementos que hacen que una casa sea blanco de allanamiento, tales como la ausencia de personas en horarios laborales o vacaciones, y la falta de mascotas o bardas limitando la propiedad.

2.2.3 Estadísticas

En términos de estadísticas, según el Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública, entre enero de 2020 y julio de 2024 se reportaron 63,751 casos de allanamiento de morada en México (Figura 2).

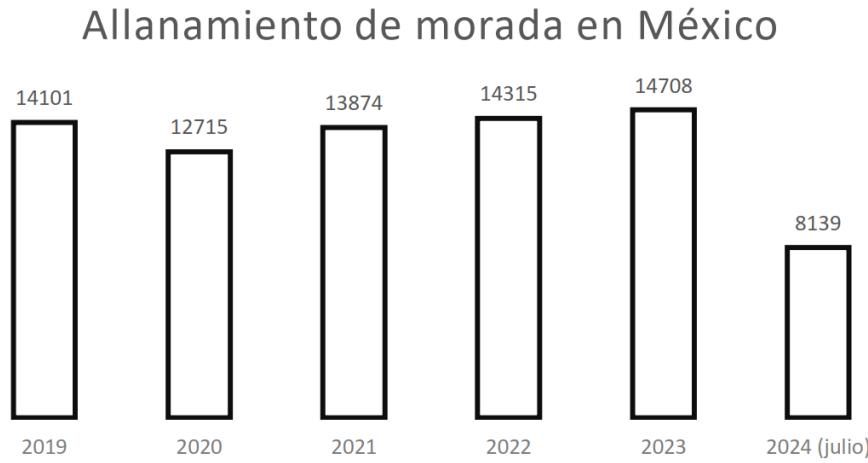


Figura 2: Estadísticas de allanamiento en México (2019 - Julio 2024)

Aunque hasta julio de 2024 se registró una disminución del 8% en comparación con el mismo período de 2023, el promedio anual de allanamientos es de 13,903 casos [25]. Por otra parte, encuestas del INEGI [9] muestran que, en el cuarto trimestre de 2023, el 18.5% de la población se sentía

2.3. PATRONES DE MOVIMIENTO

insegura en su propia casa, mientras que el 59.2% reportaba inseguridad en su área de residencia y el 48% había sido testigo de asaltos o robos en los alrededores de su hogar (Figura 3).

Seguridad Pública en México (2023)

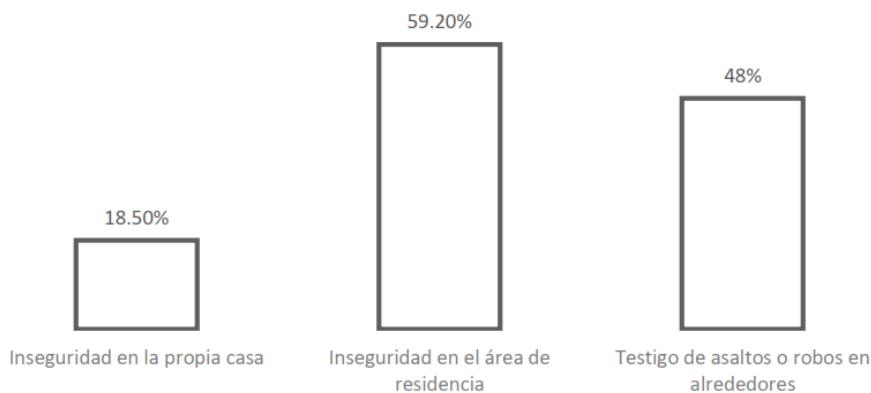


Figura 3: Estadísticas de seguridad pública en México

2.2.4 Impacto del allanamiento de morada en las personas

El impacto del allanamiento de morada va más allá de las pérdidas materiales, afectando también la percepción de seguridad de las personas y generando consecuencias emocionales y psicológicas. La Biblioteca Nacional de Medicina de Estados Unidos cuenta con un artículo sobre los efectos psicológicos que causa un allanamiento de morada [16], en ese se destaca que ser víctima de un robo o intrusión en casa es un evento traumático con efectos psicológicos duraderos. Las víctimas suelen experimentar estrés postraumático, ansiedad, miedo, culpa, irritabilidad y problemas para dormir. En muchos casos, las secuelas psicológicas tras un allanamiento pueden ser más perjudiciales que el delito en sí, afectando la calidad de vida y motivando a las personas a implementar mayores medidas de seguridad en sus hogares.

2.3 Patrones de movimiento

2.3.1 Definición

Los patrones de movimiento son formas repetitivas en las que los seres vivos o los objetos se mueven y se pueden observar en actividades como caminar, nadar o volar; estos patrones son fundamentales para entender el comportamiento de los organismos. Estos patrones de movimiento se

clasifican en tres tipos principales: locomotores, no locomotores y manipulativos. Los patrones locomotores implican la acción de mover y trasladar el cuerpo (arrastrarse, gatear, rodar, caminar, correr, brincar, etc.); los no locomotores, en cambio, son aquellos que se realizan en un solo lugar y en los que el cuerpo se mueve alrededor de un punto fijo (doblar, estirar, torcer, etc.); finalmente, los patrones manipulativos son aquellos en los que se emplean las extremidades para realizar acciones con algún objeto (lanzar, atrapar, patear, empujar, etc.) [10].

Por otro lado, cuando hablamos de movimientos relacionados con comportamientos sospechosos, es importante definir qué significa "sospechoso". Según la Real Academia Española (RAE), el verbo "sospechar" se refiere a "imaginar algo por conjeturas fundadas en apariencias o indicios; desconfiar de algo o alguien; o considerar a alguien como posible autor de un delito o una falta" [11]. Asimismo, el término "comportamiento" deriva del verbo "comportar", que se define como "actuar de una manera determinada" [11]. De esta manera, el comportamiento sospechoso puede entenderse como el conjunto de acciones o actitudes que generan desconfianza y llevan a los demás a especular sobre las intenciones de una persona o a percibirla como posible autora de un delito o falta.

2.3.2 Movimientos asociados a comportamientos sospechosos

Existen ciertos movimientos que los seres humanos asociamos instintivamente con conductas sospechosas. Por ejemplo, en el blog del estado de Guanajuato se creó una publicación orientada a ayudar a los ciudadanos a identificar y alertar a las autoridades en caso de presenciar alguna actitud inusual [6]. Entre las recomendaciones se incluyen: ruidos extraños, vidrios rotos o disparos, la presencia de personas o vehículos ajenos a la colonia, un negocio con la puerta abierta en horarios no comerciales del cual se esté retirando mercancía, o domicilios con excesivo movimiento no relacionado con una rutina familiar.

En una iniciativa similar, el gobierno de México ha dispuesto en su página oficial una sección que insta a las personas a informar a las autoridades en caso de detectar comportamientos sospechosos. Estos comportamientos incluyen la presencia de personas desconocidas en el vecindario con actitudes inusuales, domicilios con un flujo anormal de personas, ruidos inusuales o luces encendidas las 24 horas del día, individuos que parecen vigilar algún domicilio o persona, aquellos que se estacionan o permanecen demasiado tiempo frente a una propiedad, personas tomando fotografías de casas, personas o negocios en la zona, e incluso quienes solicitan información personal sobre terceros. [5]

2.3.3 Movimientos clave para detectar un allanamiento

A partir de los movimientos asociados a comportamientos sospechosos, vamos a hablar a más detalle de los que están relacionados al allanamiento de morada. Los más comunes son:

Movimientos bruscos

Estos pueden ser derivados al jalar, golpear o patear alguna puerta y/o ventana. Esto se puede ver más claro en los ejemplos mostrados en la Figura 4.



(a) Golpear la puerta



(b) Patear la puerta



(c) Jalar la puerta

Figura 4: Ejemplos de movimientos bruscos

Merodeo

El merodeo se caracteriza por el hecho de que una persona camina sin una dirección específica en un área determinada o permanece en un mismo lugar durante un período de tiempo prolongado, en este trabajo a las afueras de una casa. Debido a las características de los videos del conjunto de datos que se empleará, se considerará como merodeo cuando una persona permanezca en la visión de la cámara durante aproximadamente 15 segundos. Esto contrasta con el enfoque descrito en [18], donde se define el merodeo como permanecer en el área durante al menos 1 minuto.



Figura 5: Ejemplo de merodeo

2.4 Visión computacional

2.4.1 Conceptos generales

Definición

La visión computacional es un campo de la inteligencia artificial (IA) que utiliza el aprendizaje automático y las redes neuronales para enseñar a ordenadores y sistemas a extraer información significativa de imágenes digitales, vídeos y otras entradas visuales, y a hacer recomendaciones o tomar medidas cuando detectan defectos o problemas. [2] A continuación, se describen las principales herramientas y metodologías implementadas en el proyecto:

Flujo óptico

El flujo óptico permite estimar el movimiento de los píxeles en una secuencia de imágenes consecutivas, identificando cambios de brillo o color. Esto es esencial para detectar acciones como agacharse, jalar o golpear objetos en la escena. En este proyecto, el flujo óptico se utiliza para determinar las posiciones y movimientos de las personas en los videos de vigilancia, proporcionando la base para identificar comportamientos sospechosos [21][31].

Recuadros envolventes

Para una visualización clara del seguimiento, se utiliza un recuadro envolvente alrededor de cada persona detectada. Este rectángulo destaca la posición del objeto en el video, lo que facilita la identificación y el monitoreo en tiempo real [28].

Centroides

Los centroides representan el centro geométrico de cada objeto detectado. Este concepto es clave para:

- Rastrear trayectorias individuales
- Predecir movimientos futuros
- Diferenciar objetos del fondo de la imagen

En este proyecto, los centroides permiten identificar y analizar patrones de comportamiento asociados a actividades sospechosas, como movimientos repetitivos o permanencia inusual en una zona específica [22].

Seguimiento de objetos

El algoritmo utiliza seguimiento de objetos para rastrear la trayectoria de las personas detectadas en el video. Esto incluye:

- Detección del centroide para cada persona
- Obtención del flujo óptico a partir del centroide
- Seguimiento y análisis del flujo óptico

El seguimiento no solo ayuda a identificar individuos, sino que también permite clasificar comportamientos como el merodeo, caracterizado por movimientos erráticos o permanencia prolongada en una misma área. [27].

Procesamiento digital de imágenes

El procesamiento digital de imágenes es una rama de la visión computacional que consiste en procesar imágenes digitales mediante un ordenador digital. También podemos decir que es el uso de algoritmos y modelos matemáticos para procesar y analizar imágenes digitales. El objetivo del procesamiento digital de imágenes es mejorar la calidad de las imágenes, extraer información significativa de ellas y automatizar tareas basadas en imágenes. [12] Esto es importante debido a que la resolución en los videos de cámaras de seguridad puede ser muy bueno o malo, dependiendo de factores como el modelo de la cámara, la cantidad de luz dependiendo la hora o factores ambientales como el clima. Por esta razón el uso de técnicas que ayuden a mejorar la calidad de las imágenes resulta importante para el desarrollo este proyecto ya que la extracción de características de cada una de ellas es la parte más importante para que el modelo de aprendizaje automático pueda identificar correctamente los patrones de movimiento. Se utilizarán técnicas como:

1. Ecualización de histograma

El histograma es una función discreta que representa los niveles de intensidad en el intervalo $[0, L - 1]$ de una imagen digital. Se expresa como:

$$h(r_k) = n_k$$

donde r_k es el k -ésimo valor de intensidad y n_k es el número de píxeles de la imagen con intensidad r_k . Para optimizar la detección de comportamientos sospechosos, es crucial garantizar una buena calidad en los frames de los videos procesados. En este proyecto, se utiliza la manipulación del histograma como una herramienta para mejorar la visibilidad y el contraste de las imágenes. Esta técnica redistribuye los niveles de intensidad de la imagen, mejorando la visibilidad en escenas con baja iluminación o contraste desigual. Aunque la ecualización

de histograma puede mejorar significativamente el contraste en muchos casos, su efectividad depende de las características de la imagen. En ciertas situaciones, podría producir resultados menos deseables, como pérdida de detalles en áreas específicas. La mejora de calidad de los frames mediante la ecualización de histograma garantiza que las técnicas de visión computacional a utilizar funcionen con mayor precisión, al optimizar la claridad de las imágenes, se facilita la detección de movimientos y comportamientos sospechosos en los videos de vigilancia.

2.4.2 Aprendizaje automático para la detección de comportamientos sospechosos

Se dice que un programa aprende de la experiencia E con respecto a una clase de tareas T y una medida de rendimiento P, si su rendimiento en las tareas de T, medido por P, mejora con la experiencia E. [19] además de esto debemos identificar la clase de tareas que el programa va a aprender, la medida de rendimiento a mejorar y la fuente de experiencia mediante la cual va a aprender. Existen diversos métodos por los que un programa puede aprender, en este caso para que el programa aprenda a determinar entre una conducta normal y una conducta sospechosa se pretenden utilizar los siguientes modelos:

Perceptrón multicapa (MLP)

Antes de empezar a hablar del MLP primero sería bueno definir lo que es un perceptrón, un perceptrón es un modelo básico de red neuronal diseñado para clasificar datos linealmente separables. Funciona calculando una combinación lineal ponderada de las entradas y produciendo una salida binaria (0 o 1) en función de si la suma ponderada supera un umbral [19]. El funcionamiento del perceptrón se describe mediante la siguiente ecuación:

$$y = \begin{cases} 1 & \text{si } \mathbf{w} \cdot \mathbf{x} + b > 0 \\ 0 & \text{en otro caso} \end{cases}$$

donde \mathbf{w} es el vector de pesos, \mathbf{x} es el vector de entradas, y b es el sesgo. El aprendizaje del perceptrón se realiza ajustando los pesos mediante una regla de actualización basada en el error entre la salida predicha y la etiqueta verdadera:

$$\Delta w_i = \eta(y - \hat{y})x_i$$

donde η es la tasa de aprendizaje, y es la etiqueta verdadera, \hat{y} es la salida predicha, y x_i es la i-ésima entrada. Sin embargo, el perceptrón simple solo puede aprender funciones linealmente separables.

2.4. VISIÓN COMPUTACIONAL

Ahora que ya está definido el perceptrón simple, se puede definir el MLP, este es una extensión del perceptrón simple, diseñado para manejar problemas más complejos que no son linealmente separables [19]. Durante el desarrollo de este proyecto, el MLP recibirá una serie de vectores numéricos.

El MLP está compuesto por una capa de entrada que recibe estas características, una o más capas ocultas que aplican transformaciones no lineales mediante funciones de activación no lineales como la sigmoide o ReLU, y una capa de salida que genera una clasificación binaria o multicategoría, indicando la probabilidad de que el movimiento observado corresponda a un comportamiento sospechoso, como el merodeo o el intento de forzar un acceso.

El aprendizaje del modelo se realizará mediante el algoritmo de retropropagación del error, ajustando los pesos en función del error entre la salida predicha y la etiqueta real. Esto permitirá que el MLP capture relaciones complejas en los datos, mejorando la capacidad del sistema para detectar conductas sospechosas.

Redes Neuronales Convolucionales (CNN)

En el contexto del proyecto, las CNN juegan un papel crucial para analizar y clasificar patrones en los videos de vigilancia, este tipo de red neuronal es ampliamente utilizada en tareas de clasificación y reconocimiento de objetos en imágenes esto permitirá identificar comportamientos sospechosos. Las CNN son un tipo de aprendizaje automático supervisado compuesto por tres capas principales:

1. Capa convolucional: Extrae características básicas, como bordes, colores y texturas.
2. Capa de agrupamiento (pooling): Reduce la dimensionalidad de los datos manteniendo las características más relevantes.
3. Capa totalmente conectada (fully connected, FC): Combina la información para realizar la clasificación final.

A medida que los datos avanzan a través de estas capas, la red neuronal incrementa su capacidad de reconocer elementos más complejos, como formas o patrones específicos, hasta identificar el objeto o comportamiento deseado [3].

YOLO (You Only Look Once)

YOLO es un algoritmo de visión computacional que permite detectar y clasificar objetos en imágenes y videos de manera simultánea y en tiempo real. Esta capacidad lo hace bastante útil en sistemas de vigilancia, como el desarrollado en este proyecto, donde se requiere identificar personas en tiempo real. YOLO utiliza una CNN que divide la imagen en una cuadrícula.

Cada celda de esta cuadrícula predice un número específico de recuadros envolventes junto con una probabilidad de clase asociada. Esta probabilidad indica la certeza de que un objeto específico se encuentra dentro de un cuadro delimitador dado [24]. En la Figura 7 se muestra la arquitectura general que sigue este modelo.

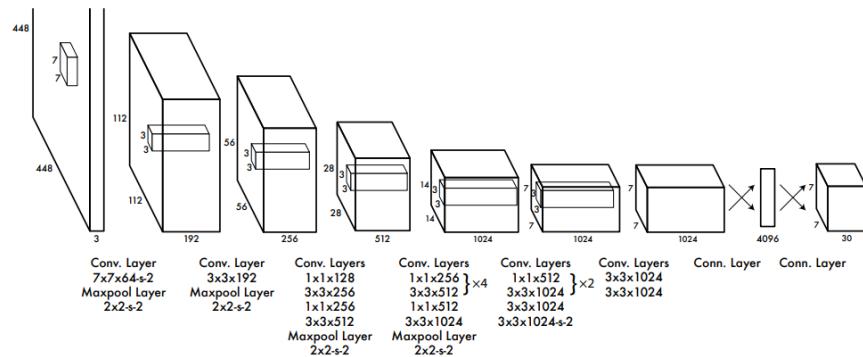


Figura 6: Arquitectura general de YOLO

Además, YOLO se puede entrenar utilizando conjuntos de datos específicos para detectar clases particulares de objetos. En este proyecto, se aplicará YOLO para identificar personas y manos en los videos de vigilancia.

CAPÍTULO 3

METODOLOGÍA

3.1 Creación del conjunto de datos

Para desarrollar el proyecto, se recolectarán videos del UCF Crime dataset [30] y de YouTube, que incluirán grabaciones de cámaras de seguridad mostrando intrusiones en hogares (como intentos de forzar puertas o ventanas) y conductas de merodeo. También se incluirán videos de actividades cotidianas, como la llegada de repartidores, con el objetivo de distinguir entre comportamientos sospechosos y normales.

Los videos tendrán una duración de entre 30 segundos y 2 minutos, con una resolución promedio de 720px. Se aplicará la ecualización del histograma a cada frame con el objetivo de mejorar el contraste y la calidad visual. Los frames mejorados reemplazarán a los originales, para que las versiones procesadas sean las que se utilicen en las siguientes etapas.

Posteriormente, los videos se dividirán en:

- Entrenamiento: Videos con mejor resolución y visibilidad
- Pruebas: Videos con menor resolución o calidad

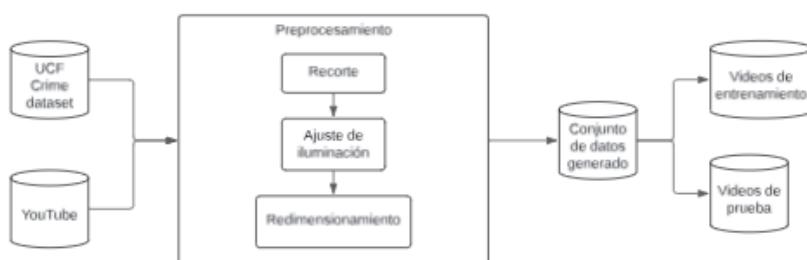


Figura 7: Diagrama de la construcción del conjunto de datos

3.2 Detección de objetos usando YOLO

Se utilizará YOLO detectar personas y manos. Aquí se generarán cuadros envolventes alrededor de las personas detectadas y para cada persona detectada, si es posible, se detectarán las manos; esto será la base para el seguimiento de movimientos y acciones.



Figura 8: Diagrama de la detección de objetos

3.3 Análisis y clasificación de movimientos

A partir de los recuadros envolventes, se calcularán los centroides de cada objeto y se calculará el flujo óptico para cada uno, se calculará en cada frame para visualizar sus trayectorias. Se realizará un análisis del movimiento del centroide para identificar el merodeo y de las manos para identificar intentos de forzar accesos.

Los datos generados se introducirán en una CNN para la clasificación de los comportamientos. Las condiciones para clasificar los comportamientos podrían ser:

- Merodeo: Se considerará merodeo si una persona permanece en la escena por más de 15 segundos.
- Forzar accesos: Se clasificará así si se detectan elementos de flujo óptico anormales, estos movimientos pueden ser de las manos o del cuerpo de la persona.



Figura 9: Diagrama de proceso de obtención de centroides y flujo óptico

3.4 Evaluación del algoritmo

El algoritmo se evaluará utilizando videos de prueba y simulaciones en entornos controlados. Se utilizarán métricas clave como: precisión, matriz de confusión, F1 score y tasa de falsos positivos

Firma de asesores para la aceptación de Reporte Técnico



Dr. Lauro Reyes Cocoletzi



Dra. Maria del Rocío Ochoa Montiel

REFERENCIAS

- [1] Código penal federal › libro segundo› título vigesimosegundo - delitos en contra de las personas en su patrimonio› capítulo i - robo› artículos 367 al 381 [Accedido el 15 noviembre 2024]. (2024). <https://mexico.justia.com/federales/codigos/codigo-penal-federal/libro-segundo/titulo-vigesimosegundo/capitulo-i/#articulo-367>
- [2] Computer vision [Accedido el 25 octubre 2024]. (2024). <https://www.ibm.com/topics/computer-vision>
- [3] Convolutional neural networks [Accedido el 17 octubre 2024]. (2024). <https://www.ibm.com/topics/convolutional-neural-networks>
- [4] de Estadística y Geografía, I. N. (s.f.). Incidencia delictiva [Recuperado el 28 de agosto de 2024]. <https://www.inegi.org.mx/temas/incidencia/>
- [5] de Gobernación, S. (n.d.). Si sospechas... ¡denuncia! <https://www.gob.mx/segob/articulos/si-sospechas-denuncia?idiom=es>
- [6] Detecta una actividad sospechosa. (n.d.). <https://efectoprevencion.guanajuato.gob.mx/al-salir-de-casa/detecta-una-actividad-sospechosa/>
- [7] El 2023 en México: A la baja la mayoría de los delitos, pero la alta impunidad subsiste [Accedido el 28 diciembre 2023]. (2023). *El Economista*. <https://www.economista.com.mx/opinion/El-2023-en-Mexico-a-la-baja-la-mayoria-de-los-delitos-pero-la-alta-impunidad-subsiste-20231228-0113.html>
- [8] El aumento de la adopción de la tecnología de inteligencia artificial (ia) genera expectación y pone de relieve la importancia de las conversaciones familiares sobre la seguridad online, según un nuevo estudio de microsoft – centro de noticias [Accedido el 6 febrero 2024]. (2024). <https://news.microsoft.com/es-es/2024/02/06/el-aumento-de-la-adopcion-dela-tecnologia-de-inteligencia-artificial-ia-genera-expectacion-y-pone-de-relieve-la-importancia-de-las-conversaciones-familiares-sobre-la-seguridad-online-segun-un-nuevo/>
- [9] Encuesta nacional de seguridad pública urbana 2023, datos al cuarto trimestre del año. (n.d.). https://www.inegi.org.mx/rnm/index.php/catalog/859/data-dictionary/F8?file_name=ENSU_CB_0623
- [10] Entrenamiento funcional patrones-mov-fund. (n.d.). http://www.saludmed.com/Entrena_II/Presentaciones/Entrenamiento_Funcional_PATRONES-MOV-FUND.pdf

- [11] Española, R. A. (2024). Diccionario de la lengua española, 23.^a ed., versión 23.7 en línea [Accedido el 23 octubre 2024]. <https://dle.rae.es/merodear#P1Q2EbV>
- [12] Gonzalez, R. C., & Woods, R. E. (2008). *Digital image processing*. Prentice Hall.
- [13] Jia Le, T. J. L. (2020). *Front yard surveillance system: Robbery scene detection* [Tesis de licenciatura]. Universiti Tunku Abdul Rahman.
- [14] Jurídicos, C. (2020). Artículo 285 del código penal – conceptos jurídicos [Accedido el 23 junio 2020]. <https://www.conceptosjuridicos.com/mx/codigo-penal-articulo-285/>
- [15] Kim, J. S., Kim, M.-G., & Pan, S. B. (2021). A study on implementation of real-time intelligent video surveillance system based on embedded module. *EURASIP Journal on Image and Video Processing*, 2021(35). <https://doi.org/10.1186/s13640-021-00576-0>
- [16] Kunst, M., & Hoek, D. (2024). Psychological distress among domestic burglary victims: A systematic review of possible risk and protective factors [PMID: 36847259; PMCID: PMC10666482]. *Trauma Violence Abuse*, 25(1), 430–447. <https://doi.org/10.1177/15248380231155525>
- [17] Lim, K. H. (2021). *Front yard robbery action detection climb-over-gate action analysis* [Tesis de licenciatura]. Universiti Tunku Abdul Rahman.
- [18] M, S., S, G., & Bharathy, S. (2022). Loitering detection in home surveillance system. *2022 10th International Conference On Emerging Trends In Engineering And Technology - Signal And Information Processing (ICETET-SIP-22)*. <https://doi.org/10.1109/icetet-sip-2254415.2022.9791651>
- [19] Mitchell, T. M. (1997). *Machine learning*. McGraw-Hill Science/Engineering/Math.
- [20] Observatorio interactivo de incidencia delictiva [Recuperado el 28 de agosto de 2024]. (n.d.). <https://delitosmexico.onc.org.mx/tendencia?unit=folders&indicator=researchFoldersRate&group=month&crime=4200&state=0&domain=>
- [21] Optical flow. (n.d.). <https://la.mathworks.com/discovery/optical-flow.html>
- [22] Parthe, R. (2024). The importance of centroid in image processing. *International Journal of Scientific Research in Engineering and Management*, 08(2). <https://doi.org/10.55041/IJSREM30775>
- [23] Police, M. (n.d.). Residential burglary facts. <https://www.met.police.uk/cp/crime-prevention/protect-home-crime/residential-burglary-facts/>
- [24] Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv*.
- [25] Secretaría de Gobernación. (n.d.). Incidencia delictiva [s. f.].
- [26] Shreyas, D., Raksha, S., & Prasad, B. (2020). Implementation of an anomalous human activity recognition system. *SN Computer Science*, 1(168). <https://doi.org/10.1007/s42979-020-00169-0>

REFERENCIAS

- [27] Singh, A. (2022, January). Top 5 object tracking methods - augmented ai - medium [Accedido el 4 enero 2022]. <https://medium.com/augmented-startups/top-5-object-tracking-methods-92f1643f8435>
- [28] Singh, R. (2024, January). A quick reference for bounding boxes in object detection [Accedido el 18 enero 2024]. <https://medium.com/@rajdeepsingh/a-quick-reference-for-bounding-boxes-in-object-detection-f02119ddb76b>
- [29] The state of ai in 2022—and a half decade in review [Accedido el 6 diciembre 2022]. (2022). <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review>
- [30] Sultani, W., Chen, C., & Shah, M. (2018). Real-world anomaly detection in surveillance videos. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [31] Szeliski, R. (2010). *Computer vision: Algorithms and applications*. Springer.
- [32] Wahyono, A., Harjoko, A., Dharmawan, F., Adhinata, F., Kosala, G., & Jo, K.-H. (2023). Loitering detection using spatial-temporal information for intelligent surveillance systems on a vision sensor. *Journal of Sensor and Actuator Networks*, 12(9). <https://doi.org/10.3390/jsan12010009>