



Facultad de Ciencias de la Administración

**Carrera de Ingeniería en Ciencias de la
Computación**

Prototipo de Modelo de Visión Computacional para
la Detección de Sospechas de Robos de Automóviles
en los Alrededores de la Universidad del Azuay

**Trabajo de titulación previo a la obtención del
grado de Ingeniero en Ciencias de la
Computación**

Autor:

Rómulo Sebastián Tapia Reyes

Director:

Marcos Patricio Orellana Cordero

Cuenca – Ecuador

Año

2026

DEDICATORIA

Quiero expresar mi más profundo agradecimiento a mis padres y hermanos, por su amor incondicional, apoyo constante y confianza en cada paso de este proceso. A mis amigos, gracias por su compañía, ánimo y por hacer más llevadero este camino. También agradezco a todas las personas especiales que conocí en el transcurso de este recorrido, quienes de una u otra forma dejaron una huella significativa en mi vida.

AGRADECIMIENTO

Expreso mi más sincero agradecimiento a mi director de tesis, Ing. Marcos Patricio Orellana Cordero, y a mi directora de taller de tesis, Ing. Catalina Verónica Astudillo Rodríguez, por su guía, apoyo y valiosos conocimientos brindados a lo largo de este proceso. A la Universidad del Azuay y a todos los miembros que hicieron posible la realización de este trabajo, por ofrecer el espacio y las herramientas necesarias para mi formación académica. De igual manera, agradezco a todos los docentes que formaron parte de mi trayectoria, por su dedicación, enseñanza y ejemplo, los cuales han sido fundamentales en mi desarrollo profesional.

Índice de Contenidos

DEDICATORIA.....	i
AGRADECIMIENTO	ii
Índice de Contenidos	iii
Índice de Figuras	iv
Índice de Tablas.....	v
RESUMEN	vi
ABSTRACT	vi
1. Introducción.....	1
2. Marco Teórico y Estado del Arte	2
2.1. Marco Teórico	2
Antecedentes de la Problemática.....	2
Fundamento de la Visión Computacional	4
Deep Learning y su Aplicación en Visión Computacional	6
Definición de Actividades Sospechosas.....	8
Factores Observables y Variables a Registrar.....	9
2.2. Estado del Arte.....	10
Casos de Éxito de Deep Learning Aplicado a Seguridad y Vigilancia.....	10
Limitaciones y Desafíos en la Detección de Comportamientos Sospechosos	12
Adaptaciones del Modelo YOLO a Entornos Vehiculares y Urbanos.....	13
3. Métodos.....	17
Preparación y Etiquetado del Dataset	19
Selección del Modelo y Entrenamiento Exploratorio.....	24
Entrenamiento con Validación Cruzada y Ampliación del Dataset.....	29
Validación con Cámaras	32
Añadido de la Variable de Tiempo, Tracking y Región de Interés (ROI).....	36
4. Resultados.....	39
5. Discusión	50
6. Conclusiones.....	57
7. Referencias	59

Índice de Figuras

Figura 1 Metodología del sistema de detección de actividades sospechosas en robo de automóviles basado en YOLO.....	18
Figura 2 Ciclo de vida de minería de datos (IBM Corporation, s. f.).....	19
Figura 3 Distribución de clases en el dataset completo - Primera versión.....	22
Figura 4 Tamaño de bounding boxes (valores normalizados 0..1) - Primera versión...	23
Figura 5 Distribución de anotaciones por imagen - Primera versión	23
Figura 6 Cámara de la Universidad del Azuay - Av. 24 de Mayo y Las Garzas	31
Figura 7 Cámara de la Universidad del Azuay - Las Garzas y Los Cisnes.....	32
Figura 8 Distribución de clases en el dataset completo - Última versión	34
Figura 9 Tamaño de bounding boxes (valores normalizados 0..1) - Última versión	35
Figura 10 Distribución de anotaciones por imagen - Última versión.....	36
Figura 11 Comparación de arquitecturas YOLO (Familias Nano y Small)	40
Figura 12 Evolución de métricas por versión (Segunda Ronda).....	42
Figura 13 Distribución de mAP50 por versión (K-Fold)	43
Figura 14 Efecto del tamaño del dataset sobre el rendimiento de detección (YOLO11s)	44
Figura 15 Dispersión general de resultados (3 rondas)	45
Figura 16 Validación específica en el campus de la Universidad del Azuay.....	47
Figura 17 Comparación de versiones (Segunda Ronda)	48
Figura 18 Comparación entre rondas (configuración con máximo volumen de datos) 49	
Figura 19 Comparativa temporal: K-Folds y Ronda Final.....	50

Índice de Tablas

Tabla 1 Arquitecturas YOLO evaluadas en la fase de selección del modelo base	25
Tabla 2 Hiperparámetros evaluados en el entrenamiento exploratorio	28
Tabla 3 Configuraciones evaluadas en el entrenamiento con validación cruzada.....	33
Tabla 4 Nomenclatura y descripción de los conjuntos de datos.....	38
Tabla 5 Comparación de experimentos de exploración - Primera Ronda (sin validación cruzada)	41
Tabla 6 Estadísticos descriptivos de la validación cruzada <i>K-Fold</i> ($K = 5$) por versión - Segunda Ronda.....	42
Tabla 7 Los diez mejores resultados individuales de las tres rondas, ordenados por mAP50.....	45
Tabla 8 Resultados de la Tercera Ronda: validación en imágenes del campus de la Universidad del Azuay	46
Tabla 9 Síntesis comparativa entre las tres rondas de experimentación.....	49

RESUMEN

La seguridad vehicular en los alrededores de los campus universitarios ecuatorianos constituye un problema creciente que los sistemas de videovigilancia tradicionales no logran atender de forma automatizada y preventiva. Los enfoques actuales basados en reconocimiento de matrículas o detección genérica de objetos alcanzan apenas cuatro de cada diez incidentes reales, lo que refleja una sensibilidad inferior al 60% en la clase de robo detectado. Ante esta necesidad, el estudio tuvo como objetivo desarrollar un prototipo de visión computacional basado en modelos YOLO para la detección automática de actividades sospechosas asociadas al robo vehicular en los alrededores de la Universidad del Azuay. Para ello, se construyó manualmente un dataset de 1,364 imágenes anotadas a partir de fuentes públicas y capturas del propio campus. Se evaluaron seis arquitecturas YOLO y se optimizó el mejor modelo mediante validación cruzada *Stratified K-Fold* ($K = 5$) en nueve configuraciones, complementándolo con un módulo de tracking con filtro de Kalman, umbral temporal de permanencia y región de interés poligonal configurable. YOLO11s alcanzó un mAP50 máximo de 94.68% (mAP50-95: 62.20%) con coeficientes de variación inferiores al 2.57% entre *folds*, lo que evidencia un desempeño estable y consistente. En la validación específica del campus, el mAP50 fue de 93.05%, superando al modelo de referencia en 1.84 puntos porcentuales y con una precisión de 91.48%. El sistema opera en hardware local, sin dependencia de la nube, genera alertas diferenciadas y puede adaptarse a nuevas cámaras, consolidándose como una solución escalable para la vigilancia preventiva en instituciones universitarias.

Palabras clave: actividades sospechosas, detección de objetos, robo vehicular, validación cruzada, visión computacional, YOLO.

ABSTRACT

Vehicle security in the surroundings of Ecuadorian university campuses has become a growing concern that traditional video surveillance systems fail to address in an automated and preventive manner. Current approaches based on license plate recognition or generic object detection capture only four out of ten real incidents, reflecting a sensitivity below 60% for the detected theft class. In response to this need, this study aimed to develop a computer vision prototype based on YOLO models for the automatic detection of suspicious activities associated with vehicle theft in the vicinity of the University of Azuay. To this end, a dataset of 1,364 annotated images was manually constructed from public sources and on-campus captures. Six YOLO architectures were evaluated, and the best-performing model was optimized using Stratified K-Fold cross-validation ($K = 5$) across nine configurations, complemented by a tracking module with a Kalman filter, a temporal persistence threshold, and a configurable polygonal region of interest. YOLO11s achieved a maximum mAP50 of 94.68% (mAP50-95: 62.20%) with coefficients of variation below 2.57% across folds, indicating stable and consistent performance. In the campus-specific validation, the mAP50 reached 93.05%, outperforming the baseline model by 1.84 percentage points and achieving a precision of 91.48%. The system operates on local hardware without cloud dependency, generates differentiated alerts, and can be adapted to new cameras, consolidating itself as a scalable solution for preventive surveillance in university institutions.

Keywords: computer vision, cross-validation, object detection, suspicious activities, vehicle theft, YOLO.