



**Inteligencia de Clientes en la Tienda Física: Un Modelo Estratégico con Visión
por Computadora para el Sector Minorista en Colombia**

Diana Paola Aldana Chaparro

Leidy Johanna Galindo Huertas

Universidad Ean

Facultad de Administración, Finanzas y Ciencias Económicas e Ingeniería

Programa

Bogotá, Colombia

Agosto/2025

**Inteligencia de Clientes en la Tienda Física: Un Modelo Estratégico con Visión por
Computadora para el Sector Minorista en Colombia**

Diana Paola Aldana Chaparro

Leidy Johanna Galindo Huertas

Trabajo de grado presentado como requisito para optar al título de:

Magister en Dirección y Administración de Empresas y Magister en Ciencia de Datos

Director (a):

PAULA ECHEVERRY PEREZ

Modalidad:

Artículo de Investigación

Universidad Ean

Facultad de Administración, Finanzas y Ciencias Económicas e Ingeniería Facultad

MBA y MBA en Ciencia de Datos

Bogotá, Colombia

Agosto/2025

Nota de aceptación:

Firma del jurado

Firma del jurado

Firma del director del trabajo de grado

Bogotá, 15/08/2025

No tienes que ser excelente en algo para
empezar, pero sí tienes que empezar para
llegar a ser excelente en algo.

Zig Ziglar.

Agradecimientos

Expresamos nuestro agradecimiento a nuestra tutora, Paula Echeverry Pérez, por su invaluable orientación y constante apoyo durante el desarrollo de esta investigación. Su riguroso seguimiento y experta tutoría fueron fundamentales para la consecución de cada uno de los objetivos planteados.

Asimismo, extendemos nuestro agradecimiento a la empresa Nalsani S.A.S. por su generosa colaboración al facilitarnos el acceso a las grabaciones de sus cámaras de seguridad. Este material fue un recurso indispensable para la validación empírica del modelo y la realización de las pruebas necesarias para este estudio.

Finalmente, expresamos el más profundo agradecimiento a nuestras familias por su paciencia, comprensión y apoyo incondicional a lo largo de todo este proceso.

Su aliento fue una fuente constante de motivación.

Resumen

El sector del comercio minorista en Colombia, en su esfuerzo por competir con la riqueza de datos del comercio electrónico, enfrenta el desafío de transformar sus tiendas físicas en centros de experiencias memorables. Sin embargo, la efectividad de las estrategias de marketing sensorial y experiencial se ve limitado por la atención de herramientas para medir su impacto real en el comportamiento del cliente.

El presente artículo aborda esta brecha analítica mediante la validación de un sistema de visión por computadora que transforma grabaciones de video en inteligencia de negocio. Utilizando los algoritmos YOLOv8 para la detección y SORT para el seguimiento anónimo de personas, el sistema automatiza la generación de métricas clave como: mapas de calor, análisis de trayectorias y permanencia en la tienda, que permiten a la gerencia cuantificar la efectividad de sus decisiones sobre el diseño de la tienda.

Más allá del desarrollo técnico, el estudio se fundamenta en un doble marco estratégico que trasciende lo técnico, aplicando el modelo de aceptación tecnológica (TAM) para evaluar la viabilidad de adopción del sistema. Se concluye que la alta utilidad y facilidad de uso percibidas son factores críticos para asegurar su implementación por parte de los gerentes y, consecuentemente, el retorno de la inversión (ROI).

Finalmente, el trabajo incorpora un análisis de costos, proyecciones de escalabilidad y las consideraciones éticas y legales, en cumplimiento con la Ley 1581 de 2012 de protección de datos en Colombia. El resultado es un modelo interdisciplinario que integra la ciencia de datos, la estrategia de retail y la teoría de la gestión del cambio para ofrecer una solución robusta y aplicable al sector.

Palabras clave: reconocimiento facial, visión por computadora, comercio minorista, comportamiento del consumidor, detección de objetos, YOLOv8, seguimiento de objetos, modelo de aceptación de tecnología (TAM), marketing sensorial, estrategia de retail.

Abstract

Colombia's retail sector faces a critical strategic-analytical gap: while striving to create memorable in-store experiences to compete with e-commerce, it lacks effective tools to measure the impact of these strategies on customer behavior. This paper addresses this challenge by designing and validating a computer vision system that converts raw video footage into actionable business intelligence.

Leveraging the YOLOv8 algorithm for detection and the advanced BoT-SORT tracker for anonymous people tracking, the system automates the generation of key performance indicators, including customer traffic heatmaps, path analysis, and dwell times. These metrics enable management to empirically assess and optimize decisions related to store layout and merchandising.

The study is distinguished by a dual strategic framework that extends beyond the technical implementation. First, it employs the technology acceptance model (TAM) to evaluate the system's adoption viability from the end-user perspective. We argue that by focusing on high perceived usefulness (addressing core business needs) and perceived ease of use (providing automated, intuitive visual reports), the system is primed for successful adoption by retail managers, ensuring a positive return on investment (ROI).

The paper concludes with a comprehensive analysis of implementation costs, scalability, and ethical considerations, ensuring compliance with local data protection laws (Law 1581). We thus present an interdisciplinary model that integrates data science, retail strategy, and change management theory to deliver a robust and applicable solution for the retail industry.

Keywords: facial recognition, computer vision, retail, consumer behavior, object detection, YOLOv8, object tracking, Technology Acceptance Model (TAM), Sensory Marketing, Retail Strategy.

1. Introducción

El sector del comercio minorista (retail) se caracteriza por su dinamismo y competencia creciente, impulsados por la globalización y la transformación digital. En este escenario, la capacidad de una empresa para entender y anticipar las necesidades de sus clientes es fundamental para su supervivencia y éxito. Históricamente, el análisis del comportamiento del consumidor ha sido un enigma, especialmente en los espacios físicos. Como señalaba Forrester (2011), la "Era del cliente" que inicio en el 2010, otorgó al consumidor un poder sin precedentes para dictar las reglas del mercado, forzando a las empresas a crear mecanismos más sofisticados para escuchar y comprender.

Esta tendencia global adquiere una relevancia estratégica particular en Colombia, donde el sector minorista es un motor clave de la economía, llegando a aportar 1.1 puntos porcentuales a la variación anual del producto interno bruto (PIB) en el segundo trimestre de 2025 (DANE, 2025). Sin embargo, la inversión tecnológica se ha concentrado en el comercio electrónico, impulsado por el crecimiento récord del sector online, que alcanzó los \$72.3 billones de pesos en 2024, dejando la modernización analítica de la tienda física en un segundo plano (CCCE, 2025).

En consecuencia, los métodos tradicionales para recolectar datos en tiendas físicas presentan limitaciones significativas. Estos métodos, que incluyen encuestas de satisfacción, entrevistas a la salida, grupos focales y la observación directa por personal encubierto, suelen ser costosos, de baja escalabilidad y susceptibles a sesgos tanto del observador como del encuestado. Además, capturan solo una pequeña y a menudo no representativa fracción de los clientes, lo que compromete la fiabilidad de los resultados (Malhotra, 2019). Si bien soluciones tecnológicas intermedias como los contadores de personas por infrarrojos han intentado abordar este problema, su alcance es limitado, ofreciendo principalmente conteos de afluencia en lugar de un análisis conductual. Estos métodos pueden decirnos cuántos clientes entran, pero no ¿qué hacen? una vez dentro,

para lo cual se abordan estudios de mercado basados en encuestas. Estos métodos suelen ser costosos, de baja escalabilidad y susceptibles a sesgos. Por ejemplo, estudios como el de Help Scout (2024) indican que la tasa de respuesta promedio de las encuestas es tan solo del 4%, lo que pone en duda la representatividad de los datos recopilados y afecta la fiabilidad de los resultados. Esta brecha de información contrasta drásticamente con la riqueza de datos disponible en el comercio electrónico, donde cada clic, vista y compra es meticulosamente registrado y analizado.

Para cerrar esta brecha de datos en el entorno físico, las tecnologías de inteligencia artificial (IA) y visión por computadora emergen como una solución prometedora, capaces de transformar las cámaras de seguridad pasivas en sensores activos de inteligencia de negocio. En este contexto, destaca el algoritmo YOLO (You Only Look Once), que optimizó la detección de objetos al procesar la imagen en una sola pasada, permitiendo su aplicación en análisis de video en tiempo real. (Redmon et al., 2016).

El éxito y la versatilidad de YOLO se fundamentan en su entrenamiento con conjuntos de datos de imágenes del mundo real, como PASCAL VOC (Pattern Analysis, Statistical Modeling and Computational Learning Visual Object Classes) (Shasha, 2023), que fue creado por un equipo de investigadores de la universidad de Oxford como parte de la serie de desafíos anuales PASCAL VOC Challenge, que se llevaron a cabo entre 2005 y 2012, con el objetivo de evaluar métodos de aprendizaje automático aplicados a imágenes reales con objetos del mundo cotidiano, lo que le ha permitido una amplia adopción en múltiples industrias. El continuo desarrollo de esta arquitectura ha dado lugar a versiones optimizadas como YOLOv8, un framework de software de Ultralytics (Jocher et al., 2023) aplicada en este estudio. Una de las fortalezas clave de YOLOv8 es su flexibilidad, ofreciendo modelos de diferente tamaño que permiten una adaptación precisa a las necesidades de cada aplicación, permitiendo un balance estratégico entre la velocidad de procesamiento y la precisión. La versión nano (n), por ejemplo, está

diseñada para un rendimiento máximo en hardware con recursos limitados, mientras que la versión small (s) ofrece un equilibrio superior para obtener detecciones fiables.

Sin embargo, la detección por sí sola no es suficiente para analizar el comportamiento a lo largo del tiempo. Por ello, la capacidad de YOLOv8 de integrarse con algoritmos de seguimiento como SORT (Simple Online and Realtime Tracking) es fundamental (Bewley, Ge, Ott, Ramos, & Upcroft, 2016). Esta combinación permite no solo detectar a las personas en cada fotograma, sino también asignarles una identidad persistente para seguir sus trayectorias. Esta sinergia entre un detector rápido y un tracker eficiente convierte las grabaciones de video en los datos estratégicos que fundamentan este artículo.

El presente estudio aborda la necesidad de herramientas de análisis objetivas en el sector del comercio minorista en Colombia, explorando la viabilidad de implementar tecnologías de visión por computadora. El problema central radica en la histórica dificultad de las tiendas físicas para obtener métricas cuantitativas y fiables sobre la interacción del cliente con el entorno comercial, desde el diseño del espacio hasta la efectividad de las estrategias de comercialización (merchandising). Para superar esta limitación, este trabajo propone y valida un sistema de visión por computadora basado en el paradigma de seguimiento por detección (tracking-by-detection). Dicho sistema combina la eficiencia del detector de objetos YOLOv8 con la robustez del algoritmo de seguimiento BoT-SORT, una mejora del algoritmo SORT que incorpora un modelo de Re-identificación (Re-ID). Este modelo le permite al sistema “recordar” a una persona a pesar de posibles oclusiones, es decir, si desaparece temporalmente detrás de una columna o de otro cliente, evitando así contarla como un nuevo visitante cuando vuelve a ser visible, lo que reduce la fragmentación de trayectorias y minimiza el sobre conteo de visitantes (Nirkin, Y. 2022).

El objetivo general de esta investigación es desarrollar y evaluar un sistema basado en visión por computadora para el análisis objetivo del comportamiento del cliente en tiendas físicas, generando inteligencia de negocio aplicable al sector. Para alcanzar este propósito, el estudio se estructura en varias fases: primero, implementar un algoritmo de detección y seguimiento de personas; segundo, se diseñar un sistema para almacenar de forma estructurada los datos de sus trayectorias; tercero, se desarrollar una herramienta de análisis para procesar dichos datos y generar indicadores clave como conteos, mapas de calor y zonas de interés; y finalmente, analizar el valor estratégico de los resultados, incluyendo el costo de implementación y el retorno de inversión esperado.

A partir del desarrollo de estos objetivos se espera responder: ¿De qué manera el desarrollo y la validación de un sistema de visión por computadora, basado en YOLOv8, permite generar inteligencia de negocio accionable y financieramente viable para optimizar la estrategia comercial y operativa en el sector minorista de Colombia?

Para responder a esta pregunta, se plantean las siguientes hipótesis de trabajo:

Basado en la teoría del marketing sensorial, que postula que el diseño del espacio influye en la atención del cliente, se plantea la siguiente hipótesis:

H1 (Correlacional): la aplicación de un sistema de análisis de video permite la identificación de zonas de alto tráfico que se correlacionan positivamente con el interés del cliente, sirviendo como un predictor para la optimización de la ubicación de productos y promociones.

Considerando las limitaciones documentadas de los métodos de recolección de datos manuales, se propone que:

H2 (Descriptiva): el análisis automatizado de las trayectorias de los clientes proporciona métricas objetivas y escalables (conteo de personas y análisis de rutas) que son más eficientes y fiables que los métodos tradicionales de observación manual para entender el flujo de visitantes.

Integrando la viabilidad técnica con el marco de aceptación tecnológica y el análisis de negocio:

H3 (Correlacional): si se implementa un sistema de análisis de video con YOLOv8 en tiendas físicas de moda, se obtendrán datos cuantitativos útiles sobre el comportamiento del cliente que contribuirán a mejorar la toma de decisiones comerciales y demostrarán viabilidad financiera a través de un retorno de inversión favorable, condicionado a una adopción gerencial efectiva de la herramienta.

2. Marco teórico

2.1. La transformación estratégica del comercio minorista físico:

El sector del comercio minorista está evolucionando desde un enfoque centrado en el producto (las 4 "P": Producto, Precio, Plaza, Promoción) hacia un modelo centrado en el cliente, donde la experiencia de compra es primordial. Autores como Sharma y Grewal (2022) argumentan que la ventaja competitiva no reside únicamente en lo que se vende, sino en cómo se vende. La tienda física ya no es solo un punto de transacción, sino un escenario para la interacción, el descubrimiento y la construcción de lealtad a la marca (Verhoef et al., 2015).

La encuesta "Por qué son importantes los análisis de clientes" realizada en el año 2015 por la consultora McKinsey a 700 ejecutivos senior de marketing y ventas, permite evidenciar la tendencia en la inversión en el análisis del cliente y que su aplicación brinda mayor probabilidad de superar a los competidores (McKinsey & Company, 2015). Para la actualidad es una de las principales iniciativas de las empresas, el poder implementar estrategias efectivas para conocer a sus clientes, sus intereses, opiniones y expectativas, con el fin de identificar oportunidades de mejora y diferenciarse en el mercado (Trejo et al, 2023), razón por la cual, el mundo empresarial ya está adoptando el uso de nuevas tecnologías como la inteligencia artificial.

Jorge Mas Velasco, CEO de la consultoría estratégica de retail CREARMAS, afirma que "La tienda del futuro es la humanización del comercio con ayuda de la tecnología" (Mas, 2020). Es decir que se requiere la capacitación de personal humano para continuar con las actividades comerciales, pero se puede hacer uso de la tecnología para analizar más rápido el comportamiento del cliente y mejorar su experiencia.

2.2. Entendiendo al cliente en la tienda, marketing experiencial y sensorial:

El actual entorno competitivo, marcado por la hegemonía del comercio electrónico, obliga a las tiendas físicas a redefinir su propuesta de valor. La supervivencia ya no

depende exclusivamente del producto, sino de la capacidad de transformar el punto de venta en un destino en sí mismo. Como señalan Marín Dueñas y Gómez Carmona (2021), la tienda física debe evolucionar "de un simple punto de venta a un espacio donde vivir experiencias". Esta transición se apoya en el experiencial de marketing, una estrategia que busca crear momentos únicos y evocar reacciones emocionales para construir una relación más profunda y duradera con el cliente.

Dentro de esta estrategia, el marketing sensorial cobra una relevancia fundamental. Se define como el conjunto de técnicas que, a través de la estimulación de los cinco sentidos, buscan mejorar la experiencia de compra e influir positivamente en el comportamiento del consumidor (Marín Dueñas & Gómez Carmona, 2021). La ventaja competitiva del comercio minorista físico reside precisamente en esta capacidad de ventaja multisensorial, especialmente en la posibilidad de "tocar y sentir el producto", un factor decisivo en sectores como la moda.

Si bien la teoría del marketing sensorial establece el "qué" y el "porqué" de la creación de ambientes estimulantes, la gerencia del comercio minorista carece históricamente de herramientas para medir el "cómo": ¿cómo impactan realmente estas estrategias en el comportamiento observable del cliente? La observación manual es sesgada y poco escalable, y las encuestas dependen de la percepción subjetiva.

Es en esta brecha donde el presente estudio posiciona su contribución. La implementación de un sistema de visión por computadora, basado en algoritmos como YOLOv8, actúa como un puente tecnológico que conecta la teoría del marketing con la inteligencia de negocio accionable. Permite cuantificar, de manera objetiva y anónima, la respuesta del consumidor a las distintas tácticas sensoriales implementadas en la tienda.

A continuación, se detalla cómo las métricas generadas por nuestro sistema validan y miden la efectividad de los componentes clave del marketing sensorial descritos en la literatura:

Marketing visual y el análisis de la distribución de la tienda: Marín Dueñas y Gómez Carmona (2021) destacan que los elementos visuales: como la iluminación, la distribución del mobiliario y el diseño interior son claves para guiar el tránsito y captar la atención. Nuestro sistema traduce esta teoría en métricas concretas: los mapas de calor ofrecen una validación empírica de la efectividad del diseño, revelando las "zonas calientes" que realmente atraen a los clientes y las "zonas frías" que pueden requerir una intervención en el diseño. El análisis de trayectorias común mapea cuantitativamente el "tránsito del consumidor", permitiendo a la gerencia optimizar el camino hacia la compra y evaluar el impacto de cambios en la disposición de los productos mediante pruebas A/B.

Marketing atmosférico y tiempo de permanencia: un objetivo central del marketing sensorial es crear un ambiente agradable (a través de la música o aromas) para "incrementar el tiempo de compra en el establecimiento" (Marín Dueñas & Gómez Carmona, 2021). La métrica de tiempo promedio de permanencia, calculada por el sistema propuesto, cuantifica directamente el éxito de estas iniciativas, permitiendo correlacionar cambios en el ambiente con el comportamiento del cliente.

Marketing háptico y la interacción con el producto: el tacto es el "punto fuerte" de las tiendas físicas de moda. Aunque en este estudio no se mide la interacción táctil directamente, ofrece un apoderado altamente efectivo. Al analizar el tiempo de detención y la densidad de personas frente a percheros o estantes específicos, se puede inferir qué productos o categorías generan un mayor interés de interacción física, proporcionando datos valiosos para las decisiones de comercialización.

Este marco teórico integra la estrategia de negocio del marketing sensorial con la capacidad analítica de la ciencia de datos. El sistema propuesto no es solo una herramienta técnica, sino un activo estratégico que permite a los gerentes de las tiendas

pasar de la intuición a la evidencia, optimizando la experiencia del cliente y, en última instancia, maximizando el retorno de la inversión de su espacio comercial.

Tabla 1 Conexión entre marketing sensorial y la analítica del sistema.

Componente del marketing sensorial	Táctica aplicada en tienda	Métrica del sistema propuesto	Inteligencia de negocio generada
Marketing Visual	Diseño del layout, iluminación, exhibiciones.	Mapa de calor, Análisis de trayectorias.	Identificación de "zonas calientes" y "frías", optimización del flujo de clientes.
Marketing Atmosférico	Música, aromas para crear un ambiente agradable.	Tiempo promedio de permanencia.	Medición del impacto del ambiente en el tiempo que los clientes pasan en la tienda.
Marketing Háptico (Tacto)	Permitir la interacción directa con productos.	Tiempo de detención y densidad de personas frente a exhibidores.	Inferencia sobre qué productos generan mayor interés de interacción física.

Nota. La tabla resume la propuesta de valor del sistema al ilustrar cómo traduce las estrategias de marketing sensorial en indicadores de comportamiento cuantificables. Cada métrica proporciona una validación objetiva del impacto de las tácticas en la tienda.

Fuente: elaboración propia.

2.3. El fundamento técnico, visión por computadora con YOLOv8:

Richard Makerson, director ejecutivo y socio gerente de BlueFletch afirma que una de las tecnologías que puede ayudar a obtener mayor cantidad de datos de la experiencia del cliente en tiendas físicas y en tiempo real es el uso de reconocimiento facial y que aplicado al sector comercial puede ayudar a detectar patrones del comportamiento del cliente en su visita en la tienda (Makerson,2024). En este contexto, el reconocimiento facial surge como una de las aplicaciones más destacadas del aprendizaje automático y la visión por computadora que son algoritmos de inteligencia artificial. Esta tecnología realiza cuatro operaciones fundamentales. La primera es la detección facial, que

identifica y delimita el área del rostro en una imagen. La segunda es la caracterización o análisis facial, donde se extraen y analizan características específicas del rostro. La tercera es la verificación, que consiste en comparar una imagen facial con otra para confirmar si ambas pertenecen a la misma persona. Finalmente, la cuarta operación es la identificación, que implica comparar la imagen con una base de datos de rostros registrados para encontrar la coincidencia más cercana y determinar su identidad (Pérez y Madrid, 2021).

La aplicación práctica de este tipo de tecnología es el trabajo desarrollado por las consultoras NtechLab y MasterTek. Ambas identificaron que las organizaciones con tiendas físicas de todo el mundo coinciden en que los datos más valiosos son aquellos que proporcionan información sobre el comportamiento y los hábitos de los clientes, aunque estos suelen ser los más difíciles de obtener. Por ello, crearon el producto Countbuddy, una solución avanzada que utiliza cámaras de vigilancia y algoritmos de detección de objetos para analizar a los visitantes. Countbuddy genera informes detallados sobre la cantidad de clientes en una tienda y la composición de edad y género de los clientes en menos de un segundo y con una precisión del 95%. Esta tecnología ha sido implementada por la empresa brasileña Unike Technologies, que utiliza estos datos para ejecutar campañas de marketing efectivas y desarrollar estrategias comerciales (NtechLab, s.f).

De manera similar a nivel internacional, la startup KSI Vision, utiliza inteligencia artificial y visión por computadora sobre cámaras de seguridad existentes para proporcionar analítica avanzada a gigantes del retail como Inditex (propietaria de Zara) y Falabella. La solución de KSI Vision, que ya opera en Colombia, permite a las tiendas físicas entender desde qué zonas captan más atención hasta las causas de una venta perdida (Piazza, 2025).

A pesar de los beneficios evidentes, el reconocimiento facial plantea desafíos éticos significativos. Uno de los aspectos más debatidos es la tensión entre la seguridad pública y los derechos a la privacidad y la autonomía individual. En particular, en las democracias liberales se destaca la importancia de equilibrar la seguridad con la preservación de los derechos civiles, ya que el uso no regulado de esta tecnología podría invadir la privacidad de las personas al recolectar y almacenar información biométrica sin su consentimiento (Smith & Miller, 2021).

A este desafío se suma el riesgo del sesgo algorítmico, el cual puede originarse tanto en datos de entrenamiento no representativos como en el diseño mismo del modelo. Como lo expone Zafar (2024), se ha comprobado que algunos sistemas de reconocimiento facial presentan tasas de error significativamente más altas en poblaciones marginadas por motivos étnicos o de género. Este tipo de sesgo no solo refuerza desigualdades existentes, sino que también debilita la confianza pública en la tecnología y puede derivar en aplicaciones discriminatorias.

Por ello, en este tipo de soluciones se recomienda adoptar el principio de ética por diseño (*ethics by design*), lo que implica disociar por completo la analítica de la identidad personal. En lugar de utilizar algoritmos para reconocer rostros, se emplean técnicas que detectan la presencia de una persona como un objeto dentro de la imagen, sin capturar, procesar ni almacenar ningún dato biométrico facial (European Commission, 2021).

Adicionalmente a los desafíos de privacidad y sesgo algorítmico, emerge un debate legal y ético más profundo en torno a la naturaleza de los datos biométricos como una forma de propiedad personal o intelectual. Decisiones judiciales y propuestas regulatorias recientes, tanto a nivel internacional como en discusiones incipientes en Colombia, exploran si los rasgos únicos de un individuo (como el rostro o la forma de caminar) constituyen un activo del cual la persona es dueña. Aunque Colombia no reconoce explícitamente estos datos como propiedad intelectual, se considera que los datos

biométricos son “datos personales sensibles” cuyo tratamiento requiere autorización explícita del titular (Ibarra Pardo, 2024).

Esta perspectiva transforma el debate: el uso de datos biométricos ya no sería solo una cuestión de consentimiento para el tratamiento de datos, sino que podría implicar la apropiación y uso comercial de una propiedad ajena. Esto abriría la puerta a que las personas pudieran exigir una compensación o licenciar el uso de su identidad biométrica. Para las empresas, esto representa un riesgo legal y financiero significativo a futuro, ya que las bases de datos biométricas que hoy recolectan podrían convertirse en una fuente de litigios por derechos de propiedad (Ibarra Pardo, 2024).

La metodología de este estudio, al evitar deliberadamente el reconocimiento facial y cualquier forma de recolección de datos biométricos, se alinea proactivamente con este principio de precaución. Al tratar a los individuos como objetos anónimos, el sistema no solo cumple con la normativa de protección de datos vigente (Ley 1581 de 2012), sino que también mitiga los riesgos legales y reputacionales asociados a los emergentes debates sobre la propiedad de la identidad digital.

Es pertinente reconocer la existencia de metodologías alternativas para el análisis de comportamiento en espacios físicos. Algunas Investigaciones han explorado el uso de señales de radiofrecuencia (RF), como las emitidas por routers Wi-Fi o balizas Bluetooth (beacons), para realizar un seguimiento de personas de manera pasiva y sin el uso de cámaras (Ciurana et al., 2022). Estos métodos aprovechan la forma en que el cuerpo humano altera las señales inalámbricas para inferir presencia y movimiento, ofreciendo ventajas significativas en términos de privacidad al no capturar imágenes.

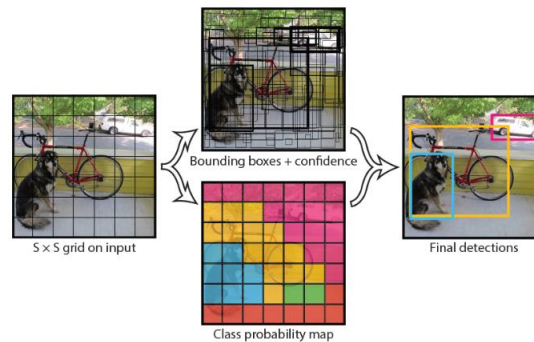
Sin embargo, a pesar de su potencial, el seguimiento basado en RF presenta limitaciones importantes para el análisis de alta granularidad que este estudio persigue. La resolución espacial de estos sistemas es inherentemente inferior a la de la visión por computadora, lo que compromete la capacidad de generar mapas de calor detallados y

de reconstruir trayectorias individuales con precisión, un desafío que se magnifica en entornos de alta densidad de personas.

Para implementar el enfoque de detección de objetos bajo las premisas de anonimato y protección de datos mencionadas, se seleccionó una de las herramientas más eficientes y robustas en el campo de la visión por computadora. El modelo escogido para este estudio es el algoritmo YOLO (You Only Look Once), cuya capacidad para identificar objetos en tiempo real y con alta precisión lo convierte en una solución idónea para cumplir con los objetivos del proyecto, sin comprometerse en ningún momento el marco ético y legal establecido.

YOLO representó un cambio de paradigma en la detección de objetos. A diferencia de los métodos anteriores de dos etapas como los algoritmos R-CNN (Regions with CNN features) y Fast R-CNN (Girshick et al., 2014; Girshick, 2015) que operaban bajo un paradigma de "propuesta y clasificación": primero, un algoritmo generaba un gran número de "regiones de interés" candidatas en la imagen y, en una segunda etapa, una red neuronal clasificaba cada una de estas regiones individualmente, lo cual generaba lentitud en el proceso y reducía sus aplicaciones en tiempo real. Por su parte, YOLO trata la detección como un único problema de regresión. Su arquitectura se basa en una sola red neuronal convolucional (CNN) que procesa la imagen completa en una sola pasada ("You Only Look Once"). Como se puede observar en la figura 1, en esta única pasada, el modelo primero redimensiona la imagen de entrada a un tamaño estándar. A continuación, una sola red neuronal convolucional procesa la imagen completa, la cual es dividida conceptualmente en una cuadrícula. Para cada celda de esta cuadrícula, la red predice simultáneamente tres elementos clave: las cajas delimitadoras (bounding boxes) que podrían contener un objeto, la puntuación de confianza para cada caja, y la probabilidad de la clase a la que pertenece el objeto detectado.

Figura 1 The model.



Nota. Adaptado de Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, real-time object detection. arXiv preprint arXiv:1506.02640

Para contextualizar estos conceptos, es útil clarificar su jerarquía. La Inteligencia Artificial (IA) es el campo general que busca emular la cognición humana. Dentro de la IA, el Aprendizaje Automático (Machine Learning) es la subdisciplina que permite a los sistemas aprender de los datos. A su vez, las redes neuronales convolucionales (CNN), que constituyen el núcleo de arquitecturas como YOLO, son un tipo avanzado de machine learning especialmente eficaz para tareas de Visión por Computadora.

En el estudio AI in Retail Forecasts 2019-2023 realizado por la consultora de mercados Juniper Research (2019) ya se anticipaba que el uso de la inteligencia artificial potenciaría significativamente la eficiencia en las operaciones del sector minorista. La aplicación de analítica avanzada en funciones como la previsión de la demanda y el marketing automatizado permitiría a los minoristas operar con mayor agilidad y mejorar sus márgenes de beneficio. El informe subraya la creación de un imperativo competitivo, afirmando que se prevé que los minoristas se enfrentarán a una carrera por la adopción de la IA, en la que los minoristas equipados con IA, que han adoptado sistemas como pioneros, desplazarán a los de menor ritmo, ofreciendo un servicio superior a precios optimizados.

A diferencia de arquitecturas de dos etapas como R-CNN, que son precisas, pero computacionalmente costosas, la familia de modelos YOLO, y en particular su versión v8, ofrece un equilibrio superior entre velocidad y precisión. Esta combinación es idónea para el sector minorista, donde se requiere analizar grandes volúmenes de video de manera eficiente sin sacrificar la fiabilidad de las detecciones.

2.4. De la aceptación a la rentabilidad, el modelo TAM como pilar de retorno de inversión (ROI)

La implementación exitosa de una innovación tecnológica en una organización trasciende la mera instalación de hardware y software. El valor real de cualquier sistema, y por ende su retorno de la inversión (ROI), no se deriva de su potencial técnico, sino de su uso efectivo y sostenido por parte de las personas a las que busca servir. La literatura sobre gestión del cambio demuestra que una de las principales causas del fracaso de proyectos tecnológicos es la resistencia y la baja adopción por parte de los usuarios finales; por ello, comprender los factores de aceptación es importante para diseñar estrategias de implantación efectivas (Yong Varela et al., 2010).

Para estructurar este análisis, el presente estudio adopta como pilar el modelo de aceptación tecnológica (TAM), propuesto originalmente por Davis (1989). A pesar de sus décadas de antigüedad, el TAM sigue siendo uno de los marcos teóricos más influyentes y empíricamente validados para predecir la aceptación de una tecnología, demostrando una notable vigencia en contextos modernos como la adopción de inteligencia artificial en la gestión de retail (Arias 2023) y el uso de analítica de datos en las organizaciones (Malik et al., 2021). El modelo postula que la intención de uso de un individuo está determinada principalmente por dos constructos psicológicos clave:

Utilidad percibida (Perceived Usefulness): el grado en que un usuario cree que el uso de una tecnología específica mejorará su desempeño laboral. En el contexto de este estudio, se materializa cuando un gerente de tienda comprende que puede pasar de la

intuición a la evidencia empírica para optimizar la distribución de la tienda y tomar decisiones estratégicas basadas en datos objetivos (Davis, 1989).

Facilidad de uso percibida (Perceived Ease of Use): el grado en que un usuario cree que el uso de dicha tecnología estará libre de esfuerzo (Davis, 1989). Para este proyecto, se traduce en la entrega de informes visuales e intuitivos (mapas de calor, conteos automáticos) que no requieren conocimientos avanzados para ser interpretados.

La conexión entre el modelo TAM y el ROI es fundamental, manifestándose en un doble impacto que influye tanto en la generación de beneficios como en la reducción de costos. Por un lado, una alta utilidad percibida asegura que los datos generados sean consultados de manera consistente para la toma de decisiones. Esto conduce a acciones comerciales más efectivas que impactan directamente los indicadores clave de rendimiento (KPIs), generando resultados medibles como el incremento en la tasa de conversión. Este es el motor que alimenta el "retorno" de la inversión.

Por otro lado, una alta facilidad de uso percibida no solo impulsa la adopción, sino que también reduce directamente los costos ocultos de la implementación, que a menudo se subestiman y pueden erosionar el ROI. Un sistema diseñado para ser intuitivo se convierte en un factor de mitigación financiera al reducir los costos de formación, minimizar la necesidad de soporte especializado y acelerar la productividad del personal (Yong Varela et al., 2010).

El diseño de este estudio, informado por el modelo de aceptación tecnológica (Davis, 1989), va más allá de la validación técnica. Se enfoca en entregar una solución que no solo es potente, sino fundamentalmente útil y fácil de consumir. Al hacerlo, se mitigan los riesgos de una baja adopción y se fortalece el caso de negocio, asegurando que el valor de los datos generados sea accesible y explotable sin barreras.

3. Metodología

La presente investigación se desarrolló bajo un protocolo metodológico diseñado para garantizar la validez, replicabilidad y precisión de los resultados. Esta sección detalla el enfoque y diseño del estudio, el procedimiento experimental en sus distintas fases, las características de la población y muestra, los instrumentos de medición y validación, y las técnicas aplicadas para el análisis de los datos.

3.1. Enfoque, diseño y alcance de la investigación:

Se trabajó con un enfoque cuantitativo, dado que su objetivo principal es la medición y el análisis numérico de variables de comportamiento del cliente. Se cuantifican métricas como el número de visitantes únicos, sus trayectorias espaciales, el tiempo promedio de permanencia y la frecuencia de visitas en zonas espaciales. Adicionalmente, se aplicó un criterio temporal para diferenciar a los asesores, considerando a cualquier persona con una permanencia superior al 70% de la duración del video como un asesor, permitiendo así aislar el análisis al comportamiento del consumidor.

Se adoptó un diseño de investigación no experimental, ya que no se realizó ninguna manipulación de variables independientes. El estudio se centró en la observación y el análisis del comportamiento de los clientes en su estado natural. El sistema desarrollado, sin embargo, está concebido como una herramienta robusta para la validación de futuros diseños experimentales, como la evaluación de estrategias de comercialización (merchandising) mediante pruebas A/B.

El alcance del estudio es descriptivo-correlacional. Es descriptivo al caracterizar los patrones de comportamiento existentes (por ejemplo, rutas de tráfico y zonas de alta densidad) y correlacional al establecer las bases para analizar futuras asociaciones entre estas métricas y variables de negocio, como las ventas por zona o el impacto de campañas promocionales. Para la validación inicial del sistema, se empleó un corte transversal, analizando un único período de tiempo continuo. No obstante, se postula que

la aplicación comercial óptima de esta metodología requiere adoptar un enfoque longitudinal, que implica recolectar datos mediante múltiples mediciones a lo largo del tiempo, lo que permite observar cambios, patrones y tendencias en lugar de capturar una única fotografía en un momento específico (enfoque transversal).

3.2. Población, muestra e instrumento de medición:

La población objetivo del estudio son los clientes que visitan la tienda física durante su horario de operación. La muestra de estudio está constituida por un conjunto de visitantes (N=555), cuyas trayectorias fueron exitosamente detectadas en una grabación de video continua de 1 hora y 15 minutos. Esta muestra fue seleccionada por conveniencia, ya que el video fue proporcionado por la tienda colaboradora y corresponde a un período de alta afluencia (sábado 10 de mayo, de 2025 de 2:00 PM - 3:15 PM), lo que garantiza la captura de una cantidad significativa de interacciones para validar el sistema.

3.3. Procedimiento metodológico:

El procedimiento se articuló en tres fases secuenciales: (1) Generación de Datos de Trayectoria, (2) Limpieza y Fusión de Datos, y (3) Análisis Cuantitativo y Visualización de Datos.

Fase I: generación y procesamiento de datos de trayectoria

El objetivo de esta fase fue la extracción de datos de trayectoria a partir de una fuente de video, estableciendo un flujo de datos robusto para la detección y el seguimiento de individuos con las siguientes condiciones:

1. **Fuente de datos y plataforma:** se utilizó como entrada una grabación de video de 1 hora y 15 minutos capturada en una tienda física, con una resolución de 1280x720 píxeles y una frecuencia de 25 fotogramas por segundo (fps). El procesamiento se realizó en el entorno Python (v. 3.11), utilizando la librería

Ultralytics y el modelo de detección YOLOv8s, seleccionado por su óptimo equilibrio entre velocidad y precisión.

2. **Protocolo de ejecución y configuración:** el proceso fue orquestado por el módulo de procesamiento 01_deteccion_tracking.py. Para maximizar la calidad del seguimiento, se implementó una configuración deliberada de técnicas e hiperparámetros, centralizados en el diccionario "CONFIG" para facilitar el ajuste por parte del usuario que desee aplicar este sistema. Para la ejecución con el video de muestra, se implementaron las siguientes decisiones metodológicas:

- A) **Definición de una región de interés (ROI):** para evitar errores de detección en los bordes del video, donde los objetos pueden ser menos estables, se definió una ROI que excluye un margen del 10% del ancho y alto del fotograma. El análisis se restringió únicamente a las detecciones cuyo centro se encontraba dentro de esta región de alta fiabilidad, reduciendo además el impacto de sombras y zonas con menor iluminación en las esquinas.
- B) **Calibración del umbral de confianza (CONFIDENCE_THRESHOLD):** este hiperparámetro establece la confianza mínima para una detección válida, implica un compromiso fundamental entre Precisión y Exhaustividad (Manning et al., 2008). Tras una validación comparativa, que consistió en ejecutar el módulo de procesamiento 01_deteccion_tracking.py. sobre varios fragmentos del video para observar la estabilidad de los track_id, la incidencia de fragmentación de trayectorias durante oclusiones y la aparición de falsos positivos, se estableció un umbral de 0.7 (70%). Esta decisión prioriza estratégicamente la precisión, aceptando la posible pérdida temporal de detecciones para evitar contaminar los datos con falsos positivos, confiando en la capacidad de Re-ID de BoT-SORT para recuperar las trayectorias.

C) **Calibración de la tasa de muestreo (FRAME_SAMPLE_RATE):** este parámetro equilibra la eficiencia computacional con la precisión del tracking ya que permite extraer una muestra de los fotogramas del video. Se determinó empíricamente que una tasa de muestreo elevada de 1 de cada 8 fotogramas (FRAME_SAMPLE_RATE = 8) resultaba en una fragmentación de trayectorias inaceptable. Por lo tanto, para el análisis final se estableció una tasa de muestreo de 1 por cada 2 fotogramas (FRAME_SAMPLE_RATE = 2). Esto implica que el sistema procesa la mitad de los fotogramas del video original, analizando la escena a una frecuencia efectiva de 12.5 fotogramas por segundo (FPS). Esta decisión representa un compromiso deliberado que prioriza la integridad y continuidad de los datos de seguimiento sobre la velocidad de procesamiento.

D) **Salidas de la fase I:** este proceso generó dos artefactos principales:

- Un archivo de datos en formato CSV (trayectorias_detectadas.csv), que contiene las detecciones y las coordenadas de estas (frame_id, track_id, y coordenadas).
- Un video de salida con las anotaciones visuales del tracking, utilizado para la validación cualitativa del rendimiento del sistema, permitiendo verificar la asignación de track_id y el manejo de oclusiones.

3.4. Evaluación del rendimiento computacional y viabilidad de escalabilidad:

Para cuantificar la eficiencia del procesamiento y establecer una base empírica para el análisis de su viabilidad comercial, se integró un módulo de medición de rendimiento. Durante la ejecución de los módulos de procesamiento, se registró el tiempo total de procesamiento. Al finalizar, se calcularon dos métricas de rendimiento fundamentales:

Latencia promedio: el tiempo medio de procesamiento por fotograma, que mide el tiempo que requiere el sistema para analizar una única imagen.

Rendimiento global (Throughput): la tasa de procesamiento promedio, expresada en Fotogramas por Segundo (FPS), que cuantifica la capacidad efectiva del sistema (software y hardware combinados) para ejecutar el ciclo completo de detección, seguimiento y anotación.

Es crucial destacar que este valor de FPS del módulo de procesamiento es una métrica de rendimiento y no debe confundirse con la frecuencia de fotogramas del video original. Esta métrica empírica es un indicador indispensable de la viabilidad técnica de la solución y constituye la línea base para la extrapolación de los requerimientos de cómputo y la estimación de costos operativos en una implementación a gran escala. Como se detallará en la sección de discusión y conclusiones, este análisis de rendimiento permite proyectar la viabilidad económica del sistema en infraestructuras de nube, como Amazon Web Services (AWS).

Fase II: Limpieza y fusión de trayectorias (post-procesamiento):

El análisis del video de salida de la fase I reveló que, a pesar del uso de un tracker avanzado, persistían desafíos inherentes al seguimiento en entornos no controlados. Se identificó la fragmentación de trayectorias como el principal problema, originada por oclusiones severas cuando un individuo pasa detrás de otro o de elementos estructurales como las columnas de la tienda por variaciones en la calidad de la detección. Este fenómeno resultaba en la asignación de múltiples identificadores (`track_id`) a una misma persona, lo que generaba un sobre conteo. Adicionalmente, se observó que el sistema asignaba trayectorias tanto a clientes como a los asesores de la tienda.

Para abordar estas limitaciones y asegurar la alta fidelidad de los datos, se diseñó e implementó el módulo de post-procesamiento (`02_limpieza_fusion_trayectorias.py`), que consta de tres etapas secuenciales de refinamiento:

Filtrado de ruido: la primera etapa se centró en eliminar las detecciones que corresponden a falsos positivos. Se estableció un umbral de persistencia

(MIN_DETECTIONS_PER_TRACK) de 15 detecciones. Todas las trayectorias que no alcanzaron este mínimo fueron descartadas, asumiendo que correspondían a ruido algorítmico o falsos positivos momentáneos y no a individuos reales.

Fusión de trayectorias fragmentadas: para reconstruir los recorridos completos de los clientes, se implementó un algoritmo iterativo de fusión. Dicho algoritmo evalúa cada par de trayectorias basándose en dos criterios: la proximidad temporal, es decir el inicio de una trayectoria debe ocurrir dentro de una ventana de tiempo máxima (MAX_FRAME_DIFFERENCE) de 30 fotogramas tras la finalización de la otra y la proximidad espacial, que hace referencia a que el punto de inicio de la segunda trayectoria debe encontrarse a una distancia euclidiana máxima (MAX_PIXEL_DISTANCE) de 100 píxeles del punto final de la primera.

Si ambas condiciones se cumplen, el sistema asume que las dos trayectorias pertenecen al mismo individuo y unifica sus identificadores bajo un único track_id.

Segmentación de clientes y asesores: con el fin de aislar el análisis al comportamiento del consumidor, se aplicó un filtro para excluir al personal de la tienda. Se calculó el porcentaje de presencia de cada trayectoria fusionada con respecto a la duración total del segmento de video analizado. Aquellas trayectorias con una presencia superior al 70% (EMPLOYEE_PRESENCE_THRESHOLD) fueron clasificadas como "Asesores" y excluidas del conjunto de datos final. Este método basado en porcentajes ofrece una mayor robustez y adaptabilidad que un umbral de tiempo fijo.

El resultado de esta fase de post-procesamiento fue un archivo CSV depurado y refinado (trayectorias_finales_clientes.csv), conteniendo únicamente las trayectorias validadas y completas de los individuos clasificados como clientes. Este conjunto de datos sirvió como la única fuente para todos los análisis de comportamiento posteriores.

Fase III: análisis cuantitativo y visualización de datos

La fase final de la metodología se centró en la transformación de los datos de trayectoria refinados (provenientes de la fase II) en métricas de negocio y visualizaciones interpretables. Para este propósito, se empleó el módulo de procesamiento de análisis (03_analisis_de_comportamiento.py) que utiliza un conjunto de librerías del ecosistema de ciencia de datos de Python, incluyendo Pandas para la manipulación de datos, NumPy para operaciones numéricas, y Matplotlib junto con Seaborn para la generación de gráficas.

El análisis se basó en técnicas de estadística descriptiva para extraer los siguientes indicadores clave de comportamiento:

Conteo final de clientes: se calculó el número de track_id únicos en el conjunto de datos final y depurado, obteniendo la métrica base del tráfico de clientes.

Análisis de densidad de tráfico: se generaron mapas de calor para representar visualmente la distribución de la actividad en el espacio físico. Esta técnica agrega la frecuencia de todas las detecciones en una cuadrícula espacial, revelando las zonas de mayor y menor concurrencia.

Identificación de zonas de interés: para complementar el análisis visual, se identificaron cuantitativamente las "zonas calientes" (hot zones) mediante la clasificación de las celdas de la cuadrícula según su índice de actividad total.

4. Resultados

En esta sección se presentan los hallazgos cuantitativos, enmarcados no solo como una validación técnica, sino como la fundamentación de una propuesta de valor de negocio robusta, escalable y financieramente atractiva.

4.1. Resultados del tracking y detección:

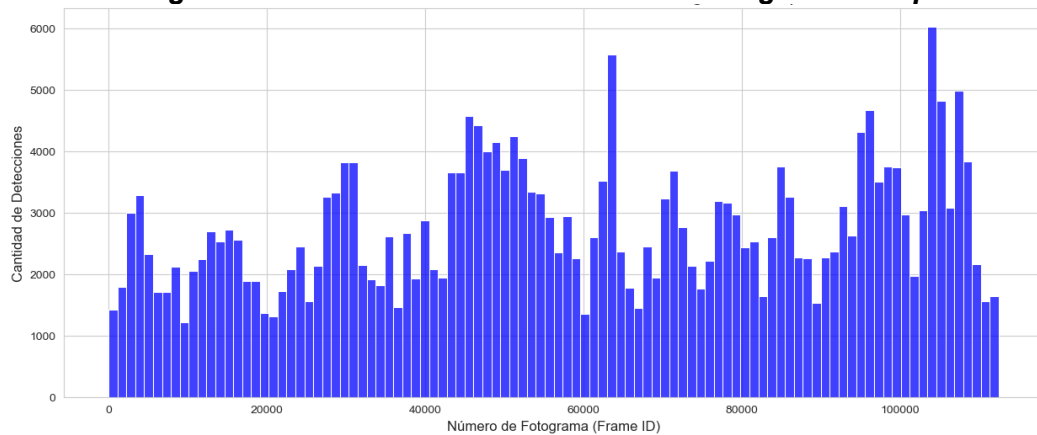
El análisis inicial se centró en la dinámica temporal del tráfico, procesando el archivo de datos del que genera el módulo de procesamiento "01_Deteccion_tracking.py" (trayectorias_detectadas.csv) para visualizar la distribución de detecciones a lo largo del tiempo. El histograma resultante, presentado en la figura 2, revela una actividad de clientes sostenida y con una alta variabilidad a lo largo de toda la grabación, lo cual es consistente con el hecho de que el video corresponde a una hora de alto tráfico de la tienda.

La gráfica muestra que el número de detecciones rara vez desciende a niveles bajos, indicando una presencia constante de personas, mientras que las fluctuaciones abruptas demuestran la naturaleza dinámica e impredecible del retail. Se identifican claramente múltiples períodos de alta densidad de tráfico, destacando dos picos principales de máxima afluencia: un pico agudo y pronunciado alrededor del fotograma 60,000, y un período de alta actividad sostenida que culmina cerca del fotograma 105,000. En estos momentos, la cantidad de detecciones acumuladas por intervalo supera las 5,000.

Más allá de cuantificar el tráfico, esta visualización actúa como una herramienta de diagnóstico fundamental. Los picos de alta detección no solo reflejan una mayor ocupación de la tienda, sino que también exponen las condiciones de máximo estrés para el algoritmo de seguimiento. En estos períodos de alta densidad, la probabilidad de oclusiones se incrementa drásticamente. Como se corroboró en la validación visual del video de salida de la fase I, estas oclusiones son el principal catalizador de la fragmentación de trayectorias y del consecuente sobre conteo de identificadores únicos.

Este hallazgo justifica empíricamente las decisiones metodológicas clave de este estudio. Primero, valida la necesidad de utilizar un tracker avanzado como BoT-SORT, cuya capacidad de re-identificación está diseñada para manejar oclusiones y demuestra la necesidad imperativa de implementar un riguroso post-procesamiento (fase II) para fusionar los fragmentos y corregir los artefactos algorítmicos generados durante estos picos de actividad, asegurando así la obtención de métricas de tráfico fiables.

Figura 2 Distribución de detecciones a lo largo del tiempo



Nota. Cada barra representa el número total de detecciones. La figura ilustra la variabilidad del tráfico de clientes a lo largo del período de grabación. Fuente: elaboración propia a partir del archivo de datos brutos trayectorias_detectadas.csv generado en la fase I.

4.2. Resultado de la limpieza y fusión de trayectorias:

La aplicación del post-procesamiento descrito en la metodología para refinar el conjunto de datos, corregir los artefactos del tracking y segmentar la población de interés generó la esperada reducción de identificadores (`track_id`). En la tabla 2 se cuantifica el impacto de cada etapa de este proceso de limpieza.

Tabla 2 Impacto de cada etapa del proceso de limpieza

Etapa de procesamiento	Track_id únicos	Reducción vs. etapa anterior
Datos brutos iniciales (sin procesar)	895	-
Después de filtrar ruido (≥ 15 detecciones)	614	-31.4%
Después de fusionar trayectorias fragmentadas	560	-8.8%
Conteo final de clientes (después de filtrar empleados)	555	-0.9%

Nota. La tabla ilustra la eficacia del módulo de procesamiento de limpieza para pasar de un conteo inflado a una estimación final y fiable del número de clientes. Fuente: elaboración propia con los datos obtenidos en la fase II.

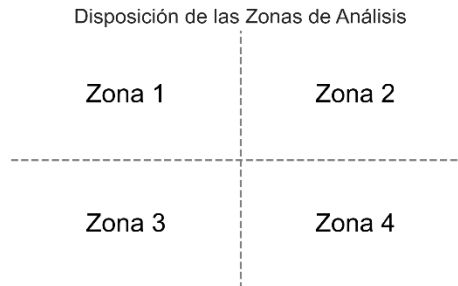
El análisis de los resultados del post-procesamiento revela el impacto decisivo de cada etapa de refinamiento. La fase inicial de filtrado de ruido fue la más significativa, eliminando 281 identificadores (una reducción del 31.4%), lo que confirma que una porción sustancial de las trayectorias podía ser atribuidas a falsos positivos o errores de detección momentáneos. Posteriormente, el algoritmo de fusión de trayectorias logró consolidar 54 trayectorias que habían sido fragmentadas por el tracker, una corrección del 8.8% que subraya la importancia de este paso para reconstruir los recorridos completos que las oclusiones impidieron seguir de manera ininterrumpida. Finalmente, la segmentación de empleados, basada en el porcentaje de presencia, identificó y excluyó a 5 individuos cuyas trayectorias eran consistentes con el comportamiento del personal de la tienda y verificable con el video de detección visual de salida. Este último paso fue crucial para aislar el conjunto de datos final, enfocándolo exclusivamente en el comportamiento del consumidor.

4.3. Resultado del análisis del comportamiento del cliente:

Con el objetivo de realizar un análisis cuantitativo de la distribución espacial de la actividad, el plano de la tienda capturado en el video fue discretizado en una cuadrícula

de 2x2, generando cuatro zonas de interés (Zoi) para el análisis, cuya disposición se presenta en la figura 3.

Figura 3 Disposición de las zonas de análisis

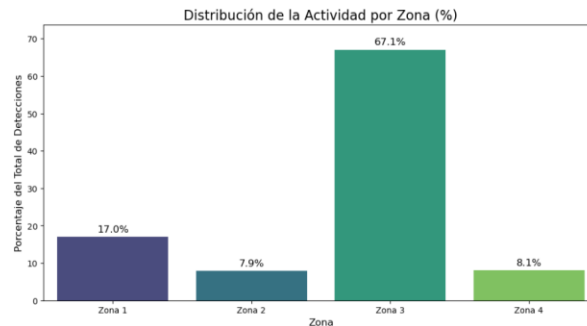


Nota. La distribución seleccionada se realiza para clasificar las zonas y diferenciar los resultados de los datos. La figura ilustra la numeración generada para diferenciar cada zona. Fuente: elaboración propia con código Python.

Cada punto de datos del conjunto de trayectorias de clientes fue geolocalizado en uno de estos cuatro cuadrantes según sus coordenadas (x_center, y_center). Esta segmentación espacial permitió la extracción de dos métricas complementarias para caracterizar el comportamiento en cada zona:

Índice de actividad por zona: para cada zona, se cuantificó el número total de detecciones. Esta métrica es proporcional al tiempo de permanencia acumulado de todos los clientes en el área y sirve para identificar las zonas de mayor congestión y actividad. Como se observa en la figura 4 la zona 3 de la tienda presenta mayor índice de actividad con un 67.1% de detecciones.

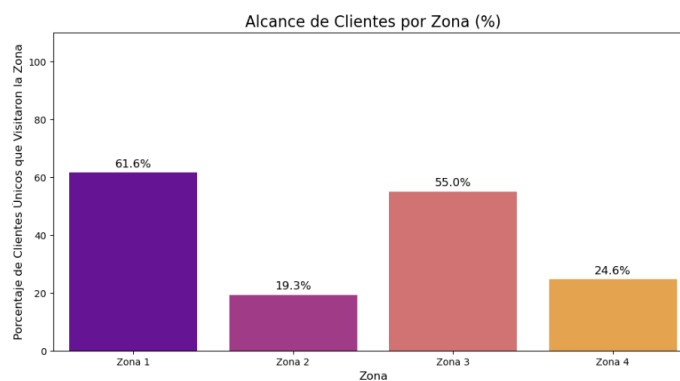
Figura 4 Índice de actividad por zona



Nota. Cada barra representa el nivel de actividad en cada una de las zonas marcadas en el espacio registrado por la cámara. Fuente: elaboración propia con los datos obtenidos en la fase II.

Alcance de clientes (Reach): se realizó un conteo de los identificadores únicos (track_id) que registraron al menos una detección dentro de cada zona. Esta segunda métrica mide la penetración de cada área, es decir, qué porcentaje del total de clientes visitó cada cuadrante. En la figura 5 se identifica que la mayor proporción de clientes ingresa a la zona 1(61.6%) y a la zona 3 (55.0%) pero no pasa a la zona 2 (19.3%) y a la zona 4(24.6%) lo que identifica que estas son las zonas de mayor interés de la tienda.

Figura 5 Alcance de clientes por zona (%)

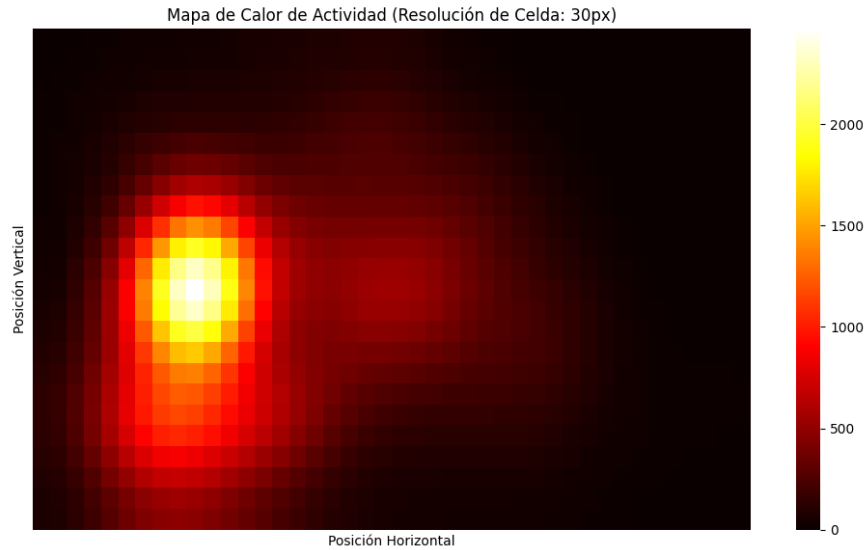


Nota. Cada barra representa el nivel de penetración de cada zona que permite identificar las zonas con mayor interés. Fuente: elaboración propia con los datos obtenidos en la fase II.

4.4. Análisis de mapa de calor:

Para comprender cómo los clientes utilizan el espacio físico de la tienda, se generó un mapa de calor que visualiza la densidad de tráfico agregada de todas las trayectorias de clientes. La figura 6 presenta esta distribución espacial de la actividad.

Figura 6 Mapa de calor de la actividad de los clientes (resolución de celda: 30px)



Nota. El mapa de calor fue generado a partir de las trayectorias finales de los clientes, con una resolución de celda de 30x30 píxeles. Las áreas más brillantes (blanco/amarillo) indican la mayor densidad de tráfico y tiempo de permanencia, mientras que las áreas oscuras (rojo/negro) representan zonas de baja actividad. Fuente: elaboración propia a partir de los datos obtenidos en la fase II.

El mapa de calor nos confirma que existe una zona dominante “Zona caliente” (el área blanca/amarilla brillante). Prácticamente toda la actividad de los clientes se concentra en una ubicación específica del lado izquierdo del plano. Esta área no solo es la más visitada, sino que es donde los clientes pasan la mayor parte de su tiempo.

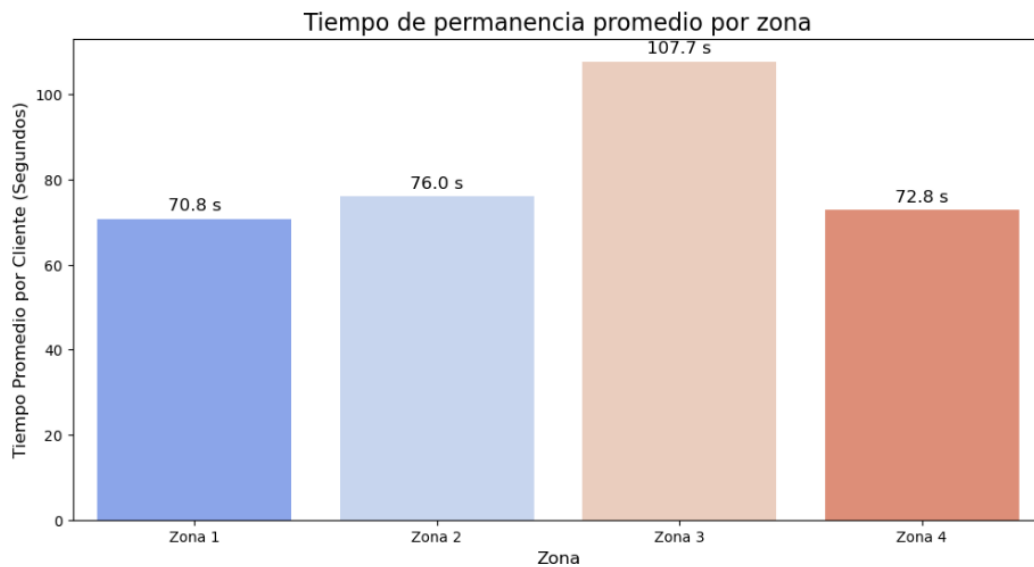
El resto del mapa es notablemente oscuro. Esto indica que la gran mayoría del espacio de la tienda (aproximadamente el 80-90% del área visible) recibe muy poco o

ningún tráfico. Existe una caída drástica de la actividad a medida que uno se aleja de la zona de entrada. Los clientes no están explorando la profundidad de la tienda.

4.5. Tiempo de permanencia promedio por zona:

Para evaluar el nivel de interés en cada zona, se realiza el cálculo de tiempo de permanencia promedio. Esta métrica se obtuvo dividiendo el tiempo total acumulado de todas las visitas en una zona entre el número de trayectorias únicas. Los resultados revelan un patrón de comportamiento significativo: la Zona 3 registró el mayor tiempo de permanencia promedio, indicando que los clientes que la visitan tienden a quedarse más tiempo. Es notable que la Zona 2, a pesar de tener un alcance menor en términos de cantidad de clientes, también mostró un tiempo de permanencia elevado.

Figura 7 Tiempo de permanencia promedio por zona



Nota: Cada barra representa el tiempo promedio por cliente en cada una de las zonas para identificar el nivel de atención capturada en cada zona. Fuente: elaboración propia con los datos obtenidos en la fase II.

4.6. Rendimiento del sistema y viabilidad comercial

Para evaluar la viabilidad de una implementación a escala, se cuantificó el rendimiento computacional de la fase I en el hardware de prueba (un sistema local sin GPU

dedicada). La ejecución sobre el video completo de 1 hora y 15 minutos arrojó una tasa de procesamiento promedio de 3.17 fotogramas por segundo (FPS). Este valor establece una línea base empírica para la eficiencia del sistema.

Con el fin de proyectar el rendimiento en un entorno de producción, se extrapoló este resultado a una instancia de computación en la nube estándar para IA, como una Amazon Web Services (AWS) g4dn.xlarge seleccionada por la relación costo- beneficio y recursos Suficientes (16 GB de VRAM y 16 GB de RAM). Aplicando un factor de aceleración conservador de 25x, que refleja la diferencia de capacidad entre una CPU y una GPU especializada como la NVIDIA T4, se estima un rendimiento operativo en la nube de aproximadamente 80 FPS.

Esta proyección permite realizar un análisis de viabilidad para un escenario comercial típico: el procesamiento de 8 horas diarias de grabación durante un mes de 30 días. Dicha carga de trabajo, que comprende un total de 10.8 millones de fotogramas a procesar (considerando una tasa de muestreo de 1 por cada 2), requeriría un tiempo de cómputo total de 37.5 horas en la instancia de AWS. Este resultado confirma que la solución es altamente escalable, ya que permite el análisis diario de grandes volúmenes de video en una fracción del tiempo de grabación.

4.7. Cuantificación del retorno de la inversión (ROI) y análisis de sensibilidad

El análisis de costos proyectado, basado en los precios públicos de Amazon Web Services (AWS), indica un gasto operativo mensual de aproximadamente \$30 USD por cámara. Si bien este es un costo marginal, el verdadero valor para un análisis de negocio reside en el "retorno" de esta inversión.

Para cuantificarlo, se modeló un escenario deliberadamente conservador para una tienda minorista hipotética. Este modelo se enfoca en el impacto sobre un único indicador clave de rendimiento (KPI): la tasa de conversión. Es importante destacar que este

cálculo no incluye otros beneficios potenciales de alto impacto, pero de más difícil cuantificación inicial, tales como:

- **Optimización de la asignación de personal:** utilizar los mapas de calor para dirigir a los asesores de ventas a las "zonas calientes" en tiempo real, maximizando las oportunidades de interacción y venta.
- **Mejora en la rotación de inventario:** identificar productos en "zonas frías" y reubicarlos en áreas de alto tráfico para aumentar su visibilidad y probabilidad de compra.
- **Validación de campañas de marketing:** medir objetivamente si una nueva vitrina o promoción logra atraer más tráfico hacia una zona específica de la tienda.

Considerando únicamente la mejora en la conversión, el modelo proyecta el siguiente retorno:

Tabla 3: Modelo de proyección de retorno de la inversión (ROI)

Métrica de negocio	Supuesto / Cálculo	Valor
Visitantes Mensuales	Tráfico promedio de la tienda	5.000
Tasa de conversión base	Porcentaje de visitantes que compran	15.0%
Compra Promedio \$	Valor de compra promedio	\$50.00 USD
Ingreso Mensual Base	(Visitantes × Tasa Conversión × Entrada)	\$37,500.00 USD
Mejora en Conversión	Supuesto conservador por optimización	+1.0% (Pasa un 16%)
Nuevo Ingreso Mensual	(5000 × 16% × \$50)	\$40,000.00 USD
Beneficio Bruto Adicional	(Nuevo Ingreso - Base Ingreso)	+\$2,500.00 USD
Costo mensual de la solución	Estimación de costos de AWS	-\$30.00 USD
Retorno Neto Mensual	(Beneficio - Costo)	+\$2,470.00 USD
ROI Mensual	(Retorno neto/costo)	8,233%

Nota: Este modelo ilustra cómo una mejora marginal y alcanzable en un solo KPI

genera un retorno exponencial sobre una inversión operativa mínima. Se utiliza el punto como separador decimal y la coma como separador de miles.

Análisis de sensibilidad del ROI

El modelo de retorno es particularmente sensible a tres variables clave: 1) la mejora porcentual en la tasa de conversión, 2) el valor de la compra promedio, y 3) el tráfico base de visitantes. Un pequeño cambio en cualquiera de estos supuestos puede alterar significativamente el resultado final.

Para ilustrar esta sensibilidad, si la optimización de la tienda lograra una mejora más modesta en la conversión, de solo +0.5% (pasando a 15.5%), el retorno neto mensual seguiría siendo de +1.220.00 USD.

Este análisis demuestra que el caso de negocio no depende de supuestos heroicos. Incluso mejoras marginales y realistas en la operación de la tienda, habilitadas por la inteligencia de negocio del sistema, son suficientes para justificar ampliamente la inversión y generar un retorno financiero sustancial.

4.8. Escalabilidad y costo total de propiedad (TCO)

Desde una perspectiva estratégica, el modelo es atractivo por su bajo costo total de propiedad (TCO). La inversión inicial es mínima al utilizar la infraestructura de cámaras existentes, y los costos operativos son predecibles y escalan de forma lineal por cámara.

Para una cadena de 20 tiendas, el costo operativo total sería de aproximadamente \$600 USD mensuales. Sin embargo, el valor de los datos generados crecimiento de forma no lineal. La gerencia podría realizar análisis comparativos entre sucursales, ejecutar pruebas A/B de diseños a gran escala y descubrir patrones de comportamiento regional, transformando la solución de una herramienta táctica por tienda a un activo de inteligencia estratégica corporativa.

5. Discusión y Conclusiones

La presente investigación alcanzó su objetivo general desarrollar y validar un sistema de visión por computadora capaz de transformar grabaciones de video de una tienda física en inteligencia de negocio estratégica. El flujo de datos, desde la detección hasta el posprocesamiento, demostró ser una solución robusta para la extracción de métricas de comportamiento fiables y accionables.

La siguiente discusión profundiza en el significado y las implicaciones de estos resultados. Se analizará la consecución de cada objetivo específico, se contrastarán los hallazgos con las hipótesis iniciales y se interpretará el valor de las métricas generadas en el contexto de los desafíos actuales del sector minorista.

5.1. Cumplimiento de objetivos y validación de hipótesis

Se realizó una evaluación directa de los objetivos e hipótesis del proyecto, cuyos resultados se detallan a continuación.

5.1.1. Evaluación de objetivos:

Objetivo 1: implementar un algoritmo de detección y seguimiento. CUMPLIDO. Se implementó exitosamente la fase I utilizando el modelo YOLOv8 para la detección y el tracker BoT-SORT para el seguimiento de personas, lo cual generó la base de datos de trayectorias a partir del video.

Objetivo 2: diseñar un sistema para almacenar datos de trayectorias. CUMPLIDO. Se alcanzó plenamente en la fase I del procedimiento. Se estableció un esquema de datos que captura la información esencial de cada detección en un formato tabular, compuesto por seis campos clave: `frame_id` (marca de tiempo discreta), `track_id` (identidad del individuo), y las coordenadas geométricas de la caja delimitadora (`x_center`, `y_center`, `width`, `height`).

La elección de almacenar estos datos en un archivo de formato CSV para este estudio se basó en su simplicidad y portabilidad, facilitando el pipeline de post-procesamiento. No

obstante, es crucial destacar que esta estructura de datos es fundamentalmente modular y escalable. El esquema diseñado puede ser implementado sin modificaciones en sistemas de bases de datos más robustos, como bases de datos relacionales para análisis estructurados para ingestas de datos a gran escala en tiempo real, lo que garantiza la viabilidad de la solución en entornos de producción.

Objetivo 3: desarrollar una herramienta de análisis para procesar los datos almacenados y generar indicadores claves. CUMPLIDO. En la fase II se obtuvo el conteo definitivo de clientes y en la fase III se identificaron las zonas con más tráfico de la tienda y el mapa de calor de actividad de los clientes que permite visualizar la densidad del tráfico en el espacio delimitado por la zona de interés.

Objetivo 4: analizar el valor estratégico y el ROI. CUMPLIDO. Se realizó un análisis detallado proyectando los costos operativos en una infraestructura de nube (AWS), se modeló un escenario de Retorno de la Inversión (ROI) y se discutió el valor estratégico de los datos para la toma de decisiones.

5.1.2. Validación de hipótesis:

H1 (Correlacional): la aplicación de un sistema de análisis de video permite la identificación de zonas de alto tráfico que se correlacionan con el interés del cliente. VALIDADA. La evidencia más contundente es el mapa de calor de actividad (Figura 6), que revela una concentración de tráfico abrumadora en las Zonas 1 y 3. Estas áreas, correspondientes a la entrada principal de la tienda, no solo registraron la mayor afluencia, sino también el mayor tiempo de permanencia acumulado, convirtiéndose en la "zona caliente" dominante del establecimiento. La clara correspondencia entre la alta densidad de tráfico y un punto de interés clave (la entrada y las primeras exhibiciones) confirma que el análisis espacial del comportamiento es un proxy fiable y cuantitativo del interés del cliente. Por lo tanto, el sistema no solo localiza dónde están los clientes, sino

que, de manera más importante, infiere dónde se centra su atención, validando la hipótesis.

H2 (Descriptiva): el análisis automatizado proporciona métricas más eficientes y fiables que los métodos tradicionales. VALIDADA. En primer lugar, la fiabilidad del sistema no se da por sentada, sino que se construye a través de un riguroso pipeline de post-procesamiento. La corrección de una sobreestimación del conteo inicial superior al 50% (de 895 a 555 IDs) es una prueba empírica de que, para ser fiable, un sistema de visión por computadora debe ir más allá de la simple detección en tiempo real e incorporar una fase de limpieza de datos. Este hallazgo es crítico, pues demuestra que la "fiabilidad" en este contexto es un resultado del diseño metodológico completo.

En segundo lugar, la capacidad de procesar de forma autónoma más de cien mil fotogramas por día, lo cual sería una tarea inviable para un observador humano, valida su eficiencia. Además, la arquitectura del pipeline, compuesta por los módulos de procesamiento secuenciales, está diseñada para una automatización completa mediante tareas programadas, permitiendo un ciclo de análisis diario sin intervención humana, lo cual es la definición de un sistema escalable.

H3 (Correlacional): la implementación con YOLOv8 en retail de moda obtendrá datos útiles y demostrará viabilidad financiera con ROI favorable. VALIDADA. El sistema, basado en YOLOv8, generó datos cuantitativos (mapas de calor y tráfico en las zonas de la tienda) que son directamente aplicables a decisiones comerciales (optimización de layout, merchandising). El análisis financiero proyectó un ROI mensual del 8,233% sobre un costo operativo mínimo de \$30 USD, confirmando su alta viabilidad financiera y cumpliendo todas las condiciones de la hipótesis.

5.2. Interpretación de los hallazgos principales

La abrumadora concentración de actividad en la zona de entrada (Figura 6) valida empíricamente los postulados de Marín Dueñas y Gómez Carmona (2021) sobre la

primacía del marketing visual para captar la atención inicial. Sin embargo, la falta de exploración hacia el interior de la tienda sugiere una ruptura en la estrategia de guiado del cliente, un área donde las métricas de trayectoria ofrecen una herramienta de diagnóstico directo

Los resultados demuestran que la solución propuesta es técnicamente viable y estratégicamente valiosa. El hallazgo central es la validación de un proceso metodológico completo que transforma datos de video en métricas de negocio fiables.

El análisis de post-procesamiento (Tabla 2) es particularmente revelador. La drástica reducción de identificadores únicos confirma que un sistema de tracking sin una fase de limpieza es propenso a generar métricas engañosas. Por otro lado, la generación de mapas de calor (Figura 6) no solo confirmó la existencia de zonas de alto interés, sino que también diagnosticó una ineficiencia crítica en el layout de la tienda, que no lograba guiar el tráfico hacia el interior.

5.3. La Velocidad del ciclo analítico como ventaja competitiva

Un hallazgo clave del proyecto, que va más allá de la viabilidad técnica, es la importancia estratégica del factor tiempo en el ciclo completo: desde la captura de datos hasta la implementación de estrategias accionables. La capacidad de acortar este ciclo gracias a la automatización del proceso es una ventaja competitiva fundamental que amplifica los beneficios.

La aceleración de este ciclo "dato-a-decisión" genera valor de manera directa. Permite a los minoristas capitalizar oportunidades y mitigar riesgos con una rapidez sin precedentes. Por ejemplo, se puede probar una nueva disposición de productos un lunes y, para el martes, disponer de un informe cuantitativo sobre su impacto en el flujo de clientes. Si la estrategia resulta exitosa, puede ser replicada de inmediato en otras tiendas; si falla, puede ser revertida antes de que genere pérdidas significativas. En mercados de moda dinámicos, donde las tendencias cambian rápidamente, esta

capacidad para explotar una oportunidad o corregir un error a gran velocidad es directamente proporcional al beneficio obtenido o la pérdida evitada.

Reducción del "costo de demora": cada día que se opera con un layout ineficiente, una promoción fallida o una mala asignación de personal, se está incurriendo en un costo de oportunidad. Un ciclo de análisis rápido minimiza este costo al reducir drásticamente el "tiempo de ceguera" operativa, el período durante el cual la empresa opera sin datos relevantes.

Aumento de la agilidad estratégica: la agilidad estratégica es la capacidad de una organización para adaptarse y crecer a pesar de las disrupciones del mercado. Un ciclo corto de "hipótesis -> acción -> medición -> aprendizaje" permite a la empresa realizar más experimentos de optimización en el mismo período. En lugar de una gran apuesta estratégica al año, se pueden realizar decenas de ajustes tácticos, aumentando exponencialmente la probabilidad de encontrar mejoras significativas en la tasa de conversión y, por ende, en el ROI.

En conclusión, el sistema propuesto no solo entrega qué está pasando en la tienda, sino que lo entrega a una velocidad que permite a la gerencia actuar de manera informada y ágil. El verdadero retorno de la inversión no proviene únicamente de la calidad de los datos, sino de la frecuencia y rapidez con la que esos datos se pueden transformar en decisiones estratégicas que impactan positivamente los resultados del negocio.

5.4. Propuesta de valor y viabilidad del modelo de negocio

La principal contribución del proyecto es un modelo que une la ciencia de datos con la estrategia de retail, fundamentado en un ciclo de inteligencia de alta velocidad. Desde una perspectiva financiera, la solución es altamente atractiva. Con un costo operativo estimado de \$30 USD mensuales por cámara, el sistema representa una inversión de bajo riesgo con un potencial de retorno exponencial, catalizado por la agilidad que

proporciona. Más allá del ROI directo, el sistema es un habilitador estratégico que cierra la brecha de información entre el retail físico y el comercio electrónico. El enfoque de ética por diseño, garantizando el anonimato, complementa la propuesta, mitigando riesgos y convirtiéndose en una ventaja competitiva.

Para optimizar aún más los costos operativos, futuras implementaciones podrían explorar estrategias de muestreo temporal. En lugar de procesar las 8 horas completas de grabación, se podría analizar una muestra representativa (p. ej., 15 minutos de cada hora). Este enfoque no solo reduce drásticamente los costos operativos, sino que también permite un análisis enfocado, donde los recursos de cómputo se pueden dirigir a los períodos de mayor relevancia estratégica (p. ej., las horas pico de los fines de semana o durante eventos promocionales), maximizando así el retorno de la inversión en la analítica.

5.5. Limitaciones y futuras líneas de investigación

Primero, el análisis se basó en un único video de una sola tienda, por lo que los patrones observados no son generalizables; esto significa que la “zona caliente” identificada es específica de ese layout y no una regla universal. Segundo, el uso de una única perspectiva de cámara limita el seguimiento a un área 2D, lo que pudo haber causado una subestimación de la actividad en zonas parcialmente ocluidas en los bordes del campo de visión.

Estas limitaciones abren claras líneas de investigación futura:

Sistemas multicámara: fusionar datos de varias cámaras para una reconstrucción 3D de las trayectorias y una cobertura completa de la tienda.

Reconocimiento de acciones: integrar modelos que detecten interacciones específicas (ej. "tomar un producto", "mirar un estante") para entender no solo el "dónde", sino también el "qué".

Integración con datos de venta (POS): vincular las trayectorias de los clientes con los datos de transacciones para calcular tasas de conversión por zona o producto, cerrando así el ciclo analítico completo.

Segmentación de perfiles de comprador: aplicar algoritmos de clustering no supervisado a las trayectorias para identificar diferentes arquetipos de comportamiento (ej. "comprador-misión", "comprador-explorador"), lo que habilitaría estrategias de personalización en el espacio físico.

Optimización para Edge Computing: adaptación de un sistema el despliegue en un dispositivo de borde de bajo costo (ej. NVIDIA Jetson) en las cámaras en la tienda para aplicar el modelo que procesa y genera datos en tiempo real.

Costos de implementación inicial: el análisis de ROI se centra en los costos operativos en la nube, asumiendo que el minorista ya cuenta con una infraestructura de cámaras de seguridad adecuada y personal con capacidades analíticas básicas. Un análisis completo del Costo Total de Propiedad (TCO) debería incluir la inversión inicial en hardware (si fuera necesario) y los costos de la integración del sistema.

Gestión del cambio organizacional: si bien el estudio se basa en el Modelo de Aceptación Tecnológica (TAM), no cuantifica los costos asociados a la gestión del cambio, como la capacitación del personal, la superación de la resistencia a nuevas herramientas y la adaptación de los procesos de toma de decisiones para incorporar esta nueva fuente de datos. Estos factores son críticos para asegurar que la herramienta se utilice eficazmente y se materialice el ROI proyectado.

Para análisis a gran escala en cadenas de retail, se recomienda la construcción y el modelado de datos de panel. Este enfoque, que combina la dimensión de series de tiempo (observaciones diarias por tienda) con la de corte transversal (múltiples tiendas observadas simultáneamente), permitiría realizar análisis comparativos entre sucursales

(benchmarking) y controlar variables específicas de cada tienda, generando inferencias estratégicas de mayor riqueza y validez.

Durante el desarrollo de este proyecto, se utilizó el modelo de inteligencia artificial (IA) Gemini de Google (versión de julio de 2025). La contribución de la IA se centró específicamente en el desarrollo y la corrección de errores en el código de Python. Sus aplicaciones incluyeron la generación de fragmentos de código para tareas específicas como la visualización de datos, la refactorización de código para mejorar la eficiencia y la legibilidad, y la identificación y solución de errores (debugging) durante las fases de prueba. Es importante destacar que toda la arquitectura lógica del sistema fue diseñada por las autoras, quienes también supervisaron, validaron y editaron todo el código generado, asumiendo la responsabilidad final por el software desarrollado y los resultados obtenidos.

Este estudio demuestra exitosamente cómo la sinergia entre la ciencia de datos, la estrategia de retail y la gestión de la innovación puede generar soluciones de alto impacto. La herramienta desarrollada y validada es un activo estratégico que dota al retail físico de la misma rigurosidad analítica que el comercio electrónico. Al ser ético por diseño, escalable y económicamente viable, el modelo propuesto se posiciona como una solución aplicable para cualquier minorista en Colombia que desee transformar las grabaciones de sus cámaras de seguridad en inteligencia de negocio.

Disponibilidad de recursos computacionales

El código fuente completo de los módulos de procesamiento creados con el lenguaje de programación Python utilizados para la detección, el post-procesamiento y el análisis de datos, está disponible de forma abierta en el siguiente repositorio de GitHub:

<https://github.com/daldana13801/YOLOv8-Customer-Tracking-Retail>

Referencias bibliográficas

- Ali, W. H., Kareem, A. A., & Jasim, M. (2019). Survey on Wireless Indoor Positioning Systems. *Cihan University–Erbil Scientific Journal*, 3(2), 42–47.
https://www.researchgate.net/publication/335439594_Survey_on_Wireless_Indoor_Positioning_Systems
- Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). Simple online and realtime tracking. En 2016 IEEE International Conference on Image Processing (ICIP) (pp. 3464–3468). IEEE. <https://doi.org/10.1109/ICIP.2016.7533003>
- Cai, Zhengying & Li, Gengze & Zhang, Jinming & Xiong, Shasha. (2023). Using an Artificial Physarum polycephalum Colony for Threshold Image Segmentation. *Applied Sciences*. 13. 11976. 10.3390/app132111976.
- Cámara Colombiana de Comercio Electrónico (CCCE). (2024). Informe de Cierre 2024 - Versión Pública. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/<https://ccce.org.co/wp-content/uploads/2017/06/Informe-de-Cierre-2024-Version-Publica-Apr.pdf>
- Congreso de Colombia. (2012). Ley 1581 de 2012. Por la cual se dictan disposiciones generales para la protección de datos personales.
<https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981>
- Departamento Administrativo Nacional de Estadística (DANE). (2023). Producto interno bruto (PIB): Información técnica. DANE.
<https://www.dane.gov.co/index.php/estadisticas-por-tema/cuentas-nacionales/cuentas-nacionales-trimestrales/pib-informacion-tecnica>
- European Commission. (2021). Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021) 206 final). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>

- Forrester. (2011) The age of the customer. Forrester Research.
- Girshick, R. (2015). Fast r-cnn. Proceedings of the IEEE international conference on computer vision. (Paper de Fast R-CNN)
- Help Scout. (2024). Customer service statistics. <https://www.helpscout.com>
- Ibarra Pardo, G. (2024, 4 de junio). Los datos biométricos: ¿Una amenaza al derecho a la intimidad? Asuntos Legales. <https://www.asuntoslegales.com.co/analisis/gabriel-ibarra-pardo-558821/los-datos-biometricos-una-amenaza-al-derecho-a-la-intimidad-3872741>
- Jocher, G., Chaurasia, A., & Qiu, J. (2023). Ultralytics YOLOv8 (Version 8.0.0) [Software]. Ultralytics. <https://doi.org/10.5281/zenodo.7636829>
- Journal of the Academy of Marketing Science, 50(6), 1189–1213. <https://doi.org/10.1007/s11747-022-00882-z>
- Juniper Research. (2019). AI spending by retailers to reach \$12 billion by 2023, driven by promise of improved margins. <https://www.juniperresearch.com/press/ai-spending-by-retailers-reach-12-billion-2023/>
- Makerson, R. (2024). How facial recognition enhances the retail experience. BlueFletch. <https://bluefletch.com/how-facial-recognition-enhances-the-retail-experience/#:~:text=Facial%20recognition%20is%20also%20gaining,recognized%20troublemaker%20enters%20the%20premises.>
- Malhotra, N. K. (2019). Marketing research: An applied orientation (7th ed.). Pearson.
- Malik, S., Chadhar, M., Vatanasakdakul, S., & Chetty, M. (2021). Factors Affecting the Organizational Adoption of Blockchain Technology: Extending the Technology–Organization–Environment (TOE) Framework in the Australian Context. *Sustainability*, 13(16), 9404. <https://doi.org/10.3390/su13169404>
- Marín Dueñas, P. P., & Gómez Carmona, D. (2021). Comportamiento de compra y marketing de los sentidos: un análisis de su influencia en los compradores de

moda. *Vivat Academia* (Alcalá de Henares), 154(154), 459–479.

<https://doi.org/10.15178/va.2021.154.e1366>

Mas Velasco, J. (2020). *La tienda del futuro: La humanización del comercio con ayuda de la tecnología*. CREARMAS.

McKinsey & Company. (2015) *Why customer analytics matter: Survey of marketing and sales executives*. McKinsey & Company.

Nirkin, Y. (2022). *BoT-SORT*. GitHub. <https://github.com/NirAharon/BoT-SORT>

NtechLab. (s.f.). *Countbuddy: Customer behavior analytics through facial recognition*.

<https://www.ntechlab.com>

Pérez, A., & Madrid, F. (2021). *Detección de objetos: Principios técnicos y aplicaciones comerciales*. Editorial TecnoAI.

Piazza. (2025, 30 de mayo). *KSI Vision: La startup uruguaya que analiza el comportamiento de tiendas y conquistó a clientes como Inditex, Mango y Falabella*. *Forbes Uruguay*. <https://www.forbesuruguay.com/negocios/ksi-vision-startup-uruguay-analiza-comportamiento-tiendas-conquisto-clientes-como-inditex-mango-falabella-n73109>

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). *You Only Look Once: Unified, real-time object detection*. arXiv preprint arXiv:1506.02640

Roy, S. K., Singh, S., & Tandon, A. (2021). *Driving customer acceptance of smart technologies in physical retail stores: An extended TAM approach*. *Journal of Business Research*, 136, 407-420. <https://doi.org/10.1016/j.jbusres.2021.07.043>

Servicios web de Amazon. (2024). *Precios de Amazon S3*. Recuperado el 14 de julio de 2025, de <https://aws.amazon.com/s3/pricing/>

Servicios web de Amazon. (2024). *Precios de Instancias EC2 On-Demand*. Recuperado el 14 de julio de 2025, de <https://aws.amazon.com/ec2/pricing/on-demand/>

Sharma, A., & Grewal, D. (2022). *The future of retail: A strategic framework for retailers*

- Smith, M., & Miller, S. (2021). The ethical application of biometric facial recognition technology. *AI & Society*, 37(1), 167-175. <https://doi.org/10.1007/s00146-021-01199-9>
- Trejo García, J. C., Soto Rosales, M. de L., & Olivares Aguayo, H. A. (2023). Optimización del Net Promoter Score (NPS) con factores de expansión, una medición de experiencia de clientes en riesgo reputacional. *Contaduría, Administración*, 68(3), 400–159. <https://doi.org/10.22201/fca.24488410e.2023.4626>
- Valencia-Arias, A., Gómez-Molina, S., Vélez-Holguín, R. M., & Cardona-Acevedo, S. (2023). Intención de uso de aprendizaje móvil (m-learning) en programas virtuales: un modelo híbrido de aceptación tecnológica (TAM) y la teoría del comportamiento planificado (TPB). *Formación Universitaria*, 16(2), 25–34. <https://doi.org/10.4067/S0718-50062023000200025>
- Verhoef, P. C., Kannan, P. K., & Inman, J. J. (2015). From multi-channel retailing to omni-channel retailing: Introduction to the special issue on multi-channel retailing. *Journal of Retailing*, 91(2), 174–181. <https://doi.org/10.1016/j.jretai.2015.02.005>
- Yong Varela, Luis Antonio, Rivas Tovar, Luis Arturo y Chaparro, Julián. (2010). Modelo de aceptación tecnológica (TAM): un estudio de la influencia de la cultura nacional y del perfil del usuario en el uso de las TIC. *Innovar*, 20 (36), 187-203. Recuperado el 13 de julio de 2025, de http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-50512010000100014&lng=en&tlng=es.
- Zafar, A. (2024). Balancing the scale: navigating ethical and practical challenges of artificial intelligence (AI) integration in legal practices. *Discover Artificial Intelligence*, 4, 27. <https://doi.org/10.1007/s44163-024-00121-8>