

Computer Vision System Proposal using Re-Identification techniques to improve Multi-Camera Vehicle Tracking Management in Trujillo, Peru

Angela A. Ticlla Córdova, Bachelor¹; Carla de La Torre Ugarte Molina, Bachelor¹; Jorge A. Zavaleta García, Bachelor¹; Juan D. Medrano Cajamune, Bachelor¹; Junior S. Cruz Cruz, Bachelor¹; Oscar A. Rodríguez Díaz, Bachelor¹; Rolando J. Berrú Beltrán, Master¹

¹Universidad Privada del Norte, Perú, n00279747@upn.pe, n00292873@upn.pe, n00274714@upn.pe, n00253364@upn.pe, n00292873@upn.pe, n00274594@upn.pe, rolando.berru@upn.edu.pe

Abstract– This research work described the issues regarding vehicle monitoring management in Trujillo, Peru, and aimed to propose a computer vision system to optimize multi-camera tracking in the year 2026. The study was descriptive-propositional with a quantitative approach; a questionnaire was applied to a non-probabilistic sample of 50 control center operators. The diagnosis revealed critical deficiencies in current operations, highlighting high discontinuity in inter-camera tracking (4.2) and visual fatigue in screen comparison (4.0), demonstrating the inefficiency of manual processes. In response, the UrbanSight technical proposal was designed, integrating the YOLOv11 model for vehicle detection and optimizing the BoT-SORT algorithm through Re-Identification (ReID) techniques trained with the Vehicle Re-Identification in Context (VRIC) dataset. It was concluded that this technological integration allows maintaining vehicle identity persistence across video surveillance networks, representing a viable technological alternative for automating traceability and potentially reducing reliance on the human factor.

Keywords– Computer vision, YOLOv11, Re-Identification, vehicle monitoring, BoT-SORT.

Propuesta de Sistema de Visión Computacional basado en técnicas de Re-Identificación para mejorar la gestión de seguimiento vehicular multi-cámara en Trujillo, Perú

Angela A. Ticlla Córdova, Bachiller¹; Carla de La Torre Ugarte Molina, Bachiller¹; Jorge A. Zavaleta García, Bachiller¹; Juan D. Medrano Cajamune, Bachiller¹; Junior S. Cruz Cruz, Bachiller¹; Oscar A. Rodríguez Díaz, Bachiller¹; Rolando J. Berrú Beltrán, Magister¹

¹Universidad Privada del Norte, Perú, n00279747@upn.pe, n00292873@upn.pe, n00274714@upn.pe, n00253364@upn.pe, n00292873@upn.pe, n00274594@upn.pe, rolando.berru@upn.edu.pe

Resumen– El presente trabajo de investigación describió la problemática de la gestión de monitoreo vehicular en Trujillo, Perú, y tuvo como objetivo proponer un sistema de visión computacional para optimizar el seguimiento multi-cámara en el año 2026. El estudio fue de tipo descriptivo-propositivo con enfoque cuantitativo; se aplicó un cuestionario a una muestra no probabilística de 50 operadores de centros de control. El diagnóstico evidenció deficiencias críticas en la operatividad actual, destacando una alta discontinuidad en el seguimiento entre cámaras (4.2) y fatiga visual en la comparación de pantallas (4.0), lo que demostró la ineficacia de los procesos manuales. En respuesta, se diseñó la propuesta técnica UrbanSight, la cual integró el modelo YOLOv11 para la detección vehicular y optimizó el algoritmo BoT-SORT mediante técnicas de Re-Identificación (ReID) entrenadas con el dataset Vehicle Re-Identification in Context (VRIC). Se concluyó que esta integración tecnológica permite mantener la persistencia de la identidad vehicular a través de diferentes nodos de videovigilancia, representando una alternativa tecnológica viable para automatizar la trazabilidad y potencialmente reducir la dependencia del factor humano.

Palabras clave– Visión computacional, YOLOv11, Re-Identificación, monitoreo vehicular, BoT-SORT.

I. INTRODUCCIÓN

A nivel global, el crecimiento acelerado del parque automotor y la complejidad de la movilidad urbana han sobrepasado la capacidad de respuesta de los sistemas de vigilancia tradicionales. La gestión de la seguridad vial se enfrenta a un escenario donde la supervisión manual resulta insuficiente para mitigar riesgos como colisiones, infracciones y congestión persistente. Según Balda y Urgilés [1], las limitaciones de los métodos convencionales radican en la incapacidad de procesar grandes volúmenes de datos de video en tiempo real, lo que puede generar brechas en el monitoreo preventivo. En este contexto, la necesidad de una vigilancia continua y precisa se ha vuelto una prioridad para el ordenamiento de las ciudades modernas, donde el retraso en la identificación de incidentes puede comprometer la integridad ciudadana [2].

En el contexto peruano, la gestión vial presenta retos estructurales debido a una marcada disparidad tecnológica entre

regiones. De acuerdo con el INEI [3], mientras que Lima Metropolitana cuenta con una infraestructura de videovigilancia superior al 50%, departamentos como Madre de Dios o regiones del interior apenas alcanzan el 9% de cobertura. Esta brecha evidencia una desprotección significativa en gran parte del territorio nacional. Además, la gestión actual en el país depende mayoritariamente del factor humano en los centros de monitoreo, lo que limita la objetividad y la capacidad de reacción ante eventos simultáneos. Esta realidad subraya la urgencia de adoptar herramientas que automaticen la supervisión para reducir la dependencia de procesos manuales propensos a la fatiga y al error [3].

A nivel de ciudades principales fuera de la capital, la implementación de estrategias tecnológicas ha demostrado ser un factor determinante para reducir la incidencia de accidentes. Investigaciones previas en entornos similares a Trujillo indican que la falta de herramientas de seguimiento automatizado impide una respuesta operativa eficaz [4]. La gestión basada únicamente en la observación pasiva de pantallas no permite una trazabilidad real de los vehículos, lo que facilita la impunidad ante infracciones y dificulta la labor de reconstrucción de incidentes viales en zonas densamente transitadas [4].

En el ámbito local, la ciudad de Trujillo presenta una gestión de monitoreo vehicular fragmentada e ineficiente. Según el Plan de Movilidad Urbana Sostenible de Trujillo [5], la infraestructura actual opera de manera aislada, sin una integración que permita el seguimiento continuo de unidades a través de diferentes cámaras. Esta situación se agrava por el crecimiento desmesurado del parque automotor en la región el cual, según registros de la SUNARP [6], ha sobrepasado la capacidad de control de los centros municipales. La carencia de un sistema que garantice la persistencia de identidades vehiculares en el tiempo y el espacio convierte a la vigilancia actual en una herramienta reactiva y limitada.

Por tal motivo, el presente trabajo se enfocó en el diseño del sistema UrbanSight, el cual propuso el uso de estrategias de visión computacional avanzada y técnicas de Re-Identificación para optimizar la gestión de monitoreo vehicular en la ciudad

de Trujillo, permitiendo una trazabilidad automatizada y eficiente.

II. ANTECEDENTES

Para fundamentar la propuesta, se revisaron investigaciones previas que abordaron la detección vehicular y el seguimiento de objetos, identificando sus aportes y limitaciones frente al contexto planteado.

Por una parte, la investigación desarrollada por Azimjonov & Özmen [7] se enfocó en superar las limitaciones de clasificación vehicular en los sistemas de monitoreo en tiempo real, para optimizar la planificación y el control del flujo en carreteras. Este estudio, realizado en Turquía, buscó transformar la gestión de monitoreo vehicular en autopistas mediante un enfoque híbrido que combinó el detector YOLO con un clasificador basado en Redes Neuronales Convolucionales (CNN) y un algoritmo de seguimiento por cuadros delimitadores. Los resultados obtenidos revelaron un incremento sustancial en la precisión de clasificación, pasando del 57% al 95.45%, lo que permitió obtener datos estadísticos más exactos para la toma de decisiones administrativas. En consecuencia, se determinó que la integración de algoritmos de alta precisión garantiza una gestión de tráfico más robusta y técnica, consolidándose como una herramienta esencial para la supervisión automatizada de infraestructuras viales complejas.

Por otro lado, la investigación de López et al. [8] se centró en optimizar la gestión de monitoreo vehicular mediante el uso de vehículos aéreos autónomos equipados con múltiples cámaras orientables de forma independiente. El estudio, desarrollado en la Universidad de Sevilla, España, planteó como objetivo generalizar una estrategia de optimización que permitiera la colaboración de diversos agentes en una misma misión de supervisión de objetivos móviles. Para ello, los autores propusieron una solución basada en una estrategia de optimización combinatorial aplicada a mecanismos de agrupación (clustering) y algoritmos de branch & bound, lo cual permitió asignar eficientemente los objetivos a las cámaras de los distintos vehículos y determinar sus posiciones óptimas en tiempo real. Tras las pruebas realizadas en el entorno de simulación Unreal Engine 5, se obtuvieron resultados que validaron la eficacia del sistema para mantener el encuadre óptimo de múltiples objetivos simultáneos, ajustando automáticamente la distancia focal y la posición de los agentes. Se concluyó que este enfoque de gestión multicámara incrementa significativamente la versatilidad y eficiencia operativa, facilitando la supervisión continua en escenarios complejos con un despliegue mínimo de recursos.

Adicionalmente, la investigación de Miguélez Sierra [9] abordó la gestión de monitoreo vehicular —extrapolada en este caso a la detección y seguimiento de objetivos en entornos multicámara— mediante la integración de información contextual para mitigar errores por oclusiones o cambios de iluminación. El estudio, desarrollado en la Universidad Autónoma de Madrid, España, se planteó como objetivo analizar cómo la combinación de detecciones de distintos

puntos de vista y el uso de la posición relativa entre dispositivos mejoran los resultados individuales de vigilancia. Para ello, el autor propuso una solución basada en la creación de un plano de suelo mediante homografías y la estimación de cilindros 3D para representar a los sujetos, lo cual permitió transferir información de una cámara a otra de forma consistente y mantener el volumen de los objetivos a pesar de los cambios de perspectiva. Los resultados obtenidos confirman una mejora en la precisión de las detecciones combinadas respecto a las cámaras independientes, reduciendo los falsos positivos y gestionando eficazmente las zonas de baja visibilidad. Se concluyó que el uso de información geométrica y contextual compartida en una red multicámara es fundamental para fortalecer la robustez y confiabilidad de los sistemas de supervisión en escenarios de alta complejidad.

Asimismo, el trabajo de investigación de Valderrama [10] propuso el desarrollo de un simulador basado en visión computacional para la detección, seguimiento y cálculo de distancia de objetos en entornos de tráfico. El estudio, realizado en Colombia, abordó la problemática de la dificultad para evaluar estrategias de control en escenarios viales reales de alto riesgo sin comprometer la seguridad. Para ello, el autor construyó una plataforma digital que utiliza el algoritmo YOLO para identificar y rastrear vehículos y peatones en secuencias de video sintéticas, permitiendo estimar distancias euclidianas y trayectorias de manera precisa. Los resultados obtenidos demostraron que el uso de simuladores de visión artificial facilita la optimización de los algoritmos antes de su despliegue físico, mejorando la robustez de la gestión de monitoreo vehicular. Se llegó a la conclusión de que la integración de estas herramientas tecnológicas permite una supervisión técnica más profunda y predictiva, esencial para la toma de decisiones en la planificación de infraestructuras viales inteligentes.

De igual manera, la investigación de Liu et al. [11] propuso un enfoque basado en aprendizaje profundo para la reidentificación progresiva de vehículos en sistemas de videovigilancia urbana. El estudio, desarrollado en China, identificó que la gestión de monitoreo vehicular convencional a menudo falla al ignorar identidades únicas (como la placa o atributos específicos) y al no poder rastrear un objetivo entre diferentes cámaras. Para mitigar esta deficiencia, los autores desarrollaron el método “PROVID”, el cual utiliza una búsqueda de lo general a lo particular en el espacio de características y una búsqueda de lo cercano a lo distante en el entorno real de vigilancia. Los resultados validaron que la integración de atributos de apariencia con la verificación de matrículas mejora significativamente la precisión en la identificación persistente de vehículos. Se concluyó que este avance es crucial para fortalecer la gestión operativa en ciudades inteligentes, permitiendo una supervisión continua y eficiente del transporte en redes de vigilancia a gran escala.

Por otra parte, el proyecto de investigación desarrollado por Zurita [12] tuvo como propósito el diseño e implementación de un sistema de monitoreo de tráfico en tiempo real basado en visión artificial y sensores de Internet de las Cosas (IoT). El

estudio, ejecutado en la ciudad de Quito, Ecuador, abordó la problemática de la gestión de monitoreo vehicular ineficiente, caracterizada por la falta de datos técnicos precisos y centralizados sobre el estado de las vías. Para solucionar esto, se construyó una arquitectura que utiliza el algoritmo YOLO para el seguimiento de unidades, integrada con una red de sensores y una plataforma web en la nube para la administración remota de la información. Los resultados demostraron que la integración de estas tecnologías permite una supervisión técnica constante y la recopilación de datos de velocidad en tiempo real, facilitando el control operativo de la movilidad urbana. Se llegó a la conclusión de que la combinación de visión computacional con infraestructura IoT robustece la gestión de monitoreo, proporcionando a las autoridades una herramienta técnica eficaz para la toma de decisiones estratégicas y la planificación del transporte.

Finalmente, el trabajo de investigación profesional de Hijar [13] tuvo como propósito el diseño de un prototipo de visión artificial basado en aeronaves no tripuladas para optimizar la precisión y los tiempos de respuesta en la identificación de automóviles. El estudio, desarrollado en la ciudad de Lima, Perú, abordó la problemática de las limitaciones de los sistemas de vigilancia fijos para la gestión de monitoreo vehicular, los cuales presentan restricciones de cobertura en áreas extensas o de difícil acceso. Para resolver este desafío, el autor construyó una plataforma tecnológica que integra cámaras aéreas con algoritmos de detección, permitiendo una supervisión dinámica y versátil. Los resultados obtenidos indicaron que el uso de drones expande significativamente el alcance de la supervisión, logrando procesar datos con alta fidelidad y en tiempos reducidos. Se llegó a la conclusión de que la incorporación de dispositivos móviles aéreos equipados con inteligencia artificial representa un avance crítico para modernizar la gestión de monitoreo, fortaleciendo la capacidad operativa y técnica en la infraestructura vial del país.

III. OBJETIVOS

A. *Objetivo General*

Describir las deficiencias de la gestión de monitoreo vehicular y los aspectos técnicos que se deben considerar de las estrategias de visión computacional avanzada para proponer el sistema UrbanSight en la ciudad de Trujillo en el año 2026.

B. *Objetivos Específicos*

- Dar a conocer las características y limitaciones actuales de la gestión de monitoreo vehicular en los centros de control de la ciudad de Trujillo en el año 2026.
- Identificar los aspectos teóricos y algoritmos de visión computacional avanzada que se puedan considerar para proponer una solución de re-identificación multicámara en la ciudad de Trujillo en el año 2026.
- Proponer el sistema de visión computacional UrbanSight para mejorar la efectividad de la gestión

de monitoreo vehicular y la trazabilidad de unidades en la ciudad de Trujillo en el año 2026.

IV. METODOLOGÍA

La presente investigación es de tipo descriptiva-propositiva, con un enfoque cuantitativo, la cual inició con un diagnóstico orientado a analizar las características y deficiencias técnicas en la gestión de monitoreo vehicular en la ciudad de Trujillo durante el año 2026. Asimismo, el estudio posee un alcance descriptivo, ya que permitió caracterizar de forma independiente las dimensiones asociadas a la capacidad de supervisión y la trazabilidad vehicular, identificando brechas en la infraestructura de vigilancia actual que limitan la eficiencia operativa de los centros de control. Con base en los resultados obtenidos en la etapa diagnóstica, se procedió al diseño de una propuesta tecnológica, sin ejecutar su implementación en un entorno real.

La población del estudio estuvo constituida por los operadores de los centros de control y agentes de tránsito encargados de la vigilancia vial en la ciudad de Trujillo durante el año 2026. Debido a restricciones institucionales y consideraciones de seguridad, no fue posible acceder a una cifra oficial exacta; sin embargo, las entidades involucradas proporcionaron un estimado poblacional aproximado de 250 operadores activos.

A partir de esta población, se seleccionó una muestra no probabilística por conveniencia, conformada por 50 operadores y agentes de tránsito que laboraban en zonas de alta congestión vehicular, lo que permitió recolectar información directamente de los responsables del proceso.

A. *Diagnóstico del problema*

Para garantizar la precisión del diagnóstico, la variable de gestión de monitoreo se operacionalizó a través de dos dimensiones principales. En primer lugar, se evaluó la dimensión Capacidad de Supervisión, enfocada en el desempeño del operador humano. Esta dimensión se midió mediante el indicador de tiempo de respuesta ante incidentes, el cual valoró la rapidez con la que el personal detecta y ejecuta acciones correctivas frente a eventos en la vía; y el indicador de nivel de fatiga visual, que permitió cuantificar el desgaste cognitivo generado por la revisión simultánea y prolongada de múltiples flujos de video.

En segundo lugar, se analizó la dimensión Trazabilidad Vehicular, orientada a examinar las capacidades técnicas del sistema actual. Para ello, se estableció el indicador de continuidad del seguimiento, que evaluó la capacidad del sistema para mantener la identidad única de un vehículo al transitar entre diferentes cámaras (re-identificación multicámara). De manera complementaria, se midió la eficacia de recuperación tras oclusiones, indicador diseñado para determinar la habilidad del sistema para retomar el rastro de un objetivo visual tras haber sido ocultado temporalmente por infraestructura u otros vehículos.

B. Conceptualización

En esta fase se obtuvo y describió el fundamento teórico de las tecnologías de Inteligencia Artificial necesarias para la solución. Se revisaron investigaciones internacionales sobre arquitecturas de detección de objetos y algoritmos de seguimiento. Con esta información previa, se determinó que la integración del modelo YOLOv11 junto con técnicas de Re-Identificación (ReID) constituyen los ejes temáticos adecuados para mejorar el proceso de seguimiento en el sistema propuesto.

C. Desarrollo de la propuesta

Basado en las deficiencias detectadas en el diagnóstico y los fundamentos de la conceptualización, se diseñó la arquitectura del sistema UrbanSight abarcando tres componentes de ingeniería claves. En primer lugar, se implementó el entrenamiento del modelo de Re-Identificación utilizando una arquitectura basada en ResNet como extractor de características, entrenado específicamente con el dataset Vehicle Re-Identification in Context (VRIC) [14] empleando dos unidades GPU NVIDIA T4; este modelo ReID tiene el propósito exclusivo de mantener la persistencia del ID del vehículo en el seguimiento multi-cámara (inter-cámara). En segundo lugar, se configuró el motor de detección YOLOv11 integrado con el algoritmo BoT-SORT, encargado de gestionar el tracking continuo de los vehículos dentro de una sola fuente de video, ajustando los hiperparámetros de confianza para reducir falsos positivos. Finalmente, la arquitectura de software se estructuró sobre un backend en FastAPI para orquestar la inferencia, utilizando PostgreSQL para la persistencia de trayectorias y Redis para la gestión eficiente de colas de video de alta concurrencia. Para validar la viabilidad técnica de la propuesta, el rendimiento del modelo de Re-Identificación se evaluó mediante las métricas mAP y Rank@k (Rank@1, Rank@5 y Rank@10). Adicionalmente, la eficacia del seguimiento integral frente al baseline se cuantificó a través de las métricas HOTA, MOTA e IDF1. Finalmente, el desempeño computacional de la aplicación se testeó en una GPU NVIDIA RTX 3060, midiendo los tiempos de inferencia, la tasa de procesamiento en cuadros por segundo (FPS), el porcentaje de utilización de la GPU y el consumo de memoria VRAM.

D. Consideraciones Éticas

La investigación se adhirió a principios éticos rigurosos, garantizando el anonimato y la confidencialidad de los datos proporcionados por los operadores encuestados. Asimismo, en el diseño técnico del sistema propuesto, se incorporaron criterios de privacidad por diseño, asegurando que el procesamiento de imágenes se centre exclusivamente en las características de los vehículos, difuminando rostros o placas que no sean objeto de interés forense, en alineación con la Ley de Protección de Datos Personales vigente.

V. RESULTADOS

A continuación, se exponen e interpretan los resultados obtenidos tanto del diagnóstico situacional de la seguridad vial en Trujillo como del diseño de la propuesta tecnológica denominada UrbanSight.

A. Diagnóstico del problema

Para identificar la problemática existente en los centros de control de tráfico, se procesaron los datos recolectados a través del cuestionario aplicado a la muestra de 50 operadores. Se evaluaron las dimensiones de Capacidad de Supervisión y Trazabilidad Vehicular utilizando una escala de Likert, donde un valor más alto indica una mayor dificultad o deficiencia percibida en el proceso actual.

TABLA I
RESULTADOS DE LA DIMENSIÓN: CAPACIDAD DE SUPERVISIÓN

Indicador	Ítem (Gestión Actual)	Definición	Prom.
Tiempo de Respuesta	1. Identificación de atributos	Rapidez para reconocer placas o rasgos únicos manualmente.	3.7
	2. Cruce de datos visuales	Tiempo en comparar un vehículo avistado con reportes previos.	3.8
	3. Verificación de identidad	Fluidez para confirmar si un vehículo es el mismo captado antes.	3.3
Promedio			3.60
Fatiga Visual	4. Comparación de pantallas	Esfuerzo por cotejar simultáneamente varias cámaras.	4.0
	5. Rastreo ocular sostenido	Concentración para no perder el "ID visual" de un auto.	3.8
	6. Saturación por flujo	Error en la re-identificación ante múltiples objetivos similares.	3.7
Promedio			3.83

En la Tabla I, se observan los resultados de la dimensión "Capacidad de Supervisión". El indicador de Fatiga Visual presenta los promedios más críticos, destacando el ítem "Comparación de pantallas" con un puntaje de 4.0, seguido por el "Rastreo ocular sostenido" con 3.8. Esto sugiere que la carga cognitiva generada por la revisión manual simultánea de múltiples monitores constituye una limitación relevante. Asimismo, el tiempo de respuesta se ve afectado, con un promedio de 3.60, lo que lo que sugiere que el proceso manual puede ser lento y propenso al error humano.

TABLA II
RESULTADOS DE LA DIMENSIÓN: TRAZABILIDAD VEHICULAR

Indicador	Ítem (Gestión Actual)	Significado	Prom.
Continuidad	1. Seguimiento inter-cámara	Complejidad de rastreo entre distintos campos visuales.	4.2
	2. Puntos ciegos	Frecuencia de pérdida del objetivo sin herramientas Re-ID.	4.0
	3. Confusión de unidades	Error al diferenciar vehículos de similar color/modelo.	3.8
Promedio			4.00
Oclusiones	4. Obstáculos urbanos	Capacidad de mantener rastro ante bloqueos físicos.	3.9
	5. Tráfico denso	Discriminación de unidades en alta congestión.	4.1
	6. Recuperación de rastro	Éxito en retomar objetivos tras ocultamientos cortos.	3.7
Promedio			3.90

Por otro lado, la Tabla II muestra los resultados de la dimensión "Trazabilidad Vehicular". Se obtuvo un promedio general alto de 4.00 en el indicador de Continuidad, siendo el ítem "Seguimiento inter-cámara" el más crítico con 4.2. Esto sugiere que los operadores presentan dificultades para mantener el rastro de los vehículos al cambiar de campo visual entre cámaras. Del mismo modo, el indicador de Oclusiones arroja un promedio de 3.90, resaltando la dificultad para discriminar unidades en condiciones de tráfico denso (4.1). Estos resultados respaldan la necesidad de una herramienta automatizada que soporte la persistencia de la identidad vehicular.

B. Definición de la propuesta

Basado en las deficiencias detectadas durante la etapa de diagnóstico, específicamente la discontinuidad en el seguimiento y la saturación cognitiva de los operadores, se diseñó la solución tecnológica denominada UrbanSight. A diferencia de los sistemas de videovigilancia pasiva tradicionales, esta propuesta integra componentes de Inteligencia Artificial avanzados para automatizar la trazabilidad vehicular.

El núcleo técnico de la propuesta reside en el entrenamiento especializado del modelo de Re-Identificación (ReID). Para superar las limitaciones de los algoritmos de seguimiento estándar, que a menudo fallan ante oclusiones o cambios de iluminación, se implementó una red neuronal basada en la arquitectura ResNet como extractor de características (feature extractor). Este modelo fue entrenado utilizando el dataset VRIC [14], lo que permitió generar embeddings o vectores numéricos únicos para cada vehículo. Gracias a esto, el sistema fue diseñado para distinguir detalles finos —como la geometría del chasis, la parrilla frontal y la parte trasera— más allá del simple color, permitiendo al algoritmo BoT-SORT asociar la

identidad de un vehículo específico y reasignarle su ID correcto incluso después de desaparecer del campo de visión y reaparecer en otra cámara.

Para soportar este procesamiento, se diseñó una arquitectura de software distribuida y escalable. El flujo de datos inicia con un módulo de ingesta que captura las señales de video (RTSP) de las cámaras urbanas. Estas señales son procesadas por el motor de IA, el cual implementa el modelo YOLOv11 para la detección precisa de objetos (bounding boxes) y el modelo ReID integrado para el seguimiento. La orquestación de estos servicios se gestionó mediante un backend desarrollado en FastAPI, utilizando Redis como gestor de colas para manejar la alta concurrencia de video y PostgreSQL para la persistencia histórica de los eventos detectados.



Fig. 1 Interfaz principal de monitoreo multicámara.

A nivel funcional, se diseñaron interfaces de usuario orientadas a mitigar la fatiga visual diagnosticada. Como se aprecia en la Fig. 1, la interfaz principal permite el procesamiento y visualización simultánea de múltiples flujos de video. Esta centralización posibilita una supervisión sincrónica de distintos tramos viales en tiempo real, proporcionando al operador una vista unificada que reduce la necesidad de alternar manualmente entre pantallas físicas.



Fig. 2 Detección vehicular con YOLOv11.

En el aspecto operativo, el sistema garantiza la detección y reconstrucción de rutas sin depender de sensores externos como

el GPS. La Fig. 2 ilustra la capacidad del modelo YOLOv11 para identificar vehículos con alta precisión dentro del flujo vehicular real. El sistema documenta visualmente el desplazamiento continuo, asignando cuadros delimitadores que siguen al objetivo mientras este permanece en el encuadre.

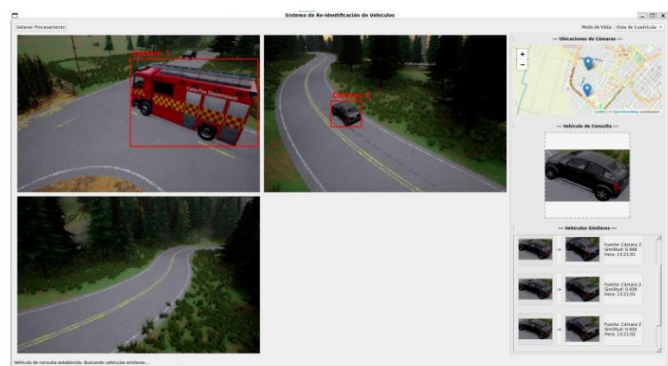


Fig. 3 Detección y seguimiento vehicular en tiempo real.

Uno de los aportes técnicos más relevantes de la propuesta es la capacidad de mantener la identidad del vehículo entre distintos puntos de vigilancia. Tal como se observa en la Fig. 3, el sistema conserva el mismo identificador numérico (ID) al pasar de una cámara a otra. Esta funcionalidad, sustentada en los embeddings visuales generados por la red ResNet, asegura la trazabilidad del vehículo a pesar de los cambios de ángulo o variaciones de luz, resolviendo directamente el problema de "Continuidad del seguimiento" evidenciado en el diagnóstico.



Fig. 4 Detección de similitud visual multi-cámara

Complementando el monitoreo activo, se integró un módulo de búsqueda forense para optimizar los tiempos de respuesta ante incidentes. La Fig. 4 presenta la herramienta de búsqueda por similitud visual, la cual permite al operador seleccionar un vehículo de interés y filtrar automáticamente todas sus apariciones en el histórico de grabaciones. Esto facilita la reconstrucción de la ruta completa del vehículo basándose únicamente en sus características físicas, eliminando la necesidad de revisar horas de video manualmente.



Fig. 5 Mapa de geolocalización de cámaras activas.

Para contextualizar geográficamente la cobertura, la Figura 5 presenta el mapa de ubicación de cámaras activas. Esta interfaz permitió visualizar las zonas monitoreadas e identificar de forma estratégica áreas críticas adicionales para la gestión de la seguridad vial en la ciudad.

C. Evaluación técnica

Para validar la eficacia del modelo ResNet entrenado con el dataset VRIC, se evaluó su capacidad de reconocimiento utilizando exclusivamente el conjunto de prueba del dataset VRIC [14] mediante las métricas de Precisión Media Promedio (mAP) y Rank@k. Los resultados obtenidos, detallados en la Tabla III, evidencian un mAP de 0.7116 (71.2%), indicando una sólida capacidad general de recuperación. Asimismo, la métrica Rank@1 alcanza un 0.6679 (66.8%), lo que significa que el modelo identifica correctamente al vehículo en su primera predicción en la mayoría de los casos, mientras que el Rank@5 asciende al 0.8619 (86.2%), garantizando una alta tasa de acierto en las primeras cinco coincidencias, algo fundamental para entornos multicámara con oclusiones severas.

TABLA III
RESULTADOS DE EVALUACIÓN DEL MODELO RE-ID

Métrica	Valor Obtenido
Rank@1	0.6679
Rank@5	0.8619
Rank@10	0.9049
mAP	0.7116

Para validar el impacto de la Re-Identificación aprendida, se comparó el sistema UrbanSight contra el método baseline de restricciones temporales en un escenario de 4 cámaras. Para asegurar una comparación equitativa y demostrar que la mejora reside únicamente en la lógica de asociación, ambos sistemas utilizaron el mismo motor de detección YOLOv11. Como se observa en la Tabla IV, la integración de ReID permite una mejora drástica en la persistencia de identidad (IDF1), superando al baseline por más de 34 puntos porcentuales. Mientras que el método tradicional falla al distinguir vehículos similares que transitan en intervalos de tiempo cercanos, UrbanSight mantiene el rastro gracias a la extracción de características visuales robustas.

TABLA IV
COMPARATIVA DE DESEMPEÑO: BASELINE VS. URBANSIGHT

Métrica	Baseline	Modelo ReID
HOTA	38.4%	58.7%
MOTA	61.2%	72.5%
IDF1	34.2%	68.4%

En cuanto al análisis de costo computacional, la aplicación fue desplegada en un entorno de pruebas utilizando una GPU NVIDIA RTX 3060 (12GB), procesando 4 entradas de video de

manera simultánea. Bajo esta configuración de hardware, el pipeline completo (que integra YOLOv11 y BoT-SORT para el seguimiento individual, apoyado por el modelo ReID para los cruces multicámara) demostró un rendimiento eficiente. Como se detalla en la Tabla V, el sistema logró procesar las cuatro fuentes concurrentes manteniendo una tasa de 22.2 FPS por canal, con un tiempo de inferencia por cuadro de 45 ms y un consumo de memoria de video (VRAM) de 5.4 GB, lo que representa apenas el 45% de la capacidad total de la tarjeta. Estos resultados demuestran que UrbanSight posee un costo computacional optimizado, siendo viable para implementaciones de borde (Edge Computing) en la infraestructura de videovigilancia existente en Trujillo sin requerir inversiones en hardware de gama alta.

TABLA V
MÉTRICAS DE RENDIMIENTO COMPUTACIONAL (4 STREAMS SIMULTÁNEOS)

Métrica de Rendimiento	Valor Promedio (RTX 3060)
Tiempo de Inferencia (Total Pipeline)	45 ms / cuadro
Tasa de Procesamiento (FPS)	22.2 FPS (por canal)
Consumo de Memoria VRAM	5.4 GB
Carga de Trabajo GPU (Utilización)	72%

VI. DISCUSIÓN

Con base en el análisis realizado durante la etapa de diagnóstico, se determinó que los operadores de los centros de control de tráfico en la ciudad de Trujillo enfrentaron una limitación crítica asociada a la discontinuidad en el seguimiento vehicular entre múltiples cámaras, lo que genera puntos ciegos en la trazabilidad y reduce la capacidad de respuesta ante incidentes viales. Estos hallazgos evidencian una alta dependencia del factor humano y de procesos manuales, los cuales resultan insuficientes frente al crecimiento del parque automotor y la complejidad del entorno urbano. Esta situación coincide con lo señalado por Paredes [4], quien identifica que la falta de integración tecnológica constituye una de las principales causas de ineficiencia en los sistemas de monitoreo vehicular tradicionales.

En relación con la variable temática, los resultados del diagnóstico justifican la necesidad de incorporar técnicas avanzadas de visión computacional, específicamente estrategias de Re-Identificación (ReID), como mecanismo para superar las limitaciones de los algoritmos de seguimiento convencionales aislados. Se observa una correlación directa entre la problemática operativa y los hallazgos técnicos: la alta 'discontinuidad en el seguimiento' (4.2) y la 'fatiga visual' (4.0) reportada por los operadores encuentran una solución cuantitativa en el IDF1 de 68.4% alcanzado por UrbanSight. Esta métrica de persistencia de identidad supera drásticamente al 34.2% obtenido por el modelo baseline; es imperativo notar que esta ventaja reside puramente en la lógica de ReID, ya que

ambos sistemas utilizaron el motor YOLOv11 para la detección. Asimismo, la robustez del modelo ResNet, validada con un mAP de 71.2% y un Rank@5 de 86.2% sobre el conjunto de prueba (test set) del dataset VRIC [14], garantiza que la pérdida de rastro inter-cámara se reduzca significativamente al mantener identidades persistentes. Estudios previos, como el desarrollado por Liu et al. [11], demostraron que los enfoques basados en aprendizaje profundo permitían mantener identidades estables incluso en escenarios con oclusiones y cambios de punto de vista. En este sentido, la adopción de modelos entrenados con datasets especializados, sumada a su bajo costo computacional validado en GPUs convencionales como la RTX 3060 procesando 4 canales en simultáneo, representa un factor clave para fortalecer la trazabilidad visual en sistemas de videovigilancia urbana sin requerir hardware de servidor inaccesible.

Desde la perspectiva propositiva, el diseño del sistema UrbanSight representa una evolución respecto a soluciones previamente reportadas en la literatura. A diferencia de propuestas que utilizan estimaciones puramente espaciales o simulaciones limitadas, esta investigación integró un modelo de Re-Identificación entrenado específicamente con el dataset VRIC, potenciando el algoritmo BoT-SORT para la identificación persistente de vehículos. Asimismo, la selección del modelo YOLOv11 para la etapa de detección se alinea con los avances recientes en detección de objetos en tiempo real, garantizando un procesamiento eficiente de múltiples flujos de video. En conjunto, la propuesta demuestra que la integración coherente de estas tecnologías permite transformar la videovigilancia pasiva en un sistema inteligente orientado al análisis forense y la toma de decisiones operativas.

VII. CONCLUSIONES

Se determinó, mediante el diagnóstico aplicado a los operadores de los centros de control de tráfico de la ciudad de Trujillo en el año 2026, que existen deficiencias significativas en la gestión de monitoreo vehicular, caracterizadas por la discontinuidad en el seguimiento entre cámaras y una alta dependencia del factor humano, lo que limita la capacidad de respuesta ante incidentes viales.

Se estableció que la incorporación de técnicas de Re-Identificación basadas en aprendizaje profundo, entrenadas con datasets especializados como VRIC, constituye un elemento clave para garantizar la trazabilidad visual de vehículos en entornos urbanos complejos, superando las limitaciones de los algoritmos de seguimiento tradicionales frente a oclusiones y variaciones visuales.

Se diseñó la propuesta del sistema UrbanSight, la cual integra el modelo YOLOv11 soportado por BoT-SORT para el tracking en video individual y la red neuronal de ReID para el seguimiento multicámara en una arquitectura escalable, sugiriendo que es técnicamente viable modernizar la gestión de monitoreo vehicular. Esta viabilidad técnica se comprobó al

lograr un rendimiento de ReID con un mAP de 71.2% y un Rank@1 de 66.8%, operando con tiempos de inferencia eficientes y un bajo consumo de VRAM sobre hardware de gama media (RTX 3060) incluso al procesar 4 cámaras en simultáneo a una tasa superior a los 20 FPS. La implementación de estas soluciones de software basadas en visión computacional avanzada representa un aporte potencial, contribuyendo al fortalecimiento de la seguridad vial y a la planificación urbana inteligente en la ciudad de Trujillo.

AGRADECIMIENTO

Se extiende un agradecimiento a la Municipalidad Provincial de Trujillo por el apoyo brindado durante la fase de diagnóstico, facilitando la información técnica necesaria y permitiendo la aplicación de los instrumentos de recolección de datos al personal de monitoreo, lo cual fue fundamental para caracterizar la problemática. Además, se agradece de manera especial al Ing. Segundo Castillo Zumarán por su contribución en la planificación y seguimiento de la propuesta elaborada.

REFERENCIAS

- [1] Balda Macias, M. A., & Urgilés Cedeño, J. (2023). Diseño de software para la Gestión Administrativa y Control Vehicular para una Terminal Terrestre. *Revista Científica Sinapsis*, 23(1). <https://doi.org/10.37117/s.v23i1.808>
- [2] Pillo Guanoluisa, D. M., Mejía Sandoval, D. R., Puetate Huera, G. H., & Lucio Vásquez, E. M. (2025). Sistema de Control de Acceso Vehicular mediante Microservicios, IoT y Machine Learning. *METANOIA: REVISTA DE CIENCIA, TECNOLOGÍA E INNOVACIÓN*, 11(2), 239–257. <https://doi.org/10.61154/metanoia.v11i2.4057>
- [3] Instituto Nacional de Estadística e Informática. (2024). Perú: Anuario Estadístico de la Criminalidad y Seguridad Ciudadana, 2019-2023. Lima, Perú.
- [4] Santos Gonzales, C. E., & Chuquista Vela, M. L. (2024). Implementación del Sistema de Monitoreo Vehicular Mecabebos y su Influencia en una Empresa de Transporte Lima, 2023. *Memorias de la Conferencia Iberoamericana en Sistemas, Cibernética e Informática*, 114–121. <https://doi.org/10.54808/CISCI2024.01.114>
- [5] Municipalidad Provincial de Trujillo. (2020). Plan de Movilidad Urbana Sostenible de Trujillo (PMUS). <https://www.munitrujillo.gob.pe/Archivosvirtual/plandet/pmus.pdf>
- [6] SUNARP. (2024). Anuario Estadístico Registral: Parque Vehicular por Zona Registral. <https://www.sunarp.gob.pe/estadisticas.asp>
- [7] Azimjonov, J., & Özmen, A. (2021). A real-time vehicle detection and a novel vehicle tracking systems for estimating and monitoring traffic flow on highways. *Advanced Engineering Informatics*, 50(101393), 101393. <https://doi.org/10.1016/j.aei.2021.101393>
- [8] López Ruiz, J. F., Rico Ranea, J., Álamo Cantarero, T., Ortega Linares, M. G., & Vargas Villanueva, M. (2024). Monitorización mediante vehículos aéreos multicámara. Caso de uso de Unreal Engine. *Jornadas de Automática*, 45. <https://doi.org/10.17979/ja-cea.2024.45.10800>
- [9] Sierra, A.M. (2016). Detección de personas en entornos multicámara utilizando información contextual. https://repositorio.uam.es/bitstream/handle/10486/669995/Migueluez_Sierra_Alejandro_tfg.pdf;sequence=1
- [10] Valderrama García, L. (2024). Simulador con visión computacional para detección, seguimiento y cálculo de distancia de objetos en movimiento. *European Public & Social Innovation Review*, 9, 1–16. <https://doi.org/10.31637/epsir-2024-812>
- [11] Liu, X., Liu, W., Mei, T., Ma, H. (2016). A Deep Learning-Based Approach to Progressive Vehicle Re-identification for Urban Surveillance. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds) *Computer Vision – ECCV 2016*.

- ECCV 2016. Lecture Notes in Computer Science(), vol 9906. Springer, Cham. https://doi.org/10.1007/978-3-319-46475-6_53
- [12] Morales, Z., & Eduardo, L. (2024). Diseño e implementación de un sistema de monitoreo de tráfico en tiempo real que utilice visión artificial y sensores IoT. Quito, Ecuador: Universidad Tecnológica Israel. <http://repositorio.uisrael.edu.ec/handle/47000/4257>
- [13] Hajar Torres, F. D. (2025). Prototipo de un sistema de visión artificial basado en un dron para mejorar la precisión y tiempo en la detección de vehículos. <https://hdl.handle.net/20.500.12867/14725>.
- [14] Kanaci, A., Zhu, X., & Gong, S. (2018). Vehicle Re-Identification in Context. In Pattern Recognition: 40th German Conference, GCPR 2018, Stuttgart, Germany, September 10-12, 2018, Proceedings (pp. 377-390). Springer. https://doi.org/10.1007/978-3-030-12939-2_26