

**PROPUESTA DE TEMA DE TESIS DE DOCTORADO
DOCTORADO EN CIENCIAS DE LA INGENIERIA**

TÍTULO DEL TEMA DE TESIS PROPUESTO

**Detección física de amenazas usando redes de videovigilancia
vecinales y vídeo análisis**

PROPONENTE

Adrián Núñez Vieyra

DIRECTOR DE LA TESIS

Doctor Juan Carlos Olivares Rojas

CODIRECTOR DE LA TESIS

Doctor Arturo Méndez Patiño

Detección física de amenazas usando redes de videovigilancia vecinales y vídeo análisis



Objetos físicos: personas y vehículos



Red de cámaras vecinales



Vídeo análisis

Los sistemas de videovigilancia ayudan a esclarecer
cierto tipo de delitos



shutterstock.com · 442920373



Los sistemas de videovigilancia requieren de personas para su operación



Sistemas de videovigilancia inteligentes



Sistemas de videovigilancia inteligentes, suelen generar saturación de alertas



Los sistemas de videovigilancia son útiles ya que el delito ha sido cometido



Nuestra propuesta

- Metodología
- Framework
- Modelo físico de amenazas

Los 3 problemas de la videovigilancia que abordamos:

- En varios procesos aún se requiere de la participación de personas
- Saturación de alertas de falsos positivos
- Son útiles hasta que se ha cometido el delito



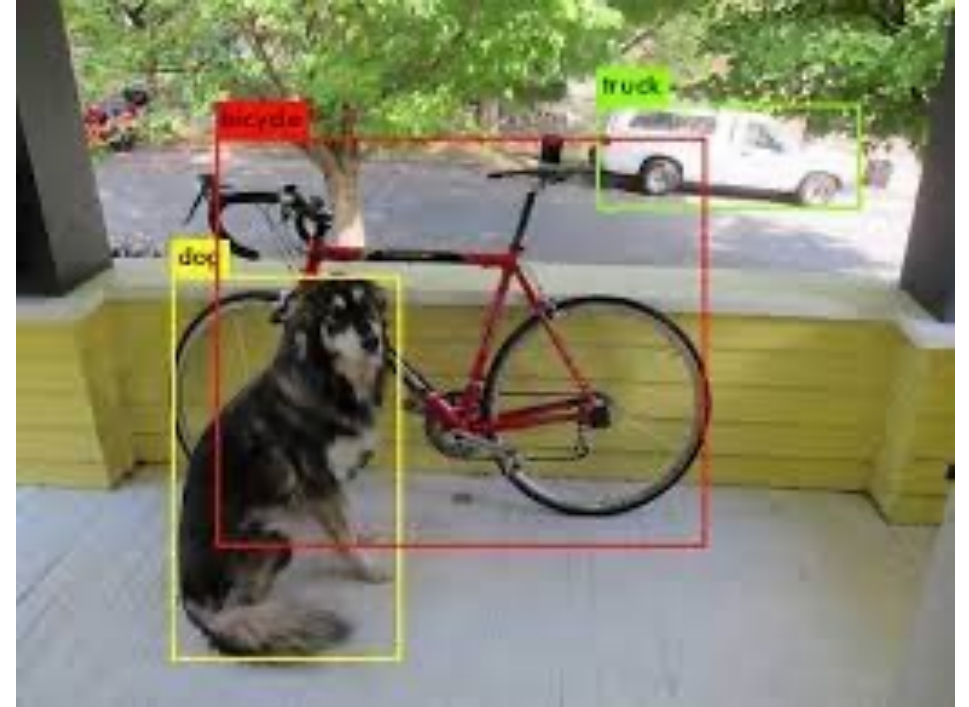
Introducción

Redes de videovigilancia vecinales



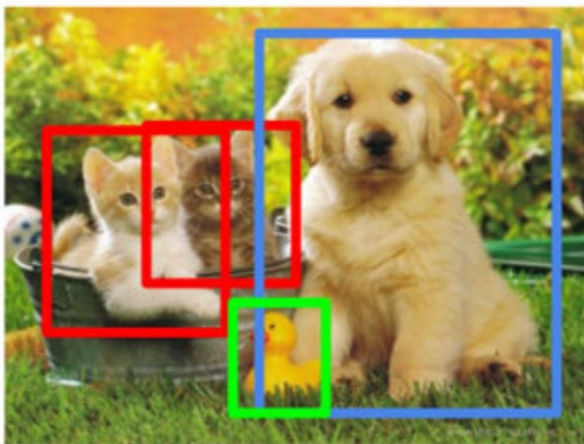
Vídeo análisis

Segmentación
Detección de objetos
Clasificación de objetos
Resumen automático de vídeo
Reconocimiento facial
Rastreo de objetos utilizando múltiples cámaras
Entendimiento o interpretación de la imagen



Vídeo análisis

Object Detection



CAT, DOG, DUCK

Classification



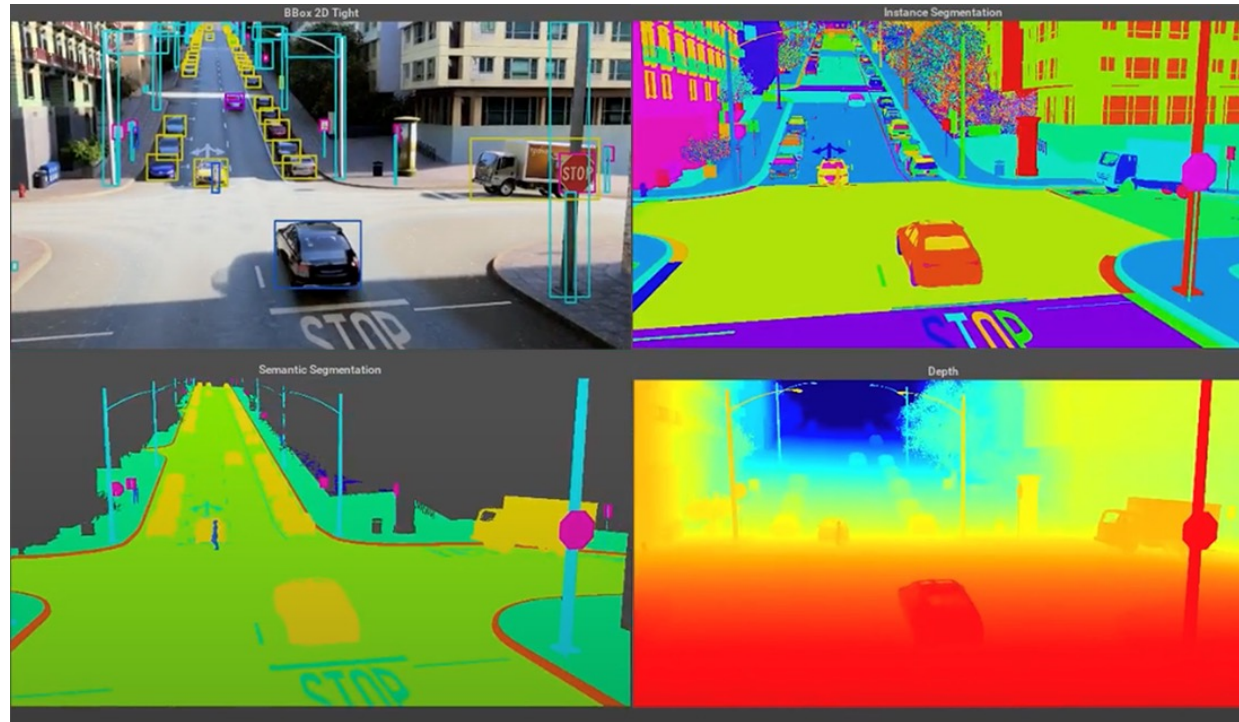
CAT

Entendimiento de una imagen

Nuestro proyecto

Semántica + Comportamiento + Contexto = Escenario

La semántica en una imagen



Comportamiento de los objetos (tracking)



Contexto

- Lugar
- Fecha
- Hora



Algunas estrategias para el vídeo análisis

	Tecnologías utilizadas para resolver las diferentes etapas del vídeo análisis					
	Detección/Clasif	Identificación	Tracking	Contexto	Procesamiento	Almacenamiento
[1]	Simplification of Background cut (min-cut)	SIFT descriptors	Geometrical analysis	Activitiy Clustering		
[2]	Viola-Jones algorithm (OpenCV)	Open CV	Optical Flow		Sony Xperia Z2, Android 5.1, Quad-core 2.3 GHz	
[3]	Color-Shape Based Object detection (OpenCV)	Deep learning	NA	Machine Learning and Big Data	Video Quality Enhancement (CLAHE)	
[4]		Person Reidentification (Chromatic and Epitomic Analyses)	Association Analysis Model (with historical data)	Pre alarming and selective storage video		Abnormal behavior database and selective storage
[5]	Neural Network	FaceNet (deep learning)			Jetson TX2	Datasets with images to train
[6]	AdaBoost algorithm	Deep Learning			Raspberry Pi and smart glasses	Storage Device (SD)
[7]	YOLOv8	YOLOv8		PHP, POST con Python. Ajax.	CUDA:0 Tesla T4	SSHD
[8]	GMM, GMG y ABL	GMM, GMG and ABL , all on TFRE approach	Temporal-Fluctuation-Reduced video-Encoding (TFRE)			
[9]	CNN, MoG	KCF (Kernelized Correlation Filter)	Sticky tracking and KCF and Kalman filter model	Specific queries on synopsis with Scene Labeling and Filtering of Tubes		Tube database
[10]	MoG in OpenCV		Predicción usando algoritmo recursivo	Trajectory-based video synopsis	Intel core i7-4770 CPU @ 3.4GHz	GRAM datasets
R3V	YOLOv8	OpenCV con generación automática de descriptores	Optical Flow Association Analysis Model	Ponderación semántica-comportamiento y datos históricos	Raspberry Pi5 y Jetson Orin Nano	Base de datos automática

Objetivos

1. Objetivo general

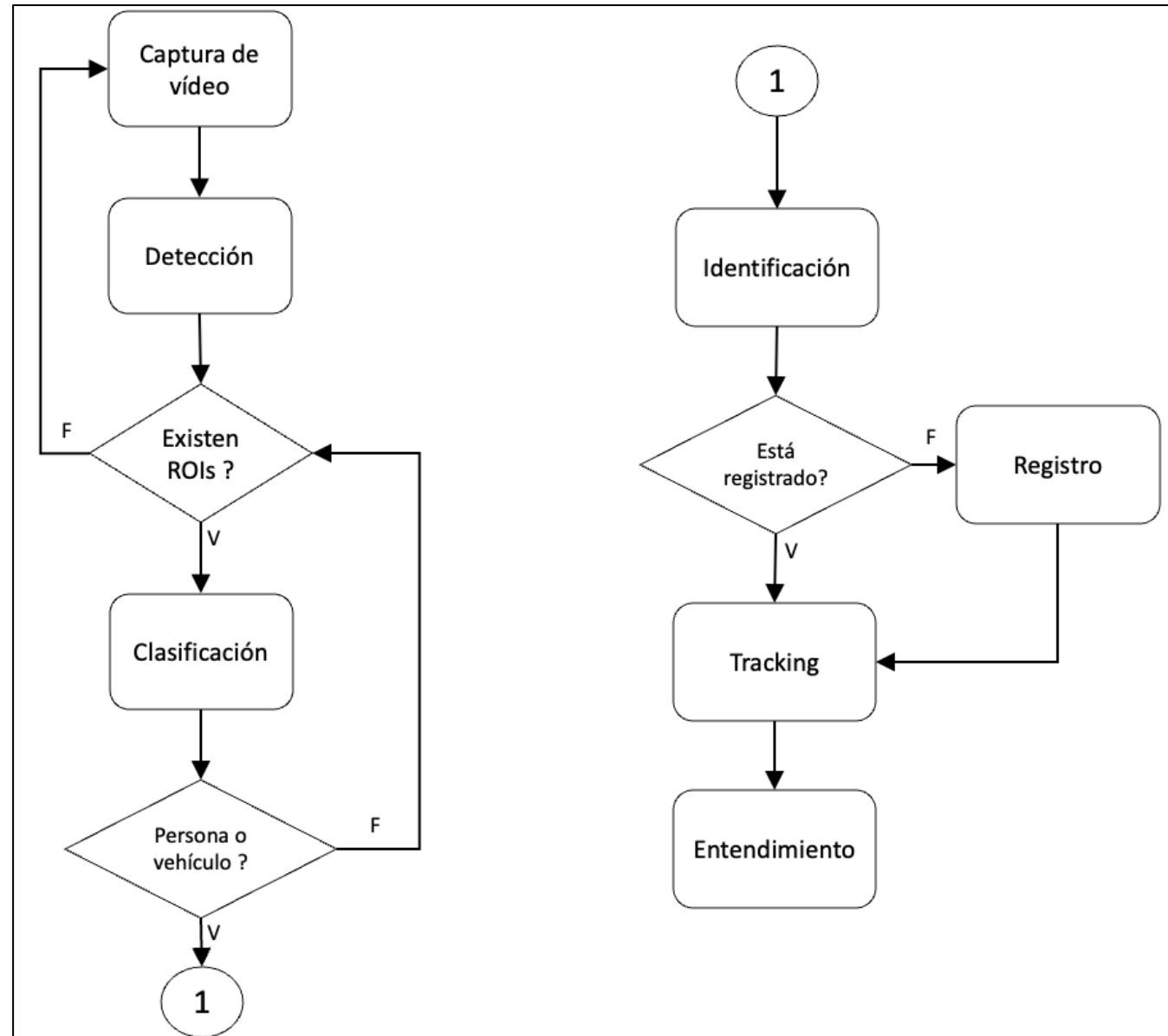
Contribuir a la seguridad pública de la ciudadanía, al desarrollar una metodología y un modelo, para detectar, prevenir e identificar potenciales amenazas físicas, utilizando una red vecinal de cámaras inteligentes

Objetivos particulares

1. Desarrollar un marco de referencia para el modelado físico de amenazas.
2. Documentar las principales técnicas de reconocimiento de amenazas físicas utilizadas en sistemas de videovigilancia.
3. Integrar los algoritmos adecuados para detectar y clasificar objetos a partir de imágenes capturadas por una cámara inteligente.
4. Desarrollar un mecanismo que asigne, de forma automática, un grado de amenaza a los objetos detectados por una red vecinal de cámaras inteligentes, en base al marco de referencia previamente establecido.
5. Desarrollar un mecanismo para determinar el grado de amenaza de un escenario en base al grado de amenaza de los objetos que lo conforman.
6. Desarrollar un sistema que permita identificar personas y matriculas de vehículos utilizando procesamiento en la nube, bases de datos locales y de terceros.
7. Desarrollar una herramienta que permita evaluar los resultados de los objetivos anteriores.

Metodología

Metodología propuesta para la detección de amenazas



Captura de vídeo



Tapo C200



Tapo C320

Detección de objetos

Algoritmos que se están probando

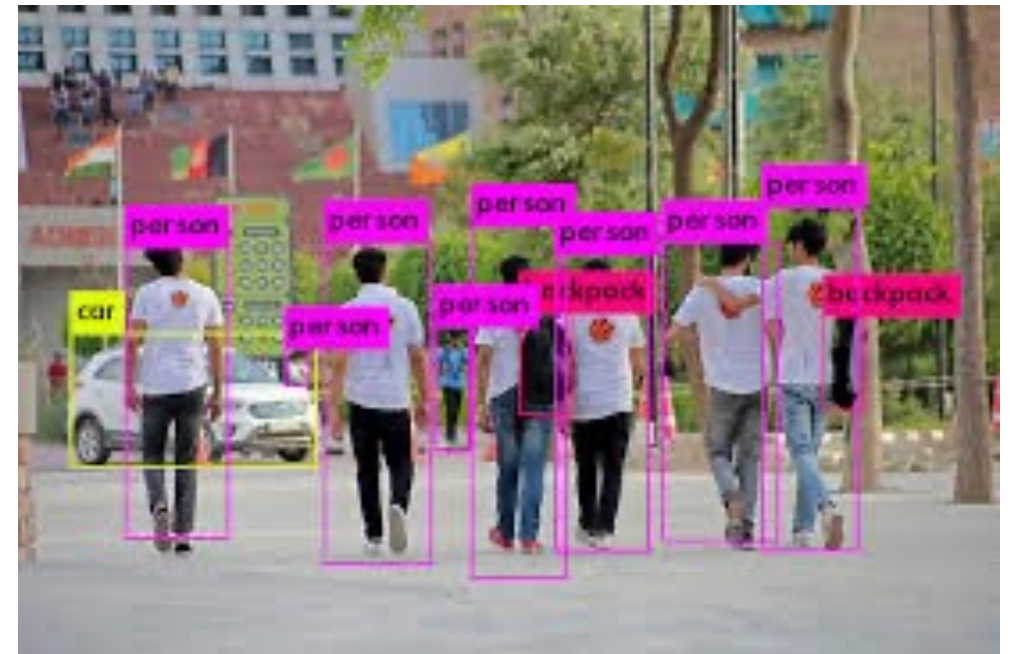
- a. Viola-Jones
- b. CNN (Convolutional Neural Network)
- c. HAAR features
- d. GMM (Gaussian Mixture Model)
- e. GMG Statistical Background Estimation
- f. ABL (Adaptive Background Learning)



Clasificación de objetos

Algoritmos que se están probando

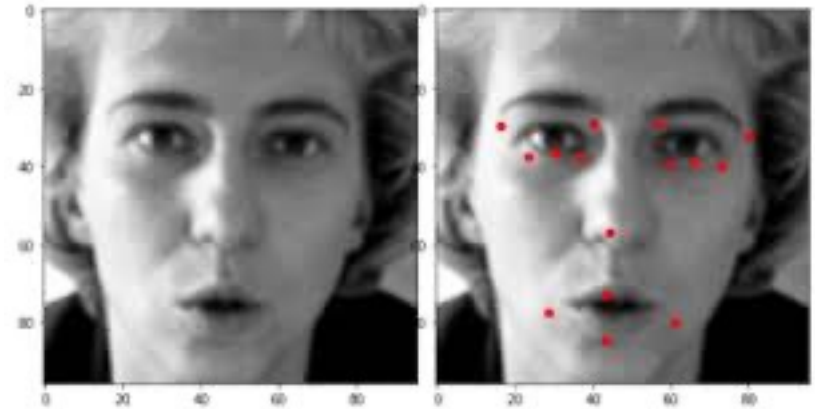
- a. CNN (Convolutional Neural Network)
- b. HOG (Histogram of Oriented Gradients)
- c. PCA (Principal Component Analysis)
- d. SURF (Speeded-Up Robust Features)
- e. ORB (Oriented FAST and Rotated BRIEF)
- f. HAAR Classifiers
- g. GMM (Gaussian Mixture Model)
- h. GMG Statistical Background Estimation
- i. ABL (Adaptive Background Learning)



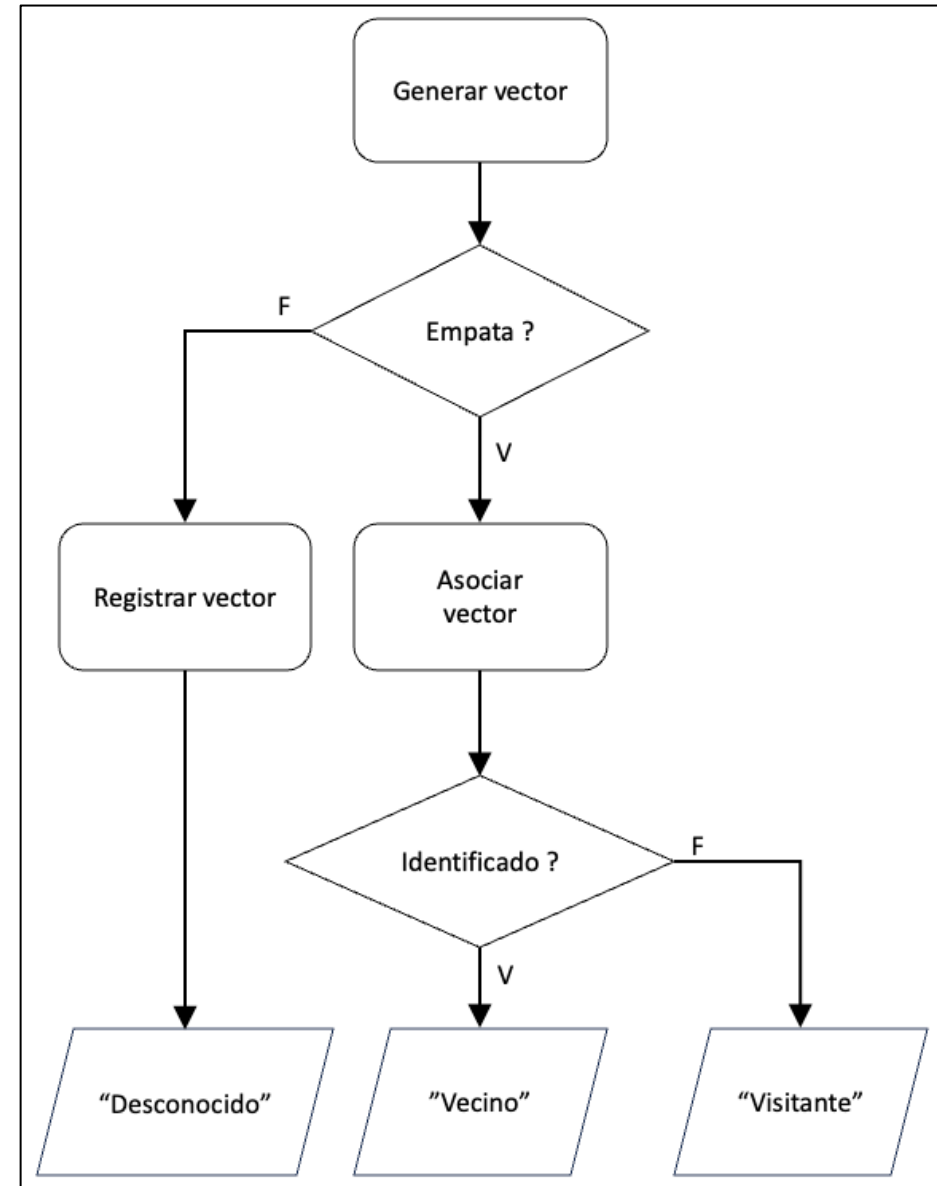
Identificación de objetos

Algoritmos que se están probando

- a. EIGEN valores
- b. HOG
- c. SVM
- d. CNN (AlexNet)
- e. CNN (FaceNet)



Etapa de registro



Re-Identificación de objetos (Tracking)

Algoritmos que se están probando

- a. Optical Flow
- b. CAMShift
- c. Sticky Tracking [9]
- d. Multi Object Tracking [34]
- e. Simple Online and Realtime Tracking [35]
- f. Association Analysis Model



Entendimiento

Common Vulnerability Scoring System (CVSS)

1. Amenazas constantes que no dependen del tiempo y el lugar.
2. Amenazas temporales, que en nuestro caso se ajustan a las condiciones contextuales de fecha y hora en que se hace la captura de vídeo.
3. Amenazas del entorno, en nuestro caso se refiere a las condiciones del lugar donde se realiza la captura de vídeo.

Entendimiento

1. Amenazas constantes = Personas y vehículos
2. Amenazas temporales = Fecha y hora de captura de las imágenes
3. Amenazas del entorno = Tipo de vecindario

Grado de amenaza de un escenario G_e

$$G_e = G_c \sum_{i=1}^n G_b(f_i) G_s(f_i)$$

$G_b(f_i)$ es el grado de amenaza por comportamiento para la i -ésima amenaza constante presente en un escenario

$G_s(f_i)$ es el grado de amenaza semántico para la i -ésima amenaza constante presente en un escenario

G_c es el grado de amenaza temporal y de entorno (amenaza contextual)

Entendimiento

Grado de amenaza contextual G_c

Afluencia vehicular anormal

$$\begin{aligned} n_v &< \mu_v - \sigma_v \\ n_v &> \mu_v + \sigma_v \end{aligned}$$

$$x_v = \frac{\sqrt{(n_v - \mu_v)^2}}{\sigma_v}$$

Afluencia peatonal anormal

$$\begin{aligned} n_p &< \mu_p - \sigma_p \\ n_p &> \mu_p + \sigma_p \end{aligned}$$

$$x_p = \frac{\sqrt{(n_p - \mu_p)^2}}{\sigma_p}$$

Normalizando con:

$$f(x) = \frac{2}{1 + e^{-x}} - 1$$

Afluencia peatonal

$$A_p = f(x_p)$$

Afluencia vehicular

$$A_v = f(x_v)$$

Fecha	Horario	Afluencia vehicular	Afluencia peatonal	Riesgo
No feriado	Mañana			
	Tarde			
	Noche			
	Madrugada			
Feriado o no laboral	Mañana			
	Tarde			
	Noche			
	Madrugada			

Definimos el grado de amenaza contextual (G_c), como el riesgo R (asignado a priori) ponderado por la afluencia vehicular y peatonal:

$$G_c = (A_v)(A_p) (R)$$

Entendimiento

Grado de amenaza semántico G_s

Tipo de objeto	Semántica del objeto	Grado de amenaza semántico (G_s)
Persona	Vecino	0.1
	Conocido	0.3
	Desconocido	0.6
	No identificable	0.9
Vehículo	Identificado	0.1
	Conocido	0.3
	Desconocido	0.6
	No identificable	0.9

Entendimiento

Grado de amenaza por comportamiento G_b

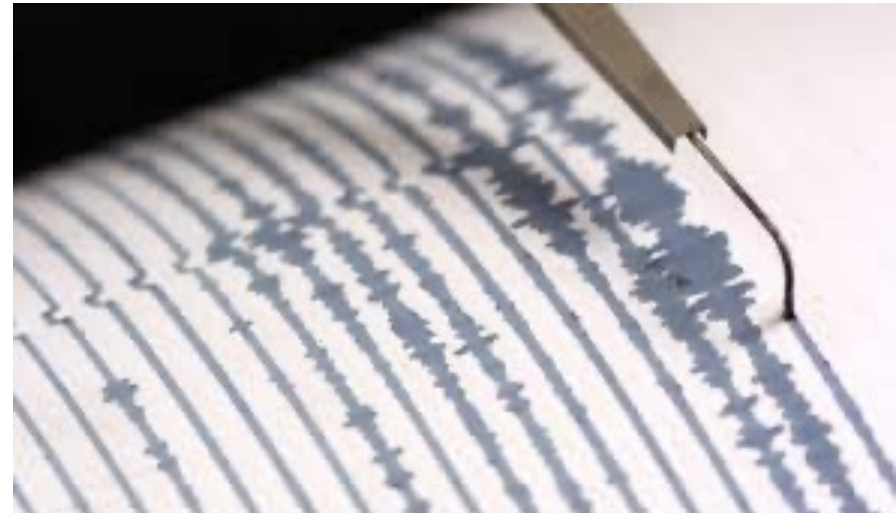
Objeto	Actividad	Grado de amenaza por comportamiento (G_b)
Persona	Caminar	0
	Correr	4
	Detenerse	2
	Invadir área privada	10
Vehículo	Avanzar lento	4
	Avanzar rápido	2
	Estacionarse	6
	Detenerse	8

Entendimiento

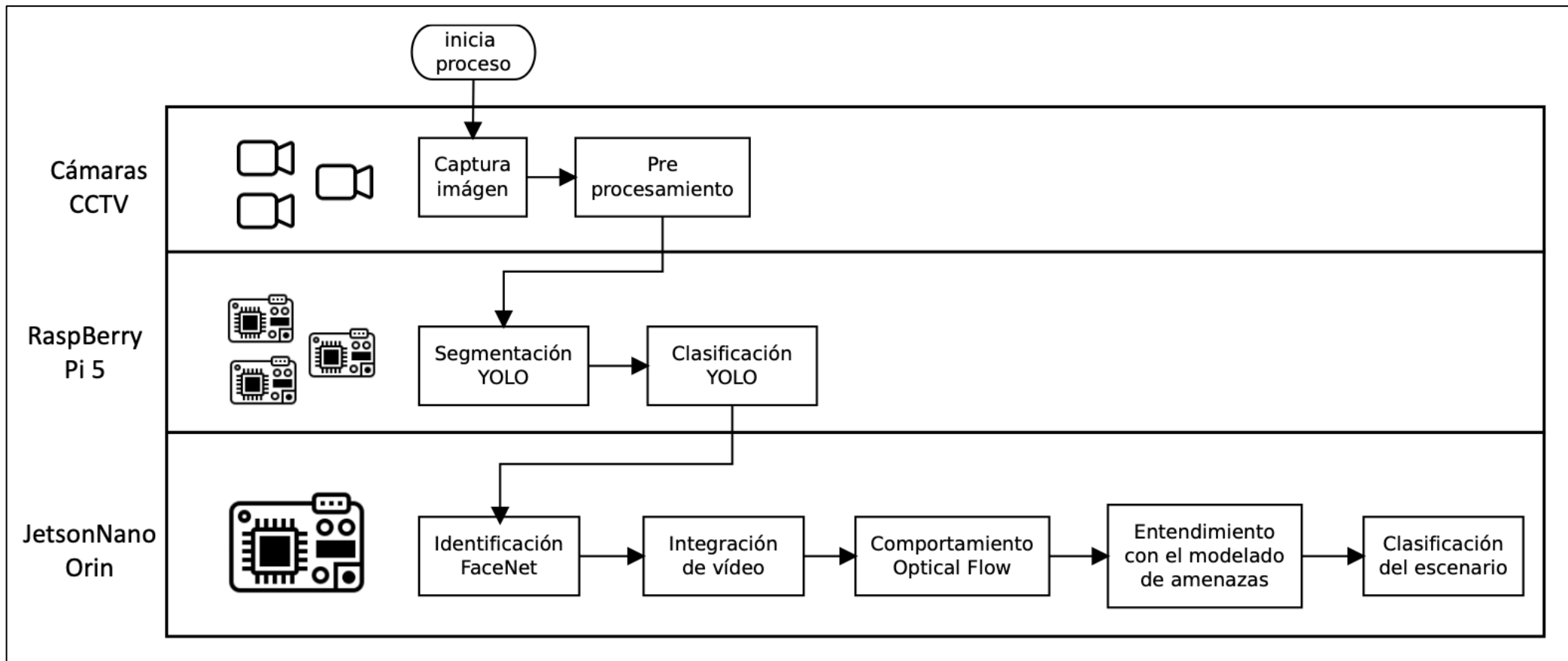
1. En cuanto se detecta una nueva persona o un vehículo se calcula su grado de amenaza semántica asociándole un ID fijo obtenido de una base de datos (en caso de ser vecino o conocido)
2. En caso de ser una amenaza etiquetada como desconocida o no identificable, se le asigna un ID temporal, aunque seguirá intentando identificarlo en las siguientes imágenes donde aparezca (tracking).
3. Para todos los objetos con un ID previamente asignado (sea temporal o fijo), el sistema intentará obtener su grado de amenaza de acuerdo a su comportamiento
4. Finalmente, en cada imagen, el sistema integrará los datos semánticos, contextuales y de comportamiento y calculará el grado de amenaza del escenario en ese momento, en base a los datos de los objetos presentes.

Grado de amenaza de un escenario G_e

$$G_e = G_c \sum_{i=1}^n G_b(f_i) G_s(f_i) \quad \approx$$



Procesamiento



Metas

1. Reducir los índices delincuenciales a través del desaliento a los malhechores y la prevención de las posibles víctimas.
2. Utilizar y mejorar tecnologías de vídeo vigilancia existentes.
3. Contribuir en la generación de bases de datos para el entrenamiento de sistemas de aprendizaje automático.
4. Contribuir en el establecimiento de marcos de referencia para la clasificación del grado de amenaza de diferentes escenarios.
5. Facilitar la labor de las autoridades, proporcionando herramientas de vídeo análisis y de predicción.

Impacto

Tecnológico-Científico

1. Una metodología aplicable a redes de vídeo vigilancia CCTV, para la detección de amenazas en el rubro del robo a casa-habitación.
2. Un marco de referencia para el modelado físico de amenazas
3. Un estudio práctico, contrastando las técnicas más recientes de vídeo análisis automático, aportando a la elección y mejoramiento de las tecnologías en vídeo vigilancia.
4. Un sistema adaptivo para llevar a cabo vídeo análisis y clasificación de escenarios amenazantes en base a criterios establecidos en el modelado físico de amenazas.

Social-Económico

1. Prevención y disminución de delitos, mejorando la percepción de seguridad en las comunidades en las que se implemente.
2. Reducción en los costos de vídeo vigilancia, al integrar distintos CCTV en redes vecinales, extendiendo su alcance, compartiendo datos y procesándolos de manera local.
3. Disminución de los índices de robo, al desalentar a los delincuentes por la alta probabilidad de quedar al descubierto, expuestos y registrados en las redes vecinales de CCTV.
4. La solución propuesta permitirá innovar y potenciar sistemas CCTV, tanto en aquellos que ya estén funcionando, como en nuevos proyectos de videovigilancia, impactando favorablemente en la economía de las comunidades vecinales.

Resumen de actividades

[illegible]

Actividades para el siguiente semestre

- Continuar con el estado del arte en técnicas y algoritmos relacionados con el vídeo análisis.
- Probar algoritmos y técnicas de detección de objetos
- Probar algoritmos y técnicas de rastreo de objetos
- Implementar una red vecinal de videovigilancia
- Integrar pruebas en ambiente de la Raspberry y Jetson Orin Nano.
- Iniciar la redacción de la tesis, la unidad 1 que se refiere a Antecedentes e Introducción.

Referencias

- [1] Y. Pritch, S. Ratovitch, A. Hendel y S. Peleg, «Clustered Synopsis of Surveillance Video,» de *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*, Genova, Italy, 2009.
- [2] L. A. Elrefaei, A. Alharthi, H. Alamoudi, S. Almutairi y F. Al-rammah, Real-time face detection and tracking on mobile phones for criminal detection, Abha, Saudi Arabia: IEE Xplore, 2017.
- [3] X. Jianyu, L. Shancang y X. Qingliang, «Video-Based Evidence Analysis and Extraction in Digital Forensic Investigation,» *IEEE Access*, vol. 7, pp. 55432-55442, 2019.
- [4] Z. Shao, J. Cai y Z. Wang, «Smart Monitoring Cameras Driven Intelligent Processing to Big Surveillance Video Data,» *EEE Transactions on Big Data*, vol. 4, nº 1, pp. 105-116, 2018.
- [5] J. Edwin, M. Greeshma, T. P. Mithun-Haridas y M. H. Supriya, «Face Recognition based Surveillance System Using FaceNet and MTCNN on Jetson TX2,» *ICACCS*, pp. 608-613, 2019.
- [6] K. Suleman, J. M. Hammad, A. Ehtasham, S. Syed y U. A. Syed, «Facial Recognition using Convolutional Neural Networks and Implementation on Smart Glasses,» de *International Conference on Information Science and Communication Technology (ICISCT)*, Karachi, Pakistan, 2019.
- [7] G. Pérez-García, E. Escobar-Gómez, J. Sarmiento-Torres, I. Juárez-Ruiz y A. d. J. Flores-García, «Intelligent weapon detection system using computer vision,» *Tecnología y Ciencia Aplicada*, vol. 6, pp. 26-31, 2023.
- [8] L. Kong y R. Dai, «Object-Detection-Based Video Compression for Wireless Surveillance Systems,» *IEEE Multimedia*, vol. 24, nº 2, pp. 76-85, 2017.
- [9] S. A. Ahmed, D. P. Dogra, S. Kar, R. Patnaik, S.-C. Lee, H. Choi, G. P. Nam y I.-J. Kim, «Query-Based Video Synopsis for Intelligent Traffic Monitoring Applications,» *EEE Transactions on Intelligent Transportation Systems*, vol. 21, nº 8, pp. 3457-3468, 2020.
- [10] A. J. Rathnarajah, S. Goonetilleke, D. Tissera, K. Balagopalan y R. Rodrigo, «Moving Object Based Collision-Free Video Synopsis,» de *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Miyazaki, Japan, 2018.
- [34] A. J. Rathnarajah, S. Goonetilleke, D. Tissera, K. Balagopalan y R. Rodrigo, «Moving Object Based Collision-Free Video Synopsis,» de *IEEE International Conference on Systems, Man, and Cybernetics*, Miyazaki, Japan, 2018.
- [35] A. a. G. Z. a. O. L. a. R. F. a. U. B. Bewley, «Simple online and realtime tracking,» de *IEEE International Conference on Image Processing (ICIP)*, Phoenix, Arizona, USA, 2016.