

**Pontificia Universidad Católica del Perú**

**Facultad de Ciencias e Ingeniería**



**DISEÑO E IMPLEMENTACIÓN DE UNA PLATAFORMA  
WEB-MÓVIL PARA GESTIÓN LOGÍSTICA EN PYMES  
CON MEDICIÓN AUTOMÁTICA DE DIMENSIONES DE  
PAQUETES MEDIANTE INTELIGENCIA ARTIFICIAL**

**Tesis para obtener el título profesional de Ingeniero de las  
Telecomunicaciones**

**Autor:**

Juan Alfonso Chapoñan Espinoza

**Asesor:**

Oscar Antonio Díaz Barriga

Lima, agosto, 2025

## INFORME DE SIMILITUD

Yo, **OSCAR ANTONIO DÍAZ BARRIGA**, docente de la Facultad de **CIENCIAS E INGENIERÍA** de la Pontificia Universidad Católica del Perú, asesor de la tesis titulada **PLATAFORMA WEB Y MÓVIL CON PROCESAMIENTO DE IMÁGENES EN LA NUBE E INTELIGENCIA ARTIFICIAL PARA LA IDENTIFICACIÓN DE PAGOS, ESTIMACIÓN DE DIMENSIONES Y OPTIMIZACIÓN LOGÍSTICA EN SERVICIOS DE DELIVERY URBANO EN LIMA**, del autor **JUAN ALFONSO CHAPOÑAN ESPINOZA**, dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de [##] %. Así lo consigna el reporte de similitud emitido por el software Turnitin el [DD/MM/AAAA].
- He revisado con detalle dicho reporte y la Tesis, y no se advierte indicios de plagio.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

**Lugar y fecha:** San Miguel, [DD] de [MES] de [AAAA].

Apellidos y nombres del asesor: <b>DÍAZ BARRIGA, OSCAR ANTONIO</b>	
DNI: <b>71477000</b>	Firma
ORCID: 0000-0002-9930-5984	

## DEDICATORIA

A mis padres.

## **AGRADECIMIENTOS**

[Por completar]

## RESUMEN

[Por completar]

# ÍNDICE

DEDICATORIA . . . . .	I
AGRADECIMIENTOS . . . . .	II
RESUMEN . . . . .	III
ÍNDICE . . . . .	VII
ÍNDICE DE TABLAS . . . . .	VIII
ÍNDICE DE FIGURAS . . . . .	IX
GLOSARIO . . . . .	X
INTRODUCCIÓN . . . . .	1
Capítulo 1. Presentación del problema de ingeniería. . . . .	2
1.1 Identificación temática y motivación personal . . . . .	2
1.1.1 Área de especialización en telecomunicaciones . . . . .	2
1.1.2 Relación con los estudios realizados . . . . .	4
1.1.3 Motivación personal y experiencia profesional . . . . .	5
1.2 Descripción y características del problema . . . . .	6
1.2.1 Contexto del problema en la industria de <i>delivery</i> . . . . .	6
1.2.2 Desafío técnico desde la perspectiva de telecomunicaciones . . . . .	7
1.2.3 Características específicas del problema . . . . .	7
1.3 Importancia del problema y su solución . . . . .	8
1.3.1 Perspectiva técnica . . . . .	8
1.3.2 Perspectiva económica y financiera . . . . .	8
1.3.3 Perspectiva social y cultural . . . . .	9
1.3.4 Perspectiva ambiental y de sostenibilidad . . . . .	10
1.3.5 Perspectiva legal y reglamentaria . . . . .	10
1.3.6 Perspectiva ética . . . . .	11
1.4 Impactos previstos y beneficiarios . . . . .	12
1.4.1 Impactos operacionales . . . . .	12

1.4.2	Impactos tecnológicos . . . . .	14
1.4.3	Impactos económicos . . . . .	15
1.4.4	Beneficiarios directos . . . . .	16
1.4.5	Beneficiarios indirectos . . . . .	17
<b>Capítulo 2. Estado del arte o de la cuestión, alternativas de solución al problema o desafío a resolver. . . . .</b>		<b>19</b>
2.1	Antecedentes de solución semejantes o similares al desafío de ingeniería	19
2.1.1	Sistemas de análisis visual con inteligencia artificial multimodal	19
2.1.2	Aplicaciones de LLMs multimodal en logística y comercio electrónico . . . . .	27
2.2	Características de soluciones semejantes o similares . . . . .	28
2.2.1	Arquitecturas de procesamiento predominantes . . . . .	28
2.2.2	Capacidades de interpretación dimensional . . . . .	28
2.2.3	Precisión y confiabilidad según implementación . . . . .	29
2.3	Compendio de tecnologías, herramientas, métodos, modelos utilizados con éxito . . . . .	29
2.3.1	Modelos de lenguaje multimodal predominantes . . . . .	29
2.3.2	Frameworks de visión por computadora académicos . . . . .	29
2.3.3	Tecnologías de infraestructura en la nube . . . . .	30
2.3.4	Herramientas de desarrollo e integración . . . . .	30
2.3.5	Metodologías de prompt engineering y optimización . . . . .	31
2.4	Conjunto de características y especificaciones para la solución óptima .	31
2.4.1	Arquitectura tecnológica óptima . . . . .	31
2.4.2	Precisión y confiabilidad requeridas . . . . .	32
2.4.3	Limitaciones y alcances reconocidos . . . . .	32
<b>Capítulo 3. Diseño de la solución al desafío de ingeniería . . . . .</b>		<b>33</b>
3.1	Introducción metodológica – Enfoque Design Thinking . . . . .	33
3.1.1	Justificación del enfoque centrado en el usuario . . . . .	33
3.1.2	Fases de Design Thinking aplicadas en este capítulo . . . . .	34
3.2	Fase 1: Empatizar – Análisis del contexto y usuarios . . . . .	34

3.2.1	Contexto del problema y metodología de investigación . . . . .	34
3.2.2	Síntesis de hallazgos . . . . .	36
3.3	Fase 2: Definir – Reformulación del problema técnico . . . . .	36
3.3.1	Declaración del problema de ingeniería . . . . .	36
3.3.2	Requerimientos funcionales del sistema . . . . .	36
3.3.3	Requisitos no funcionales . . . . .	38
3.3.4	Matriz de trazabilidad: necesidades vs. requerimientos . . . . .	40
3.3.5	Restricciones de diseño . . . . .	40
3.3.6	Arquitectura conceptual del sistema . . . . .	42
3.4	Fase 3: Idear – Generación y selección de alternativas . . . . .	43
3.4.1	Introducción a la generación de alternativas . . . . .	43
3.4.2	Criterios de evaluación . . . . .	44
3.4.3	Alternativas consideradas . . . . .	45
3.4.4	Matriz de evaluación multicriterio . . . . .	48
3.4.5	Selección de la solución óptima . . . . .	49
3.4.6	Selección de componentes específicos dentro del ecosistema Firebase	50
3.4.7	Conclusión de la fase de ideación . . . . .	51
3.5	Fase 4: Prototipar – Diseño técnico detallado . . . . .	51
3.5.1	Introducción al diseño de la solución . . . . .	51
3.5.2	Arquitectura general del sistema . . . . .	52
3.5.3	Flujo crítico del sistema: Creación de pedido con procesamiento de IA . . . . .	54
3.5.4	Módulo de IA para estimación de dimensiones . . . . .	56
3.5.5	Interfaces funcionales del sistema . . . . .	57
3.5.6	Especificaciones de seguridad y cumplimiento normativo . . . . .	58
3.5.7	Conclusión de la fase de prototipado . . . . .	59
	<b>Conclusiones . . . . .</b>	<b>60</b>
	<b>REFERENCIAS BIBLIOGRAFICAS . . . . .</b>	<b>63</b>
	<b>Anexo A. Caracterización detallada de actores del sistema . . . . .</b>	<b>64</b>
	<b>Anexo B. Síntesis de hallazgos de la fase de empatía . . . . .</b>	<b>66</b>



Anexo C. Matriz de trazabilidad completa . . . . .	68
Anexo D. Modelo de datos detallado . . . . .	70
Anexo E. Configuración de seguridad y autenticación . . . . .	77
Anexo F. Justificación de puntuaciones en matriz multicriterio . . . . .	80

## ÍNDICE DE TABLAS

1.1	Arquitectura IoT [2]. . . . .	3
1.2	Proyecciones del Mercado de <i>delivery</i> en Perú [5]. . . . .	6
3.1	Requisitos funcionales del sistema . . . . .	37
3.2	Requisitos no funcionales del sistema . . . . .	38
3.3	Criterios de evaluación de alternativas tecnológicas. . . . .	44
3.4	Matriz de evaluación multicriterio de alternativas. . . . .	49
3.5	Componentes seleccionados del ecosistema Firebase. . . . .	50
3.6	Protocolos de comunicación del sistema. . . . .	54
3.7	Especificaciones de seguridad en transporte. . . . .	58
3.8	Cumplimiento de Ley N.º 29733. . . . .	59
C.1	Matriz de trazabilidad: Necesidades vs. Requerimientos. . . . .	68

## ÍNICE DE FIGURAS

1.1	Gasto en tecnologías de la información - América Latina. Fuente: [6] . .	7
1.2	Gasto en tecnologías de la información - América Latina. Fuente: [10] .	9
2.1	Consulta a GPT-4 para análisis de múltiples imágenes. Fuente [23] . . .	20
2.2	Identificación de objetos visualmente usando modelos de Claude 3. Fuente: [24] . . . . .	22
2.3	Solicitud de reconocimiento de una imagen y reorganización en formato JSON.Fuente: [24] . . . . .	23
2.4	Se solicita a Gemini reconocer las imágenes y encontrar una relación entre ellas. . . . .	25
3.1	Arquitectura de capas del sistema . . . . .	43
3.2	Arquitectura de capas del sistema con tecnologías seleccionadas . . . .	53
3.3	Diagrama de flujo de entrega del sistema de delivery propuesto . . . .	55

**GLORASIO**

[Por completar]

# INTRODUCCIÓN

[Por completar]

## Capítulo 1. Presentación del problema de ingeniería

La transformación digital en el sector logístico demanda soluciones innovadoras que aprovechen tecnologías emergentes para resolver desafíos operacionales críticos. Esta investigación aborda la problemática de la inexactitud en la determinación de dimensiones de paquetes en la industria de *delivery*, factor que genera ineficiencias significativas en la planificación de rutas y utilización de capacidad vehicular. Mediante la convergencia de aplicaciones IoT, procesamiento de imágenes con inteligencia artificial y arquitecturas distribuidas en la nube, se propone una solución integral que mejora la precisión operacional y democratiza el acceso a tecnologías sofisticadas para empresas de diferentes escalas.

### 1.1 Identificación temática y motivación personal

#### 1.1.1 Área de especialización en telecomunicaciones

##### Aplicaciones IoT (*Internet of Things*)

Las aplicaciones IoT constituyen un ecosistema tecnológico integral que integra inteligencia artificial, redes de comunicación y automatización para crear una infraestructura de conectividad ubicua entre objetos físicos y sistemas digitales [1]. Se define como la implementación de una arquitectura de cinco capas interconectadas:

Tabla 1.1: Arquitectura IoT [2].

Capa	Descripción
Capa de Aplicación	Capa superior que implementa soluciones específicas basadas en la integración de recursos de información de toda la infraestructura IoT.
Capa de Gestión de Servicios	Nivel de integración que combina recursos de información de las capas inferiores para formular soluciones específicas a problemas concretos en campos especializados.
Capa de Internet	Capa de procesamiento que filtra, clasifica e integra los recursos de información transmitidos desde la capa de acceso, construyendo una plataforma de red confiable y eficiente.
Capa de Acceso	Infraestructura de comunicación que facilita la transmisión eficiente de datos percibidos hacia la red de Internet mediante tecnologías de comunicación móvil, redes satelitales y LAN inalámbricas.
Capa de Percepción	Capa fundamental que emplea diversos tipos de sensores (RFID, infrarrojos, láser, etc.) para identificar, capturar y procesar información sobre atributos, comportamiento, estado y entorno de los objetos físicos.

### Servicios de Telecomunicaciones para Logística

Los servicios de telecomunicaciones para logística se definen como el conjunto de tecnologías y protocolos de comunicación que operan principalmente en la capa de red del ecosistema IoT, actuando como puente crítico entre la percepción de datos y su procesamiento funcional. Estos servicios garantizan la transmisión eficiente y segura de información tanto estática como móvil durante todas las fases del proceso logístico [3].

## **Convergencia Tecnológica**

La tesis desarrollada en esta investigación representa la convergencia de múltiples disciplinas dentro de la ingeniería de telecomunicaciones:

- **Arquitecturas Distribuidas en la Nube:** Para el procesamiento remoto y almacenamiento escalable
- **Inteligencia Artificial:** Específicamente visión por computadora para el procesamiento automatizado de imágenes

Esta integración tecnológica permite abordar desafíos reales del sector logístico mediante soluciones que aprovechan las capacidades de procesamiento remoto, almacenamiento distribuido y comunicaciones continuas, características fundamentales de los sistemas modernos de telecomunicaciones en el contexto de la transformación digital.

### **1.1.2 Relación con los estudios realizados**

La presente investigación se fundamenta en una progresión curricular especializada que abarca desde los fundamentos del desarrollo web hasta la implementación de soluciones IoT avanzadas, estableciendo una relación directa y sistemática con tres cursos clave que proporcionan las competencias técnicas necesarias para el diseño, desarrollo e implementación de sistemas de telecomunicaciones aplicados a la optimización logística.

#### **TEL131 Ingeniería Web para Telecomunicaciones**

Este curso proporciona la base tecnológica para desarrollar la capa de aplicación e interfaces de gestión en soluciones de logística inteligente. Se abordan fundamentos de programación, desarrollo web con conexión a bases de datos, y modelado relacional con SQL, aplicables en dashboards para monitoreo en tiempo real, interfaces de control y manejo de datos IoT. La arquitectura web moderna (HTML5, CSS3, servidores, servlets, MVC), junto con nociones de seguridad, confiabilidad y despliegue en la nube, permite construir aplicaciones logísticas escalables, accesibles y seguras.

#### **TEL137 Gestión de Servicios de TICs**

Este curso se enfoca en la gestión de servicios dentro del ecosistema IoT, brindando competencias para desarrollar infraestructuras seguras, escalables y robustas que soporten aplicaciones logísticas avanzadas. A través de frameworks modernos y servicios



web, se construyen sistemas capaces de optimizar rutas y asignar recursos inteligentemente. Se abordan despliegues en la nube (IaaS, PaaS, FaaS), esenciales para alojar componentes distribuidos como módulos de análisis o monitoreo con IA. Además, se cubre la implementación de servicios REST, SOAP y websockets para integrar sistemas heterogéneos con visibilidad en tiempo real. Las capacidades en seguridad avanzada aseguran la protección frente a accesos no autorizados y amenazas cibernéticas.

## **1TEL05 Servicios y Aplicaciones para IoT**

Este curso se centra en la construcción de la capa de aplicación IoT y su integración con la nube, facilitando el desarrollo de soluciones logísticas orientadas al usuario final. Se abordan competencias en aplicaciones móviles conectadas a servicios SaaS, útiles para rastreo de productos, alertas inteligentes y control remoto de condiciones ambientales. Además, se utiliza arquitectura de microservicios y bases de datos NoSQL (Firebase) para garantizar escalabilidad y resiliencia en el manejo de grandes volúmenes de datos IoT. También se desarrollan habilidades en sensores (GPS), captura y procesamiento de imágenes, esenciales en la recopilación de datos desde la capa de percepción.

### **1.1.3 Motivación personal y experiencia profesional**

La motivación personal para abordar esta problemática combina una vocación por la automatización de procesos con un interés social en mejorar la eficiencia empresarial, especialmente en sectores que influyen en la calidad de vida. A lo largo de los cursos mencionados, se desarrolló una afinidad por crear soluciones tecnológicas que simplifican tareas rutinarias y complejas, visualizando en ello una vía para contribuir al bienestar social. La experiencia profesional durante la carrera, desde almacenero hasta asistente logístico, brindó una comprensión directa de los desafíos operacionales, destacando la importancia del cumplimiento de tiempos en cada etapa del proceso logístico. Además, la participación en el sector de aplicativos de transporte, resolviendo problemas asociados a viajes de taxi, mostró el potencial transformador de tecnologías como GPS, captura de imágenes y análisis de datos, reforzando el papel de las telecomunicaciones en el monitoreo remoto de operaciones distribuidas. Estas vivencias demostraron cómo un buen diseño permite automatizar procesos que funcionan de manera autónoma, liberando a los responsables de intervenciones constantes. Esta visión se alinea con los principios adquiridos durante la formación académica. La fusión entre teoría y práctica permitió identificar oportunidades reales donde las telecomunicaciones pueden generar impactos positivos, sostenibles y medibles en sectores clave.

## 1.2 Descripción y características del problema

### 1.2.1 Contexto del problema en la industria de *delivery*

La industria de servicios de *delivery* y logística de última milla ha experimentado un crecimiento exponencial en los últimos años, impulsada por el auge del comercio electrónico y los cambios en los hábitos de consumo de la población [4]. Sin embargo, este crecimiento acelerado ha puesto en evidencia múltiples ineficiencias operacionales que afectan tanto la rentabilidad de las empresas como la calidad del servicio ofrecido a los usuarios finales. Uno de los principales desafíos que enfrentan las empresas de *delivery* es obtener información precisa sobre las dimensiones y características de los paquetes que deben recoger y entregar.

Tabla 1.2: Proyecciones del Mercado de *delivery* en Perú [5].

Indicador	Valor	Año
Valor actual del mercado	US\$ 1.942 millones	2024
Valor de crecimiento actual	11,03 % anual	2029
Tasa de crecimiento esperada	15,6 %	2026
Ingresos proyectados	US\$ 2.951 millones	2029

Actualmente, este proceso se realiza de forma tradicional, basándose en mediciones manuales o estimaciones visuales por parte del remitente o del personal de la empresa, lo que genera diversos problemas operativos. La falta de datos exactos sobre el tamaño de los paquetes dificulta la planificación eficiente de rutas, la asignación adecuada de recursos de transporte y la optimización de la capacidad de carga de los vehículos motorizados. Esta problemática se agrava cuando se considera que las empresas de *delivery* manejan volúmenes crecientes de paquetes con características muy diversas, desde documentos pequeños hasta paquetes voluminosos con requisitos especiales de manipulación. La ausencia de un sistema automatizado para la determinación de medidas genera incertidumbre en la planificación operacional, resultando en situaciones donde los motorizados llegan a puntos de recojo sin la capacidad suficiente para transportar los paquetes, o, por el contrario, subutilizan su capacidad de carga al no tener información precisa sobre las dimensiones reales de los envíos.

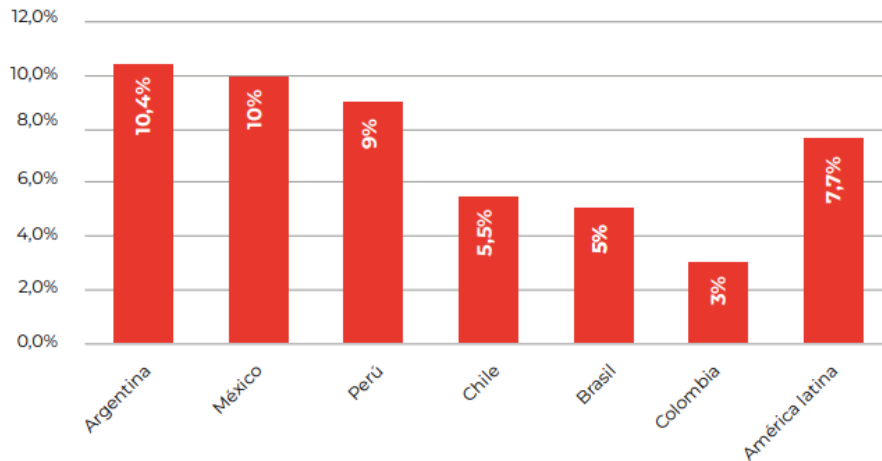


Figura 1.1: Gasto en tecnologías de la información - América Latina. Fuente: [6]

### 1.2.2 Desafío técnico desde la perspectiva de telecomunicaciones

Desde la perspectiva de la ingeniería de telecomunicaciones, el problema central radica en la necesidad de desarrollar un sistema distribuido que permita el procesamiento automatizado de imágenes capturadas por dispositivos móviles para la determinación precisa de las dimensiones de paquetes. Este desafío implica la integración eficiente y confiable de múltiples componentes tecnológicos.

Actualmente, las telecomunicaciones facilitan la captura y el almacenamiento de información, como fotografías y ubicaciones GPS, que pueden contribuir al seguimiento en tiempo real de los procesos logísticos. No obstante, el avance tecnológico permite ir más allá del simple almacenamiento de datos, habilitando el procesamiento inteligente de la información para extraer datos adicionales que generen ahorros significativos en horas-hombre y mejoren sustancialmente la eficiencia operativa.

El desafío específico consiste en desarrollar una fuente confiable para la determinación automatizada de las medidas de los paquetes, que permita optimizar tanto las rutas de distribución como la carga asignada a cada motorizado durante los procesos de recojo y entrega. Esta solución debe aprovechar las capacidades de procesamiento en la nube para realizar análisis complejos de imágenes sin requerir dispositivos móviles con especificaciones premium, garantizando así la escalabilidad y accesibilidad del sistema.

### 1.2.3 Características específicas del problema

Desde la perspectiva de la ingeniería de telecomunicaciones, el problema central radica en la necesidad de desarrollar un sistema distribuido que permita el procesamiento automatizado de imágenes capturadas por dispositivos móviles, con el fin de determi-

nar con precisión las dimensiones de los paquetes. Este desafío implica la integración eficiente y confiable de múltiples componentes tecnológicos.

Actualmente, las telecomunicaciones facilitan la captura y el almacenamiento de información, como fotografías y ubicaciones GPS, que pueden contribuir al seguimiento en tiempo real de los procesos logísticos. No obstante, el avance tecnológico permite ir más allá del simple almacenamiento de datos, habilitando el procesamiento inteligente de la información para extraer datos adicionales que generen ahorros significativos en horas-hombre y mejoren sustancialmente la eficiencia operativa.

El desafío específico consiste en desarrollar una fuente confiable para la determinación automatizada de las medidas de los paquetes, que permita optimizar tanto las rutas de distribución como la carga asignada a cada motorizado durante los procesos de recojo y entrega. Esta solución debe aprovechar las capacidades de procesamiento en la nube para realizar análisis complejos de imágenes sin requerir dispositivos móviles con especificaciones premium, garantizando así la escalabilidad y accesibilidad del sistema.

### **1.3 Importancia del problema y su solución**

#### **1.3.1 Perspectiva técnica**

Resolver este problema es clave debido al uso de tecnologías emergentes que están transformando las telecomunicaciones y la computación distribuida. El procesamiento en la nube ha madurado lo suficiente como para realizar análisis complejos de imágenes sin necesidad de ejecución local en los dispositivos. Esto democratiza el acceso a capacidades avanzadas de procesamiento, permitiendo que empresas con recursos limitados accedan a soluciones sofisticadas sin requerir infraestructuras costosas.

Los algoritmos de visión por computadora en entornos de nube distribuida escalan dinámicamente el procesamiento, optimizando tanto el rendimiento como los costos. La integración del Internet de las Cosas (IoT) con telecomunicaciones avanzadas habilita aplicaciones inteligentes capaces de procesar datos en tiempo real y proporcionar retroalimentación inmediata. Esta convergencia resulta vital en la transformación digital de diversos sectores económicos, donde la capacidad de analizar grandes volúmenes de datos automatizados representa una ventaja competitiva [7], [8].

#### **1.3.2 Perspectiva económica y financiera**

La automatización en la medición de paquetes contribuye directamente a mejorar la rentabilidad y la competitividad en los procesos logísticos, al generar ahorros significativos y optimizar el uso de recursos de transporte. El procesamiento en la nube

permite reducir la inversión en dispositivos de alto costo, facilitando el acceso a tecnologías avanzadas para pequeñas y medianas empresas (pymes). Asimismo, elimina la necesidad de infraestructura local y personal técnico especializado, transformando los costos fijos en gastos operacionales escalables [9].

La optimización de rutas y la maximización de la carga útil permiten reducir el consumo de combustible y los tiempos operativos, lo que se traduce en un aumento de la productividad y una mejora en los márgenes de rentabilidad.

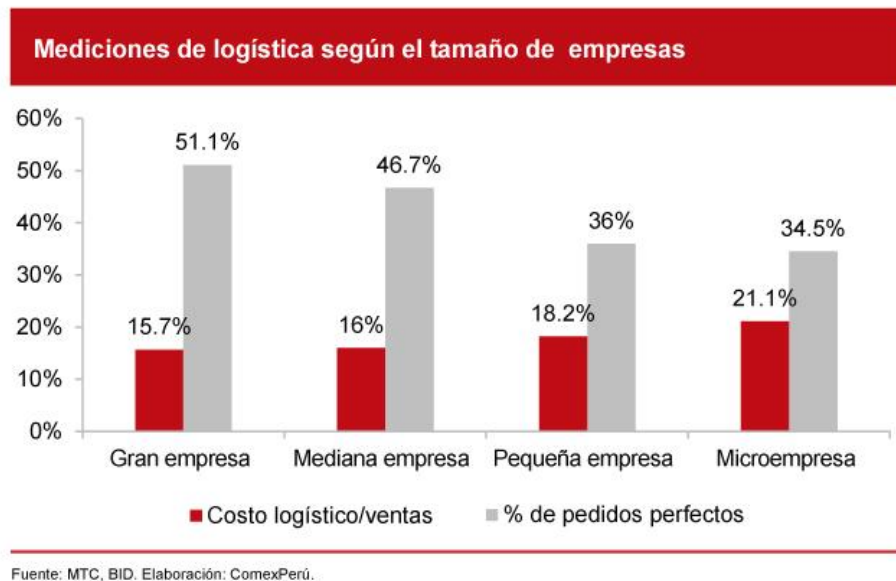


Figura 1.2: Gasto en tecnologías de la información - América Latina. Fuente: [10]

### 1.3.3 Perspectiva social y cultural

Socialmente, la solución propuesta contribuye a mejorar la calidad de vida tanto de los trabajadores logísticos como de los usuarios finales. La automatización de tareas repetitivas permite redirigir los recursos humanos hacia actividades de mayor valor agregado, lo que favorece la satisfacción laboral y el desarrollo profesional.

La mayor precisión y confiabilidad en los procesos de entrega fortalece la confianza en los servicios digitales, facilitando el acceso oportuno a bienes, especialmente en poblaciones que dependen del *delivery* como canal principal de abastecimiento. Además, la democratización de tecnologías avanzadas permite que las pequeñas y medianas empresas compitan en igualdad de condiciones con grandes corporaciones, fomentando la diversidad empresarial, la innovación, el empleo técnico especializado y el desarrollo local en sectores vinculados a tecnologías emergentes.

### 1.3.4 Perspectiva ambiental y de sostenibilidad

Abordar este problema es clave para reducir emisiones de gases de efecto invernadero y usar recursos energéticos eficientemente. La optimización de rutas, basada en medidas precisas de paquetes, permite planificar trayectos más cortos, disminuyendo consumo de combustible y emisiones de CO<sub>2</sub>.

La eficiencia en recojo y entrega, junto con rutas óptimas y verificación automatizada, hace las operaciones más cortas y seguras, minimizando el impacto ambiental y eliminando viajes innecesarios [11].

Los ahorros energéticos incluyen también la reducción del trabajo manual y procesos administrativos repetitivos, optimizando recursos y contribuyendo a la sostenibilidad empresarial y sectorial.

### 1.3.5 Perspectiva legal y reglamentaria

#### Cumplimiento de normativas de protección de datos

En Perú, la medición automatizada de paquetes con IoT debe cumplir la Ley N.º 29733, que exige consentimiento previo, finalidad clara, calidad y seguridad en el tratamiento de datos [12]. La solución debe proteger información sensible de remitentes, destinatarios, vehículos, empleados y clientes. Dado el gran volumen y flujo constante de datos, se requieren sólidas medidas de seguridad para garantizar el cumplimiento ético y legal.

#### Estándares de calidad en servicios logísticos

La logística demanda precisión, confiabilidad y transparencia. La automatización mediante sensores *IoT* y algoritmos de *Machine Learning* permite optimizar inventarios y rutas, proporcionando datos objetivos y verificables. El monitoreo continuo genera alertas tempranas ante anomalías, fortaleciendo el control de calidad y la conformidad normativa. La trazabilidad mejora la visibilidad en tiempo real, cumpliendo con las expectativas comerciales y las responsabilidades empresariales [13], [14].

#### Regulaciones de telecomunicaciones e *IoT*

Las normativas garantizan seguridad, interoperabilidad y conectividad. La Estrategia Nacional de IA en Perú impulsa el desarrollo de infraestructura digital y el despliegue de redes 5G, facilitando la adopción masiva de *IoT*. La Red Dorsal Nacional de Fibra Óptica, con más de 13,500 km operativos, permite la conexión de dispositivos a gran escala. Además, el procesamiento distribuido mediante el Centro Nacional de

Computación de Alto Rendimiento y los centros de datos en la nube fortalecen el *edge computing* y mejoran el desempeño de aplicaciones logísticas inteligentes.

### **Marco legal para la inteligencia artificial**

Se busca integrar la inteligencia artificial en los procesos empresariales mediante sistemas *IoT* para realizar análisis en tiempo real, utilizando algoritmos capaces de detectar riesgos o anomalías. La gestión de *Big Data* con sensores *IoT* permite tomar decisiones estratégicas, respaldadas por políticas que fomentan el desarrollo de talento en computación paralela y procesamiento de señales. Asimismo, se promueve el desarrollo de aplicaciones móviles logísticas, aprovechando la alta penetración del internet móvil y los *smartphones* para optimizar las entregas y monitorear las cargas en tiempo real [14].

#### **1.3.6 Perspectiva ética**

##### **Responsabilidad Social y Laboral**

La automatización en la medición plantea un dilema ético entre la eficiencia y el empleo. Esta investigación propone un enfoque de complementariedad tecnológica, en el cual la automatización potencia las habilidades humanas en lugar de generar desplazamiento laboral. Es necesario diseñar interfaces que faciliten la reconversión profesional hacia tareas de mayor valor, como el análisis, la gestión de excepciones y la supervisión. Asimismo, se requieren programas de capacitación que permitan a los trabajadores adaptarse a nuevos roles, asegurando que la tecnología promueva el desarrollo humano sin generar exclusión social en los sectores logísticos.

##### **Equidad Digital y Democratización Tecnológica**

El diseño ético debe contribuir al cierre de brechas digitales, evitando que el avance tecnológico incremente las desigualdades entre empresas. Para ello, se requieren arquitecturas escalables, interfaces intuitivas y modelos de precios accesibles que favorezcan la adopción tecnológica por parte de pequeñas y medianas empresas. Asimismo, es fundamental considerar las diferencias en infraestructura regional, diseñando soluciones que funcionen adecuadamente tanto en zonas con buena conectividad como en aquellas que presentan limitaciones.

## **Sostenibilidad Ambiental y Responsabilidad Climática**

La ética tecnológica incluye la evaluación de la huella de carbono generada por el procesamiento en la nube frente a los beneficios derivados de la optimización logística. Aunque la nube ofrece escalabilidad, su consumo energético debe ser balanceado con la reducción de emisiones lograda mediante un mejor aprovechamiento de espacios y la disminución de desplazamientos físicos. Es fundamental seleccionar proveedores comprometidos con el uso de energías renovables y desarrollar algoritmos eficientes que minimicen el consumo de recursos sin comprometer la precisión de los resultados.

## **Privacidad de Datos y Soberanía Informacional**

El manejo ético de datos va más allá del cumplimiento legal, al enfocarse en la protección de información sensible relacionada con operaciones comerciales y logísticas. Es fundamental aplicar el principio de *privacy by design*, incorporando mecanismos de anonimización, *end-to-end encryption* y políticas estrictas de retención que limiten el almacenamiento únicamente al tiempo necesario.

## **Responsabilidad y Rendición de Cuentas**

Es vital establecer mecanismos claros para gestionar errores, compensar daños económicos y permitir apelaciones frente a decisiones automatizadas. Asimismo, deben realizarse evaluaciones periódicas del impacto social, económico y ambiental, incorporando retroalimentación continua para maximizar los beneficios y mitigar los efectos negativos.

Esta visión ética integral garantiza que la tecnología respete los valores humanos y contribuya al desarrollo sostenible en el ámbito logístico, estableciendo un ejemplo responsable para futuras innovaciones.

### **1.4 Impactos previstos y beneficiarios**

#### **1.4.1 Impactos operacionales**

La implementación de la solución propuesta tendrá impactos operacionales significativos, mejorando la eficiencia de los procesos logísticos mediante tecnologías *IoT*, inteligencia artificial y visión artificial. Estas permiten recopilar, transmitir y analizar datos en tiempo real, automatizando tareas clave.



## **Reducción de tiempos de procesamiento**

La medición automatizada de paquetes elimina procesos manuales, reduciendo el tiempo requerido para registrar, verificar y procesar solicitudes. Gracias al uso de tecnologías *IoT* y al procesamiento masivo de datos, se agiliza la toma de decisiones y la selección de servicios adecuados, mejorando la experiencia del usuario [7], [15], [16].

## **Optimización de rutas de entrega**

Con datos precisos sobre las dimensiones de los paquetes, los sistemas pueden planificar rutas más eficientes considerando la capacidad de carga, el tiempo y la distancia. El uso de algoritmos de ruteo inteligente y planificación asistida por *IA* permite reducir costos, tiempos y emisiones, contribuyendo además al desarrollo sostenible [17].

## **Precisión en estimaciones de capacidad**

La información confiable sobre los paquetes permite ajustar de manera más precisa la carga por vehículo, evitando excesos. El análisis de datos históricos y predictivos posibilita una logística anticipada, apoyada en componentes analíticos que procesan grandes volúmenes de información *IoT* para mejorar la planificación operativa [11], [18].

## **Automatización de verificaciones y trazabilidad**

La automatización proporciona evidencia visual y documental confiable en cada transacción, mejorando la seguridad y la resolución de disputas. Sensores, cámaras y dispositivos de rastreo permiten la supervisión y verificación en tiempo real a lo largo de toda la cadena logística [7].

## **Detección de anomalías**

El análisis en tiempo real permite identificar desvíos o condiciones anómalas en el transporte, activando alertas preventivas que fortalecen la seguridad y la confiabilidad del servicio [7]. En conjunto, estos impactos representan una transformación operativa integral, orientada a una mayor precisión, automatización y capacidad de respuesta en las operaciones logísticas.

### 1.4.2 Impactos tecnológicos

La implementación de esta solución tendrá un impacto tecnológico significativo en el ámbito de las telecomunicaciones y las tecnologías *IoT*, estableciendo nuevos paradigmas en el procesamiento de imágenes aplicadas a la logística y demostrando el potencial de las arquitecturas distribuidas en la nube.

#### Procesamiento distribuido para análisis de imágenes

Utilizando modelos de lenguaje multimodal (*LLMs*) especializados en análisis visual-semántico, la nube manejará grandes volúmenes de datos visuales generados por dispositivos *IoT*. Esta arquitectura escalable facilitará el análisis en tiempo real de millones de imágenes para medir automáticamente dimensiones de paquetes. Esto reducirá costos al reemplazar mediciones manuales, mejorará la precisión y permitirá el monitoreo remoto continuo, optimizando la logística automatizada.

#### Integración con telecomunicaciones y aplicaciones móviles

La propuesta combina procesamiento intensivo con interfaces móviles accesibles, aprovechando redes 4G/5G para transmitir imágenes en alta resolución de manera eficiente. La arquitectura soporta miles de dispositivos simultáneamente, garantizando escalabilidad para grandes implementaciones empresariales. Además, se conecta con infraestructura existente para reducir la latencia y mejorar los tiempos de respuesta y la experiencia del usuario.

Así, se habilita el análisis de grandes volúmenes de datos visuales, sentando las bases para soluciones *Big Data* que cumplan con precisión, rapidez y confiabilidad en la medición automatizada de paquetes.

#### Innovación en aplicaciones móviles para logística

Se desarrollarán aplicaciones fáciles de usar para que usuarios no técnicos capturen dimensiones complejas mediante fotografías. La aplicación permitirá el monitoreo en tiempo real y el seguimiento preciso. Integrada con bases *NoSQL* y microservicios, la solución asegura escalabilidad y despliegue eficiente en la nube, aprovechando tecnologías *BaaS* para optimizar el rendimiento. La automatización transformará los centros de distribución, reduciendo tiempos y mejorando la precisión, además de facilitar la comunicación directa entre operadores y sistemas de gestión.

## **Impacto transformacional en logística**

Estos avances promoverán una transformación profunda, mejorando la eficiencia, precisión y transparencia. La solución habilitará sistemas inteligentes adaptables a las demandas del mercado, ofreciendo mayor visibilidad y control en la medición y clasificación de paquetes. Su adopción facilitará decisiones en tiempo real, optimizando recursos e inventarios, y sentará precedentes para aplicar soluciones similares en otros sectores que requieran procesamiento automatizado de datos visuales y mediciones precisas con tecnologías móviles y en la nube.

### **1.4.3 Impactos económicos**

Los impactos económicos se reflejan desde la productividad individual hasta la competitividad sectorial, aprovechando el procesamiento de imágenes con *IA* en la nube para optimizar las entregas urbanas.

#### **Maximización de productividad de motorizados**

Se optimiza la capacidad de transporte y se reducen tiempos improductivos mediante ruteo inteligente con algoritmos que asignan rutas más cortas y eficientes, disminuyendo tiempos y consumo de combustible, lo que permite realizar más entregas por viaje [11]. El monitoreo en tiempo real facilita decisiones rápidas y reduce intentos fallidos, mejorando la experiencia del cliente.

#### **Reducción de costos operacionales**

Incluye ahorro en combustible, tiempo de personal, comunicación y administración. El ruteo optimizado minimiza kilómetros recorridos y desgaste vehicular, generando hasta un 15 % de ahorro. La automatización reduce costos de mano de obra y mantenimiento. El monitoreo previene pérdidas y daños al activar alertas ante anomalías. La infraestructura en la nube asegura procesamiento eficiente y escalable, evitando inversiones locales costosas [7].

#### **Mejora en la rentabilidad empresarial**

La combinación de menores costos y mayores ingresos permite procesar más pedidos con la misma infraestructura, aumentando la rentabilidad y competitividad. Se estima una reducción de costos logísticos superior al 30 % y una mejora operativa de hasta el 40 %. La satisfacción del cliente se incrementa gracias a la visibilidad y las entregas

puntuales, fomentando la lealtad y la repetición de compra. El análisis de datos fortalece las decisiones estratégicas en demanda, inventarios y gestión de riesgos.

## **Generación de nuevo valor agregado**

Se crean servicios diferenciados mediante el uso de datos detallados para modelos de estimación de precios precisos. El *Quick Commerce* se potencia con entregas en menos de 90 minutos, mejorando la eficiencia en la última milla. El análisis en tiempo real habilita la personalización y la transparencia. La optimización de rutas favorece la sostenibilidad, reduciendo emisiones y reforzando la imagen ecológica como una ventaja competitiva clave [19].

### **1.4.4 Beneficiarios directos**

#### **Empresas de *delivery* y logística**

Son los principales beneficiarios, mejorando la eficiencia y la rentabilidad gracias al procesamiento de imágenes con *IA* en la nube. La optimización del ruteo reduce kilómetros recorridos y consumo de combustible, generando hasta un 15 % de ahorro en costos y desgaste vehicular. La infraestructura en la nube proporciona recursos escalables, evitando inversiones locales costosas. La rentabilidad mejora al gestionar más pedidos con los mismos recursos, reduciendo los costos logísticos en más del 30 % y aumentando la eficiencia hasta en un 40 %. Además, la logística verde contribuye a reducir emisiones y fortalecer la imagen de marca.

#### **Motorizados y personal operativo**

Mejoran su productividad y condiciones laborales mediante rutas optimizadas que permiten realizar más entregas en menos tiempo. La visión artificial facilita la identificación del tamaño de los paquetes, optimizando la carga y agilizando la operación. Esto reduce el estrés y la incertidumbre, mejorando los ingresos y las condiciones laborales. El monitoreo en tiempo real incrementa la seguridad al alertar sobre riesgos y desvíos.

#### **Clientes emisores de paquetes**

Experimentan mayor fiabilidad y transparencia gracias a rutas inteligentes que optimizan las entregas y el recojo. La captura fotográfica automatiza la medición del paquete, acelerando el proceso. La inspección automatizada mediante visión artificial garantiza calidad y confianza. La trazabilidad en tiempo real, combinando tecnologías *IoT* e

imágenes en la nube, permite rastrear los paquetes en cualquier momento [20].

## **Destinatarios de entregas**

Reciben un servicio más rápido y predecible, cumpliendo con los estándares de *quick commerce* en menos de 90 minutos. La comunicación en tiempo real ofrece seguimiento del paquete y contacto directo con el repartidor, reduciendo los intentos fallidos. La experiencia mejora gracias a la rapidez, precisión y transparencia, fomentando la lealtad. El monitoreo ambiental previene daños, asegurando la llegada en óptimas condiciones [11].

### **1.4.5 Beneficiarios indirectos**

#### **Para el sector de telecomunicaciones**

Este sector verá un aumento en la demanda de servicios especializados y mejoras en infraestructura. La solución incrementará la necesidad de conectividad fija y móvil de alta velocidad, impulsando la adopción de tecnologías 4G y 5G para procesar grandes volúmenes de datos en tiempo real. La arquitectura en la nube motivará inversiones en servicios como *IaaS*, *PaaS* y *FaaS*, además de optimizar redes y detectar anomalías, estimulando inversiones en centros de datos locales.

#### **Para la industria de desarrollo de software**

Se generarán nuevas oportunidades y mayor demanda de talento especializado en aplicaciones móviles, bases *NoSQL*, microservicios y despliegue en la nube. La solución promoverá mejores prácticas para aplicaciones *IoT* integradas con la nube, fomentará la investigación y el desarrollo local, y fortalecerá la colaboración academia-industria, alineándose con la Estrategia Nacional de Inteligencia Artificial (ENIA) del Perú [13].

#### **Para el medio ambiente y la sociedad**

La optimización de rutas reducirá las emisiones de CO<sub>2</sub> y el consumo de combustible, contribuyendo a un entorno urbano más saludable [5]. Menor congestión vehicular y menos viajes innecesarios mejorarán la movilidad en Lima, incentivando el uso de vehículos eléctricos y bicicletas para la última milla. Estos beneficios respaldan la *Logística 5.0*, que integra innovación tecnológica con sostenibilidad ambiental [11].

## Para el ecosistema de innovación tecnológica

La solución fortalecerá la innovación al demostrar cómo las tecnologías emergentes pueden resolver problemas reales. Actuará como catalizador para la madurez tecnológica y la formación en inteligencia artificial, alineándose con la ENIA, y fomentará la colaboración entre academia, industria y emprendedores. El proyecto puede mejorar la posición del Perú en los índices globales de *IA* y servir como modelo para futuras implementaciones públicas y privadas [13].

La implementación de sistemas automatizados de medición de paquetes mediante tecnologías *IoT* y procesamiento de imágenes en la nube representa un avance significativo hacia la logística inteligente del futuro. Los impactos multidimensionales de esta solución demuestran el potencial transformador de la convergencia tecnológica en telecomunicaciones, beneficiando directamente a empresas de *delivery*, motorizados y usuarios finales, mientras fortalece el ecosistema de innovación tecnológica nacional. Este proyecto establece un precedente para futuras aplicaciones de *IoT* en sectores críticos y contribuye al desarrollo de una infraestructura digital robusta que posiciona al país en la vanguardia de la transformación logística global.

## **Capítulo 2. Estado del arte o de la cuestión, alternativas de solución al problema o desafío a resolver**

### **2.1 Antecedentes de solución semejantes o similares al desafío de ingeniería**

#### **2.1.1 Sistemas de análisis visual con inteligencia artificial multimodal**

##### **Caso 1: GPT-4 Vision para análisis de objetos (OpenAI, Estados Unidos, 2023)**

GPT-4 Vision, lanzado en septiembre de 2023, representa el primer modelo de lenguaje de gran escala con capacidades multimodal nativas de OpenAI. El sistema demuestra capacidades avanzadas de interpretación visual con un 95 % de precisión en reconocimiento de objetos comunes y 78.5 % de efectividad en análisis de gráficos complejos según evaluaciones técnicas independientes [21].

En términos de estimación dimensional, GPT-4V utiliza razonamiento contextual para comparar tamaños relativos entre objetos, empleando elementos conocidos como referencias de escala. Las evaluaciones técnicas reportan una precisión de  $\pm 15\text{--}25\%$  de error en estimaciones dimensionales cuando se proporcionan referencias visuales adecuadas, mejorando a  $\pm 10\text{--}20\%$  bajo condiciones de iluminación controlada [22].



GPT-4      The image shows a package for a "Lightning Cable" adapter with three panels.

Panel 1: A smartphone with a VGA connector (a large, blue, 15-pin connector typically used for computer monitors) plugged into its charging port.

Panel 2: The package for the "Lightning Cable" adapter with a picture of a VGA connector on it.

Panel 3: A close-up of the VGA connector with a small Lightning connector (used for charging iPhones and other Apple devices) at the end.

The humor in this image comes from the absurdity of plugging a large, outdated VGA connector into a small, modern smartphone charging port.

Figura 2.1: Consulta a GPT-4 para análisis de múltiples imágenes. Fuente [23]

La arquitectura se basa en un *transformer* multimodal que integra un codificador visual especializado con el modelo de lenguaje GPT-4, estimado en 1.76 trillones de parámetros. El sistema procesa imágenes de hasta  $2048 \times 2048$  píxeles en formatos JPEG, PNG, GIF y WebP, utilizando técnicas de atención cruzada para correlacionar información visual con conocimiento lingüístico [23].

El modelo emplea un enfoque de *vision-language understanding* que permite no solo identificar objetos, sino también razonar sobre sus propiedades físicas, relaciones espaciales y características dimensionales mediante una interpretación contextual similar al razonamiento humano [21].

Las evaluaciones técnicas del sistema revelan un rendimiento variable según el contexto



de aplicación:

- *MMMU Benchmark*: 56.8 % en tareas multimodal complejas
- *MathVista*: 49.9 % en razonamiento visual-matemático
- *AI2D*: 78.2 % en interpretación de diagramas técnicos
- *ChartQA*: 78.5 % en análisis de gráficos y visualizaciones

Para tareas específicas de estimación dimensional, el sistema muestra mejor rendimiento en cajas rectangulares estándar, objetos con formas geométricas simples y productos comerciales conocidos, alcanzando precisiones útiles para la categorización logística y la clasificación de tarifas de envío [22].

Las evaluaciones técnicas identifican limitaciones significativas en aplicaciones que requieren precisión cuantitativa absoluta:

- **Sin calibración externa:** incremento del error a  $\pm 30\text{--}50\%$  en ausencia de referencias de escala conocidas.
- **Sensibilidad ambiental:** degradación de precisión con variaciones en iluminación, ángulos de captura y condiciones visuales.
- **Limitaciones de escala:** rendimiento reducido en objetos menores a 5 cm o con formas irregulares complejas.
- **Objetos problemáticos:** dificultades con elementos transparentes, reflectivos o sin texturas distintivas.

El *GPT-4V System Card* oficial de OpenAI reconoce explícitamente que “el modelo no está optimizado para mediciones precisas y puede proporcionar estimaciones aproximadas que requieren validación adicional para aplicaciones críticas” [21].

## Caso 2: Claude 3 Vision

Claude 3, lanzado en marzo de 2024 en sus variantes Haiku, Sonnet y Opus, establece un nuevo estándar en interpretación multimodal con un enfoque específico en razonamiento avanzado sobre contenido visual complejo. El sistema demuestra capacidades superiores a GPT-4V en interpretación de gráficos y documentos técnicos, alcanzando un 86.8 % en el benchmark MMLU y un 60.1 % en problemas matemáticos complejos con componentes visuales [24].

La arquitectura del modelo incorpora una ventana de contexto extendida de 200,000 tokens que incluye contenido visual, permitiendo el análisis simultáneo de múltiples imágenes y documentos dentro de una sola conversación. Esta capacidad es particularmente relevante para el análisis de inventarios complejos, donde se requiere correlacionar información entre múltiples fuentes visuales [25].

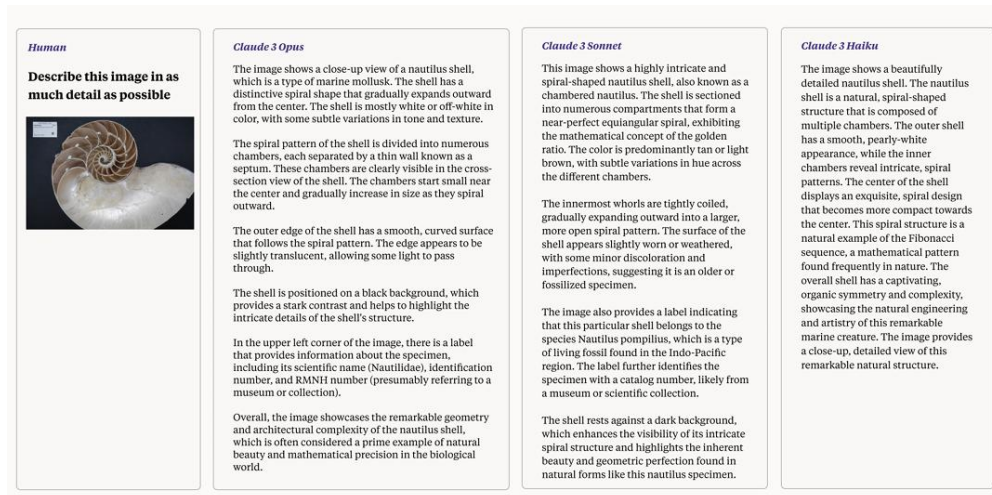


Figura 2.2: Identificación de objetos visualmente usando modelos de Claude 3. Fuente: [24]

Claude 3 destaca en el reconocimiento óptico de caracteres (*OCR*) en imágenes complejas, el procesamiento de documentos técnicos con *layouts* sofisticados y la comprensión contextual de diagramas industriales, superando consistentemente a modelos anteriores en *benchmarks* de comprensión documental [24].

El modelo exhibe capacidades avanzadas de razonamiento espacial que superan a sus predecesores en tareas que requieren comprensión de relaciones geométricas y propiedades físicas de objetos. Las evaluaciones técnicas independientes reportan un rendimiento superior en tareas de razonamiento espacial, con particular fortaleza en la interpretación de *layouts* complejos y relaciones proporcionales entre elementos.

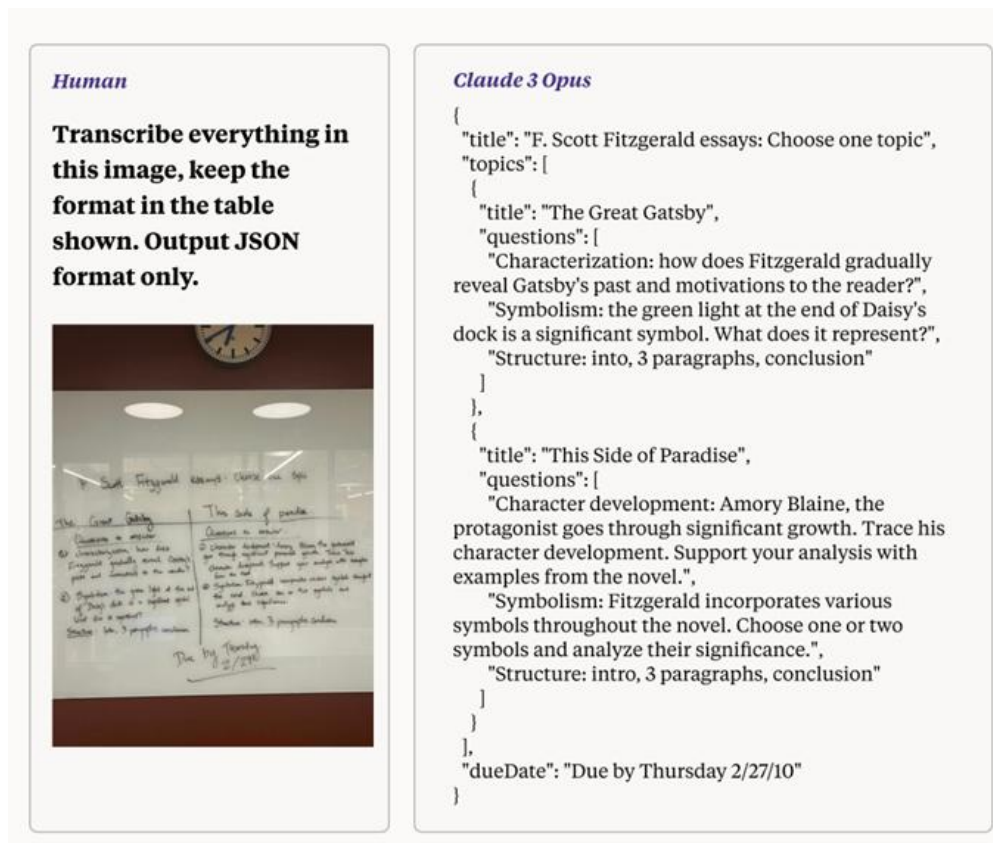


Figura 2.3: Solicitud de reconocimiento de una imagen y reorganización en formato JSON.Fuente: [24]

Para estimación dimensional, Claude 3 utiliza razonamiento contextual sofisticado que combina el reconocimiento de objetos conocidos con el análisis proporcional de elementos en la imagen. El sistema puede interpretar planos técnicos con dimensiones especificadas y extrapolar esta información para estimar las dimensiones de objetos fotografiados, logrando precisiones de  $\pm 20\text{--}30\%$  en estimaciones sin referencias calibradas externas [25].

La capacidad de procesamiento de documentos técnicos permite al modelo analizar especificaciones de productos, diagramas de embalaje y planos industriales, proporcionando estimaciones dimensionales basadas en la información contextual disponible en los documentos [24].

Claude 3 ofrece integración empresarial a través de la *Anthropic API v1*, que proporciona *endpoints RESTful* con autenticación basada en *API keys* y límites de 4,000 tokens por minuto para todas las variantes del modelo. La *API* soporta imágenes de hasta 20 MB en formatos JPEG, PNG, GIF, WebP y PDF, facilitando la integración con sistemas existentes de gestión documental.

La arquitectura de la *API* permite respuestas en múltiples formatos estructurados, incluyendo *JSON*, texto estructurado y *Markdown*, lo que facilita la integración con

sistemas *ERP*, plataformas de *e-commerce* y aplicaciones de gestión logística. El sistema soporta procesamiento *batch* para el análisis de grandes volúmenes de documentos y fotografías de inventario.

Las capacidades de *streaming* permiten respuestas en tiempo real para aplicaciones interactivas, mientras que el modelo de *pricing* por token de entrada y salida ofrece predictibilidad en los costos operacionales para implementaciones empresariales a gran escala.

Las implementaciones documentadas de Claude 3 en sectores logísticos y manufactureros incluyen aplicaciones específicas que aprovechan sus capacidades multimodales avanzadas:

- **Análisis de especificaciones técnicas:** procesamiento automatizado de documentos de productos que incluyen planos técnicos con dimensiones especificadas, permitiendo extraer automáticamente información dimensional para sistemas de gestión de inventarios.
- **Interpretación de diagramas de embalaje:** análisis de documentos de *packaging* que especifican configuraciones de empaque, permitiendo optimizar la utilización de espacio en contenedores y vehículos de transporte basándose en la interpretación visual de diagramas complejos.
- **Auditoría visual de inventarios:** procesamiento de fotografías de almacenes y centros de distribución para identificar discrepancias entre el inventario físico y los registros digitales, utilizando capacidades de reconocimiento de objetos y análisis espacial.
- **Análisis de documentos comerciales:** interpretación de facturas, órdenes de compra y documentos de envío que incluyen especificaciones de productos, automatizando la extracción de información dimensional crítica para procesos logísticos [24].

Las ventajas competitivas identificadas incluyen la ventana de contexto extendida (200K frente a 128K tokens de GPT-4), mejor comprensión de documentos complejos, razonamiento espacial más avanzado y menor tendencia a alucinaciones en datos técnicos críticos [24].

Las limitaciones operacionales incluyen una precisión dimensional comparable a GPT-4V ( $\pm 20\text{--}30\%$ ), dependencia de la calidad de imagen para un *OCR* efectivo, costos por token potencialmente superiores a los de alternativas, y un ecosistema menos maduro de herramientas de terceros en comparación con OpenAI.

### Caso 3: Gemini 2.0 Flash

Gemini 2.0 Flash, lanzado en diciembre de 2024, representa la segunda generación de modelos multimodales de Google con optimizaciones específicas para velocidad de procesamiento y análisis visual en tiempo real. El modelo incorpora una arquitectura *transformer* multimodal de segunda generación optimizada para respuestas rápidas (denominación *Flash*), logrando velocidades hasta  $2\times$  superiores a Gemini 1.0 en procesamiento multimodal [26], [27].

El sistema soporta modalidades múltiples incluyendo texto, imagen, audio y video, con capacidad de procesamiento de imágenes de hasta 30 MB en formatos *JPEG*, *PNG*, *GIF*, *WebP*, *PDF*, *SVG* y *HEIC*. La ventana de contexto se expande masivamente a 2 millones de tokens, permitiendo el análisis simultáneo de múltiples documentos visuales complejos dentro de una sola sesión.

Las mejoras en análisis visual incluyen mejor razonamiento espacial, capacidades de análisis en tiempo real optimizadas y procesamiento simultáneo de hasta 20 objetos en una imagen individual, superando significativamente las limitaciones de 8–10 objetos de Gemini 1.0 [27].

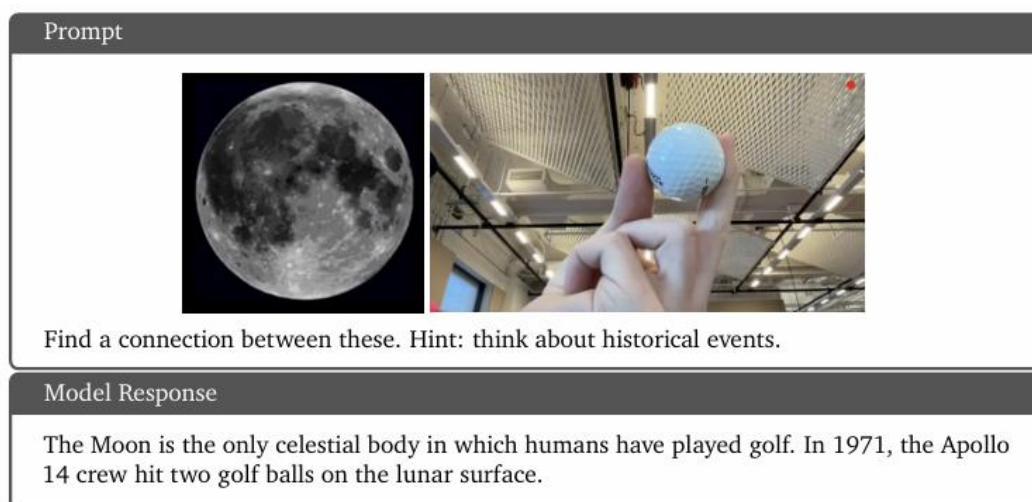


Figura 2.4: Se solicita a Gemini reconocer las imágenes y encontrar una relación entre ellas.

Gemini 2.0 Flash integra *Google Lens* como módulo nativo, eliminando la arquitectura de integración externa de generaciones anteriores. Esta fusión completa permite acceso directo al *Knowledge Graph* de Google y a la base de datos de productos de *Google Shopping* para la identificación automática de objetos de referencia conocidos.

El sistema demuestra precisión mejorada en estimaciones dimensionales, alcanzando márgenes de  $\pm 8\text{--}15\%$  con objetos de referencia en productos estándar,  $\pm 5\text{--}12\%$  en

condiciones de laboratorio controlado, y  $\pm 12\text{--}25\%$  en uso típico de usuarios reales. Estas métricas representan una mejora del  $30\text{--}40\%$  comparado con Gemini 1.0 [27].

Las capacidades incluyen detección automática de objetos de referencia conocidos, comparación automática con elementos de escala estándar y mejor adaptación a condiciones variables de iluminación y ángulos de captura. El tiempo de procesamiento promedio se reduce a 2–4 segundos, comparado con 5–8 segundos de versiones anteriores.

Las evaluaciones técnicas de Gemini 2.0 Flash revelan variabilidad significativa en precisión según condiciones operacionales. En condiciones controladas de laboratorio con iluminación estándar y ángulos óptimos, el sistema alcanza precisiones de  $\pm 5\text{--}12\%$ , competitivas con sistemas de visión por computadora tradicionales para aplicaciones no críticas.

En condiciones reales de uso, incluyendo variaciones de iluminación, ángulos subóptimos y *backgrounds* complejos, la precisión se degrada a  $\pm 12\text{--}25\%$ , manteniendo utilidad para categorización logística y estimaciones aproximadas. La adaptación a condiciones variables representa una mejora sustancial comparado con generaciones anteriores.

Implementaciones específicas en logística incluyen análisis de inventario en tiempo real para *Google Shopping*, medición automatizada de productos para *Google Merchant Center*, optimización de embalaje en centros de distribución de Google y análisis de fotografías de productos para *Google Lens Shopping*, demostrando aplicabilidad práctica en operaciones comerciales reales.

Las ventajas competitivas específicas incluyen velocidad significativamente superior comparado con GPT-4V debido a optimizaciones *Flash*, integración nativa con *Google Lens* y *Knowledge Graph*, mejor manejo de múltiples objetos simultáneos y acceso privilegiado a la base de datos de productos de *Google Shopping* para identificación automática.

Las limitaciones operacionales identificadas incluyen dependencia del ecosistema Google para rendimiento máximo, menor precisión que sensores dedicados de hardware, variabilidad en rendimiento entre dispositivos Android de diferentes fabricantes y documentación técnica detallada aún limitada debido al reciente lanzamiento.

El modelo establece un nuevo estándar en velocidad de procesamiento multimodal, posicionándose como solución óptima para aplicaciones logísticas que requieren análisis visual rápido con precisión moderada, particularmente en entornos que ya utilizan infraestructura Android y servicios Google.

### 2.1.2 Aplicaciones de LLMs multimodal en logística y comercio electrónico

#### Caso 4: Vision Transformer (ViT) para reconocimiento y dimensionado de productos

##### Aplicaciones académicas de Vision Transformers en estimación dimensional

*Vision Transformer* (ViT), introducido por Dosovitskiy et al. en Google Research (2020) y publicado en ICLR 2021, revolucionó el campo de *computer vision* al demostrar que arquitecturas *transformer* puras pueden superar las redes neuronales convolucionales tradicionales en tareas de reconocimiento de imágenes. El modelo alcanza un 94.2% de precisión en *ImageNet classification* cuando se preentrena en conjuntos de datos suficientemente grandes [28].

La arquitectura divide imágenes en *patches* de  $16 \times 16$  píxeles que se procesan como secuencias, similar al procesamiento de tokens en modelos de lenguaje. Esta metodología permite un *transfer learning* eficiente para nuevas categorías de productos, facilitando la adaptación a catálogos específicos de *e-commerce* sin necesidad de reentrenamiento completo [29].

Las implementaciones académicas posteriores han demostrado aplicabilidad directa en *retail analytics*. DINOv2 (Meta AI Research, 2023) introduce *self-supervised learning* que elimina la dependencia de conjuntos de datos etiquetados masivos, logrando representaciones visuales robustas especialmente efectivas para el reconocimiento de productos con variaciones en ángulos, iluminación y *background* [27], [30].

Para estimación dimensional, las investigaciones académicas reportan precisiones de  $\pm 12\text{--}18\%$  en objetos con referencias visuales, utilizando *visual reasoning* sobre *features transformer* que capturan relaciones espaciales complejas entre elementos en la imagen. Esta precisión es competitiva con sistemas comerciales, manteniendo la ventaja de una metodología académicamente verificable [29], [30].

Florence (Microsoft Research, 2021) establece un *framework* multimodal que combina *vision transformers* con capacidades de lenguaje natural, demostrando aplicabilidad directa para tareas que requieren comprensión semántica de productos y estimación de propiedades físicas. El modelo utiliza *contrastive learning* entre modalidades visual y textual [31].

Las evaluaciones académicas independientes confirman escalabilidad para el procesamiento de millones de productos, con implementaciones que mantienen eficiencia computacional mediante técnicas de *patch embedding* optimizadas y *sparse attention mechanisms* adaptados para imágenes de alta resolución típicas en fotografía comercial [32].

Estudios recientes sobre *Scaling Vision Transformers to Gigapixel Images* (2022) demuestran aplicabilidad en *retail analytics* de alta resolución, donde se requiere análisis detallado de texturas, materiales y características dimensionales no evidentes en resoluciones estándar. Estas implementaciones utilizan *hierarchical attention mechanisms* que procesan progresivamente desde características globales hasta detalles locales [32].

Las adaptaciones académicas incluyen *fine-tuning* específico para categorías de productos, donde modelos preentrenados se especializan en dominios particulares (electrónicos, textiles, alimentos), manteniendo la robustez general del *framework ViT* mientras optimizan la precisión para características particulares de cada categoría [28].

## 2.2 Características de soluciones semejantes o similares

### 2.2.1 Arquitecturas de procesamiento predominantes

Las soluciones analizadas convergen en arquitecturas de procesamiento que combinan múltiples enfoques tecnológicos. Los sistemas comerciales líderes utilizan *transformers* multimodales que integran *vision encoders* especializados (CLIP, ALIGN, ViT) con *large language models* para interpretación semántica avanzada. Las implementaciones actuales emplean *attention mechanisms cross-modal* que permiten la correlación directa entre características visuales y comprensión textual, facilitando la estimación dimensional mediante razonamiento contextual [28], [29].

### 2.2.2 Capacidades de interpretación dimensional

Las capacidades de interpretación dimensional se fundamentan en el reconocimiento automático de referencias de escala, donde los modelos identifican objetos conocidos (monedas, tarjetas de crédito, manos humanas) para establecer proporciones relativas. El análisis de perspectiva y profundidad visual utiliza *depth estimation* implícito derivado de características de textura, sombras y relaciones espaciales capturadas por *attention mechanisms* [31], [33].

Los sistemas más avanzados incorporan razonamiento contextual sofisticado que combina el reconocimiento de objetos conocidos con el análisis proporcional de elementos en la imagen. Estos modelos pueden interpretar planos técnicos con dimensiones especificadas y extrapolar esta información para estimar las dimensiones de objetos fotografiados, logrando precisiones variables según la implementación.



### 2.2.3 Precisión y confiabilidad según implementación

La precisión varía significativamente según la implementación y las condiciones operacionales. Los *LLMs* multimodales alcanzan márgenes de error de  $\pm 10\text{--}25\%$  en condiciones óptimas, los *Vision Transformers* académicos reportan precisiones de  $\pm 12\text{--}18\%$ , mientras que los sistemas comerciales con hardware especializado pueden alcanzar márgenes de  $\pm 2\text{--}5\text{ mm}$  para objetos regulares bajo condiciones controladas [28], [29], [33].

La degradación de precisión ante iluminación deficiente, ángulos subóptimos y *back-grounds* complejos representa un desafío común. Los sistemas empresariales mitigan estas limitaciones mediante estaciones de medición controladas y múltiples puntos de captura, mientras que las soluciones basadas únicamente en inteligencia artificial dependen de técnicas de *data augmentation* y *prompt engineering* avanzado.

## 2.3 Compendio de tecnologías, herramientas, métodos, modelos utilizados con éxito

### 2.3.1 Modelos de lenguaje multimodal predominantes

Los modelos de gran escala comerciales incluyen *GPT-4 Vision* (OpenAI), con una arquitectura estimada en 175 mil millones de parámetros multimodal; *Claude Sonnet 4* (Anthropic), optimizado para razonamiento visual avanzado; *Gemini 2.0 Flash* (Google), con optimizaciones de velocidad para procesamiento en tiempo real; y implementaciones académicas como *LLaVA*, que proporcionan alternativas *open-source* para investigación.

Las arquitecturas especializadas predominantes incluyen *CLIP* (*Contrastive Language-Image Pre-training*), que establece *embeddings joint* para modalidades visual y textual; *BLIP-2* (*Bootstrapped vision-language pre-training*), optimizado para comprensión visual detallada; *InstructBLIP*, adaptado para seguimiento de instrucciones específicas; y *MiniGPT-4*, diseñado para eficiencia en recursos computacionales limitados [28].

### 2.3.2 Frameworks de visión por computadora académicos

#### Modelos académicos y arquitecturas híbridas en visión computacional

*Vision Transformers* (ViT) representan el estado del arte en análisis visual académico, con implementaciones que incluyen *ViT-Base* (86 millones de parámetros), *ViT-Large* (307 millones de parámetros) y *ViT-Huge* (632 millones de parámetros), según los requerimientos de precisión frente a eficiencia computacional [31].

Los modelos híbridos combinan fortalezas de arquitecturas convolucionales y *transformer*, incluyendo *DeiT* (*Data-efficient image Transformers*), *Swin Transformer*, optimizado para imágenes de alta resolución, y *EfficientViT*, diseñado para implementaciones móviles y *edge computing*.

### 2.3.3 Tecnologías de infraestructura en la nube

#### Infraestructura de inteligencia artificial como servicio y computación distribuida

Las plataformas de inteligencia artificial como servicio incluyen *OpenAI API* con *GPT-4 Vision endpoints*, *Claude API* de Anthropic para análisis multimodal vía *REST*, *Google Vertex AI* con modelos *Gemini* integrados, *Azure Cognitive Services* combinando capacidades de *Computer Vision* y lenguaje natural, y *AWS Bedrock*, que proporciona acceso unificado a múltiples modelos *foundation* [28].

Los servicios de computación distribuida utilizan *AWS Lambda* para procesamiento *serverless* con escalamiento automático, *Google Cloud Run* para *containerización* y despliegue simplificado, *Azure Container Instances* para escalabilidad elástica, y *Kubernetes* para orquestación de microservicios en implementaciones híbridas *cloud-on-premises* [28], [30].

### 2.3.4 Herramientas de desarrollo e integración

#### Bibliotecas de integración y SDKs para desarrollo móvil

Las bibliotecas de integración predominantes incluyen *LangChain* como *framework* integral para aplicaciones basadas en *LLM*, *LlamaIndex* optimizado para *Retrieval-Augmented Generation* con datos multimodales, *Haystack* proporcionando *pipelines* de procesamiento *end-to-end*, y *Transformers* (HuggingFace), que ofrece acceso unificado a modelos preentrenados [28], [31].

Los *SDKs* para desarrollo móvil facilitan la integración nativa, incluyendo *React Native* con *plugins* especializados para *APIs* de inteligencia artificial, *Flutter* con *packages* optimizados para *computer vision*, *Swift/Kotlin* con *SDKs* nativos de proveedores principales, y *Progressive Web Apps* para acceso universal *cross-platform* [29].

### 2.3.5 Metodologías de prompt engineering y optimización

#### Técnicas avanzadas de diseño de *prompts* y optimización de modelos

Las técnicas avanzadas de diseño de *prompts* incluyen el uso de *few-shot learning* para abordar tareas específicas de medición dimensional mediante la provisión de ejemplos representativos; *chain-of-thought prompting* para fomentar el razonamiento paso a paso en el análisis de características espaciales; *structured output formatting* para asegurar consistencia y precisión en la generación de datos cuantitativos; y *multi-turn conversation* para permitir la refinación progresiva de estimaciones a través de interacciones iterativas.

Por otro lado, las estrategias de optimización del modelo contemplan el *fine-tuning* con conjuntos de datos específicos del dominio logístico; la aplicación de *Retrieval-Augmented Generation (RAG)* para integrar conocimiento empresarial especializado durante la generación de respuestas; el uso de *in-context learning* con ejemplos calibrados de medición para mejorar la adaptación contextual del modelo; y técnicas de *parameter-efficient fine-tuning*, como *LoRA* y *Adapters*, que permiten una personalización eficiente en costos sin requerir el ajuste completo de todos los parámetros del modelo [31].

## 2.4 Conjunto de características y especificaciones para la solución óptima

### 2.4.1 Arquitectura tecnológica óptima

#### Propuesta de arquitectura óptima

Basándose en el análisis de soluciones existentes, la arquitectura óptima debe combinar modelos de lenguaje multimodal de última generación con infraestructura *cloud* escalable. El sistema debe utilizar un modelo multimodal como motor principal de análisis, capaz de interpretar simultáneamente entradas visuales y textuales para tareas como estimación dimensional, clasificación logística y auditoría documental.

Esta integración permite aprovechar capacidades avanzadas de razonamiento contextual, atención cruzada entre modalidades y generación estructurada de respuestas, mientras se garantiza escalabilidad operativa mediante servicios distribuidos en la nube.

### 2.4.2 Precisión y confiabilidad requeridas

Para aplicaciones logísticas comerciales, el sistema debe alcanzar precisiones de  $\pm 15\text{--}20\%$  en condiciones reales de uso, mejorando a  $\pm 10\text{--}15\%$  con referencias de escala adecuadas. Esta precisión es suficiente para la categorización de tarifas de envío, la optimización básica de carga vehicular y las estimaciones de costos logísticos sin requerir inversión en hardware especializado de sistemas *enterprise*.

### 2.4.3 Limitaciones y alcances reconocidos

El sistema está diseñado para realizar estimaciones dimensionales aproximadas, adecuadas para la categorización logística, pero no para aplicaciones que requieran precisión milimétrica. La solución es óptima para paquetes regulares de *e-commerce*, dentro del rango de 5 cm a 2 m en cualquier dimensión.

Las limitaciones incluyen pérdida de precisión con objetos transparentes, altamente reflectivos o de formas extremadamente irregulares. El sistema requiere una iluminación mínima adecuada y no está optimizado para condiciones de muy baja luminosidad sin asistencia de *flash*.

La implementación está inicialmente orientada al mercado peruano.

## Capítulo 3. Diseño de la solución al desafío de ingeniería

El capítulo describe el diseño de la solución técnica al problema de la estimación manual e inexacta de dimensiones de paquetes en pymes de *delivery* en Lima. Basado en la metodología Design Thinking, el proceso asegura que cada decisión técnica responda a las necesidades reales de los usuarios y al contexto local.

Se siguen cuatro fases del método aplicadas al diseño: Empatizar, Definir, Idear y Prototipar, dejando la fase de Evaluación para el siguiente capítulo donde se analizará la viabilidad técnica, económica y social de la solución diseñada.

Cada etapa genera entregables que reflejan las competencias del Ingeniero de Telecomunicaciones de la PUCP, vinculando el diseño con el perfil profesional. El capítulo cierra con una reflexión crítica sobre fortalezas, limitaciones y mejoras futuras, sentando la base para la implementación del sistema.

### 3.1 Introducción metodológica – Enfoque Design Thinking

#### 3.1.1 Justificación del enfoque centrado en el usuario

La aplicación de Design Thinking en este proyecto de ingeniería de telecomunicaciones permite equilibrar la complejidad técnica con las necesidades reales de las pymes de *delivery* en Lima. Este enfoque integra el rigor de la ingeniería —infraestructura cloud,

latencia, disponibilidad, escalabilidad y optimización en dispositivos móviles— con la comprensión de las condiciones operativas reales, evitando soluciones avanzadas pero inviables o desconectadas del contexto.

Su carácter iterativo facilita ajustes basados en retroalimentación, garantizando una solución técnicamente robusta, económicamente viable y adoptable, alineada con las competencias profesionales del ingeniero de telecomunicaciones y los Objetivos de Desarrollo Sostenible.

### 3.1.2 Fases de Design Thinking aplicadas en este capítulo

Este capítulo desarrolla cuatro de las cinco fases del Design Thinking, enfocándose en el diseño de la solución:

- **Empatizar:** Se busca comprender las necesidades reales de los usuarios, identificando los problemas operativos que enfrentan las empresas de *delivery* al estimar manualmente las dimensiones de paquetes.
- **Definir:** Se establecen los requerimientos funcionales y no funcionales que debe cumplir el sistema de gestión logística a partir de los hallazgos obtenidos.
- **Idear:** Se evalúan múltiples alternativas tecnológicas mediante criterios de escalabilidad, costo y viabilidad para seleccionar la solución óptima.
- **Prototipar:** Se materializa el diseño técnico detallado, incluyendo la arquitectura de sistema, modelo de datos, módulo de IA y especificaciones de red.

## 3.2 Fase 1: Empatizar – Análisis del contexto y usuarios

### 3.2.1 Contexto del problema y metodología de investigación

#### Contexto geográfico e institucional

El proyecto se desarrolla en el contexto de las pequeñas y medianas empresas de *delivery* que operan en Lima, Perú, una ciudad de alta densidad urbana. El mercado logístico local se caracteriza por la predominancia de procesos manuales, la alta informalidad, el uso intensivo de motorizados en motocicletas para entregas de último kilómetro y una creciente demanda de servicios de *delivery* impulsada por el comercio electrónico.

Estas empresas operan sin sistemas tecnológicos integrados, dependiendo de la comunicación por WhatsApp, los registros en hojas de cálculo y la coordinación telefónica, lo que genera ineficiencias, falta de trazabilidad y dificultades para escalar sus operaciones.

## Técnicas de investigación aplicadas

Para comprender el contexto y las necesidades de los usuarios, se aplicaron dos técnicas principales:

1. **Análisis del flujo logístico:** Se estudió el proceso desde el registro de pedidos hasta la entrega final en pequeñas empresas de *delivery*, identificando inefficiencias y oportunidades de automatización.
2. **Análisis comparativo de soluciones:** Se evaluaron las herramientas tecnológicas existentes, tanto informales como comerciales, analizando sus funciones, limitaciones y costos. Este análisis evidenció vacíos en las herramientas actuales y oportunidades para integrar tecnologías emergentes como la inteligencia artificial multimodal.

## Actores clave del sistema

El análisis identificó cinco actores clave en el sistema:

- **Clientes emisores:** Solicitan envíos y requieren transparencia en costos y trazabilidad de sus pedidos.
- **Administradores:** Gestionan la operación logística y necesitan herramientas para supervisión centralizada, asignación eficiente de recursos y toma de decisiones basada en datos.
- **Motorizados:** Ejecutan las entregas en campo y demandan información clara sobre asignaciones, rutas optimizadas y herramientas para registrar evidencias.
- **Destinatarios:** Reciben los paquetes y esperan puntualidad e información sobre el estado de su entrega.

La tabla del Anexo A detalla el rol específico de cada actor, sus necesidades principales y los problemas operativos actuales que enfrentan, información fundamental para la definición de requerimientos del sistema. Esta caracterización exhaustiva de actores permite asegurar que el diseño técnico considere las perspectivas de todos los usuarios que interactuarán con el sistema, desde la interfaz móvil optimizada para operación en campo por motorizados hasta el dashboard web con análisis en tiempo real para administradores.

### 3.2.2 Síntesis de hallazgos

La fase de empatía reveló hallazgos críticos organizados en tres categorías que fundamentan el diseño técnico posterior: necesidades identificadas, restricciones del contexto y oportunidades de mejora mediante tecnología. Estos hallazgos, desarrollados en el Anexo B, representan la transformación de observaciones cualitativas en insumos estructurados que guiarán las decisiones de arquitectura, selección de tecnologías y priorización de funcionalidades en las fases subsecuentes del proceso de diseño.

Estos hallazgos revelan que el desafío técnico trasciende la simple automatización de mediciones, representando una oportunidad para transformar integralmente la gestión logística de pequeñas empresas mediante tecnologías de telecomunicaciones modernas que hasta ahora estaban reservadas para empresas con grandes presupuestos tecnológicos.

## 3.3 Fase 2: Definir – Reformulación del problema técnico

### 3.3.1 Declaración del problema de ingeniería

Con base en los hallazgos de la fase de empatía, se formula el siguiente problema técnico:

Las pequeñas y medianas empresas de *delivery* en Lima enfrentan ineficiencias operativas derivadas de la estimación manual e inexacta de dimensiones de paquetes, la falta de trazabilidad en tiempo real de las entregas, y la gestión desarticulada mediante herramientas informales (WhatsApp, hojas de cálculo). Esta problemática genera costos ocultos por viajes innecesarios, insatisfacción de clientes por reprogramaciones, y limitaciones para escalar operaciones.

Desde la perspectiva de ingeniería de telecomunicaciones, se requiere diseñar una arquitectura distribuida que integre procesamiento de imágenes con inteligencia artificial en la nube y aplicaciones móviles, todo bajo un modelo económicamente viable para empresas emergentes y cumpliendo con normativas de protección de datos personales.

### 3.3.2 Requerimientos funcionales del sistema



Tabla 3.1: Requisitos funcionales del sistema

Categoría	Código	Descripción
Gestión de usuarios y roles	RF1.1	El sistema debe permitir el registro y autenticación de usuarios.
	RF1.2	El sistema debe permitir que el Administrador gestione pedidos, usuarios y visualice motorizados en tiempo real.
	RF1.3	El sistema debe permitir que el Cliente cree pedidos, registre datos y suba fotos.
	RF1.4	El sistema debe permitir que el Motorizado reciba y actualice el estado de pedidos asignados.
Gestión de pedidos	RF2.1	El cliente debe poder crear pedidos en el sistema, incluyendo dirección de recojo, dirección de entrega, cliente, foto, celular y distrito.
	RF2.2	Al crear un pedido, el cliente debe subir fotos del paquete.
	RF2.3	El sistema debe procesar las fotos usando IA multimodal para detectar objetos de referencia y calcular dimensiones.
	RF2.4	El sistema debe guardar las dimensiones junto con la información del pedido y permitir su edición manual.
	RF2.5	El administrador debe poder asignar pedidos a un motorizado.
Gestión de motorizados	RF3.1	El motorizado debe visualizar en la aplicación móvil los pedidos asignados.
	RF3.2	El motorizado debe poder capturar foto en el recojo del pedido (evidencia).
	RF3.3	El motorizado debe poder capturar foto en la entrega (evidencia).
	RF3.4	El sistema debe cambiar automáticamente el estado del pedido al capturar fotos de evidencia.
	RF3.5	El motorizado puede visualizar los pedidos en un mapa.
Dashboard Web (administrador)	RF4.1	El administrador debe visualizar en un dashboard la lista de pedidos, estados, dimensiones, fotos y motorizado asignado.
Continúa en la siguiente página		

Tabla 3.1 – continuación de la página anterior

<b>Categoría</b>	<b>Código</b>	<b>Descripción</b>
	RF4.2	El dashboard debe permitir filtrar pedidos por estado, cliente o motorizado.
	RF4.3	El administrador debe poder visualizar la ubicación en tiempo real de los motorizados en un mapa dentro del dashboard web.
Notificaciones y sincronización	RF5.1	El motorizado debe recibir notificaciones push automáticas cuando se le asigne un nuevo pedido.
	RF5.2	El cliente debe recibir notificaciones push sobre cambios de estado del pedido.
Seguridad y consistencia	RF6.1	El sistema debe asegurar el acceso por roles (cada usuario solo ve lo que le corresponde).
	RF6.2	El sistema debe cifrar las comunicaciones mediante TLS y proteger el acceso a datos mediante reglas de seguridad de Firebase.

### 3.3.3 Requisitos no funcionales

Tabla 3.2: Requisitos no funcionales del sistema

<b>Categoría</b>	<b>Código</b>	<b>Descripción</b>
Rendimiento	RNF1.1	El sistema debe procesar la estimación de dimensiones de un paquete mediante IA en menos de 10 segundos después de subir la foto.
	RNF1.2	El dashboard web debe cargar la información principal en menos de 5 segundos.
Disponibilidad	RNF2.1	El sistema debe mantener una disponibilidad mínima del 99.5 % aprovechando la infraestructura cloud de Firebase.
	RNF2.2	Debe existir backup automático de datos en Firebase para evitar pérdida de información.
Seguridad	RNF3.1	Toda comunicación entre cliente y servidor debe usar cifrado SSL/TLS (implementado por defecto en Firebase).

Continúa en la siguiente página

Tabla 3.2 – continuación de la página anterior

<b>Categoría</b>	<b>Código</b>	<b>Descripción</b>
	RNF3.2	Las contraseñas deben almacenarse de forma segura mediante Firebase Auth.
	RNF3.3	El acceso a datos debe estar protegido mediante reglas de seguridad basadas en roles.
	RNF4.1	La aplicación móvil debe ser compatible con dispositivos Android 12.0 (API level 31) o superior, considerando el uso de terminales de gama media a baja.
	RNF4.2	El dashboard web debe funcionar correctamente en navegadores modernos (Chrome, Firefox, Edge).
Compatibilidad	RNF4.1	La aplicación móvil debe ser compatible con dispositivos Android 12.0 (API level 31) o superior, considerando el uso de terminales de gama media a baja.
	RNF4.2	El dashboard web debe funcionar correctamente en navegadores modernos (Chrome, Firefox, Edge).
Usabilidad	RNF5.1	La interfaz debe ser intuitiva y permitir completar las tareas principales sin necesidad de capacitación extensa.
	RNF5.2	Los mensajes de error deben ser claros y comprensibles.
Fiabilidad del Modelo de IA	RNF6.1	El modelo de IA debe tener un error promedio menor al 20 % en la estimación de dimensiones cuando se proporciona un objeto de referencia reconocible.
	RNF6.2	El sistema debe permitir corrección manual de las dimensiones estimadas.
Cumplimiento Legal	RNF7.1	El sistema debe cumplir con la Ley N.º 29733 - Ley de Protección de Datos Personales del Perú.
	RNF7.2	Debe implementarse un consentimiento informado al registrar usuarios, indicando qué datos personales se recopilan, para qué se utilizan, quién tendrá acceso y los derechos ARCO.
	RNF7.3	Las fotografías de paquetes y evidencias deben almacenarse de forma segura y solo ser accesibles por usuarios autorizados.
	RNF7.4	Los datos de ubicación de motorizados solo deben ser visibles durante entregas activas.
	RNF7.5	Debe existir una Política de Privacidad clara y accesible en la aplicación.
Mantenibilidad	RNF8.1	El código debe estar documentado para facilitar su comprensión y futuras modificaciones.
	RNF8.2	El sistema debe generar logs básicos de errores críticos.

Continúa en la siguiente página

Tabla 3.2 – continuación de la página anterior

Categoría	Código	Descripción
Eficiencia de recursos	RNF9.1	Las imágenes capturadas deben comprimirse automáticamente a un tamaño máximo de 500KB para minimizar el uso de datos móviles.
	RNF9.2	El procesamiento intensivo (análisis de IA) debe ejecutarse en el backend (Cloud Functions) para no sobrecargar dispositivos de recursos limitados.
	RNF9.3	La aplicación debe utilizar thumbnails (máximo 30KB) en listas y cargar imágenes completas solo cuando el usuario lo solicite explícitamente.

### 3.3.4 Matriz de trazabilidad: necesidades vs. requerimientos

La tabla en el anexo C establece la trazabilidad entre las necesidades identificadas en la fase de empatía y los requerimientos funcionales y no funcionales definidos, asegurando que cada necesidad del usuario tenga una respuesta técnica específica en el diseño.

### 3.3.5 Restricciones de diseño

Las restricciones identificadas en la fase de empatía condicionan las decisiones técnicas del diseño de la solución y establecen los límites operativos, normativos y éticos dentro de los cuales debe funcionar el sistema.

#### Restricciones económicas

Las microempresas de *delivery* requieren tecnologías de bajo costo que respalden su operación sin comprometer su sostenibilidad financiera. Dado su limitado presupuesto y la necesidad de costos operativos escalables, no pueden asumir gastos fijos elevados por infraestructura o licencias. Por ello, esta tesis emplea las capas gratuitas de servicios cloud (Firebase Free Tier) para desarrollar, prototipar y validar la solución sin incurrir en costos significativos durante la fase académica.

#### Restricciones técnicas

Los motorizados operan con conectividad móvil variable (3G/4G intermitente) en los distintos distritos de Lima, dependiendo del operador contratado y la cobertura disponible en cada zona. Esta condición externa no está bajo control del presente desarrollo,

ya que responde a la infraestructura de telecomunicaciones existente. Dado el uso predominante de dispositivos Android de gama media a baja, el sistema debe contemplar capacidades de procesamiento en el borde (edge), permitiendo la compresión y reducción de imágenes antes de su envío a la nube, con el fin de optimizar el uso de datos y garantizar la viabilidad operativa.

### **Restricciones operativas**

Los motorizados requieren flujos de interacción simples y ágiles durante las operaciones de recojo y entrega, realizadas en exteriores con condiciones variables. La aplicación debe funcionar eficientemente en dispositivos móviles estándar, sin depender de capacitación extensa, considerando la alta rotación de personal en el sector logístico. Esta restricción se aborda desde el diseño funcional del sistema, sin profundizar en aspectos de interfaz gráfica.

### **Restricciones normativas**

El sistema gestiona información sensible como teléfonos, direcciones, razón social, imágenes y evidencias de entrega, por lo que debe cumplir estrictamente con la Ley N.º 29733 de Protección de Datos Personales. Esto incluye consentimiento informado al registro, definición clara de uso y acceso a datos, y respeto a los derechos ARCO (Acceso, Rectificación, Cancelación, Oposición).

Las imágenes procesadas por IA se almacenan en Firebase Storage con reglas de acceso por rol, y los datos de ubicación de motorizados solo se visualizan durante entregas activas, sin persistencia innecesaria.

### **Restricciones éticas**

Las imágenes procesadas por el módulo de inteligencia artificial se utilizarán exclusivamente para estimar dimensiones de paquetes, garantizando un uso ético y limitado de esta información sensible. No se emplearán para entrenar modelos externos, compartir con terceros ni otros fines distintos a los declarados. El sistema debe informar claramente a los usuarios sobre el análisis automatizado de sus fotografías y aplicar políticas de minimización de datos para proteger la privacidad de personas que pudieran aparecer incidentalmente.

## **Restricciones de alcance y tiempo**

El proyecto se desarrollará en un plazo estimado de cinco meses, centrado en construir un prototipo funcional que permita validar la propuesta técnica. Se priorizan funcionalidades esenciales como gestión de usuarios y roles, seguimiento de pedidos, medición automática de dimensiones con IA, asignación a motorizados, registro de evidencias y trazabilidad básica en tiempo real.

No se incluyen optimizaciones avanzadas ni características empresariales como reportes analíticos, integraciones con ERP, facturación o analítica predictiva, las cuales se consideran fuera del alcance comprometido y posibles extensiones futuras. El enfoque se limita a una experiencia operativa suficiente para evaluar la viabilidad técnica del sistema.

### **3.3.6 Arquitectura conceptual del sistema**

Basándose en los requerimientos funcionales y no funcionales definidos, el sistema requiere una arquitectura distribuida compuesta por:

- **Capa de presentación:** App móvil y App web
- **Capa de lógica de negocio:** Procesamiento serverless, funciones serverless
- **Capa de datos:** Base de datos, almacenamiento de archivos, autenticación
- **Capa de servicios:** Geolocalización, notificaciones, procesamiento IA

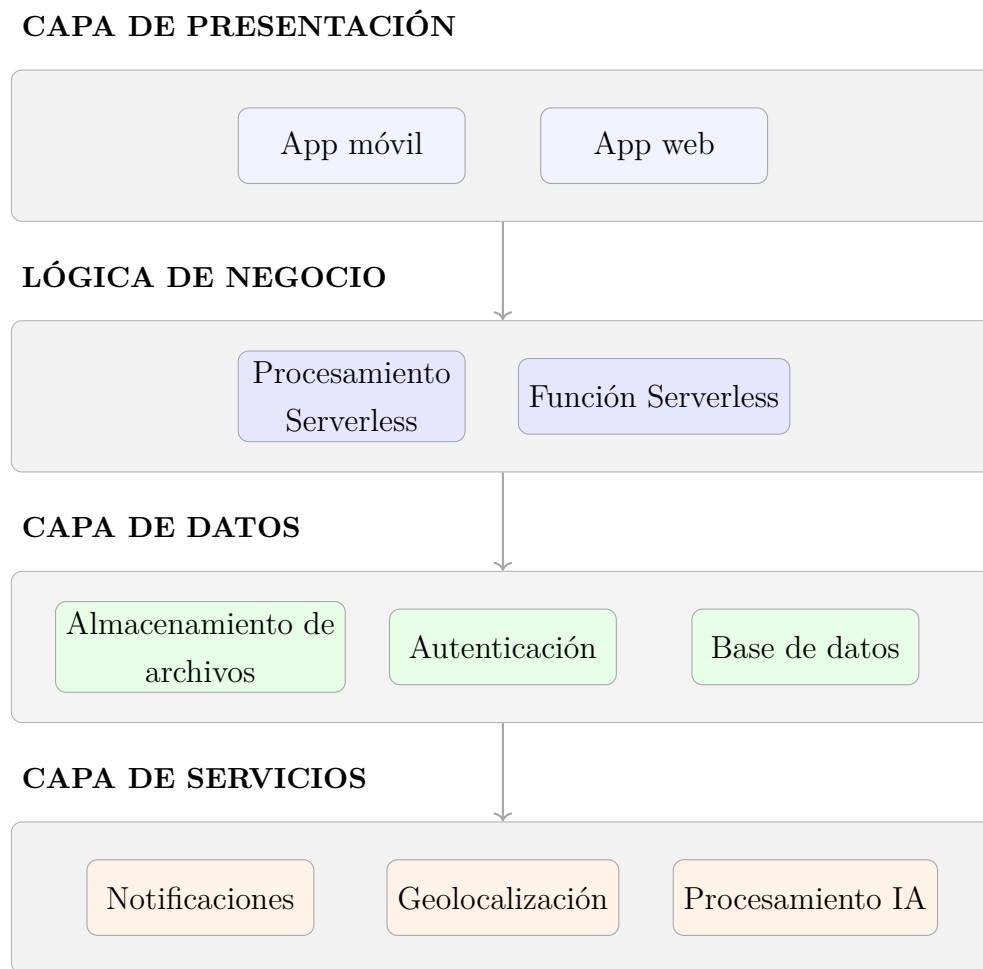


Figura 3.1: Arquitectura de capas del sistema

*Fuente:* Elaboración propia

Esta arquitectura conceptual será materializada en la Fase 3 mediante la evaluación de alternativas tecnológicas específicas.

### 3.4 Fase 3: Idear – Generación y selección de alternativas

#### 3.4.1 Introducción a la generación de alternativas

Con los requerimientos definidos, la fase de ideación evalúa alternativas tecnológicas que respondan a las necesidades identificadas. Se comparan sistemáticamente opciones de arquitectura, telecomunicaciones, plataformas cloud y enfoques de implementación mediante criterios técnicos, económicos y operativos. Este análisis multicriterio permite justificar la solución seleccionada con rigor ingenieril, evitando decisiones basadas en preferencias subjetivas o familiaridad tecnológica.

### 3.4.2 Criterios de evaluación

Para la evaluación de alternativas se establecen seis criterios principales que reflejan los requerimientos y restricciones identificados en la fase anterior:

Tabla 3.3: Criterios de evaluación de alternativas tecnológicas.

Criterio	Descripción	Peso	Justificación
Escalabilidad	Capacidad del sistema para crecer automáticamente con la demanda sin requerir reconfiguración manual significativa	20 %	Fundamental para pequeñas empresas que planean crecer progresivamente sin inversiones disruptivas
Costo	Inversión inicial, costos operativos mensuales y modelo de pago (fijo vs. variable)	25 %	Criterio crítico dado el presupuesto limitado de las empresas objetivo y el uso de capas gratuitas para el desarrollo de tesis
Facilidad de integración	Complejidad de integrar los diferentes componentes del sistema y tiempo de desarrollo requerido	15 %	Importante considerando el plazo de 5 meses para el desarrollo de la tesis
Mantenimiento	Esfuerzo requerido para mantener la infraestructura, actualizar componentes y gestionar la seguridad	15 %	Relevante para empresas pequeñas con recursos técnicos limitados
Rendimiento	Latencia, throughput y tiempo de respuesta del sistema en condiciones reales de uso	15 %	Necesario para garantizar experiencia de usuario aceptable y sincronización en tiempo real
Ecosistema y soporte	Disponibilidad de documentación, comunidad de desarrolladores, SDKs y herramientas de desarrollo	10 %	Facilita la resolución de problemas y reduce riesgos durante la implementación



### 3.4.3 Alternativas consideradas

Se identifican tres alternativas principales de arquitectura que cumplen con los requerimientos funcionales establecidos.

#### **Alternativa 1: Ecosistema Firebase/Google Cloud**

**Descripción técnica:** Arquitectura completamente basada en servicios de Google Cloud Platform, utilizando Firebase como Backend-as-a-Service (BaaS). Los componentes principales incluyen: Firebase Firestore (base de datos NoSQL en tiempo real), Firebase Authentication (gestión de usuarios y roles), Firebase Storage (almacenamiento de imágenes), Firebase Cloud Functions (lógica serverless), Firebase Cloud Messaging (notificaciones push), Firebase Hosting (dashboard web), Google Maps API (geo-localización), y integración con modelos de IA multimodal (Gemini, Claude, GPT-4 Vision) mediante Cloud Functions.

#### **Arquitectura de telecomunicaciones:**

- Sincronización en tiempo real mediante WebSockets gestionados por Firestore
- Protocolo HTTPS/TLS para todas las comunicaciones
- CDN global de Google para distribución de contenido estático
- Latencia típica de 50-200ms según ubicación geográfica
- Escalamiento automático horizontal sin configuración

#### **Ventajas:**

- Integración nativa entre todos los componentes del ecosistema
- Capa gratuita generosa para desarrollo y validación de tesis
- Sincronización en tiempo real nativa (offline-first)
- Escalabilidad automática sin gestión de servidores
- SDKs optimizados para Android y Web
- Reglas de seguridad declarativas integradas con autenticación
- Despliegue simplificado con Firebase CLI

#### **Desventajas:**

- Vendor lock-in con Google Cloud Platform
- Menor flexibilidad en configuraciones avanzadas
- Costos pueden incrementarse significativamente al superar la capa gratuita
- Modelo de datos NoSQL puede requerir desnormalización

#### **Estimación de costos:**

- Desarrollo (capa gratuita): \$0/mes
- Operación inicial (<1000 usuarios activos/mes): \$0-25/mes
- Operación media (1000-5000 usuarios): \$50-150/mes

#### **Alternativa 2: Amazon Web Services (AWS)**

**Descripción técnica:** Arquitectura basada en servicios de AWS utilizando: Amazon RDS o DynamoDB (base de datos), AWS Cognito (autenticación), Amazon S3 (almacenamiento), AWS Lambda (funciones serverless), Amazon SNS (notificaciones), Amazon CloudFront (CDN), Amazon API Gateway (exposición de APIs), y EC2 o Elastic Beanstalk para el dashboard web. Integración con servicios de IA mediante Amazon Rekognition o APIs externas.

#### **Arquitectura de telecomunicaciones:**

- API REST mediante API Gateway con throttling configurable
- Sincronización mediante polling o WebSockets (AWS AppSync)
- Protocolo HTTPS con certificados SSL de AWS Certificate Manager
- Latencia variable según configuración de regiones
- Escalamiento mediante Auto Scaling Groups

#### **Ventajas:**

- Ecosistema maduro con amplia documentación
- Flexibilidad máxima en configuraciones avanzadas
- Servicios especializados para diversos casos de uso
- Posibilidad de usar bases de datos relacionales (RDS) o NoSQL (DynamoDB)

- Control granular sobre infraestructura de red
- Amplio soporte empresarial

#### **Desventajas:**

- Mayor complejidad de configuración e integración
- Curva de aprendizaje pronunciada
- Capa gratuita limitada a 12 meses
- Requiere más tiempo de desarrollo por configuraciones manuales
- Sincronización en tiempo real no nativa (requiere configuración adicional)
- Costos más difíciles de predecir

#### **Estimación de costos:**

- Desarrollo (primeros 12 meses con capa gratuita): \$10-30/mes
- Operación inicial: \$50-100/mes
- Operación media: \$150-300/mes

### **Alternativa 3: Solución híbrida con backend propio**

**Descripción técnica:** Backend personalizado desarrollado con Node.js/Express o Python/Django desplegado en servicios de hosting como Heroku, DigitalOcean o Railway. Base de datos PostgreSQL o MongoDB gestionada. Autenticación con JWT. Almacenamiento en AWS S3 o Cloudinary. Integración directa con APIs de IA multimodal. Notificaciones mediante servicios de terceros (OneSignal, Firebase Cloud Messaging). Frontend web con React deployado en Vercel o Netlify.

#### **Arquitectura de telecomunicaciones:**

- API REST con autenticación mediante tokens JWT
- Sincronización mediante polling o WebSockets con Socket.io
- Protocolo HTTPS configurado manualmente
- Latencia dependiente del proveedor de hosting
- Escalamiento mediante configuración manual de instancias

**Ventajas:**

- Control total sobre la lógica de negocio
- No hay vendor lock-in (portabilidad entre proveedores)
- Posibilidad de usar tecnologías específicas según preferencia
- Costos predecibles con planes de hosting fijos
- Integración flexible con múltiples servicios de terceros

**Desventajas:**

- Mayor esfuerzo de desarrollo e integración
- Requiere gestionar múltiples servicios de diferentes proveedores
- Sincronización en tiempo real requiere implementación manual compleja
- Mayor responsabilidad en seguridad y mantenimiento
- Escalabilidad manual (no automática)
- Tiempo de desarrollo significativamente mayor (2-3x)
- Costos de mantenimiento y actualización recaen en el desarrollador

**Estimación de costos:**

- Desarrollo: \$15-40/mes (hosting + base de datos + almacenamiento)
- Operación inicial: \$40-80/mes
- Operación media: \$100-200/mes + tiempo de mantenimiento

**3.4.4 Matriz de evaluación multicriterio**

La siguiente tabla evalúa cada alternativa según los criterios establecidos, utilizando una escala de 1 a 5 donde 5 representa el mejor desempeño en ese criterio. La justificación detallada de cada puntuación asignada se encuentra en el Anexo G.

Tabla 3.4: Matriz de evaluación multicriterio de alternativas.

Criterio	Peso	Alt 1: Firebase		Alt 2: AWS		Alt 3: Propio	
		Punt.	Pond.	Punt.	Pond.	Punt.	Pond.
Escalabilidad	20 %	5	1.00	4	0.80	2	0.40
Costo	25 %	5	1.25	3	0.75	3	0.75
Facilidad integración	15 %	5	0.75	3	0.45	2	0.30
Mantenimiento	15 %	5	0.75	3	0.45	2	0.30
Rendimiento	15 %	4	0.60	4	0.60	3	0.45
Ecosistema y soporte	10 %	5	0.50	5	0.50	3	0.30
<b>TOTAL</b>	<b>100 %</b>		<b>4.85</b>		<b>3.55</b>		<b>2.50</b>

### 3.4.5 Selección de la solución óptima

Con base en la evaluación multicriterio, se selecciona la alternativa del ecosistema Firebase/Google Cloud como la arquitectura óptima para el desarrollo de este proyecto, obteniendo una puntuación ponderada de 4.85/5.00, significativamente superior a las alternativas evaluadas.

### Justificación técnica de la selección

Firebase ofrece una infraestructura distribuida en la nube que satisface los requerimientos críticos del sistema desde una perspectiva de ingeniería de telecomunicaciones. Su arquitectura serverless permite escalabilidad automática sin necesidad de provisión previa, con sincronización en tiempo real mediante WebSockets (latencias de 50-200ms), alta disponibilidad y replicación multi-región. La seguridad está integrada mediante OAuth 2.0, reglas de acceso por rol y cifrado TLS, cumpliendo con los requisitos RNF1.2, RNF2.1 y RNF3.x.

Además, los SDKs nativos para Android están optimizados para dispositivos de gama media-baja, con soporte offline-first, lo que permite una operación eficiente en condiciones de conectividad variable. Esta solución se alinea con las restricciones operativas identificadas: bajo costo (Spark Plan gratuito), bajo mantenimiento (infraestructura gestionada), cumplimiento normativo (Ley N.º 29733) y plazo limitado de desarrollo (5 meses), gracias a la integración nativa entre componentes.

Se reconoce el riesgo de vendor lock-in con Google Cloud Platform, pero se considera aceptable en el contexto de un prototipo académico. Este riesgo se mitiga mediante el uso de patrones de diseño como Repository Pattern, que facilitan la portabilidad

futura del código de negocio.

### 3.4.6 Selección de componentes específicos dentro del ecosistema Firebase

Una vez seleccionada la arquitectura general, se detallan los componentes específicos que se utilizarán:

Tabla 3.5: Componentes seleccionados del ecosistema Firebase.

Componente del sistema	Tecnología seleccionada		Justificación
Base de datos en tiempo real	Firebase	Firestore	NoSQL en tiempo real con sincronización automática, escalabilidad horizontal, modelo flexible para datos logísticos variables
Servicio de Autenticación	Firebase	Authentication	Gestión de sesiones segura, soporte de roles mediante custom claims, integración nativa con Firestore Security Rules
Servicio de almacenamiento	Firebase	Storage	Almacenamiento escalable de imágenes con CDN global, reglas de seguridad integradas, URLs firmadas para acceso controlado
Lógica de negocio serverless	Firebase	Cloud Functions	Ejecución de código bajo demanda, triggers automáticos ante eventos en Firestore/Storage, integración con APIs de IA externas
Aplicación móvil	Android (Kotlin)	Nativo	Rendimiento óptimo, acceso completo a APIs del dispositivo (cámara, GPS), SDKs de Firebase nativos, amplia documentación
App Web	React.js + Firebase Hosting		Framework moderno con alto rendimiento, despliegue global automático con CDN, certificado SSL incluido

*Continúa en la siguiente página*

Tabla 3.5 – *Continuación*

Componente del sistema	Tecnología seleccionada		Justificación
Notificaciones push	Firestore Messaging	Cloud	Integración nativa con Android, envío masivo eficiente, segmentación por roles, <i>delivery</i> garantizado
Geolocalización	Google Maps API		Precisión superior en ubicaciones de Lima, integración directa con ecosistema Google, APIs de rutas y geocodificación
Procesamiento de IA	GPT-4 / Gemini / Claude		Modelos multimodales capaces de analizar imágenes y estimar dimensiones mediante razonamiento visual contextual

### 3.4.7 Conclusión de la fase de ideación

La selección del ecosistema Firebase/Google Cloud se fundamenta en un análisis multicriterio riguroso que pondera aspectos técnicos, económicos y operativos relevantes para el contexto del proyecto. Esta decisión equilibra las necesidades de pequeñas empresas de *delivery* con las restricciones de un proyecto de tesis, proporcionando una base tecnológica robusta, escalable y económicamente viable para el desarrollo del prototipo y su eventual adopción comercial.

## 3.5 Fase 4: Prototipar – Diseño técnico detallado

### 3.5.1 Introducción al diseño de la solución

La fase de prototipado materializa la alternativa seleccionada en un diseño técnico completo que especifica la arquitectura del sistema, los componentes de telecomunicaciones, el modelo de datos, las interfaces de usuario y los protocolos de comunicación. Este diseño debe ser lo suficientemente detallado para permitir su implementación y validación posterior, asegurando que todos los requerimientos funcionales y no funcionales establecidos en la fase de definición sean satisfechos mediante decisiones técnicas justificadas.

El prototipo conceptual incluye diagramas de arquitectura, especificaciones de red,

diseño de base de datos, flujos de información y mockups de interfaces, proporcionando una visión integral de la solución propuesta desde la perspectiva de ingeniería de telecomunicaciones.

### **3.5.2 Arquitectura general del sistema**

La arquitectura del sistema se basa en un modelo cliente-servidor distribuido en la nube con tres capas principales que interactúan mediante protocolos estándar de telecomunicaciones.

#### **Diagrama de arquitectura de alto nivel**

##### **Capa de presentación (Frontend):**

- Aplicación móvil Android nativa para clientes y motorizados
- Dashboard web React.js para administradores
- Interfaces responsive con Material Design

##### **Capa de lógica de negocio (Backend):**

- Firebase Cloud Functions (serverless)
- Procesamiento de imágenes con IA
- Reglas de negocio y validaciones
- Integración con servicios externos (Google Maps, APIs de IA)

##### **Capa de datos (Data Layer):**

- Firebase Firestore (base de datos NoSQL)
- Firebase Storage (almacenamiento de imágenes)
- Firebase Authentication (gestión de usuarios)



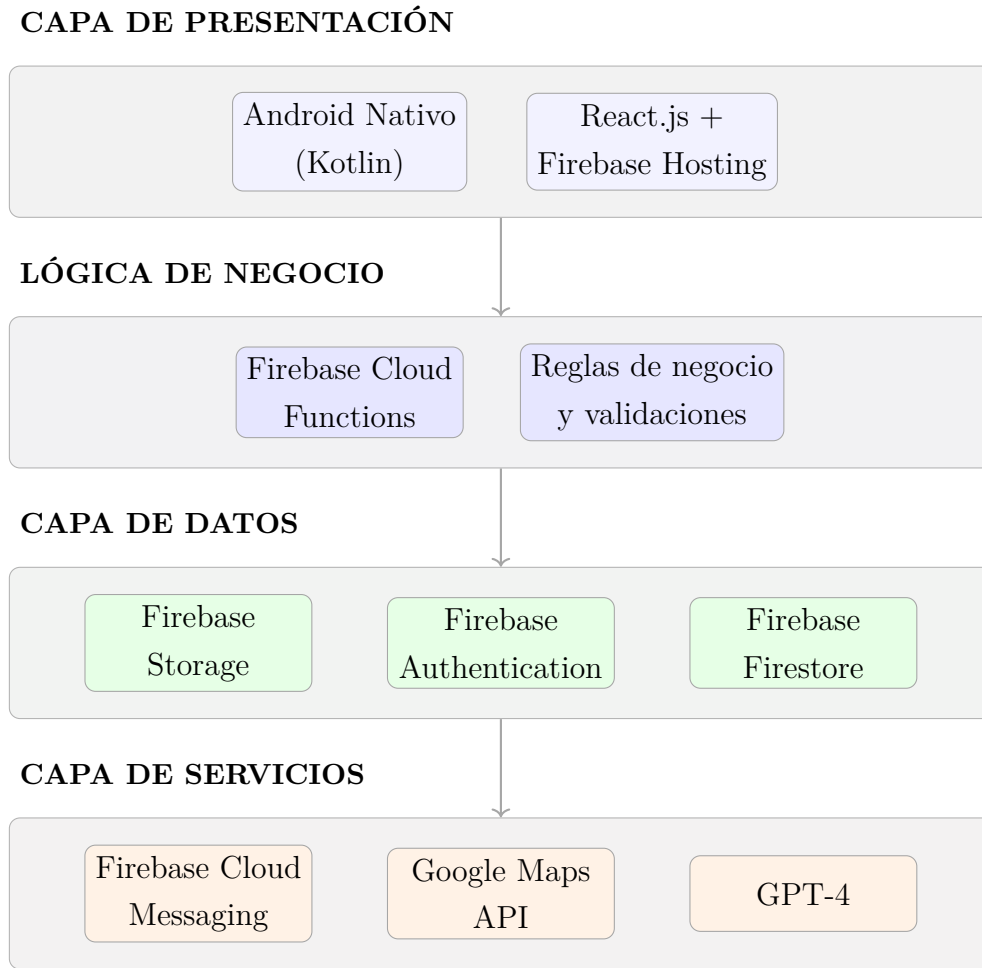


Figura 3.2: Arquitectura de capas del sistema con tecnologías seleccionadas

*Fuente:* Elaboración propia

### Flujo de datos y comunicación

El sistema opera mediante comunicación asíncrona basada en eventos:

1. **Autenticación:** Firebase Authentication maneja login/logout con tokens JWT
2. **Sincronización en tiempo real:** Firestore utiliza WebSockets para actualizaciones bidireccionales
3. **Almacenamiento de archivos:** Firebase Storage recibe imágenes con URLs firmadas
4. **Procesamiento serverless:** Cloud Functions se activan mediante triggers automáticos
5. **Notificaciones:** Firebase Cloud Messaging envía push notifications a dispositivos registrados

## Protocolos de comunicación utilizados:

Tabla 3.6: Protocolos de comunicación del sistema.

Protocolo	Uso	Justificación técnica
HTTPS (TLS 1.3)	Todas las comunicaciones cliente-servidor	Cifrado extremo a extremo, autenticación de servidor
WebSockets sobre TLS	Sincronización en tiempo real de Firestore	Comunicación bidireccional de baja latencia
OAuth 2.0	Autenticación y autorización	Estándar de la industria, secure token management
REST API	Integración con servicios externos	Interoperabilidad con Google Maps y APIs de IA

### 3.5.3 Flujo crítico del sistema: Creación de pedido con procesamiento de IA

El siguiente diagrama describe el flujo completo desde que un cliente captura la foto de un paquete hasta que recibe las dimensiones estimadas:

Para incluirlo como figura formal en tu tesis, aquí está el código con el entorno figure, caption y label para referencias cruzadas:

1. Cliente captura foto del paquete con objeto de referencia
2. App móvil comprime imagen localmente (máx. 500KB)
3. Imagen se sube a Firebase Storage
4. Cloud Function se activa automáticamente por trigger
5. Function descarga imagen y envía a API de IA multimodal
6. Modelo IA procesa imagen y estima dimensiones
7. Dimensiones se guardan en Firestore
8. Cliente recibe actualización en tiempo real vía WebSocket
9. Sistema permite edición manual si confianza es baja

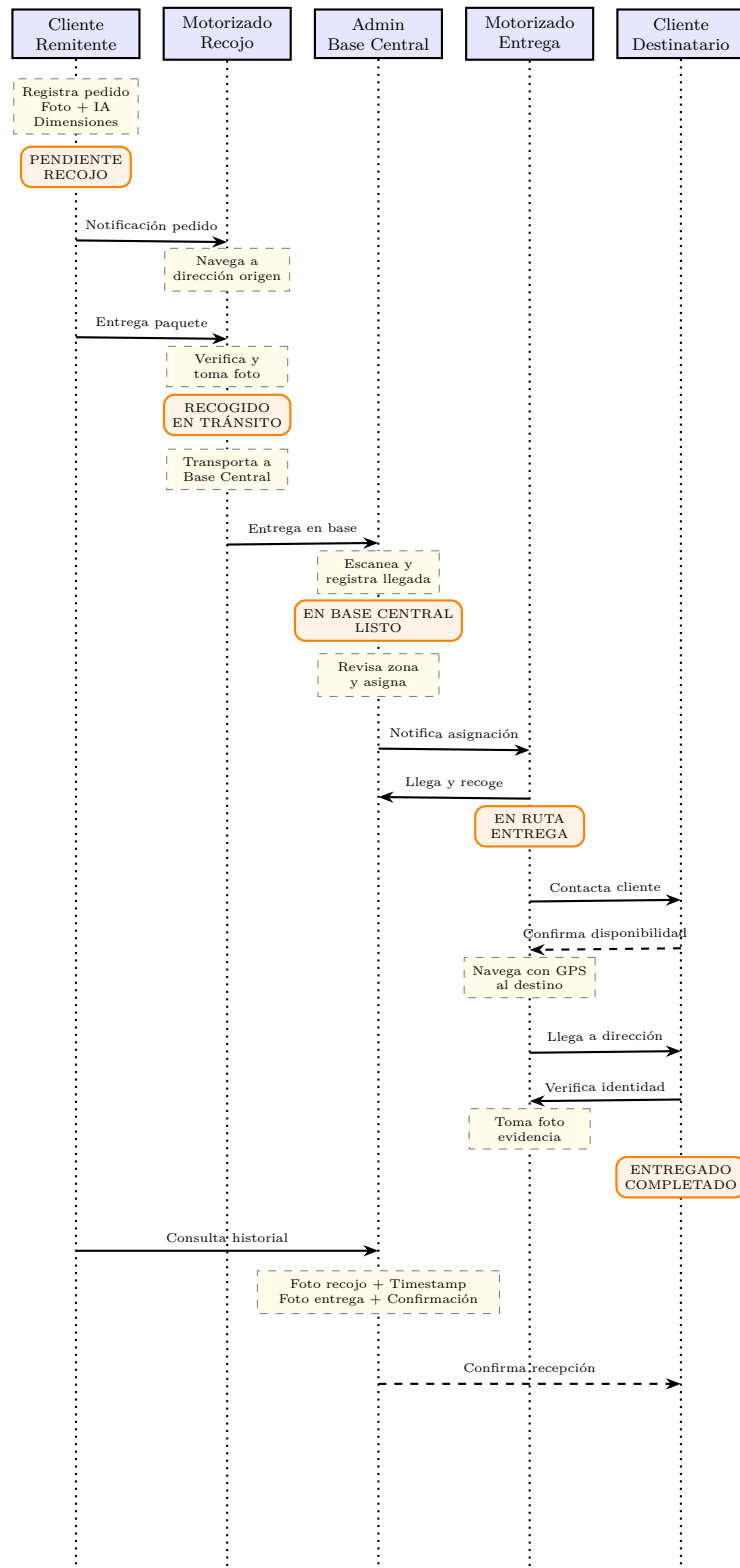


Figura 3.3: Diagrama de flujo de entrega del sistema de delivery propuesto

### 3.5.4 Módulo de IA para estimación de dimensiones

#### Enfoque de prompt engineering

El módulo de IA utiliza modelos de lenguaje multimodal (Gemini 2.0 Flash, Claude Sonnet 4) mediante técnicas de prompt engineering estructurado para estimar dimensiones de paquetes a partir de fotografías.

#### Estructura del prompt:

- **Rol del sistema:** “Eres un asistente especializado en logística que estima dimensiones de paquetes...”
- **Contexto:** “El motorizado está en el domicilio del remitente fotografiando un paquete para *delivery* en Lima, Perú...”
- **Tarea específica:** “Analiza la imagen y estima las dimensiones (alto, ancho, largo) en centímetros...”
- **Formato de salida:** JSON estructurado

```
{  
  "alto_cm": número,  
  "ancho_cm": número,  
  "largo_cm": número,  
  "confianza": "alta|media|baja",  
  "observaciones": "texto"  
}
```

- **Restricciones:** “Si no puedes estimar con confianza razonable, indica confianza ‘baja’ y explica el motivo...”

#### Validación y edición manual

El sistema permite corrección manual de las dimensiones estimadas:

- **Confianza alta (>80 %):** Dimensiones se aceptan automáticamente pero son editables
- **Confianza media (50-80 %):** Sistema sugiere verificación manual
- **Confianza baja (<50 %):** Requiere medición manual obligatoria

### 3.5.5 Interfaces funcionales del sistema

El sistema requiere tres interfaces funcionales que permitan demostrar el flujo de comunicación y procesamiento de datos diseñado. El enfoque está en la funcionalidad técnica, no en el diseño visual, ya que el objetivo es validar la arquitectura de telecomunicaciones propuesta.

#### Aplicación móvil (Cliente y Motorizado)

##### Funcionalidades técnicas implementadas:

- Autenticación mediante Firebase Authentication (OAuth 2.0)
- Captura de imagen y compresión local en el dispositivo
- Upload de archivos a Firebase Storage mediante HTTPS
- Recepción de datos en tiempo real vía WebSocket (Firestore)
- Envío y recepción de geolocalización (Google Maps API)
- Recepción de notificaciones push (Firebase Cloud Messaging)

**Propósito:** Demostrar la comunicación entre dispositivo móvil y servicios cloud, validando latencias, consumo de datos y sincronización en tiempo real.

#### Dashboard web (Administrador)

##### Funcionalidades técnicas implementadas:

- Visualización de datos en tiempo real desde Firestore (WebSocket)
- Renderización de mapa con ubicaciones geográficas (Google Maps API)
- Consulta y filtrado de información desde base de datos NoSQL
- Visualización de imágenes almacenadas en Firebase Storage
- Gestión de usuarios mediante Firebase Authentication

**Propósito:** Demostrar sincronización bidireccional, actualización en tiempo real de múltiples clientes conectados simultáneamente y consultas eficientes a base de datos distribuida.

### 3.5.6 Especificaciones de seguridad y cumplimiento normativo

El diseño incorpora mecanismos de seguridad que cumplen con la Ley N.º 29733 de Protección de Datos Personales de Perú.

#### Seguridad en capa de transporte

Tabla 3.7: Especificaciones de seguridad en transporte.

Mecanismo	Especificación	Justificación
Cifrado	TLS 1.3	Versión más reciente, elimina vulnerabilidades de TLS 1.2
Certificados	Automáticos (Let's Encrypt)	Renovación automática sin intervención manual
Perfect Forward Secrecy	Habilitado	Protege comunicaciones pasadas si clave se compromete

#### Especificaciones técnicas consolidadas

- **OAuth 2.0:** Autenticación delegada con tokens JWT
- **Roles y permisos:** Administrador, Cliente, Motorizado
- **Security Rules de Firestore:** Control granular de acceso a nivel de documento
- **Sesiones:** Expiración automática tras 24 horas de inactividad

#### Protección de datos personales

Según Ley N.º 29733:

Tabla 3.8: Cumplimiento de Ley N.º 29733.

Requisito legal	Implementación técnica
Consentimiento informado	Aceptación explícita en registro
Derechos ARCO (Acceso, Rectificación, Cancelación, Oposición)	Opciones en configuración de cuenta
Seguridad técnica	Cifrado TLS + reglas de acceso
Plazo de conservación	Eliminación automática tras 90 días de inactividad

### 3.5.7 Conclusión de la fase de prototipado

El diseño técnico propuesto combina robustez tecnológica, viabilidad económica y eficiencia operativa mediante una arquitectura serverless basada en Firebase, que ofrece escalabilidad automática, costos proporcionales al uso, sincronización en tiempo real, seguridad normativa y desarrollo ágil. Los componentes detallados se documentan en el Anexo D, permitiendo que este capítulo se centre en la justificación de alto nivel.

El diseño técnico propuesto establece una base sólida que será evaluada en el Capítulo 4 mediante análisis de viabilidad técnica, económica y de sostenibilidad, validando su capacidad para cumplir los requerimientos establecidos en el contexto de las pymes de *delivery* en Lima.

## Conclusiones

[Por completar]



## REFERENCIAS BIBLIOGRAFICAS

- [1] «IoT-based TPL whole supply chain logistics information system model,» *Proceedings - International Conference on Machine Learning and Cybernetics*, vol. 4, págs. 1758-1762, jul. de 2013, ISSN: 21601348. DOI: 10.1109/ICMLC.2013.6890882. dirección: <https://ieeexplore.ieee.org/document/6890882>.
- [2] *What Is IoT Architecture? / MongoDB*. dirección: <https://www.mongodb.com/resources/basics/cloud-explained/iot-architecture>.
- [3] «Applications of Computer Internet of Things Technologies in Logistics,» *2nd IEEE International Conference on Advanced Technologies in Intelligent Control, Environment, Computing and Communication Engineering, ICATIECE 2022*, 2022. DOI: 10.1109/ICATIECE56365.2022.10047270. dirección: <https://ieeexplore.ieee.org/document/10047270>.
- [4] R. TLW, *Desafíos y soluciones en la distribución de última milla: ¿Cómo adaptarse al futuro?* Feb. de 2025. dirección: <https://thelogisticsworld.com/logistica-y-distribucion/desafios-y-soluciones-en-la-distribucion-de-ultima-milla-en-entornos-urbanos/>.
- [5] *Perú proyecta un crecimiento del 11,03 % en delivery e ingresos por US\$ 2 951 millones hacia 2029*. dirección: <https://lacamara.pe/peru-proyecta-un-crecimiento-del-1103-en-delivery-e-ingresos-por-us-2-951-millones-hacia-2029/>.
- [6] A. García-Arranz, «Cristina Gallego Gómez Asistentes de investigación: Adriana Lizalde-Litza Jiménez-Ana Paula Dorantes Internet de las Cosas Comparativa Internacional, España y Comunidades Autónomas,»
- [7] R. TLW, *Supervisión de seguridad en tiempo real en la cadena de suministro: Innovación con IoT*, nov. de 2024. dirección: <https://thelogisticsworld.com/tecnologia/el-uso-de-iot-para-la-supervision-de-seguridad-en-tiempo-real-en-la-cadena-de-suministro/>.
- [8] C. Juárez, *Diferencias en el uso de la IoT y la IIoT en la logística*, ago. de 2023. dirección: <https://thelogisticsworld.com/tecnologia/iot-e-iiot-en-la-logistica-como-sacarle-provecho/>.
- [9] J. Krysińska, *On-premise vs. cloud: definition & differences / NordLayer*, ago. de 2024. dirección: <https://nordlayer.com/blog/on-premise-vs-cloud-differences/>.

- [10] «LOS COSTOS LOGÍSTICOS DE LAS EMPRESAS EN EL PAÍS SON DEL 16 % EN PROMEDIO, PERO UN 21.1 % PARA LAS MICROEMPRESAS,» dirección: <https://www.comexperu.org.pe/articulo/los-costos-logisticos-de-las-empresas-en-el-pais-son-del-16-en-promedio-pero-un-211-para-las-microempresas>.
- [11] *La carrera de la última milla: entregas exprés en el Día de las Madres.* dirección: <https://thelogisticsworld.com/logistica-comercio-electronico/carrera-ultima-milla-entregas-expres-dia-madres-ecommerce-logistica/>.
- [12] E. Perú, «REGLAMENTO DE LA LEY N° 29733 LEY DE PROTECCIÓN DE DATOS PERSONALES,» 2973. dirección: <https://cdn.www.gob.pe/uploads/document/file/272360/Ley%20N%C2%BA%2029733.pdf.pdf?v=1618338779>.
- [13] «ESTRATEGIA NACIONAL DE INTELIGENCIA ARTIFICIAL Documento de Trabajo para la Participación de la Ciudadanía 2021-2026 Secretaría de Gobierno y Transformación Digital PRESIDENCIA DEL CONSEJO DE MINISTROS,»
- [14] M. Cipriano, «La regulación para la consolidación de las redes y los servicios que soporten el desarrollo de las TICs y los macro datos,» dirección: <https://repositorio.osiptel.gob.pe/bitstream/handle/20.500.12630/103/presentacion-reg-redes-desarrollo.pdf?sequence=1&isAllowed=y>.
- [15] P. Sun y Z. Zhang, «Design and Implementation of an Agricultural Logistics Control System Based on IoT Techology,» *2024 International Conference on Advanced Control Systems and Automation Technologies (ACSAT)*, págs. 40-43, nov. de 2024. DOI: 10.1109/ACSAT63853.2024.10823707. dirección: <https://ieeexplore.ieee.org/document/10823707/>.
- [16] «Research on Cloud Logistics-based One-stop Service Platform for Logistics Center,» *Proceedings of the 2012 IEEE 16th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2012*, págs. 558-563, 2012. DOI: 10.1109/CSCWD.2012.6221873.
- [17] *La entrega a tiempo, el gran reto de la última milla.* dirección: <https://thelogisticsworld.com/transporte/ultima-milla-el-desafio-de-la-entrega-rapida/>.
- [18] «Green and Network-Aware Smart IoT Logistics Applications,» *2023 International Conference on Information Technology: Cybersecurity Challenges for Sustainable Cities, ICIT 2023 - Proceeding*, págs. 441-446, 2023. DOI: 10.1109/ICIT58056.2023.10225953.
- [19] *Entregas limpias en la última milla: innovaciones ecológicas que transforman la logística B2B.* dirección: <https://thelogisticsworld.com/logistica-y-distribucion/soluciones-ecologicas-para-la-entrega-de-ultima-milla-en-ciudades-mexicanas/>.

- [20] *Visión artificial en logística: Retos, beneficios y futuro en México.* dirección: <https://thelogisticsworld.com/logistica-y-distribucion/como-la-industria-logistica-mexicana-lidera-con-vision-artificial-e-iot/>.
- [21] O. September, «GPT-4V(ision) System Card,»
- [22] «MM-Vet: Evaluating Large Multimodal Models for Integrated Capabilities,» 2024. dirección: <https://huggingface..>
- [23] «GPT-4 Technical Report,»
- [24] «The Claude 3 Model Family: Opus, Sonnet, Haiku Anthropic,» dirección: <https://docs.anthropic.com/>.
- [25] Anthropic, *Home - Anthropic.* dirección: <https://docs.anthropic.com/en/home>.
- [26] G. Team, «Gemini 2.5: Pushing the Frontier with Advanced Reasoning, Multimodality, Long Context, and Next Generation Agentic Capabilities.,» 2025.
- [27] «Gemini: A Family of Highly Capable Multimodal Models,» 2025.
- [28] «AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE,» dirección: <https://github.com/>.
- [29] Anthropic, «System Card: Claude Opus 4 & Claude Sonnet 4,» 2025.
- [30] «DINOv2: Learning Robust Visual Features without Supervision,» dirección: <https://openreview.net/forum?id=a68SUt6zFt>.
- [31] «Florence: A New Foundation Model for Computer Vision,»
- [32] «Scaling Vision Transformers to Gigapixel Images via Hierarchical Self-Supervised Learning,» dirección: <https://github.com/mahmoodlab/HIPT..>
- [33] «DINOv2: Learning Robust Visual Features without Supervision,» dirección: <https://openreview.net/forum?id=a68SUt6zFt>.

## Anexo A. Caracterización detallada de actores del sistema

### Actor: Clientes (Emisores)

<b>Rol en el sistema</b>	Usuarios que solicitan el servicio de envío de paquetes
<b>Necesidades principales</b>	Necesitan enviar paquetes de manera rápida y confiable sin invertir tiempo en mediciones manuales. Requieren conocer el estado de sus pedidos en tiempo real y tener transparencia en los costos del servicio.
<b>Puntos de dolor actuales</b>	Actualmente deben medir manualmente los paquetes o proporcionar estimaciones que pueden no corresponder con la realidad. Experimentan falta de trazabilidad durante la entrega y sufren reprogramaciones cuando los paquetes están mal estimados, generando incertidumbre sobre las condiciones del servicio.

### Actor: Administradores

<b>Rol en el sistema</b>	Gestores de la operación logística de la empresa
<b>Necesidades principales</b>	Buscan maximizar la eficiencia operativa atendiendo el mayor número de pedidos posible. Necesitan asignar pedidos adecuadamente a motorizados, supervisar entregas en tiempo real y poder rechazar pedidos que excedan la capacidad operativa del servicio.
<b>Puntos de dolor actuales</b>	Enfrentan una gestión manual de pedidos mediante WhatsApp y hojas de Excel que no escala. Carecen de visibilidad en tiempo real de la operación y tienen dificultad para validar las dimensiones reportadas por los clientes, lo que genera pérdida de eficiencia por aceptar paquetes sobredimensionados.

### Actor: Motorizados (Repartidores)

<b>Rol en el sistema</b>	Operadores de campo que recogen y entregan paquetes
<b>Necesidades principales</b>	Requieren recibir información clara sobre los pedidos asignados para optimizar el espacio en su mochila o cajas de reparto. Necesitan conocer rutas y ubicaciones con precisión, registrar evidencias de recojo y entrega fácilmente, y evitar viajes innecesarios por paquetes que no pueden transportar.
<b>Puntos de dolor actuales</b>	Experimentan sorpresas frecuentes con paquetes más grandes de lo reportado, lo que genera pérdida de tiempo y combustible en recojos no viables. Carecen de herramientas adecuadas para comunicar el estado del pedido en tiempo real y enfrentan dificultad para gestionar múltiples entregas simultáneas de manera eficiente.

### Actor: Destinatarios (Receptores)

<b>Rol en el sistema</b>	Usuarios finales que reciben los paquetes
<b>Necesidades principales</b>	Esperan recibir sus paquetes en el tiempo estimado y conocer cuándo llegará el motorizado. Necesitan tener confianza en el servicio y sentirse informados durante el proceso de entrega.
<b>Puntos de dolor actuales</b>	Sufren reprogramaciones frecuentes debido a problemas logísticos en la cadena de entrega. Experimentan falta de información sobre el estado real de su pedido y quedan insatisfechos cuando el servicio se retrasa sin explicación clara.

## Anexo B. Síntesis de hallazgos de la fase de empatía

### B.1 Necesidades identificadas

- **Automatización de procesos manuales:** La medición de dimensiones de paquetes consume tiempo y genera estimaciones inexactas, mientras que la gestión de pedidos mediante WhatsApp y hojas de cálculo resulta ineficiente para escalar operaciones.
- **Visibilidad y trazabilidad en tiempo real:** Clientes requieren conocer el estado de sus envíos, administradores necesitan supervisar la operación completa, y motorizados demandan información clara sobre asignaciones y rutas.
- **Optimización de recursos limitados:** Las empresas necesitan soluciones de bajo costo con modelos de pago progresivos, y los motorizados requieren maximizar el aprovechamiento del espacio de transporte para cumplir múltiples entregas eficientemente.

### B.2 Restricciones del contexto

- **Restricción económica:** Presupuestos limitados para inversión tecnológica inicial y necesidad de evitar costos fijos elevados que comprometan la viabilidad del negocio.
- **Restricción técnica:** El uso predominante de dispositivos Android de gama media a baja condiciona las decisiones de diseño hacia arquitecturas que minimicen el procesamiento local. Se implementan estrategias básicas de eficiencia (compresión de imágenes, procesamiento en cloud, uso de SDKs nativos).
- **Restricción operativa:** Conectividad variable en diferentes zonas de Lima que requiere estrategias de sincronización y manejo de estados offline para garantizar continuidad del servicio.
- **Restricción normativa:** Cumplimiento obligatorio de la Ley N.º 29733 de Protección de Datos Personales en el manejo de ubicaciones, fotografías de evidencia y datos personales de usuarios.

### B.3 Oportunidades de mejora mediante tecnología

- **Inteligencia artificial para visión por computadora:** Automatización de estimación de dimensiones mediante fotografías con objetos de referencia, eliminando tiempo de medición manual y reduciendo errores humanos.
- **Infraestructura cloud escalable (Firebase):** Sincronización en tiempo real entre dispositivos móviles y dashboard web sin requerir servidores físicos costosos, proporcionando trazabilidad completa de operaciones.
- **Notificaciones push (Firebase Cloud Messaging):** Comunicación instantánea de cambios de estado a todos los actores, mejorando coordinación operativa y experiencia de usuario.
- **Geolocalización (Google Maps API):** Optimización de rutas de entrega y seguimiento en tiempo real de motorizados, aumentando eficiencia del servicio y reduciendo tiempos de entrega.
- **Modelo de costos por uso:** Servicios cloud con pago progresional que se alinea con el crecimiento de empresas pequeñas, eliminando barreras de entrada por inversiones iniciales prohibitivas.

## Anexo C. Matriz de trazabilidad completa

Tabla C.1: Matriz de trazabilidad: Necesidades vs. Requerimientos.

<b>Necesidad identificada</b>	<b>Req. funcionales</b>	<b>Req. no funcionales</b>	<b>Justificación técnica</b>
Automatización de medición de dimensiones	RF2.3, RF2.4	RNF1.1, RNF6.2, RNF7.1, RNF7.2	El módulo de IA con Cloud Functions procesa imágenes automáticamente, reduciendo tiempo de medición manual y errores humanos, con compresión optimizada para minimizar consumo de datos.
Trazabilidad en tiempo real de pedidos	RF2.1, RF2.5, RF3.4, RF4.1, RF5.1, RF5.2	RNF1.2, RNF2.1	La sincronización en tiempo real de Firestore combinada con notificaciones push proporciona visibilidad completa del estado de pedidos a todos los actores.
Gestión centralizada para administradores	RF4.1, RF4.2, RF4.3	RNF1.2, RNF4.2, RNF5.1	El dashboard web con React permite supervisión completa, filtrado eficiente y visualización geográfica de operaciones desde cualquier navegador.
Herramientas operativas para motorizados	RF3.1, RF3.2, RF3.3, RF3.5	RNF4.1, RNF5.1, RNF6.1	La aplicación Android nativa optimizada proporciona acceso eficiente a funciones del dispositivo (cámara, GPS) con bajo consumo de batería.
Solución de bajo costo escalable	RF1.1, RF6.1, RF6.2	RNF2.1, RNF3.1, RNF9.1, RNF9.2	La arquitectura serverless de Firebase elimina costos de servidores físicos y escala automáticamente con modelo de pago por uso.



*Continuación de la tabla*

<b>Necesidad identificada</b>	<b>Req. funcionales</b>	<b>Req. no funcionales</b>	<b>Justificación técnica</b>
Seguridad de información sensible	RF6.1, RF6.2	RNF3.1, RNF3.2, RNF3.3, RNF7.1-RNF7.5	Firestore Auth, reglas de seguridad de Firestore, cifrado SSL/TLS y cumplimiento de Ley N.º 29733 protegen datos personales y operativos.
Optimización de espacio de transporte	RF2.3, RF2.4, RF3.5	RNF6.1	Las dimensiones precisas permiten a motorizados planificar mejor la capacidad de carga y a administradores asignar pedidos compatibles.
Compatibilidad con dispositivos limitados	RF3.1, RF3.2, RF3.3	RNF4.1, RNF9.1, RNF9.2	El diseño considera dispositivos Android de gama media-baja, optimizando procesamiento, memoria y consumo de batería.

## Anexo D. Modelo de datos detallado

### D.1 Colección: USUARIOS

Ruta: usuarios/{uid}

```
1 {
2   uid: string,           // ID unico de Firebase Auth
3   email: string,        // Correo electronico
4   phone: string,        // Telefono de contacto
5   nombre: string,       // Nombre del usuario
6   apellido: string,     // Apellido del usuario
7   rol: string,          // 'administrador' | 'cliente' | '
   motorizado'
8
9   // Datos especificos para clientes empresariales (opcional)
10  datosEmpresariales: {
11    nombreEmpresa: string,
12    ruc: string,
13    razonSocial: string
14  } | null,
15
16  // Datos especificos para motorizados (opcional)
17  datosMotorizado: {
18    estadoServicio: string, // 'disponible' | 'ocupado' | '
   desconectado'
19    vehiculo: {
20      placa: string,
21      tipo: string,        // 'motocicleta' | 'bicicleta' | '
   auto'
22      modelo: string,
23      anio: number
24    },
25    estadisticas: {
26      pedidosCompletados: number,
27      calificacionPromedio: number,
28      tiempoPromedioEntrega: number // en minutos
29    }
30  } | null,
```

```

31
32 // Ultima ubicacion conocida (para motorizados)
33 ultimaUbicacionConocida: {
34     latitud: number,
35     longitud: number,
36     timestamp: timestamp,
37     distrito: string
38 } | null,
39
40 // Metadatos
41 creadoEn: timestamp,
42 actualizadoEn: timestamp,
43 activo: boolean,
44 ultimaConexion: timestamp
45 }

```

Listing D.1: Estructura de documento Usuario

## D.2 Colección: PEDIDOS

Ruta: pedidos/{pedidoId}

```

1 {
2   id: string, // ID unico del pedido
3
4   // Estructura de asignacion (recojo y entrega pueden ser
5   // diferentes)
6   asignacion: {
7     recojo: {
8       estado: string, // 'pendiente' | 'asignado' | '
9       completado'
10      rutaId: string | null,
11      rutaNombre: string | null,
12      motorizadoUid: string | null,
13      motorizadoNombre: string | null,
14      asignadaEn: timestamp | null,
15      razonPendiente: string | null
16    },
17    entrega: {
18      // Estructura similar a recojo
19    }
20  },
21 }

```

```

19
20 // Control de visibilidad por rol
21 visibilidad: {
22     motorizadoRecojo: boolean,
23     motorizadoEntrega: boolean,
24     administrador: boolean
25 },
26
27 // Ciclo operativo del pedido
28 cicloOperativo: {
29     diaCreacion: string,      // YYYY-MM-DD
30     diaAsignado: string,
31     cerradoPorAdmin: boolean,
32     fechaCierreAdmin: timestamp | null,
33     horaLimiteRecojo: timestamp,
34     permiteEntregaAntes13h: boolean
35 },
36
37 // Indices para consultas eficientes
38 indices: {
39     requiereAsignacionManual: boolean,
40     motorizadoRecojoUid: string | null,
41     motorizadoEntregaUid: string | null,
42     rutaRecojoId: string | null,
43     rutaEntregaId: string | null,
44     distritoRecojo: string,
45     distritoEntrega: string
46 }
47 }

```

Listing D.2: Estructura de documento Pedido (parte 1)

```

1 {
2     // Informacion del proveedor (quien envia)
3     proveedor: {
4         uid: string,
5         nombre: string,
6         correo: string,
7         telefono: string,
8         direccion: {
9             link: string,      // URL de Google Maps
10            distrito: string,

```

```

11     coordenadas: {
12         latitud: number,
13         longitud: number
14     }
15 }
16 },
17
18 // Informacion del destinatario (quien recibe)
19 destinatario: {
20     nombre: string,
21     telefono: string,
22     direccion: {
23         link: string,
24         distrito: string,
25         coordenadas: {
26             latitud: number,
27             longitud: number
28         }
29     }
30 },
31
32 // Informacion del paquete
33 paquete: {
34     detalle: string,           // Descripcion del contenido
35     observaciones: string,
36     dimensiones: {
37         alto: number,         // en centimetros
38         ancho: number,
39         largo: number,
40         volumen: number,      // calculado automaticamente
41         peso: number | null,  // opcional
42         estimadoPorIA: boolean,
43         editadoManualmente: boolean,
44         excedeMaximo: boolean,
45         confianzaIA: number | null // 0-1, nivel de confianza
46     },
47     fotos: {
48         recojo: {
49             url: string,
50             thumbnail: string,

```

```

51     timestamp: timestamp,
52     procesadaPorIA: boolean
53 },
54 entrega: {
55     url: string,
56     thumbnail: string,
57     timestamp: timestamp
58 },
59 comprobantePago: {
60     url: string,
61     thumbnail: string,
62     timestamp: timestamp
63 }
64 }
65 }
66 }

```

Listing D.3: Estructura de documento Pedido (parte 2)

```

1 {
2     // Informacion de pago
3     pago: {
4         seCobra: boolean,
5         metodoPago: string,      // 'efectivo' | 'transferencia' | '
6         yape' | 'plin'
7         monto: number,
8         comision: number,
9         montoTotal: number,
10        estadoPago: string,      // 'pendiente' | 'pagado' | '
11        por_cobrar'
12        billeteraUsada: {
13            id: string,
14            propietarioTipo: string, // 'proveedor' | 'empresa'
15            metodo: string,
16            titular: string
17        }
18    },
19
20    // Informacion del motorizado asignado (desnormalizado)
21    motorizado: {
22        uid: string,
23        nombre: string,

```

```

22     asignadoEn: timestamp,
23     aceptadoEn: timestamp | null
24 } | null,
25
26 // Fechas del ciclo de vida del pedido
27 fechas: {
28     creacion: timestamp,
29     entregaProgramada: timestamp,
30     recojo: timestamp | null,
31     entrega: timestamp | null,
32     anulacion: timestamp | null
33 },
34
35 // Control de version para actualizaciones concurrentes
36 actualizadoEn: timestamp,
37 version: number
38 }

```

Listing D.4: Estructura de documento Pedido (parte 3)

### D.3 Colección: BILLETERAS

Ruta: billeteras/{billeteraId}

```

1 {
2     id: string,
3     propietario: {
4         tipo: string,           // 'proveedor' | 'empresa'
5         uid: string | null,     // Si es proveedor, referencia a
                                // usuario
6         nombreEmpresa: string
7     },
8     alias: string,             // Nombre descriptivo
9     metodo: string,           // 'yape' | 'plin' | '
                                // transferencia_bancaria'
10    cuenta: {
11        titular: string,
12        telefono: string,       // Para Yape/Plin
13        numeroDocumento: string | null, // DNI/RUC
14        tipoDocumento: string | null    // 'DNI' | 'RUC'
15    },
16    qr: {

```

```

17     urlImagen: string,           // URL de Firebase Storage
18     urlImagenThumbnail: string | null,
19     hash: string | null         // Para validacion
20 },
21 config: {
22     activo: boolean,
23     porDefecto: boolean,
24     soloEfectivo: boolean,
25     limiteMaximo: number | null,
26     comisionPorcentaje: number,
27     prioridad: number           // Para ordenar opciones
28 },
29 verificado: boolean,
30 verificadoPor: {
31     uid: string,
32     nombre: string,
33     fecha: timestamp
34 },
35 creadoEn: timestamp,
36 actualizadoEn: timestamp,
37 ultimoUso: timestamp | null,
38 totalTransacciones: number
39 }

```

Listing D.5: Estructura de documento Billetera

### D.3.1 Índices compuestos necesarios

- proveedor.uid + fechas.creacion (pedidos por cliente)
- motorizado.uid + fechas.creacion (pedidos por motorizado)
- asignacion.recojo.estado + cicloOperativo.diaCreacion (pedidos pendientes del día)
- indices.districtoEntrega + fechas.entregaProgramada (optimización de rutas)



## Anexo E. Configuración de seguridad y autenticación

### E.1 Estructura de custom claims

```
1 {  
2   rol: 'administrador' | 'cliente' | 'motorizado',  
3   permisos: ['crear_pedido', 'asignar_motorizado', '  
4     ver_dashboard'],  
5   empresaId: string | null // Para clientes empresariales  
6 }
```

Listing E.1: Custom claims en Firebase Authentication

### E.2 Reglas de seguridad de Firestore

```
1 rules_version = '2';  
2 service cloud.firestore {  
3   match /databases/{database}/documents {  
4     // Funcion helper para verificar autenticacion  
5     function isAuthenticated() {  
6       return request.auth != null;  
7     }  
8  
9     // Funcion helper para verificar rol  
10    function hasRole(role) {  
11      return isAuthenticated() &&  
12        request.auth.token.rol == role;  
13    }  
14  
15    // Usuarios: pueden leer/actualizar su propio perfil  
16    match /usuarios/{userId} {  
17      allow read: if isAuthenticated() &&  
18        (request.auth.uid == userId ||  
19          hasRole('administrador'));  
20      allow write: if isAuthenticated() &&  
21        request.auth.uid == userId;  
22      allow create: if hasRole('administrador');  
23    }  
24  
25    // Pedidos: acceso segun rol  
26    match /pedidos/{pedidoId} {
```

```

27     allow read: if isAuthenticated() && (
28         hasRole('administrador') ||
29         (hasRole('cliente') &&
30             resource.data.proveedor.uid == request.auth.uid) ||
31         (hasRole('motorizado') &&
32             resource.data.motorizado.uid == request.auth.uid)
33     );
34
35     allow create: if isAuthenticated() &&
36                     (hasRole('administrador') ||
37                     hasRole('cliente'));
38
39     allow update: if isAuthenticated() && (
40         hasRole('administrador') ||
41         (hasRole('motorizado') &&
42             resource.data.motorizado.uid == request.auth.uid)
43     );
44
45     allow delete: if hasRole('administrador');
46 }
47
48 // Billeteras: solo administradores y propietarios
49 match /billeteras/{billeteraId} {
50     allow read: if isAuthenticated();
51     allow write: if hasRole('administrador') ||
52                   resource.data.propietario.uid == request.auth.
53 uid;
54 }
55 }

```

Listing E.2: Firestore Security Rules

### E.3 Reglas de seguridad de Storage

```

1 rules_version = '2';
2 service firebase.storage {
3     match /b/{bucket}/o {
4         // Helper functions
5         function isAuthenticated() {
6             return request.auth != null;
7         }
8
9         function hasRole(role) {
10             return isAuthenticated() &&
11                 request.auth.token.role == role;

```

```

12     }
13
14     // Imagenes de pedidos
15     match /pedidos/{pedidoId}/{tipo}/{filename} {
16         allow read: if isAuthenticated();
17         allow write: if isAuthenticated() && (
18             hasRole('administrador') ||
19             hasRole('cliente') ||
20             hasRole('motorizado')
21         );
22     }
23
24     // Fotos de perfil
25     match /usuarios/{userId}/{filename} {
26         allow read: if isAuthenticated();
27         allow write: if isAuthenticated() &&
28                     request.auth.uid == userId;
29     }
30
31     // QR de billeteras
32     match /billeteras/{billeteraId}/{filename} {
33         allow read: if isAuthenticated();
34         allow write: if hasRole('administrador');
35     }
36 }
37 }

```

Listing E.3: Firebase Storage Security Rules

## Anexo F. Justificación de puntuaciones en matriz multicriterio

### F.1 Escalabilidad

- **Firebase (5):** Escalamiento automático sin configuración, maneja desde 0 hasta millones de usuarios transparentemente.
- **AWS (4):** Excelente escalabilidad pero requiere configuración de Auto Scaling y gestión de capacidad.
- **Backend propio (2):** Requiere planificación manual y redimensionamiento de servidores.

### F.2 Costo

- **Firebase (5):** Capa gratuita generosa, ideal para desarrollo de tesis y etapas iniciales de empresas pequeñas.
- **AWS (3):** Capa gratuita limitada a 12 meses, costos variables pero competitivos.
- **Backend propio (3):** Costos fijos predecibles pero presentes desde el inicio.

### F.3 Facilidad de integración

- **Firebase (5):** Componentes nativamente integrados, SDKs optimizados, despliegue con un comando.
- **AWS (3):** Requiere integrar múltiples servicios manualmente, configuración compleja.
- **Backend propio (2):** Integración completamente manual de cada componente.

### F.4 Mantenimiento

- **Firebase (5):** Gestión de infraestructura completamente delegada a Google.
- **AWS (3):** Gestión parcial, requiere actualización de componentes y monitoreo.

- **Backend propio (2)**: Responsabilidad total del mantenimiento, actualizaciones de seguridad, backups.

## F.5 Rendimiento

- **Firebase (4)**: Excelente para operaciones en tiempo real, latencia baja en sincronización.
- **AWS (4)**: Rendimiento configurable según recursos asignados, muy bueno con optimización.
- **Backend propio (3)**: Depende del proveedor de hosting y configuración manual.

## F.6 Ecosistema y soporte

- **Firebase (5)**: Documentación excelente, amplia comunidad, tutoriales especializados.
- **AWS (5)**: Documentación exhaustiva, soporte empresarial, ecosistema maduro.
- **Backend propio (3)**: Depende de múltiples fuentes, documentación fragmentada.