

Grado en Ciencia e Ingeniería de Datos

Trabajo Fin de Grado

Procesamiento automático de cupones precinto de farmacia por medio de técnicas de aprendizaje profundo

Andrea Mayor Gómez

Tutores

Javier Sánchez Pérez

Idafen Santana Pérez

Las Palmas de Gran Canaria
25 de julio de 2025

SOLICITUD DE DEFENSA DE TRABAJO DE FIN DE TÍTULO

D./D^a **Andrea Mayor Gómez**, autor del trabajo **Procesamiento automático de cupones precinto de farmacia por medio de técnicas de aprendizaje profundo**, correspondiente a la titulación **Grado en Ciencia e Ingeniería de Datos**.

SOLICITA

que se inicie el procedimiento de defensa del mismo, para lo que se adjunta la documentación requerida, haciendo constar que

☒ se autoriza / ☐ no se autoriza la grabación en audio de la exposición y turno de preguntas.

Asimismo, con respecto al registro de la propiedad intelectual/industrial del TFT, declara que:

☐ Se ha iniciado o hay intención de iniciarlo (defensa no pública).

☒ No está previsto.

Y para que así conste firma la presente. (fecha en firma electrónica)

El/La estudiante

Fdo. Andrea Mayor Gómez

A rellenar y firmar **obligatoriamente** por el/la/los/las tutores

En relación con la presente solicitud, se informa: (firmar donde corresponda)

Positivamente (en caso de detección de copia, esta firma quedará invalidada)

Fdo. Javier Sánchez Pérez, Idafen Santana Pérez

Negativamente (justificación en TFT05)

Fdo. Javier Sánchez Pérez, Idafen Santana Pérez

DIRECTORA DE LA ESCUELA DE INGENIERÍA INFORMÁTICA

Procesamiento automático de cupones precinto de farmacia por medio de técnicas de aprendizaje profundo

Trabajo Fin de Grado

Andrea Mayor Gómez

Tutores

Javier Sánchez Pérez

Idafen Santana Pérez

Grado en Ciencia e Ingeniería de Datos

Escuela de Ingeniería Informática

Universidad de Las Palmas de Gran Canaria

Las Palmas de Gran Canaria

25 de julio de 2025

Índice general

Índice de figuras	vi
Índice de tablas	vii
Agradecimientos	viii
Resumen	ix
Abstract	x
1. Introducción	1
1.1. Contexto y relevancia del trabajo	2
1.2. Motivaciones	2
1.3. Objetivos	3
1.3.1. Objetivos específicos tecnológicos	3
1.3.2. Objetivos específicos metodológicos	4
1.3.3. Objetivos específicos académicos	4
1.4. Diferenciación respecto a trabajos previos	4
1.5. Alcance y limitaciones	5
1.6. Organización del documento	6
2. Estado del arte	8
2.1. Modelos YOLO para detección de objetos	8
2.2. Procesamiento de imágenes y códigos de barras	10
2.3. Herramientas de etiquetado y preparación de datos	10
2.4. Modelos de segmentación avanzados	11
2.5. APIs de inteligencia artificial multimodal	11
2.6. Plataformas de desarrollo web especializadas	12
2.7. Aplicaciones farmacéuticas de visión por computador	13
2.8. Métodos basados en gradientes	14
2.9. Análisis comparativo y posicionamiento tecnológico	14
2.10. Conclusiones del estado del arte	15
3. Sistema de gestión de cupones precinto	16
3.1. Marco normativo y legal del sistema	16
3.2. Anatomía del cupón precinto farmacéutico	17
3.2.1. Origen del cupón: de la caja del medicamento	17
3.2.2. Estructura y componentes del cupón individual	18
3.2.3. Organización en hojas de cupones precinto	19

3.3.	Proceso administrativo completo	20
3.3.1.	Flujo desde prescripción hasta facturación	20
3.3.2.	Interacción con colegios farmacéuticos	21
3.3.3.	Sistemas informáticos de gestión	21
3.4.	Impacto económico y desafíos del sistema actual	21
3.4.1.	Costes del sistema manual	22
3.4.2.	Tiempo dedicado por farmacéuticos	22
3.4.3.	Fallos y su impacto	22
3.4.4.	Casos de digitalización exitosa	23
3.5.	Perspectivas de futuro y tendencias	23
3.5.1.	Tendencias hacia la digitalización completa	23
3.5.2.	Integración con sistemas de salud digital	23
3.5.3.	Impacto en la práctica farmacéutica	24
4.	Metodología y planificación del proyecto	25
4.1.	Adaptación de CRISP-DM al contexto del proyecto	25
4.1.1.	Fase 1: Comprensión del negocio	26
4.1.2.	Fase 2: Comprensión de los datos	27
4.1.3.	Fase 3: Preparación de los datos	27
4.1.4.	Fase 4: Modelado	27
4.1.5.	Fase 5: Evaluación	28
4.1.6.	Fase 6: Despliegue	28
4.2.	Planificación temporal del proyecto	28
4.2.1.	Cronograma general	28
4.2.2.	Hitos principales	29
4.2.3.	Gestión de riesgos	29
4.3.	Herramientas y tecnologías empleadas	30
4.3.1.	Entorno de desarrollo	30
4.3.2.	Bibliotecas de visión por computador	30
4.3.3.	Herramientas de etiquetado y evaluación	31
4.3.4.	Desarrollo de interfaz web	31
4.4.	Metodología de evaluación comparativa	31
4.4.1.	Métricas de evaluación técnica	31
4.4.2.	Métricas de evaluación práctica	32
4.4.3.	Protocolos experimentales	32
5.	Conjunto de datos de cupones precinto	33
5.1.	Características generales del conjunto de datos	33
5.1.1.	Composición y estructura de las imágenes	33
5.1.2.	Análisis de elementos detectables	34
5.2.	Análisis exploratorio visual de los datos	35
5.2.1.	Distribución de tamaños de elementos	35
5.2.2.	Casos problemáticos y atípicos	36
5.3.	Proceso de etiquetado y validación	37
5.3.1.	Control de calidad y validación	37
5.4.	División del conjunto de datos y estrategia evolutiva	38
5.4.1.	Justificación metodológica	38
5.5.	Consideraciones sobre aumento de datos	39
5.6.	Limitaciones y oportunidades de mejora	39

6. Métodos empleados	40
6.1. Modelos YOLO para detección de objetos	40
6.1.1. Evolución arquitectónica: YOLO10 y YOLO11	40
6.1.2. Fundamentos arquitectónicos de YOLO	42
6.1.3. Ventajas para detección múltiple	42
6.2. Análisis de gradientes para detección de códigos	42
6.2.1. Principios matemáticos del método	42
6.2.2. Aplicación a códigos de barras farmacéuticos	44
6.3. APIs multimodales	44
6.3.1. GPT-4 Vision: capacidades de comprensión visual	44
6.3.2. Gemini: procesamiento nativo multimodal	45
6.4. SAM2 para segmentación universal	45
6.4.1. Arquitectura transformer para segmentación	45
6.4.2. Capacidades zero shot	45
6.5. Análisis comparativo y selección metodológica	46
7. Configuración de experimentos	48
7.1. Configuración de detectores basados en redes neuronales	48
7.2. Configuración del detector basado en análisis de gradientes	50
7.3. Configuración de APIs	50
7.3.1. Desarrollo y optimización de prompts	50
7.4. Segmentación automática con SAM2	52
7.5. Protocolo experimental y framework de evaluación	52
7.6. Consideraciones metodológicas	53
8. Resultados experimentales	55
8.1. Resultados del detector basado en YOLO	55
8.1.1. Análisis de la discrepancia observada	57
8.1.2. Selección del modelo óptimo	57
8.2. Resultados basados en análisis de gradiente	60
8.2.1. Capacidad de detección básica	60
8.2.2. Capacidad de detección básica	62
8.2.3. Algoritmo de merge inteligente	63
8.2.4. Técnicas de preprocesamiento	63
8.2.5. Análisis de rendimiento	64
8.2.6. Resultados cuantitativos	67
8.3. Resultados de detectores basados en APIs	67
8.3.1. Rendimiento comparativo	67
8.4. Resultados del detector basado en SAM2	69
8.4.1. Arquitectura y requisitos computacionales	69
8.4.2. Proceso de segmentación	69
8.4.3. Limitaciones identificadas y decisión de discontinuación	70
8.4.4. Decodificación y exactitud	71
8.4.5. Evaluación y conclusiones	72
8.5. Análisis comparativo global de métodos	72
8.5.1. Rendimiento cuantitativo	72
8.5.2. Análisis comparativo por categorías	73
8.5.3. Estrategias de implementación recomendadas	73

9. Desarrollo de la aplicación web	75
9.1. Arquitectura del orquestador híbrido	75
9.2. Sistema de corrección posicional	76
9.3. Diseño de la interfaz de usuario	76
9.3.1. Carga y Procesamiento	76
9.3.2. Visualización y Edición de Resultados	77
9.3.3. Validación regulatoria	80
9.4. Valor generado por la aplicación	80
10. Conclusiones y trabajo futuro	82
10.1. Síntesis de resultados y contribuciones del trabajo	82
10.2. Impacto práctico y relevancia del trabajo desarrollado	83
10.3. Lecciones aprendidas y limitaciones identificadas	84
10.4. Direcciones futuras y oportunidades de desarrollo	84
10.5. Reflexiones finales sobre el trabajo realizado	85
A. Competencias específicas cubiertas	87
A.1. Evaluación de competencias específicas	87
A.2. Competencias transversales y desarrollo profesional	87
A.3. Contribución al perfil profesional	88
A.4. Reflexión sobre el aprendizaje	89
B. Objetivos de desarrollo sostenible	90
Lista de Acrónimos	92
Glosario de Términos	93
Bibliografía	99

Índice de figuras

3.1. Elementos del envase farmacéutico [Agencia Española de Medicamentos y Productos Sanitarios, 2024].	17
3.2. Ejemplo real de cupón precinto farmacéutico mostrando la estructura completa del documento con todas sus líneas de información textual, matriz de códigos de barras EAN-13 y elementos de seguridad anti-falsificación. . . .	18
3.3. Hojas de cupones precinto mostrando la organización matricial antes del recorte individual. La imagen izquierda muestra la cara delantera con organización de 6 filas por 4 columnas (24 cupones). La imagen derecha presenta la cara trasera con organización de 7 filas por 4 columnas que permite 26 códigos útiles.	20
4.1. Metodología CRISP-DM adaptada al proyecto de procesamiento automático de cupones precinto farmacéuticos.	26
5.1. Distribución de elementos por clase en el conjunto completo de datos. La relación entre codes (2879), headers (67) y barcodes (124) refleja la estructura típica de los cupones precinto procesados.	34
5.2. Distribución de tamaños de caja delimitadora por clase. Los headers muestran mayor variabilidad en dimensiones, mientras que los códigos de barras presentan características más consistentes.	35
5.3. Ejemplos de casos problemáticos identificados en el conjunto de datos: perspectiva distorsionada, códigos de barras alejados, y posicionamiento incierto por códigos ausentes.	36
5.4. Ejemplo de etiquetado con LabelMe mostrando la delimitación de regiones header (rojo), barcode (verde) y code (amarillo) en cupones precinto. La imagen izquierda presenta el etiquetado de la cara delantera, mientras que la imagen derecha muestra el etiquetado de la cara trasera. Este proceso permite crear anotaciones precisas mediante polígonos que delimitan las regiones de interés para el entrenamiento de modelos de detección.	37
5.5. Distribución de elementos etiquetados por conjunto de datos.	38
6.1. Evolución arquitectónica: comparación paralela entre YOLO10 y YOLO11.	41
6.2. Pipeline del detector basado en análisis de gradientes.	43
6.3. Arquitectura de procesamiento para APIs multimodales.	44
6.4. Arquitectura SAM2 con componentes principales y flujo de información. . . .	46
7.1. Determinación automática de <i>disposición</i> según detección de header SOE. . . .	49
7.2. Framework experimental para evaluación comparativa de detectores.	52

8.1.	Resultado del detector YOLO sobre cupones precinto mostrando detección de códigos individuales, headers manuscritos y áreas generales de códigos .	56
8.2.	Análisis detallado del rendimiento del detector YOLOv10 sin correspondencia posicional mostrando métricas de precisión, distribución de errores y correlación con ground truth	58
8.3.	Análisis detallado del detector YOLOv10 con correspondencia posicional exacta incluyendo métricas de ordenamiento	59
8.4.	Resultado del detector de gradientes mostrando códigos detectados con numeración posicional.	61
8.5.	Configuraciones de sensibilidad del detector de gradientes: (superior) configuración sensible con umbral reducido para mayor cobertura; (inferior) configuración selectiva con umbral elevado para mayor precisión.	62
8.6.	Ejemplo de merge inteligente: reconstrucción automática de códigos de barras fragmentados mediante algoritmos de proximidad y alineación geométrica.	63
8.7.	Distribución de efectividad de las técnicas de preprocesamiento implementadas en el detector de gradientes.	64
8.8.	Análisis de precisión del detector de gradientes por imagen individual.	64
8.9.	Análisis detallado del rendimiento del detector de gradientes sin correspondencia posicional mostrando métricas de precisión y distribución de errores.	65
8.10.	Análisis detallado del detector de gradientes con correspondencia posicional exacta.	66
8.11.	Métricas detalladas de rendimiento con ground truth mostrando precisión, exhaustividad, F1-Score y análisis por imagen y detector.	68
8.12.	Pipeline de SAM2 aplicado a cupones precinto.	70
8.13.	Cupones recortados y códigos decodificados mediante SAM2 + pyzbar.	71
9.1.	Sección de carga de imágenes con interfaz drag-and-drop.	77
9.2.	Sección de resultados con visualización dividida y editor de códigos.	78
9.3.	Comparación de accuracy por imagen entre los tres métodos implementados, evidenciando la superioridad del sistema híbrido.	79
9.4.	Sección de validación CIMA con verificación oficial de códigos farmacéuticos.	80

Índice de tablas

2.1. Comparación crítica de tecnologías para procesamiento de cupones precinto.	15
5.1. División del conjunto de datos según desarrollo iterativo del proyecto. . . .	38
6.1. Comparación de innovaciones clave entre YOLO10 y YOLO11.	41
6.2. Comparación de métodos según criterios específicos para cupones precinto.	47
7.1. Configuración comparativa de detectores YOLO10 y YOLO11	49
7.2. Parámetros del detector basado en análisis de gradientes	50
7.3. Parámetros de configuración para GPT-4 Vision y Google Gemini	51
8.1. Comparación cuantitativa entre YOLOv10 y YOLOv11	57
8.2. Resultados cuantitativos del detector YOLOv10	59
8.3. Resultados cuantitativos del detector de gradientes.	67
8.4. Comparación cuantitativa entre detectores multimodales.	68
8.5. Comparación comprehensiva de todos los métodos evaluados.	73
A.1. Evaluación del nivel de adquisición de competencias específicas.	88
B.1. Tabla con la valoración del grado de relación del Trabajo con los ODS, según una escala de 0 a 3 (no procede, bajo, medio, alto).	91

Agradecimientos

No podría empezar estos agradecimientos sin poner en primer lugar a mis padres, quienes desde pequeña me dieron las alas para comerme el mundo y me apoyaron desde el primer momento que elegí esta carrera, siendo mi pilar fundamental y creyendo en mí en todo momento.

Mis amigos, mi segunda familia, han sido mi refugio y mi fortaleza durante todos estos años, celebrando conmigo cada pequeño logro y acompañándome en los momentos más difíciles.

A mis compañeros de clase, porque no me podría haber tocado mejores acompañantes de viaje en esta aventura universitaria. Gracias por vuestra ayuda constante, por los momentos de estudio compartidos, por las risas en los descansos y por crear un ambiente de apoyo mutuo que ha hecho que estos años sean mucho más llevaderos y enriquecedores.

Quiero dar un enorme agradecimiento a mi tutor, por su atención, paciencia y por la tranquilidad que me ha transmitido en todo este proceso. Gracias por su compromiso, dedicación y por guiarme.

A todo el equipo de Farmalitics, por confiar en mí para la realización de este trabajo y brindarme la oportunidad de desarrollarlo en un entorno profesional tan enriquecedor, que me ha permitido crecer tanto a nivel académico como personal.

Por fin, he llegado al final de este camino. Han sido años de aprendizaje, de crecimiento y de experiencias inolvidables. Terminó esta etapa con la misma emoción con la que la comencé, pero ahora con la satisfacción del deber cumplido y la certeza de haber elegido el camino correcto.

Gracias a todos por ser parte de este viaje.

Resumen

La gestión manual de cupones precinto en farmacia es lenta y propensa a errores. Este Trabajo de Fin de Grado desarrolla una solución automática basada en visión por computador y aprendizaje profundo para digitalizar y analizar estos documentos. Utilizando modelos como YOLO y técnicas avanzadas de lectura de códigos de barras, el sistema detecta y extrae la información clave de imágenes reales, agilizando la gestión y reduciendo errores administrativos. Validado con datos de una empresa colaboradora, el sistema demuestra su utilidad práctica y está preparado para integrarse con futuros asistentes conversacionales. Esta propuesta impulsa la modernización y eficiencia en la gestión documental del sector farmacéutico.

Abstract

Manual management of pharmacy coupon labels is slow and error-prone. This Bachelor's Thesis presents an automated solution based on computer vision and deep learning to digitize and analyze these documents. Leveraging models like YOLO and advanced barcode reading techniques, the system detects and extracts key information from real images, streamlining document handling and minimizing administrative errors. Validated with data from an industry partner, the system proves its practical value and is ready for integration with future conversational assistants. This approach drives modernization and efficiency in pharmaceutical document management.

Capítulo 1

Introducción

La gestión documental en las farmacias españolas continúa dependiendo del uso del cupón precinto: un documento físico que se adhiere a la receta médica para justificar la dispensación y facturación de determinados medicamentos. Este sistema representa uno de los últimos obstáculos en la transformación digital completa del sector farmacéutico español, donde el procesamiento manual consume tiempo valioso y recursos humanos.

El proceso manual de archivo y gestión de cupones genera una carga administrativa significativa para los profesionales farmacéuticos. Estos deben dedicar tiempo considerable a tareas automatizables que podrían destinarse a actividades de mayor valor añadido, como la atención farmacéutica personalizada y el asesoramiento especializado. La gestión manual es propensa a errores humanos, afectando la facturación, el control de medicamentos y el seguimiento de los productos farmacéuticos [Consejo General de Colegios Oficiales de Farmacéuticos, 2024].

En el contexto actual de transformación digital del sector salud, la persistencia de procesos manuales representa una barrera para la implementación de sistemas integrados de gestión sanitaria. La digitalización completa de la gestión farmacéutica es fundamental para mejorar el seguimiento de los medicamentos, facilitar la interoperabilidad entre sistemas sanitarios y permitir análisis de datos que contribuyan a la mejora de la atención sanitaria [Farmaindustria, 2024].

Los avances en visión por computador han alcanzado niveles de precisión que permiten su aplicación práctica en entornos reales. Las técnicas de detección de objetos y segmentación de imágenes han evolucionado significativamente en los últimos años, especialmente con el desarrollo de arquitecturas de aprendizaje profundo.

Los modelos de aprendizaje profundo, especialmente las arquitecturas de detección más recientes, demuestran capacidades excepcionales para la identificación y localización de objetos en imágenes complejas. Estas tecnologías ofrecen la base técnica necesaria para automatizar el procesamiento de documentos farmacéuticos con alta precisión [Ultralytics, 2024a].

Este trabajo propone una solución integral que combina múltiples tecnologías de vanguardia para automatizar completamente el procesamiento de cupones precinto. El sistema integra modelos de detección de objetos como las versiones más recientes de arquitecturas neuronales, técnicas tradicionales de procesamiento mediante OpenCV, métodos de detección por gradiente, y tecnologías emergentes de IA como herramientas de segmentación avanzada y servicios de inteligencia artificial multimodal. La extracción de información se realiza mediante un enfoque multicapa que combina la biblioteca especializada `pyzbar` con técnicas exhaustivas de preprocesamiento de imágenes, garantizando la

máxima probabilidad de éxito en la lectura de código de barras.

A diferencia de soluciones existentes que se centran en documentos estandarizados o controlados, este trabajo aborda específicamente los desafíos únicos de los cupones precinto farmacéuticos españoles. Estos documentos presentan características particulares que los distinguen de otros tipos de documentos procesados automáticamente. Los cupones precinto incluyen múltiples códigos de barras en disposiciones complejas, requieren procesamiento en condiciones de iluminación variable, presentan distorsiones físicas por manipulación, reflejos por la cinta adhesiva de protección, y condiciones de captura no controladas en entornos operativos reales. Estas características demandan enfoques específicos que no se encuentran en trabajos previos sobre procesamiento automático de documentos [Instituto de Ingeniería del Conocimiento, 2024].

El sistema multicapa propuesto, que integra modelos de aprendizaje profundo con técnicas tradicionales de procesamiento de imágenes, representa una contribución metodológica significativa al combinar la fiabilidad de los métodos clásicos con la capacidad adaptativa de las redes neuronales. Esta integración permite aprovechar las fortalezas complementarias de diferentes tecnologías para maximizar la precisión en condiciones adversas.

1.1. Contexto y relevancia del trabajo

El sector farmacéutico español se encuentra en un proceso de modernización acelerado, impulsado por las demandas de eficiencia operativa y los requisitos normativos establecidos en la Estrategia de la Industria Farmacéutica 2024-2028. La gestión de cupones precinto representa uno de los últimos procesos analógicos que requieren sistematización digital [Ministerio de Sanidad, 2024].

Los procesos manuales representan un coste significativo tanto en tiempo como en recursos económicos. Las experiencias pioneras en la Comunidad Valenciana han demostrado ahorros de hasta 4,7 millones de euros anuales mediante la eliminación del cupón precinto [Conselleria de Sanidad, 2024].

La automatización del procesamiento de cupones precinto contribuye directamente a múltiples objetivos estratégicos del sector. En primer lugar, mejora la seguridad del paciente al reducir errores humanos en la identificación y registro de medicamentos. En segundo lugar, facilita el cumplimiento normativo mediante la mejora del seguimiento y el registro automático de transacciones. Finalmente, libera recursos humanos especializados para tareas de mayor valor clínico.

La sistematización de estos datos abre nuevas posibilidades para el análisis de patrones de dispensación, la detección de anomalías en el consumo de medicamentos y la optimización de la gestión de inventarios. Estos análisis pueden contribuir a la mejora de la atención sanitaria mediante la identificación de tendencias en el uso de medicamentos y la optimización de los procesos de dispensación.

1.2. Motivaciones

Este trabajo surge de la convergencia entre una necesidad real del sector farmacéutico español y las oportunidades que ofrecen las tecnologías emergentes de inteligencia artificial, creando un escenario ideal para el desarrollo de soluciones innovadoras con impacto social directo.

Desde la perspectiva técnica y científica, el proyecto permite explorar la aplicación práctica de tecnologías de vanguardia en visión por computador a un problema real con características únicas. Los cupones precinto presentan desafíos específicos que no se encuentran en otros documentos estandarizados: variabilidad en la calidad de impresión, distorsiones por manipulación física y la necesidad de procesar múltiples códigos de barras en disposiciones complejas. El desarrollo de un sistema robusto para estas condiciones requiere la integración de múltiples enfoques tecnológicos y contribuye al conocimiento científico sobre la aplicabilidad de diferentes técnicas de visión por computador en entornos operativos reales [Grupo de Investigación en Visión Artificial, 2024].

Profesionalmente, el proyecto aborda una necesidad documentada del sector sanitario español, donde los procesos manuales representan un coste significativo tanto en tiempo como en recursos económicos. La colaboración empresarial garantiza que el desarrollo se base en datos reales y requisitos auténticos del sector, contribuyendo directamente a la digitalización como prioridad estratégica nacional y europea [Farmalitics, 2024].

Académicamente, el trabajo proporciona una experiencia formativa integral que permite aplicar conocimientos de visión por computador, aprendizaje profundo y desarrollo de aplicaciones web a un caso de estudio con impacto práctico real [CRISP-DM Consortium, 2024].

1.3. Objetivos

El objetivo principal de este trabajo es desarrollar un sistema automático integral para la sistematización y procesamiento de cupones precinto farmacéuticos, capaz de detectar, segmentar y extraer información relevante de imágenes tomadas en condiciones reales de trabajo, con niveles de precisión y fiabilidad adecuados para su implementación práctica en entornos farmacéuticos.

1.3.1. Objetivos específicos tecnológicos

Los objetivos tecnológicos se centran en el desarrollo e implementación de soluciones técnicas fiables:

- Implementar y evaluar múltiples enfoques de detección de objetos para la identificación de zonas clave en cupones precinto, incluyendo modelos de detección de última generación y técnicas tradicionales de procesamiento de imágenes
- Desarrollar un sistema fiable de preprocesamiento de imágenes que maneje eficazmente variaciones en iluminación, perspectiva y calidad de imagen típicas de entornos operativos reales
- Integrar bibliotecas especializadas para maximizar la tasa de éxito en la lectura de código de barras EAN-13, combinando enfoques complementarios para garantizar la máxima cobertura
- Diseñar e implementar una interfaz web intuitiva que facilite la adopción por parte de profesionales farmacéuticos sin conocimientos técnicos especializados
- Establecer la integración con la API de CIMA de la Agencia Española de Medicamentos y Productos Sanitarios AEMPS para la validación automática de datos

extraídos y la verificación de la coherencia de la información [Agencia Española de Medicamentos y Productos Sanitarios, 2024]

1.3.2. Objetivos específicos metodológicos

Los objetivos metodológicos se enfocan en la evaluación rigurosa y la validación científica de las soluciones desarrolladas:

- Realizar una evaluación comparativa exhaustiva de diferentes enfoques tecnológicos mediante métricas objetivas de rendimiento, incluyendo precisión, recall, F1-score y tiempo de procesamiento
- Desarrollar protocolos de evaluación que consideren las condiciones reales de trabajo en farmacias, incluyendo variaciones en iluminación, calidad de imagen y tipos de cupones
- Establecer criterios de calidad y fiabilidad apropiados para la implementación práctica del sistema, considerando los requisitos operativos del sector farmacéutico
- Crear un conjunto de datos de referencia con ground truth validado manualmente para futuras investigaciones en el procesamiento automático de documentos farmacéuticos

1.3.3. Objetivos específicos académicos

Los objetivos académicos se relacionan con la formación integral y la contribución al conocimiento científico:

- Aplicar conocimientos de visión por computador, aprendizaje profundo y desarrollo de aplicaciones web a un caso de estudio real con impacto social directo
- Desarrollar competencias en la evaluación comparativa de tecnologías emergentes y la toma de decisiones técnicas fundamentadas en evidencia empírica
- Adquirir experiencia práctica en todas las fases del ciclo de vida de un proyecto de ingeniería de datos, desde la definición del problema hasta la implementación de la solución
- Contribuir al conocimiento científico mediante la documentación rigurosa y la potencial publicación de resultados en el ámbito de la visión por computador aplicada al sector sanitario

1.4. Diferenciación respecto a trabajos previos

Este trabajo se diferencia significativamente de las soluciones existentes en el procesamiento automático de documentos por su enfoque específico en las características únicas de los cupones precinto farmacéuticos españoles. Mientras que la mayoría de sistemas de procesamiento documental se desarrollan para documentos estandarizados con formatos controlados, este proyecto aborda los desafíos específicos de documentos que presentan alta variabilidad y condiciones operativas complejas.

Los trabajos previos sobre reconocimiento óptico de caracteres y procesamiento de códigos de barras se centran en documentos con características uniformes, iluminación controlada y orientación estándar. Los cupones precinto, por el contrario, presentan múltiples códigos de barras en disposiciones complejas, variabilidad en la calidad de impresión por diferentes proveedores, distorsiones físicas por manipulación, reflejos por la cinta adhesiva de protección, y condiciones de captura no controladas en entornos operativos reales.

El sistema multicapa desarrollado, que integra modelos de aprendizaje profundo con técnicas tradicionales de procesamiento de imágenes, representa una contribución metodológica al combinar la fiabilidad de los métodos clásicos con la capacidad adaptativa de las redes neuronales. Esta integración permite aprovechar las fortalezas complementarias de diferentes tecnologías para maximizar la precisión en condiciones adversas.

Además, la validación directa con sistemas operativos reales del sector farmacéutico español, mediante la integración con CIMA y la colaboración empresarial, proporciona un nivel de validación práctica que supera las evaluaciones académicas típicas basadas únicamente en datasets sintéticos o controlados [Instituto de Ingeniería del Conocimiento, 2024].

1.5. Alcance y limitaciones

El alcance de este trabajo se centra en el desarrollo y evaluación de un sistema de procesamiento automático de cupones precinto, utilizando imágenes reales proporcionadas por la empresa colaboradora. El sistema aborda específicamente la detección de zonas de interés y la extracción de código de barras, con validación mediante la interfaz de CIMA.

Los principales desafíos técnicos están relacionados con el procesamiento de imágenes en condiciones reales. La variabilidad en iluminación afecta la uniformidad del procesamiento y requiere técnicas adaptativas de preprocesamiento. La distorsión de perspectiva dificulta la detección automática, ya que los cupones frecuentemente aparecen con orientaciones no ideales.

La fragmentación de códigos presenta retos adicionales, pues la detección de códigos de barras parcialmente visibles o dañados requiere algoritmos sofisticados. El ordenamiento posicional demanda establecer la correspondencia exacta entre códigos detectados y su posición real en la cuadrícula del cupón, aspecto crítico para la interpretación correcta de la información.

Las restricciones del conjunto de datos incluyen su tamaño reducido de 31 imágenes, que puede no capturar completamente la variabilidad del ecosistema farmacéutico español. La dependencia de un único proveedor puede restringir la generalización a otros formatos de cupones, mientras que la calidad variable de algunas imágenes presenta códigos parcialmente ilegibles que afectan las métricas de evaluación.

El trabajo se enfoca exclusivamente en cupones precinto españoles sin abordar otros tipos de documentos farmacéuticos. La verificación se realiza únicamente con la interfaz de CIMA, sin considerar otras fuentes de contraste. El sistema no incluye funcionalidades de gestión documental completa ni integración con sistemas de gestión farmacéutica existentes.

1.6. Organización del documento

El presente documento se estructura siguiendo una progresión lógica desde el contexto teórico hasta los resultados prácticos obtenidos.

El **Capítulo 2** presenta una revisión exhaustiva del estado del arte, analizando las tecnologías más relevantes para el procesamiento automático de documentos farmacéuticos. Se examinan desde los modelos de detección más recientes hasta las interfaces de inteligencia artificial multimodal, estableciendo las bases teóricas para la evaluación comparativa posterior. Este capítulo proporciona el contexto científico necesario para comprender las decisiones tecnológicas adoptadas en el proyecto.

El **Capítulo 3** proporciona una descripción detallada del funcionamiento real del sistema de gestión de cupones precinto en farmacias españolas. Se explica el proceso administrativo completo, la interacción con los colegios farmacéuticos y se justifica la necesidad de automatización desde el punto de vista del usuario final. Este capítulo contextualiza el problema desde la óptica operativa y establece los requisitos funcionales del sistema.

El **Capítulo 4** detalla la metodología y planificación del proyecto, presentando la adaptación de CRISP-DM al contexto específico del proyecto. Se documenta la planificación temporal evolutiva, las herramientas empleadas y el enfoque iterativo adoptado. Se incluyen las decisiones metodológicas que guiaron el desarrollo y las desviaciones respecto a la planificación inicial.

El **Capítulo 5** describe las características del conjunto de datos utilizado, incluyendo el proceso de etiquetado con LabelMe y las técnicas de preprocesamiento implementadas. Se presenta la estrategia de división evolutiva adoptada y se analizan las restricciones del conjunto de datos y su impacto en la generalización de los resultados. Se incluyen ejemplos visuales y análisis estadísticos que ayudan a comprender la naturaleza de los datos.

El **Capítulo 6** aborda los aspectos técnicos centrales, explicando los fundamentos teóricos de los cuatro enfoques implementados: modelos de detección neuronal, análisis de gradientes, herramientas de segmentación avanzada e interfaces multimodales. Se justifica la selección metodológica y se describe la complementariedad entre enfoques. Se incluyen diagramas para facilitar la comprensión de los métodos implementados.

El **Capítulo 7** especifica la configuración experimental desarrollada para la evaluación comparativa. Se detallan los parámetros específicos de cada detector, los protocolos experimentales y el framework de evaluación implementado. Se presentan las métricas utilizadas con sus fórmulas explícitas y se establecen los criterios de comparación.

El **Capítulo 8** presenta los resultados obtenidos con cada enfoque, proporcionando un análisis detallado del rendimiento mediante métricas objetivas y evaluación cualitativa. Se incluyen ejemplos de resultados exitosos y fallidos, análisis de errores frecuentes y un análisis comparativo global que identifica las fortalezas y limitaciones de cada método en diferentes escenarios operativos.

El **Capítulo 9** describe el desarrollo de una aplicación web que implementa un sistema híbrido YOLO-gradientes, superando las limitaciones de los métodos individuales mediante un orquestador inteligente y corrección posicional avanzada. Se presenta la interfaz web farmacéutica y se valida experimentalmente la efectividad del enfoque integrado.

El **Capítulo 10** recoge las conclusiones y líneas de trabajo futuro, junto con la evaluación del cumplimiento de objetivos y el impacto del proyecto. Se incluye una reflexión personal sobre el proceso de aprendizaje y las lecciones adquiridas durante el desarrollo del trabajo.

Los anexos incluyen las competencias específicas trabajadas, la contribución a los Objetivos de Desarrollo Sostenible, y resultados detallados complementarios que amplían la información presentada en los capítulos principales.

A lo largo del documento se emplean términos técnicos y acrónimos cuya definición puede consultarse en el glosario incluido al final.

Capítulo 2

Estado del arte

La digitalización automática de documentos farmacéuticos mediante técnicas de visión por computador ha experimentado un crecimiento exponencial en los últimos años, impulsado por la necesidad urgente de automatizar procesos administrativos en el sector sanitario [Redmon et al., 2016, Ultralytics, 2024a]. El procesamiento de cupón precinto representa un desafío específico que requiere la convergencia de múltiples disciplinas tecnológicas: detección de objetos, reconocimiento de código de barras, preprocesamiento de imágenes y desarrollo de interfaces de usuario intuitivas.

La relevancia de este campo se ha intensificado con la reciente eliminación del cupón precinto en la Comunidad Valenciana, donde se ha implementado un sistema de código digital que ha generado un ahorro estimado de 4,7 millones de euros anuales [Alacantí TV, 2025]. Este cambio normativo no solo demuestra la viabilidad económica de la digitalización, sino que también establece un precedente para el resto de España, donde el coste total del sistema manual se estima entre 80-90 millones de euros anuales [El Global Farmacia, 2024].

El mercado de IA farmacéutica, valorado en 4.350 millones de USD (dólares estadounidenses) en 2025 y proyectado a alcanzar 25.730 millones de USD en 2030, refleja la importancia estratégica de estas tecnologías [Ultralytics, 2025]. Esta transformación digital no es meramente una optimización de procesos existentes, sino una reconfiguración fundamental de cómo se gestiona la información farmacéutica en el siglo XXI.

Este capítulo presenta una revisión crítica y comparativa de las tecnologías relevantes para el desarrollo de sistemas de procesamiento automático de cupones precinto, analizando no solo sus capacidades técnicas, sino también sus limitaciones prácticas, costes de implementación y viabilidad en entornos farmacéuticos reales. Se examina cómo cada tecnología contribuye al ecosistema global de digitalización farmacéutica y se identifican las oportunidades de innovación que permanecen inexploradas.

2.1. Modelos YOLO para detección de objetos

Los modelos You Only Look Once (YOLO) han revolucionado el campo de la detección de objetos en tiempo real, estableciendo un nuevo paradigma que equilibra velocidad y precisión de manera óptima para aplicaciones prácticas [Redmon et al., 2016, Wang et al., 2024a]. Su impacto en el sector farmacéutico ha sido particularmente significativo, donde la necesidad de procesar grandes volúmenes de documentos con alta precisión y velocidad constante hace que estos modelos sean especialmente atractivos.

La arquitectura fundamental de YOLO se basa en un enfoque unificado que realiza

detección y clasificación en una sola pasada por la red neuronal, eliminando la complejidad computacional de métodos anteriores como Region-based Convolutional Neural Networks (R-CNN) que requerían múltiples etapas de procesamiento [Girshick et al., 2014]. Esta eficiencia arquitectónica es crucial en aplicaciones farmacéuticas donde el tiempo de procesamiento impacta directamente en la productividad de las farmacias.

El funcionamiento de YOLO se fundamenta en la división de la imagen de entrada en una cuadrícula regular, donde cada celda predice simultáneamente las coordenadas de la caja delimitadora, las probabilidades de clase y las puntuaciones de confianza [Redmon and Farhadi, 2018]. Esta aproximación permite un procesamiento extremadamente eficiente, ya que toda la información necesaria se obtiene en una única evaluación de la red neuronal, contrastando favorablemente con métodos tradicionales que requerían múltiples pasadas y post-procesamiento complejo.

La evolución de YOLO desde sus primeras versiones hasta las implementaciones más recientes como YOLO10 y YOLO11 refleja una optimización continua para aplicaciones de procesamiento documental [Wang et al., 2024a, Ultralytics, 2024c]. YOLO10 introduce optimizaciones para hardware moderno y mejoras en la detección de objetos pequeños, características especialmente importantes para la identificación precisa de códigos de barras y elementos textuales en cupones precinto. Sus innovaciones incluyen el Consistent Dual Assignments que elimina la necesidad de Non-Maximum Suppression (NMS) durante la inferencia, reduciendo significativamente la latencia [Wang et al., 2024a].

YOLO11 representa una evolución significativa que incorpora mecanismos de atención más sofisticados y técnicas de fusión de características que mejoran la detección en condiciones desafiantes [Ultralytics, 2024c]. La introducción del bloque Cross Stage Partial with Partial Self-Attention (C2PSA) mejora la capacidad de modelado global con un coste computacional mínimo. Para el procesamiento de documentos farmacéuticos, estas mejoras se traducen en mayor *solidez* frente a variaciones de iluminación, perspectiva y calidad de imagen, condiciones comunes en entornos farmacéuticos reales donde las condiciones de captura no siempre son ideales.

La aplicación específica de YOLO en farmacia ha demostrado resultados prometedores en múltiples contextos. La detección automatizada de píldoras utilizando YOLO11 ha mostrado precisión superior al 90 % en clasificación farmacéutica, acelerando procesos que tradicionalmente eran lentos, repetitivos y propensos a errores humanos [Ultralytics, 2025]. Esta capacidad se extiende naturalmente al procesamiento de cupones precinto, donde la detección simultánea de múltiples elementos (códigos de barras, texto, áreas específicas) requiere exactamente las capacidades multiclase que YOLO proporciona de manera nativa.

El análisis comparativo de YOLO frente a otras arquitecturas de detección revela ventajas específicas para aplicaciones farmacéuticas. La velocidad de procesamiento, es de entre 30-60 FPS en hardware moderno, permite implementaciones en tiempo real que son fundamentales para la adopción práctica en farmacias donde la eficiencia operativa es crítica [Ultralytics, 2024a]. Esta velocidad contrasta favorablemente con métodos como R-CNN o Faster R-CNN que, aunque pueden ofrecer mayor precisión en algunos casos, requieren tiempos de procesamiento que los hacen *inviables* para uso farmacéutico rutinario [Ren et al., 2015].

Sin embargo, YOLO presenta limitaciones importantes que deben considerarse en el contexto farmacéutico. La detección de objetos muy pequeños o densamente agrupados puede ser problemática, especialmente relevante cuando los cupones precinto contienen códigos de barras de tamaño reducido o cuando múltiples cupones se superponen en la imagen [Redmon and Farhadi, 2018]. Además, la dependencia de conjuntos de datos de

entrenamiento específicos del dominio significa que la implementación en nuevos contextos farmacéuticos requiere procesos de fine-tuning que pueden ser técnicamente complejos y costosos [Yosinski et al., 2014].

2.2. Procesamiento de imágenes y códigos de barras

El reconocimiento automático de código de barras constituye el componente técnico más crítico en el procesamiento de cupones precinto, requiriendo bibliotecas especializadas capaces de manejar la variabilidad inherente en documentos farmacéuticos reales. Esta variabilidad incluye diferencias en calidad de impresión, condiciones de iluminación, perspectiva de captura y estado físico de los documentos, factores que pueden comprometer significativamente la tasa de éxito en la lectura automática [Gallo and Manduchi, 2011].

La biblioteca `pyzbar` se ha establecido como una solución *fiable* para la decodificación de códigos de barras en Python, ofreciendo soporte nativo para múltiples formatos incluyendo Code 128, Code 39, European Article Number de 13 dígitos (EAN-13), y códigos QR [Hudson, 2024]. Su implementación permite integración directa con pipelines de procesamiento basados en Open Source Computer Vision Library (OpenCV), facilitando el desarrollo de sistemas completos de procesamiento documental [OpenCV Development Team, 2024]. La efectividad de `pyzbar` puede optimizarse significativamente mediante técnicas de preprocesamiento específicas que incluyen binarización adaptativa, corrección de inclinación, ajustes de contraste y brillo, y filtrado morfológico.

Para optimizar la lectura se aplican técnicas como ajustes de contraste y brillo dinámicos, ecualización adaptativa de histograma (CLAHE), filtros gaussianos/medianos y operaciones morfológicas, corrección gamma y máscaras de enfoque, y redimensionado multi-escala con interpolación cúbica.

El uso secuencial de estas técnicas, junto a estrategias de *fallback* (por ejemplo combinando `pyzbar`), maximiza la tasa de lectura, crucial para la trazabilidad y facturación.

2.3. Herramientas de etiquetado y preparación de datos

La creación de conjuntos de datos de alta calidad representa un factor determinante en el éxito de sistemas de visión por computador aplicados a documentos farmacéuticos. Las herramientas de etiquetado manual no solo proporcionan la base para el entrenamiento de modelos especializados, sino que también determinan el límite superior de rendimiento que estos modelos pueden alcanzar [Russell et al., 2008].

LabelMe, desarrollado originalmente por el Laboratorio de Ciencias de la Computación e Inteligencia Artificial del MIT en 2008, se ha consolidado como la herramienta estándar para anotación manual de imágenes en aplicaciones de investigación y desarrollo comercial [Russell et al., 2008]. Su capacidad para generar anotaciones en formato JSON compatible con frameworks de aprendizaje profundo, combinada con una interfaz intuitiva para la creación de máscaras de segmentación de imágenes precisas, la convierte en una herramienta indispensable para el desarrollo de sistemas farmacéuticos [Innovatiana, 2024].

Para el etiquetado específico de cupones precinto, LabelMe permite la anotación precisa de regiones de interés complejas incluyendo cabeceras, áreas de códigos de barras individuales y códigos específicos. La herramienta facilita la creación de polígonos de segmentación que definen con precisión los contornos de cada elemento, proporcionando la

resolución necesaria para entrenar detectores *fiabiles* capaces de manejar la variabilidad presente en documentos farmacéuticos reales. Su capacidad para manejar anotaciones *versátiles* incluyendo cuadros delimitadores, polígonos, rectángulos, círculos, líneas y puntos la hace especialmente adaptable para documentos farmacéuticos complejos [Innovatiana, 2024].

La consistencia en el proceso de etiquetado emerge como un factor crítico que determina la calidad del conjunto de datos resultante. LabelMe proporciona funcionalidades específicas para mantener esta consistencia, incluyendo plantillas de etiquetas predefinidas, sistemas de validación de anotaciones y herramientas de revisión colaborativa que permiten verificar la calidad antes del uso en entrenamiento. Su compatibilidad multiplataforma (Ubuntu, macOS y Windows) facilita la colaboración en equipos distribuidos, aspecto importante en proyectos farmacéuticos que pueden involucrar múltiples instituciones [Russell et al., 2008].

2.4. Modelos de segmentación avanzados

Los avances recientes en modelos de segmentación universal han introducido nuevas posibilidades teóricas para el procesamiento de documentos complejos, aunque su implementación práctica presenta desafíos significativos que deben evaluarse cuidadosamente en el contexto de aplicaciones farmacéuticas reales.

El modelo Segment Anything Model 2 (SAM2) representa un avance significativo en la segmentación de imágenes, ofreciendo capacidades de segmentación universal que teóricamente pueden adaptarse a diferentes tipos de objetos sin entrenamiento específico [Ravi et al., 2024]. Su arquitectura permite la segmentación de objetos mediante prompts visuales o textuales, proporcionando flexibilidad conceptual para aplicaciones diversas. SAM2 combina capacidades de segmentación de imagen y vídeo en un único modelo, alcanzando velocidades de inferencia de aproximadamente 44 fotogramas por segundo en hardware especializado [Ultralytics, 2024b].

Sin embargo, la evaluación práctica de SAM2 para el procesamiento de cupones precinto revela limitaciones significativas que cuestionan su viabilidad en aplicaciones farmacéuticas reales. Los requisitos computacionales son sustancialmente mayores que los de modelos YOLO optimizados, requiriendo hardware especializado y tiempos de procesamiento que pueden no ser adecuados para aplicaciones en tiempo real. Aunque logra cortes muy precisos, sus requerimientos de GPU y tiempo de procesamiento lo hacen poco *escalable* en farmacias comunitarias [Yosinski et al., 2014].

La evaluación comparativa directa entre SAM2 y modelos YOLO en tareas de procesamiento de cupones precinto demuestra que, aunque SAM2 puede lograr segmentaciones más precisas en casos ideales, los beneficios marginales no justifican la complejidad adicional, los requisitos de recursos y los costes de implementación para la aplicación específica de procesamiento farmacéutico. Esta conclusión es consistente con estudios previos que sugieren que modelos más simples y especializados pueden superar a modelos generales complejos en dominios específicos [Ruder, 2017].

2.5. APIs de inteligencia artificial multimodal

Las Application Programming Interfaces (API) de inteligencia artificial multimodal representan la frontera tecnológica actual en el procesamiento de documentos complejos,

ofreciendo capacidades avanzadas que combinan análisis visual y textual en sistemas integrados. Estas tecnologías proporcionan alternativas o métodos de validación para enfoques tradicionales, aunque introducen dependencias y consideraciones que deben evaluarse cuidadosamente en contextos farmacéuticos.

GPT-4 Vision extiende las capacidades de procesamiento de lenguaje natural al dominio visual, permitiendo análisis y comprensión de imágenes mediante procesamiento de lenguaje natural [OpenAI, 2024a]. Esta integración multimodal permite formular consultas complejas sobre contenido visual utilizando instrucciones en lenguaje natural, eliminando teóricamente la necesidad de desarrollar pipelines de procesamiento específicos. La arquitectura subyacente combina encoders visuales especializados con modelos de lenguaje transformer, permitiendo la traducción de información visual a representaciones textuales procesables por mecanismos de razonamiento del modelo de lenguaje.

Para aplicaciones farmacéuticas, esto se traduce en capacidades de Optical Character Recognition (OCR) avanzado, identificación de objetos, análisis de layout y razonamiento visual que pueden interpretar la estructura y organización de estos documentos de manera contextual. Sin embargo, la implementación práctica de GPT-4 Vision en entornos farmacéuticos presenta limitaciones significativas. La dependencia de conectividad constante y disponibilidad del servicio introduce riesgos operativos inaceptables en aplicaciones críticas. Los costes variables por uso pueden volverse impredecibles en aplicaciones de alto volumen, mientras que las consideraciones de privacidad relacionadas con el envío de información farmacéutica a servicios externos plantean cuestiones de cumplimiento normativo.

Google Gemini representa una aproximación arquitectónica diferente al procesamiento multimodal, con diseño nativo para manejar múltiples modalidades de forma integrada [Google DeepMind, 2024]. Esta integración fundamental permite procesamiento más eficiente y coherente de información multimodal, evitando limitaciones de enfoques que combinan modelos especializados desarrollados independientemente. La capacidad de Gemini para procesar documentos complejos con layouts estructurados lo hace conceptualmente relevante para aplicaciones farmacéuticas, donde los cupones precinto presentan organizaciones específicas que requieren comprensión tanto visual como estructural.

No obstante, Gemini presenta limitaciones similares a GPT-4 Vision en términos de dependencia externa, control limitado sobre el procesamiento interno, costes variables y consideraciones de privacidad. Adicionalmente, las limitaciones del plan gratuito (50 solicitudes diarias) restringen severamente su aplicabilidad en entornos de producción farmacéutica [Google DeepMind, 2024].

2.6. Plataformas de desarrollo web especializadas

El desarrollo de interfaces de usuario accesibles e intuitivas constituye un factor determinante en la adopción exitosa de sistemas de procesamiento automático de documentos farmacéuticos. Las plataformas modernas de desarrollo web deben equilibrar funcionalidad avanzada con simplicidad de uso, considerando que los usuarios finales son profesionales farmacéuticos sin necesidad de conocimientos técnicos especializados.

Streamlit se ha establecido como una plataforma líder para el desarrollo rápido de aplicaciones web interactivas centradas en ciencia de datos y machine learning [Streamlit Inc., 2024]. Su filosofía de “escribir menos código, crear más aplicaciones” permite a desarrolladores crear interfaces funcionales directamente desde scripts de Python, eliminando la complejidad tradicional del desarrollo web que requiere conocimientos especializados en múltiples tecnologías [Kanaries, 2023].

Para aplicaciones específicas de procesamiento de documentos farmacéuticos, Streamlit ofrece componentes nativos que facilitan la carga de archivos, visualización de imágenes procesadas y presentación de resultados en formatos tabulares. Su capacidad para manejar sesiones de usuario y mantener estado entre interacciones lo hace especialmente adecuado para aplicaciones que requieren procesamiento de múltiples documentos de forma secuencial.

La integración nativa con bibliotecas de machine learning como PyTorch, OpenCV y pandas permite crear pipelines de procesamiento completos dentro de la misma aplicación [Streamlit Inc., 2024]. Streamlit también proporciona widgets interactivos para ajustar parámetros de procesamiento en tiempo real, facilitando la experimentación y optimización de algoritmos directamente por parte de usuarios finales.

2.7. Aplicaciones farmacéuticas de visión por computador

La aplicación de técnicas de visión por computador en el sector farmacéutico ha experimentado una expansión significativa, impulsada por la necesidad de automatizar procesos que tradicionalmente requerían intervención humana intensiva. Estas aplicaciones van desde el control de calidad en producción hasta la gestión de inventarios en farmacias hospitalarias, demostrando la versatilidad y potencial de estas tecnologías [Pharmaceutical Technology, 2023].

La visión por computador se utiliza extensivamente para verificar que los envases de productos farmacéuticos estén etiquetados correctamente, incluyendo fechas de caducidad, códigos de lote e información de trazabilidad [International Journal of Pharmaceutical Sciences, 2023]. Esta aplicación es particularmente crítica en el contexto de cupones precinto, donde la verificación automática de códigos de barras e información textual puede prevenir errores en la facturación y garantizar el cumplimiento normativo.

Los sistemas de visión artificial permiten inspección de calidad en tiempo real, control de etiquetado y envasado, automatización de procesos de clasificación, análisis de datos, detección de patrones anómalos, seguimiento en tiempo real de productos, identificación y verificación automática de productos y optimización de la logística farmacéutica [Pharmaceutical Executive, 2024]. Estas capacidades se traducen directamente en beneficios para el procesamiento de cupones precinto, donde la automatización puede reducir errores humanos y acelerar procesos administrativos.

La aplicación de visión por computador en farmacia hospitalaria ha demostrado resultados prometedores en la gestión automatizada de inventarios [American Journal of Health-System Pharmacy, 2023]. Los sistemas pueden contabilizar automáticamente el número de unidades de medicamentos disponibles mediante análisis de imágenes, reduciendo significativamente las horas de trabajo del personal hospitalario dedicadas a tareas de inventario.

Esta experiencia en farmacia hospitalaria proporciona insights valiosos para la aplicación en farmacias comunitarias, donde el procesamiento automático de estos documentos puede liberar tiempo del personal farmacéutico para actividades de mayor valor clínico. La transferibilidad de estas tecnologías sugiere que los sistemas desarrollados para cupones precinto pueden expandirse a otras aplicaciones de gestión farmacéutica.

2.8. Métodos basados en gradientes

Los enfoques basados en gradientes detectan códigos EAN-13 aprovechando las características visuales intrínsecas de sus bordes verticales mediante un proceso de análisis direccional. Este método utiliza operadores de gradiente para generar mapas de energía que resaltan estructuras verticales características de los códigos de barras mientras suprimen elementos horizontales.

La principal ventaja de este enfoque es su completa independencia de datos de entrenamiento, lo que elimina la necesidad de conjuntos de datos específicos del dominio farmacéutico. Además, ofrece tiempos de procesamiento muy rápidos, típicamente entre 1 y 2 segundos por imagen, lo que lo hace especialmente atractivo para aplicaciones en tiempo real.

Sin embargo, el método presenta sensibilidad a condiciones de bajo contraste, lo que puede afectar significativamente su rendimiento en cupones con calidad de impresión deficiente. Adicionalmente, la ausencia de contexto semántico limita su capacidad para reordenar automáticamente celdas erróneas o manejar situaciones donde la estructura esperada del cupón presenta irregularidades.

2.9. Análisis comparativo y posicionamiento tecnológico

El análisis comparativo de las tecnologías revisadas revela un paisaje tecnológico complejo donde diferentes enfoques ofrecen ventajas específicas según el contexto de aplicación, los recursos disponibles y los requisitos operativos. Esta diversidad tecnológica no representa una fragmentación problemática, sino una riqueza de opciones que permite optimizar soluciones según necesidades específicas.

Los modelos YOLO emergen como la opción más equilibrada para aplicaciones farmacéuticas, ofreciendo un balance óptimo entre precisión, velocidad y requisitos computacionales. Su madurez tecnológica, amplia documentación y comunidad de desarrollo activa los convierten en una opción de bajo riesgo para implementaciones comerciales [Ultralytics, 2024a]. La capacidad de fine-tuning específico del dominio permite adaptación a las particularidades de cupones precinto sin requerir desarrollo desde cero [Yosinski et al., 2014].

En contraste, los modelos de segmentación avanzados como SAM2, aunque técnicamente superiores en capacidades de segmentación, presentan barreras de entrada significativas en términos de requisitos computacionales, complejidad de implementación y costes operativos [Ravi et al., 2024]. Su aplicación se justifica principalmente en contextos de investigación o aplicaciones donde la precisión de segmentación es crítica y los recursos no son limitantes.

Las APIs multimodales ofrecen capacidades impresionantes con mínima inversión en desarrollo, pero introducen dependencias externas y costes operativos que pueden ser problemáticos en aplicaciones de producción [OpenAI, 2024a, Google DeepMind, 2024]. Su utilidad principal radica en prototipado rápido, validación de conceptos y aplicaciones donde la dependencia externa es aceptable.

La selección de tecnologías para aplicaciones farmacéuticas debe considerar factores que van más allá de las capacidades técnicas puras. La fiabilidad operativa, los costes a largo plazo, la facilidad de mantenimiento y la escalabilidad son factores determinantes que

Tabla 2.1: Comparación crítica de tecnologías para procesamiento de cupones precinto.

Tecnología	Precisión	Velocidad	Complejidad	Coste
YOLO10/11	Alta	Muy Alta	Media	Bajo
Gradiente	Media	Alta	Baja	Muy Bajo
GPT-4 Vision	Alta	Baja	Baja	Variable
Gemini	Alta	Media	Baja	Limitado
SAM2	Muy Alta	Media	Muy Alta	Alto

pueden superar consideraciones de rendimiento técnico marginal [Drug Discovery Today, 2024].

El análisis del estado del arte revela varias oportunidades de innovación que permanecen inexploradas. La integración de técnicas de procesamiento de imágenes tradicionales con modelos de aprendizaje profundo puede ofrecer soluciones más fiables que cualquier enfoque individual [IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024]. El desarrollo de modelos específicamente entrenados para documentos farmacéuticos españoles puede mejorar significativamente el rendimiento comparado con modelos generales.

2.10. Conclusiones del estado del arte

El análisis exhaustivo del estado del arte revela un ecosistema tecnológico maduro con múltiples opciones viables para el procesamiento automático de cupones precinto farmacéuticos. Los modelos YOLO se posicionan como la opción más práctica para implementaciones comerciales, ofreciendo el mejor balance entre capacidades técnicas, facilidad de implementación y sostenibilidad operativa.

Las técnicas de procesamiento de imágenes tradicionales mantienen relevancia como componentes complementarios que pueden mejorar significativamente el rendimiento de sistemas basados en aprendizaje profundo. Las APIs multimodales, aunque impresionantes en capacidades, presentan limitaciones prácticas que restringen su aplicabilidad en entornos de producción farmacéutica.

La tendencia hacia sistemas híbridos que combinan múltiples tecnologías representa la dirección más prometedora para el desarrollo futuro, permitiendo aprovechar las fortalezas de diferentes enfoques mientras se mitigan sus limitaciones individuales. Esta aproximación híbrida es particularmente relevante en el contexto farmacéutico, donde la fiabilidad y precisión son requisitos no negociables.

El estado actual de la tecnología permite el desarrollo de soluciones prácticas y efectivas para el procesamiento automático de estos documentos, con potencial para generar ahorros significativos y mejorar la eficiencia operativa en el sector farmacéutico español. La clave del éxito radica en la selección cuidadosa de tecnologías apropiadas y su integración inteligente en sistemas que prioricen la usabilidad, fiabilidad y sostenibilidad operativa.

Capítulo 3

Sistema de gestión de cupones precinto

El sistema de gestión de cupones precinto constituye uno de los pilares fundamentales del control farmacéutico en España, representando el último eslabón analógico en la cadena de trazabilidad de medicamentos financiados por el SNS. Este capítulo proporciona una descripción exhaustiva del funcionamiento real del sistema, analizando desde el origen físico de los cupones hasta los procesos administrativos complejos que sustentan la facturación farmacéutica nacional.

La comprensión detallada de este sistema resulta fundamental para el desarrollo de soluciones de automatización efectivas, ya que cada elemento del cupón precinto tiene una función específica en el proceso de control y facturación que debe preservarse en cualquier implementación digital. La complejidad del sistema actual, desarrollado a lo largo de décadas, refleja las necesidades específicas del sector farmacéutico español y las exigencias normativas de la Agencia Española de Medicamentos y Productos Sanitarios (AEMPS) [Agencia Española de Medicamentos y Productos Sanitarios, 2024].

El análisis presentado se basa en la colaboración directa con Farmalitics y el estudio de documentación oficial de colegios farmacéuticos, proporcionando una visión práctica y actualizada de los desafíos reales que enfrentan las farmacias en su gestión diaria [Farmalitics, 2024]. Esta perspectiva práctica es esencial para desarrollar soluciones tecnológicas que no solo sean técnicamente viables, sino también adoptables en el contexto operativo real de las farmacias españolas.

3.1. Marco normativo y legal del sistema

El sistema de cupones precinto se rige por un marco normativo complejo que establece las bases legales para el control de medicamentos financiados. El Real Decreto 1348/2003 establece los fundamentos del sistema de clasificación de medicamentos, mientras que la normativa específica regula el uso de cupones precinto para medicamentos con aportación reducida [Boletín Oficial del Estado, 2003].

La normativa actual exige que todos los medicamentos financiados por el Sistema Nacional de Salud incorporen cupones precinto que permitan la identificación unívoca del producto dispensado [Ministerio de Sanidad, 2024]. Esta exigencia responde a la necesidad de garantizar el seguimiento completo desde la prescripción hasta la dispensación, cumpliendo con las directivas europeas sobre seguridad farmacéutica [Consejo General de Colegios Oficiales de Farmacéuticos, 2024].

La AEMPS actúa como organismo regulador central del sistema farmacéutico español, estableciendo las especificaciones técnicas y los protocolos de seguridad que deben cumplir todos los medicamentos financiados por el Sistema Nacional de Salud [Agencia Española de Medicamentos y Productos Sanitarios, 2024]. Su función abarca desde la autorización de comercialización hasta la supervisión del cumplimiento normativo en toda la cadena de distribución farmacéutica.

El marco legal también establece las responsabilidades de cada actor en la cadena: los laboratorios farmacéuticos deben generar cupones que cumplan especificaciones técnicas estrictas, las farmacias deben conservar y gestionar estos cupones siguiendo protocolos específicos, y los colegios farmacéuticos actúan como intermediarios en el proceso de facturación y control [Portal Farma, 2024].

3.2. Anatomía del cupón precinto farmacéutico

El cupón precinto constituye un documento técnico complejo que integra múltiples elementos de información en un formato estandarizado. Su diseño responde a décadas de evolución del sistema farmacéutico español, incorporando elementos de seguridad, información regulatoria y códigos de identificación que facilitan tanto el control manual como el procesamiento automatizado.

3.2.1. Origen del cupón: de la caja del medicamento

Los cupones precinto tienen su origen en las cajas de medicamentos, donde se adhieren como elementos de control y trazabilidad. Como se muestra en la Figura 3.1, cada envase farmacéutico incluye un cupón específico que debe ser separado durante el proceso de dispensación para justificar la entrega del medicamento al paciente [Observatorio de Tecnologías Sanitarias, 2024].



Figura 3.1: Elementos del envase farmacéutico [Agencia Española de Medicamentos y Productos Sanitarios, 2024].

3.2.2. Estructura y componentes del cupón individual

Una vez separado de la caja, el cupón precinto revela una estructura compleja que maximiza la densidad de información en un espacio reducido (véase Figura 3.2). La organización jerárquica de los elementos responde a décadas de evolución del sistema farmacéutico español, optimizando tanto la legibilidad humana como la capacidad de procesamiento automático.

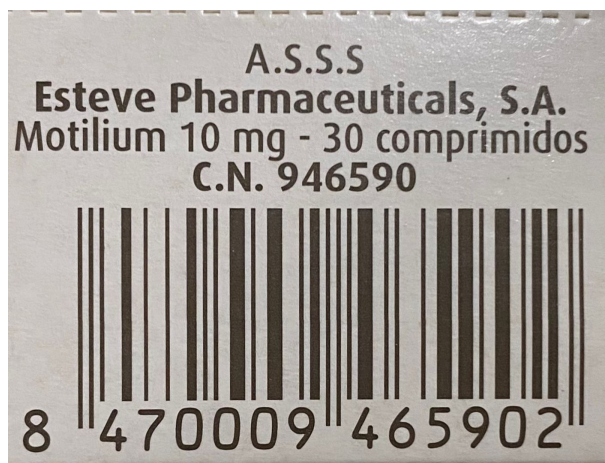


Figura 3.2: Ejemplo real de cupón precinto farmacéutico mostrando la estructura completa del documento con todas sus líneas de información textual, matriz de códigos de barras EAN-13 y elementos de seguridad anti-falsificación.

Línea 1: ASSS (Área Sanitaria del Sistema de Salud)

La primera línea del cupón contiene el código ASSS, que identifica el área geográfica de dispensación para efectos de facturación autonómica [Consejo General de Colegios Oficiales de Farmacéuticos, 2024].

Línea 2: Laboratorio fabricante

La segunda línea especifica el laboratorio fabricante mediante su denominación comercial y código asignado por la AEMPS [Farmaindustria, 2024]. Esta información es crucial para el seguimiento en caso de alertas sanitarias, retiradas de medicamentos o investigaciones de farmacovigilancia.

Línea 3: Nombre del medicamento

La tercera línea contiene la denominación del medicamento, que puede incluir tanto el nombre comercial como el principio activo según la denominación común internacional [Agencia Española de Medicamentos y Productos Sanitarios, 2024]. Esta información textual proporciona verificación legible que facilita la confirmación manual de la correspondencia entre la prescripción médica y el medicamento dispensado.

La presentación de esta información sigue estándares específicos que incluyen la concentración del principio activo, la forma farmacéutica y el número de unidades por envase. Esta estandarización facilita la identificación inequívoca del medicamento y reduce fallos de dispensación.

Línea 4: CN (Código Nacional)

La cuarta línea presenta el Código Nacional (CN) del medicamento, un identificador único asignado por la AEMPS durante el proceso de registro sanitario [Agencia Española de Medicamentos y Productos Sanitarios, 2024]. Este código de seis dígitos identifica unívocamente el medicamento en el mercado español y constituye la base para la generación del código EAN-13 correspondiente.

El CN es fundamental para la integración con bases de datos farmacéuticas como CIMA, permitiendo la verificación automática de información del medicamento, precios de referencia, condiciones de financiación y alertas de seguridad. Los sistemas de gestión farmacéutica utilizan este código para realizar validaciones automáticas que garantizan la coherencia de la dispensación.

Códigos EAN-13: estructura y función

Los códigos de barras EAN-13 constituyen el elemento central del sistema de seguimiento farmacéutico. Cada código sigue la estructura específica 847-000-XXXXXX-X, donde cada segmento tiene una función definida en el sistema de control farmacéutico español.

Los primeros tres dígitos (847) identifican España como país de origen según la numeración internacional GS1, facilitando el control aduanero y el seguimiento internacional de medicamentos. Los dígitos 4-6 son siempre "000" en el mercado farmacéutico español, estableciendo un patrón fijo que permite la identificación automática de códigos farmacéuticos.

Los dígitos 7-12 corresponden directamente al Código Nacional (CN) del medicamento, creando una correspondencia directa entre la identificación textual y la codificación de barras. Esta correspondencia facilita la verificación y reduce fallos de interpretación durante el procesamiento automático.

El dígito 13 funciona como dígito de control, calculado mediante el algoritmo estándar EAN que utiliza una suma ponderada de los 12 dígitos anteriores. Este mecanismo de verificación permite detectar automáticamente incidencias de lectura o transcripción, proporcionando una primera línea de validación de la integridad del código.

3.2.3. Organización en hojas de cupones precinto

Los cupones individuales se organizan inicialmente en hojas estructuradas que optimizan tanto la producción masiva por parte de los laboratorios como el posterior procesamiento en las farmacias. La Figura 3.3 ilustra esta organización matricial que responde a criterios técnicos específicos que equilibran la densidad de información con la facilidad de manipulación.

Cara delantera: 6 filas × 4 columnas

La cara delantera de las hojas utiliza una organización estándar de 6 filas por 4 columnas, proporcionando 24 cupones por hoja [Observatorio del Medicamento, 2024]. Esta configuración incluye espacio para la cabecera con información de la farmacia: número SOE, día, hora y número de orden.



Figura 3.3: Hojas de cupones precinto mostrando la organización matricial antes del recorte individual. La imagen izquierda muestra la cara delantera con organización de 6 filas por 4 columnas (24 cupones). La imagen derecha presenta la cara trasera con organización de 7 filas por 4 columnas que permite 26 códigos útiles.

Cara trasera: 7 filas × 4 columnas

La cara trasera utiliza una configuración de 7 filas por 4 columnas, proporcionando 26 códigos útiles [Federación Empresarial de Farmacéuticos Españoles, 2024]. Al no requerir espacio para cabecera, se maximiza el aprovechamiento del papel para cupones adicionales.

3.3. Proceso administrativo completo

El proceso de gestión de cupones precinto abarca desde la prescripción médica hasta la facturación final, involucrando múltiples actores y sistemas que deben coordinarse para garantizar el seguimiento y el control económico del gasto farmacéutico [Consejo General de Colegios Oficiales de Farmacéuticos, 2024].

3.3.1. Flujo desde prescripción hasta facturación

El proceso comienza con la prescripción médica, donde el médico especifica el medicamento mediante código nacional o denominación común internacional. La receta electrónica incluye información que debe coincidir con los datos del cupón precinto, estableciendo la primera línea de validación entre la prescripción digital y el documento físico.

En la farmacia, el farmacéutico verifica la correspondencia entre la prescripción y el medicamento dispensado, utilizando tanto la información textual del cupón como los códigos de barras para confirmar la identidad del producto. Este proceso de verificación

es crítico para la seguridad del paciente y el cumplimiento normativo, requiriendo la confirmación de múltiples elementos informativos.

La dispensación genera un registro electrónico que incluye los códigos de barras leídos del cupón precinto, estableciendo el seguimiento entre la prescripción, el medicamento dispensado y el paciente [Observatorio de Tecnologías Sanitarias, 2024]. Este registro se transmite automáticamente a los sistemas de facturación del colegio farmacéutico correspondiente, iniciando el proceso de reembolso.

El cupón físico se conserva en la farmacia siguiendo protocolos específicos de archivo que facilitan auditorías posteriores y verificaciones de calidad. Los cupones se organizan cronológicamente y se conservan durante períodos establecidos por la normativa vigente, permitiendo la reconstrucción completa del historial de dispensaciones.

3.3.2. Interacción con colegios farmacéuticos

Los colegios farmacéuticos actúan como intermediarios entre las farmacias individuales y los organismos pagadores, gestionando la validación de cupones, la agregación de datos de facturación y la resolución de incidencias [Consejo General de Colegios Oficiales de Farmacéuticos, 2024]. Esta intermediación es fundamental para el funcionamiento del sistema descentralizado español.

El proceso de validación colegial incluye verificaciones automáticas de coherencia entre los códigos de barras reportados y las bases de datos de medicamentos autorizados. Los sistemas colegiales detectan automáticamente discrepancias que requieren investigación adicional o corrección manual, garantizando la integridad del proceso de facturación.

La agregación de datos permite generar estadísticas de consumo farmacéutico que facilitan la planificación sanitaria y la detección de patrones anómalos de prescripción o dispensación [Centro Nacional de Epidemiología, 2024]. Estos análisis contribuyen al uso racional de medicamentos y la optimización del gasto sanitario a nivel autonómico y nacional.

3.3.3. Sistemas informáticos de gestión

Las farmacias utilizan sistemas informáticos especializados que integran la gestión de cupones precinto con otros procesos farmacéuticos como control de inventarios, atención al paciente y facturación [Farmalitics, 2024]. Estos sistemas incluyen funcionalidades específicas para la lectura automática de códigos de barras y la validación de cupones.

Los sistemas de gestión farmacéutica se integran con bases de datos centrales como CIMA para verificar automáticamente la validez de los códigos detectados y obtener información adicional sobre medicamentos. Esta integración mejora la precisión del proceso y reduce incidencias de dispensación mediante validaciones automáticas.

La conectividad con sistemas de receta electrónica permite verificaciones bidireccionales que comprueban la correspondencia entre la prescripción digital y los cupones físicos procesados. Esta validación mejora la seguridad del proceso y reduce la carga administrativa manual, liberando tiempo del personal farmacéutico para actividades de mayor valor clínico.

3.4. Impacto económico y desafíos del sistema actual

El sistema manual de gestión de cupones precinto genera costes significativos que afectan tanto a farmacias individuales como al conjunto del sistema sanitario español. El

análisis de estos costes es fundamental para justificar las inversiones en automatización y evaluar el retorno esperado de las soluciones tecnológicas.

3.4.1. Costes del sistema manual

Las estimaciones del sector indican que el procesamiento manual de cupones precinto consume aproximadamente el 15 % del tiempo administrativo total en farmacias comunitarias [Observatorio del Medicamento, 2024]. Para una farmacia promedio que dispensa 300 recetas diarias, esto representa entre 45-60 minutos diarios dedicados exclusivamente a la gestión de cupones.

El coste laboral asociado, considerando el salario promedio del personal farmacéutico, se estima en 2.500-3.500 euros anuales por farmacia [Federación Empresarial de Farmacéuticos Españoles, 2024]. Extrapolando a las aproximadamente 22.000 farmacias españolas, el coste total del sistema manual alcanza entre 55-77 millones de euros anuales solo en costes laborales directos.

Los costes indirectos incluyen incidencias de procesamiento que requieren corrección manual, pérdida de cupones que necesitan reposición, y tiempo adicional dedicado a la resolución de problemas con colegios farmacéuticos. Estos costes indirectos pueden representar un 20-30 % adicional sobre los costes directos.

3.4.2. Tiempo dedicado por farmacéuticos

El análisis detallado del tiempo dedicado revela que el procesamiento de cupones incluye múltiples actividades: separación física de cupones de las recetas, verificación visual de correspondencia con medicamentos dispensados, archivo cronológico para facilitar auditorías posteriores, y gestión de cupones problemáticos o ilegibles.

El tiempo promedio por cupón oscila entre 15-25 segundos, dependiendo de la experiencia del personal y la calidad del cupón [Observatorio del Medicamento, 2024]. Los cupones con problemas de legibilidad o formato pueden requerir hasta 2-3 minutos adicionales para su procesamiento o resolución.

La variabilidad estacional afecta significativamente la carga de trabajo, con picos durante períodos de mayor consumo farmacéutico que pueden duplicar el tiempo dedicado a gestión de cupones durante semanas específicas. Esta variabilidad complica la planificación de recursos y puede generar cuellos de botella operativos.

3.4.3. Fallos y su impacto

Las incidencias en el procesamiento manual de cupones generan múltiples tipos de impacto económico y operativo [Grupo de Procesamiento de Imágenes, 2024]. Los problemas de archivo pueden resultar en pérdida de cupones que requieren procesos de reposición complejos y costosos que involucran a laboratorios, colegios farmacéuticos y organismos pagadores.

Las discrepancias de correspondencia entre cupones y medicamentos dispensados pueden generar inconsistencias en la facturación que requieren investigación y corrección manual. Estas incidencias pueden detectarse semanas después de la dispensación, complicando la resolución y aumentando los costes administrativos.

La pérdida de seguimiento debido a fallos de procesamiento puede tener implicaciones graves en caso de alertas sanitarias o retiradas de medicamentos, requiriendo procesos

manuales de investigación que consumen recursos significativos y pueden comprometer la seguridad del paciente.

3.4.4. Casos de digitalización exitosa

La experiencia pionera de la Comunidad Valenciana, que implementó en 2024 un sistema de código digital eliminando completamente los cupones físicos, ha demostrado la viabilidad económica de la automatización con ahorros documentados de 4,7 millones de euros anuales [Conselleria de Sanidad, 2024]. Esta implementación, que requirió una inversión inicial de 2,1 millones de euros y logró un retorno en aproximadamente 5 meses, establece un precedente para el resto de España donde el coste del sistema manual se estima entre 80-90 millones de euros anuales.

Otras comunidades autónomas han desarrollado iniciativas complementarias que confirman la tendencia irreversible hacia la digitalización completa del sector farmacéutico español, implementando soluciones basadas en tecnologías emergentes que demuestran la viabilidad técnica y económica de la automatización.

3.5. Perspectivas de futuro y tendencias

La evolución del sistema de cupones precinto hacia soluciones completamente digitales representa una tendencia irreversible que se acelerará en los próximos años. Las tecnologías emergentes ofrecen oportunidades para mejoras adicionales en eficiencia, seguridad y seguimiento que van más allá de la simple automatización del proceso actual.

3.5.1. Tendencias hacia la digitalización completa

La integración de tecnologías de IA y machine learning permitirá sistemas adaptativos que mejoren automáticamente su rendimiento mediante el análisis de patrones de uso y detección de anomalías [Sociedad Española de Farmacia Hospitalaria, 2024]. Estos sistemas podrán identificar automáticamente medicamentos falsificados, detectar patrones de prescripción anómalos, y optimizar la gestión de inventarios farmacéuticos.

La implementación de tecnología blockchain para el seguimiento de medicamentos proporcionará registros inmutables que mejoren la seguridad y transparencia del proceso. Esta tecnología facilitará la detección de medicamentos falsificados y mejorará la respuesta ante alertas sanitarias o retiradas de medicamentos.

3.5.2. Integración con sistemas de salud digital

La futura digitalización permitirá conectar la gestión de cupones con otros sistemas sanitarios como historiales clínicos electrónicos y receta electrónica [Ministerio de Asuntos Económicos y Transformación Digital, 2024]. Esta conexión proporcionará un seguimiento completo del tratamiento del paciente desde la prescripción hasta la dispensación.

Los sistemas automatizados facilitarán la detección de problemas como incumplimiento terapéutico o interacciones medicamentosas, mejorando la seguridad y eficacia del tratamiento farmacológico.

3.5.3. Impacto en la práctica farmacéutica

La automatización completa del procesamiento de cupones liberará tiempo del personal farmacéutico para actividades de mayor valor clínico, como servicios de atención farmacéutica, seguimiento farmacoterapéutico, y educación sanitaria de pacientes [Asociación de Distribuidores Farmacéuticos, 2024].

Los sistemas automatizados proporcionarán datos en tiempo real sobre consumo farmacéutico que facilitarán la gestión proactiva de inventarios, la detección temprana de desabastecimientos, y la optimización de la cadena de suministro farmacéutica.

La digitalización del sistema de cupones precinto representa una oportunidad de mejora significativa para el sector farmacéutico español. El conocimiento profundo del sistema actual es esencial para desarrollar soluciones tecnológicas que no solo automaticen los procesos existentes, sino que también introduzcan mejoras en eficiencia, precisión y calidad asistencial mediante el aprovechamiento de tecnologías emergentes.

Capítulo 4

Metodología y planificación del proyecto

El desarrollo de un sistema automático para el procesamiento de cupones precinto farmacéuticos requiere una metodología estructurada que garantice tanto la calidad técnica como la viabilidad práctica de la solución. Este capítulo presenta la estrategia metodológica adoptada, basada en la adaptación de CRISP-DM al contexto específico del proyecto, junto con la planificación temporal y las herramientas empleadas durante el desarrollo.

La elección metodológica se fundamenta en la naturaleza experimental del proyecto, que combina investigación aplicada con desarrollo tecnológico práctico. La necesidad de evaluar múltiples enfoques tecnológicos de forma comparativa requiere un marco metodológico flexible que permita iteraciones rápidas y evaluación continua de resultados [CRISP-DM Consortium, 2024].

La colaboración con Farmalitics como empresa del sector proporciona acceso a datos reales y requisitos auténticos, condicionando la metodología hacia un enfoque pragmático que priorice la aplicabilidad práctica de los resultados obtenidos [Farmalitics, 2024]. Esta orientación práctica se refleja en todas las fases del proyecto, desde la comprensión del negocio hasta la evaluación final de resultados.

4.1. Adaptación de CRISP-DM al contexto del proyecto

La metodología CRISP-DM (Cross Industry Standard Process for Data Mining) proporciona un marco estructurado para proyectos de ciencia de datos que organiza el trabajo en seis fases interconectadas e iterativas. Esta metodología facilita la gestión sistemática de proyectos complejos de análisis de datos, desde la comprensión inicial del problema de negocio hasta el despliegue efectivo de soluciones en entornos reales. Su naturaleza cíclica permite la mejora continua mediante iteraciones que refinan tanto la comprensión del problema como la calidad de las soluciones desarrolladas.

Como se observa en la Figura 4.1, la adaptación de CRISP-DM a este proyecto de procesamiento automático de cupones precinto farmacéuticos mantiene la estructura fundamental de la metodología mientras se especializa en las particularidades del contexto farmacéutico español. La metodología se adapta perfectamente a las características de este trabajo debido a su capacidad para gestionar la evaluación continua de múltiples enfoques tecnológicos, aspecto fundamental en un proyecto de investigación comparativa.

La adaptación realizada enfatiza especialmente la fase de modelado, donde se implementan y evalúan cuatro enfoques diferentes: modelos YOLO, detectores de gradientes, SAM2 e interfaces de IA multimodal. Esta multiplicidad de enfoques requiere una planificación cuidadosa de experimentos y una evaluación sistemática que permita comparaciones objetivas entre métodos, beneficiándose directamente del marco estructurado que proporciona CRISP-DM.



Figura 4.1: Metodología CRISP-DM adaptada al proyecto de procesamiento automático de cupones precinto farmacéuticos.

4.1.1. Fase 1: Comprensión del negocio

La comprensión del contexto farmacéutico español constituye el fundamento del proyecto. Esta fase incluye el análisis del sistema actual de gestión de cupones precinto, la identificación de stakeholders clave y la definición de objetivos específicos alineados con las necesidades reales del sector.

La colaboración con Farmalitics facilita el acceso directo a profesionales del sector, permitiendo una comprensión profunda de los flujos de trabajo actuales, los puntos de dolor específicos y las restricciones operativas que debe considerar cualquier solución tecnológica. Esta comprensión directa del contexto operativo es fundamental para el éxito del proyecto.

Los objetivos de negocio se centran en la automatización del procesamiento manual, la reducción de errores humanos y la liberación de tiempo del personal farmacéutico para actividades de mayor valor clínico. Estos objetivos se traducen en métricas específicas de rendimiento que guían el desarrollo y evaluación de las soluciones propuestas.

4.1.2. Fase 2: Comprensión de los datos

La fase de comprensión de datos abarca tanto el análisis del conjunto de imágenes proporcionado por Farmalitics como la exploración de las características específicas de los cupones precinto farmacéuticos. El conjunto de datos de 31 imágenes, aunque limitado en tamaño, proporciona una muestra representativa de la variabilidad encontrada en entornos reales.

El análisis exploratorio revela desafíos específicos como variaciones en iluminación, distorsiones de perspectiva, calidad variable de impresión y fragmentación de códigos de barras. Estos desafíos informan directamente las decisiones de preprocesamiento y la selección de enfoques tecnológicos apropiados.

La documentación detallada de las características del conjunto de datos incluye análisis estadísticos de distribución de tamaños, variabilidad en condiciones de captura y categorización de tipos de problemas encontrados. Esta documentación sirve como base para la evaluación posterior de la generalización de los resultados obtenidos.

4.1.3. Fase 3: Preparación de los datos

La preparación de datos incluye tanto el etiquetado manual con LabelMe como el desarrollo de técnicas de preprocesamiento específicas para cupones precinto. El proceso de etiquetado sigue protocolos estrictos que garantizan la consistencia y calidad del ground truth utilizado para entrenamiento y evaluación.

Las técnicas de preprocesamiento desarrolladas abordan específicamente los desafíos identificados durante la fase de comprensión. Estas incluyen corrección de perspectiva, normalización de iluminación, mejora de contraste y filtrado de ruido. La efectividad de estas técnicas se evalúa mediante métricas objetivas de calidad de imagen.

La estrategia de división del conjunto de datos considera tanto la limitación en el número de imágenes como la necesidad de mantener representatividad en los conjuntos de entrenamiento, validación y prueba. Esta división evolutiva se ajusta según los resultados obtenidos en iteraciones sucesivas del proyecto.

4.1.4. Fase 4: Modelado

La fase de modelado constituye el núcleo técnico del proyecto, donde se implementan y evalúan los cuatro enfoques seleccionados. Cada enfoque se desarrolla de forma independiente pero siguiendo protocolos de evaluación consistentes que permiten comparaciones objetivas.

Los modelos YOLO incluyen tanto YOLO10 como YOLO11, con fine-tuning específico para la detección de códigos de barras en cupones precinto. Los experimentos incluyen variaciones en arquitectura, parámetros de entrenamiento y técnicas de aumento de datos.

Los detectores de gradientes se desarrollan basándose en análisis de gradientes direccionales que aprovechan las características específicas de los códigos de barras. Estos detectores se optimizan para detectar patrones verticales característicos de códigos EAN-13.

La adaptación del modelo de segmentación universal SAM2 explora su capacidad para segmentar códigos de barras mediante prompts visuales. Los experimentos evalúan diferentes estrategias de prompting y post-procesamiento de máscaras de segmentación.

La integración con servicios como GPT-4 Vision y Gemini explora las capacidades de análisis visual de modelos de IA generalistas. Los experimentos incluyen optimización de prompts y estrategias de post-procesamiento de respuestas.

4.1.5. Fase 5: Evaluación

La evaluación comparativa utiliza métricas objetivas que consideran tanto la precisión de detección como la viabilidad práctica de implementación. Las métricas incluyen precisión, recall, tiempo de procesamiento, estabilidad ante variaciones y facilidad de implementación.

La evaluación considera específicamente los requisitos del contexto farmacéutico, donde la precisión es crítica pero debe equilibrarse con consideraciones prácticas como velocidad de procesamiento, costes operativos y facilidad de mantenimiento. Esta evaluación multidimensional proporciona una base sólida para recomendaciones de implementación.

Los protocolos de evaluación incluyen tanto pruebas automatizadas como validación manual por expertos del sector. Esta validación cruzada garantiza que los resultados técnicos se traduzcan en beneficios prácticos reales para los usuarios finales.

4.1.6. Fase 6: Despliegue

La fase de despliegue incluye el desarrollo de una interfaz web funcional que demuestra la viabilidad práctica de la solución seleccionada. Esta interfaz, desarrollada con Streamlit, proporciona una experiencia de usuario intuitiva que facilita la adopción por parte de profesionales farmacéuticos.

El prototipo de despliegue incluye funcionalidades completas de carga de imágenes, procesamiento automático, visualización de resultados y validación con la API de CIMA. Esta implementación completa demuestra la viabilidad técnica y operativa de la solución propuesta.

La documentación de despliegue incluye guías de instalación, configuración y mantenimiento que facilitan la adopción por parte de organizaciones del sector. Esta documentación considera tanto aspectos técnicos como operativos, proporcionando una base sólida para implementaciones futuras.

4.2. Planificación temporal del proyecto

La planificación temporal del proyecto sigue un enfoque iterativo que permite ajustes basados en resultados intermedios y descubrimientos durante el desarrollo. Esta flexibilidad es fundamental en un proyecto de investigación aplicada donde los resultados de cada fase informan las decisiones de las fases posteriores.

4.2.1. Cronograma general

El proyecto se estructura en cuatro fases principales distribuidas a lo largo de 16 semanas, con solapamiento entre fases para maximizar la eficiencia y permitir iteraciones rápidas basadas en resultados intermedios.

La primera fase abarca las semanas 1 a 4 y se centra en la investigación y comprensión del problema, incluyendo el análisis del estado del arte en visión por computador aplicada a documentos farmacéuticos y la comprensión profunda del sistema de cupones precinto. Durante este período se realiza también el análisis exploratorio del conjunto de datos proporcionado por Farmalitics y se definen las métricas de evaluación y protocolos experimentales que guiarán el resto del proyecto.

La segunda fase se desarrolla entre las semanas 5 y 8, enfocándose en la preparación de datos y desarrollo inicial de las soluciones. Esta etapa incluye el etiquetado completo

del conjunto de datos utilizando LabelMe, el desarrollo de técnicas de preprocesamiento específicas para cupones precinto, la implementación inicial de los cuatro enfoques tecnológicos seleccionados y la validación de la calidad del ground truth generado.

Las semanas 9 a 12 constituyen la tercera fase, dedicada a la implementación completa y experimentación exhaustiva. Durante este período se desarrollan completamente todos los enfoques tecnológicos, se realiza experimentación exhaustiva con diferentes configuraciones y parámetros, se lleva a cabo la evaluación comparativa inicial con refinamiento de métodos y se desarrolla la interfaz web de demostración.

La fase final, que abarca las semanas 13 a 16, se concentra en la evaluación final y documentación del proyecto. Esta etapa incluye la evaluación comparativa completa con métricas objetivas, la validación con expertos del sector farmacéutico, el desarrollo de recomendaciones de implementación y la redacción de la memoria final junto con la preparación de la presentación.

4.2.2. Hitos principales

Los hitos principales del proyecto proporcionan puntos de control que permiten evaluar el progreso y realizar ajustes en la planificación según sea necesario.

El primer hito se alcanza en la semana 4 con la comprensión completa del problema, momento en el cual se debe tener documentado y analizado el estado del arte, claramente definidos los requisitos del sistema, completamente caracterizado el conjunto de datos y establecidas las métricas de evaluación que guiarán el desarrollo posterior.

El segundo hito importante se sitúa en la semana 8, cuando se completa la preparación de datos y la implementación inicial de métodos. En este punto se debe contar con un ground truth de calidad validado, técnicas de preprocesamiento optimizadas para las características específicas de los cupones precinto, implementación inicial funcional de los cuatro enfoques seleccionados y protocolos experimentales completamente validados.

La semana 12 marca el tercer hito con la finalización de la evaluación comparativa inicial. Para este momento se requiere tener todos los enfoques completamente desarrollados y optimizados, experimentación exhaustiva realizada con documentación completa de resultados, resultados comparativos documentados y analizados, y una interfaz de demostración completamente funcional que permita la validación práctica de las soluciones.

El hito final se alcanza en la semana 16 con la completación del proyecto

4.2.3. Gestión de riesgos

La planificación incluye identificación y mitigación de riesgos específicos que podrían afectar el desarrollo del proyecto. Esta gestión proactiva de riesgos es fundamental para garantizar la finalización exitosa dentro de los plazos establecidos.

Los riesgos técnicos principales incluyen las limitaciones del conjunto de datos que podrían afectar la generalización de resultados, dificultades técnicas en la implementación de alguno de los enfoques seleccionados, y problemas de rendimiento que podrían impedir la evaluación comparativa efectiva entre métodos.

Para mitigar esos riesgos técnicos se implementan estrategias específicas como el desarrollo de técnicas de aumento de datos para ampliar el conjunto de datos efectivo y mejorar la capacidad de generalización, la implementación de múltiples enfoques que proporcionen alternativas sólidas en caso de que alguno presente dificultades insuperables,

y la optimización continua del rendimiento durante todo el desarrollo para evitar cuellos de botella que comprometan la evaluación.

Los riesgos de recursos constituyen otra categoría importante e incluyen limitaciones computacionales que podrían afectar el entrenamiento de modelos de aprendizaje profundo, disponibilidad limitada de expertos del sector para validación de resultados, y restricciones de tiempo que podrían impedir la evaluación exhaustiva de todos los enfoques.

Las estrategias de mitigación para estos riesgos incluyen la utilización de recursos computacionales en la nube cuando sea necesario para garantizar capacidad de procesamiento adecuada, coordinación temprana con Farmalitics para garantizar disponibilidad de expertos del sector en momentos críticos del proyecto, y planificación flexible que permita priorizar actividades críticas en caso de restricciones temporales.

4.3. Herramientas y tecnologías empleadas

La selección de herramientas y tecnologías se basa en criterios de eficiencia, compatibilidad y facilidad de uso, priorizando soluciones que faciliten tanto el desarrollo como la reproducibilidad de resultados. La combinación de herramientas seleccionadas proporciona un entorno de desarrollo completo y eficiente.

4.3.1. Entorno de desarrollo

Python 3.9+: Lenguaje principal de desarrollo, seleccionado por su amplio ecosistema de bibliotecas para visión por computador y aprendizaje profundo. La versión 3.9+ garantiza compatibilidad con las bibliotecas más recientes y características de rendimiento optimizadas.

Jupyter Notebooks: Entorno de desarrollo interactivo que facilita la experimentación rápida y la documentación del proceso de desarrollo. Los notebooks permiten combinar código, visualizaciones y documentación en un formato que facilita la reproducibilidad y el análisis de resultados.

Visual Studio Code: Editor principal para desarrollo de código de producción, con extensiones específicas para Python, Git y desarrollo web. La integración con herramientas de debugging y control de versiones facilita el desarrollo colaborativo y la gestión de código.

Git y GitHub: Sistema de control de versiones que garantiza la trazabilidad del desarrollo y facilita la colaboración. La utilización de GitHub proporciona backup automático y facilita la documentación del proceso de desarrollo.

4.3.2. Bibliotecas de visión por computador

OpenCV 4.8+: Biblioteca fundamental para operaciones básicas de procesamiento de imágenes, incluyendo lectura, escritura, transformaciones geométricas y filtrado. Las funcionalidades de OpenCV proporcionan la base para todas las operaciones de preprocesamiento.

Ultralytics: Implementación optimizada de modelos YOLO que facilita tanto el entrenamiento como la inferencia. La biblioteca incluye herramientas de visualización y evaluación que aceleran el desarrollo y análisis de resultados.

PyZbar: Biblioteca especializada en decodificación de códigos de barras que proporciona funcionalidades complementarias a los métodos de visión por computador. Su

integración permite validación cruzada de resultados y implementación de sistemas de fallback.

NumPy y SciPy: Bibliotecas fundamentales para operaciones numéricas y científicas que sustentan todos los cálculos de procesamiento de imágenes y análisis de resultados.

4.3.3. Herramientas de etiquetado y evaluación

LabelMe: Herramienta de etiquetado manual que facilita la creación de ground truth de alta calidad. Su interfaz intuitiva y formatos de exportación estándar facilitan la integración con pipeline de entrenamiento.

Matplotlib y Seaborn: Bibliotecas de visualización que facilitan el análisis exploratorio de datos y la presentación de resultados. Las visualizaciones generadas son fundamentales para la comprensión y comunicación de resultados.

Pandas: Biblioteca para manipulación y análisis de datos que facilita la gestión de metadatos, resultados de experimentos y métricas de evaluación. Su integración con otras herramientas de análisis acelera el procesamiento de resultados.

4.3.4. Desarrollo de interfaz web

Streamlit: framework para desarrollo rápido de aplicaciones web interactivas que facilita la demostración de resultados y la validación con usuarios finales. Su simplicidad permite enfocar el desarrollo en la funcionalidad core sin complejidades de desarrollo web tradicional.

Pillow (PIL): Biblioteca para manipulación de imágenes en el contexto web, facilitando la carga, procesamiento y visualización de imágenes en la interfaz de usuario.

Requests: Biblioteca para integración con API externas, específicamente para la validación de códigos de barras con la API de CIMA. Su simplicidad facilita la implementación de validaciones automáticas.

4.4. Metodología de evaluación comparativa

La evaluación comparativa de los cuatro enfoques tecnológicos requiere una metodología rigurosa que garantice la objetividad y reproducibilidad de los resultados. Esta metodología considera tanto métricas técnicas como criterios prácticos de implementación.

4.4.1. Métricas de evaluación técnica

Las métricas de evaluación se basan en la clasificación estándar de resultados de detección:

- **True Positives (TP):** Códigos de barras correctamente detectados y presentes en la imagen
- **False Positives (FP):** Detecciones incorrectas (códigos detectados que no existen)
- **False Negatives (FN):** Códigos de barras presentes pero no detectados
- **True Negatives (TN):** Regiones correctamente identificadas como no conteniendo códigos

Precisión (Precision): Mide la proporción de detecciones correctas respecto al total de detecciones realizadas:

Esta métrica evalúa la calidad de las detecciones, penalizando las falsas alarmas. Una precisión alta indica que la mayoría de las detecciones son correctas.

Recall (Sensibilidad): Mide la proporción de códigos reales que fueron detectados correctamente:

Esta métrica evalúa la capacidad del sistema para encontrar todos los códigos presentes, siendo crítica en aplicaciones donde la pérdida de información tiene consecuencias importantes.

F1-Score: Combina precisión y recall en una única métrica mediante su media armónica:

El F1-Score proporciona una medida equilibrada del rendimiento, especialmente útil cuando existe un trade-off entre precisión y recall.

Precisión de decodificación: Porcentaje de códigos correctamente decodificados respecto al total de códigos detectados. Esta métrica evalúa la calidad de la lectura más allá de la simple localización.

Tiempo de procesamiento: Tiempo promedio requerido para procesar una imagen completa, incluyendo preprocesamiento, detección y post-procesamiento. Esta métrica es fundamental para evaluar la viabilidad práctica de cada enfoque.

4.4.2. Métricas de evaluación práctica

Estabilidad ante variaciones: Capacidad de mantener rendimiento consistente ante variaciones en iluminación, perspectiva y calidad de imagen. Esta evaluación utiliza subconjuntos específicos del conjunto de datos que representan condiciones desafiantes.

Facilidad de implementación: Evaluación cualitativa de la complejidad de implementación, incluyendo requisitos de infraestructura, dependencias externas y complejidad de configuración.

Costes operativos: Análisis de costes asociados con la implementación y operación de cada enfoque, incluyendo costes computacionales, licencias y mantenimiento.

Escalabilidad: Capacidad de manejar volúmenes crecientes de procesamiento sin degradación significativa del rendimiento.

4.4.3. Protocolos experimentales

Los experimentos siguen protocolos estandarizados que garantizan la comparabilidad de resultados entre diferentes enfoques. Estos protocolos incluyen procedimientos específicos para la preparación de datos, configuración de experimentos y recopilación de resultados.

Validación cruzada: Utilización de múltiples divisiones del conjunto de datos para garantizar la solidez de los resultados y evaluar la variabilidad del rendimiento.

Repetición de experimentos: Cada configuración se evalúa múltiples veces para caracterizar la variabilidad de los resultados y proporcionar intervalos de confianza estadísticamente válidos.

Condiciones controladas: Todos los experimentos se realizan en condiciones controladas que permiten aislar el efecto de las variables de interés y minimizar factores de confusión.

Capítulo 5

Conjunto de datos de cupones precinto

El desarrollo de un sistema eficaz de procesamiento automático de cupones precinto requiere un conjunto de datos que represente bien la variabilidad real de estos documentos en entornos farmacéuticos. Este capítulo describe el proceso de creación, organización y validación del conjunto de datos utilizado en el proyecto, detallando las decisiones metodológicas adoptadas para garantizar la calidad y que los datos de entrenamiento y evaluación sean representativos.

El conjunto de datos se basa en 31 imágenes reales de cupones precinto proporcionadas por la empresa colaboradora Farmalitics, las cuales han sido procesadas mediante técnicas de etiquetado manual para crear un conjunto de datos adecuado para el entrenamiento de modelos de detección de objetos. La selección de utilizar exclusivamente datos reales, en lugar de generar imágenes sintéticas, garantiza que los modelos desarrollados sean capaces de manejar las condiciones y variabilidades presentes en entornos farmacéuticos reales.

5.1. Características generales del conjunto de datos

Las 31 imágenes que constituyen el conjunto representan una muestra diversa de cupones precinto reales utilizados en farmacias españolas. El análisis cuantitativo del conjunto revela una estructura significativamente desbalanceada que refleja la naturaleza real de los cupones precinto, con un total de 3,070 elementos etiquetados distribuidos entre las tres clases principales de detección.

5.1.1. Composición y estructura de las imágenes

Cada imagen representa una **hoja completa** (DIN-A4) de cupones precinto. Dentro de cada hoja existe:

- **Una única zona barcode** que delimita el área completa donde se encuentran todos los cupones de la hoja.
- Entre **24 y 26 cupones individuales**, cada uno con su código EAN-13 correspondiente (**codes**).
- **Headers opcionales** que contienen información textual del medicamento (no todas las hojas incluyen header).

La estructura real es: *hoja* $\rightarrow$ *1 barcode* $\rightarrow$ *24-26 codes (+ headers opcionales)*.

El conjunto de datos contiene 124 barcodes para 31 hojas debido a técnicas de aumento de datos aplicadas durante el procesamiento, lo que explica que el número total de elementos sea superior al esperado con la estructura base ($31 \text{ hojas} \times 1 \text{ barcode} = 31 \text{ barcodes teóricos}$). De manera similar, los 2,879 codes y 67 headers reflejan tanto la estructura original como las variaciones generadas por el aumento de datos.

Esta organización explica por qué los **codes** constituyen la clase mayoritaria (representando los cupones individuales), mientras que los **headers** son la clase minoritaria (ya que no todos los cupones incluyen información textual claramente delimitada).

5.1.2. Análisis de elementos detectables

Cada hoja contiene tres tipos principales de elementos que requieren detección automática, cuya distribución y características se han analizado en detalle para comprender la complejidad del problema de detección.

Como se observa en la Figura 5.1, las regiones denominadas **code** constituyen la clase mayoritaria con 2,879 elementos identificados, representando el 93.8 % del total.

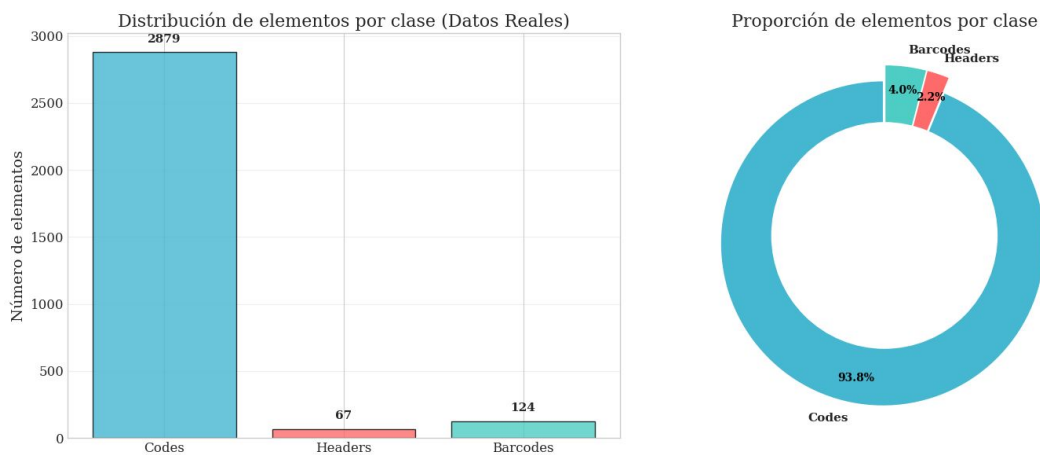


Figura 5.1: Distribución de elementos por clase en el conjunto completo de datos. La relación entre codes (2879), headers (67) y barcodes (124) refleja la estructura típica de los cupones precinto procesados.

Esta alta proporción mostrada en la Figura 5.1 se debe a que cada hoja contiene múltiples cupones precintos, siendo estos los elementos más numerosos y críticos para el proceso de facturación farmacéutica. Las regiones **barcode** totalizan 124 elementos, representando el 4.0 % del conjunto. Estas áreas delimitan las zonas completas que contienen todos los códigos de barras de cada hoja, incluyendo el espacio circundante necesario para la lectura automática. El número superior a las 31 hojas originales se debe a las técnicas de aumento de datos aplicadas durante el procesamiento.

Las regiones **header** constituyen la clase minoritaria con 67 elementos identificados, representando el 2.2 % del total. Esta cifra refleja que no todos los cupones precinto en las imágenes procesadas contienen cabeceras claramente delimitadas o que algunas cabeceras pueden estar parcialmente cortadas o ser ilegibles debido a las condiciones de captura.

5.2. Análisis exploratorio visual de los datos

El análisis exploratorio de los datos proporciona información crucial para comprender las características dimensionales de los elementos y los desafíos específicos que presentan para los algoritmos de detección automática.

5.2.1. Distribución de tamaños de elementos

La distribución de dimensiones de los elementos detectables proporciona información crucial para la configuración de los modelos de detección y la definición de anclas apropiadas para los algoritmos YOLO.

Como se puede observar en la Figura 5.2, los elementos **code** presentan la distribución más concentrada, con la mayoría de elementos agrupados en un rango estrecho de tamaños, lo que es consistente con el formato estandarizado de los códigos EAN-13. Esta uniformidad facilita su detección automática al presentar características dimensionales predecibles.

Los elementos **header** muestran una variabilidad considerablemente mayor en sus dimensiones, reflejando las diferencias en la cantidad de información textual y el diseño específico de cada medicamento. Esta variabilidad requiere que los modelos sean capaces de detectar elementos de tamaños muy diversos, desde headers compactos hasta cabeceras extensas que incluyen información detallada del producto.

Los elementos **barcode** presentan una distribución intermedia, con dimensiones que varían según el área total de cada hoja y las condiciones de captura. Al representar la zona completa de cupones de cada hoja, estos elementos tienden a tener dimensiones más grandes y consistentes que los códigos individuales.

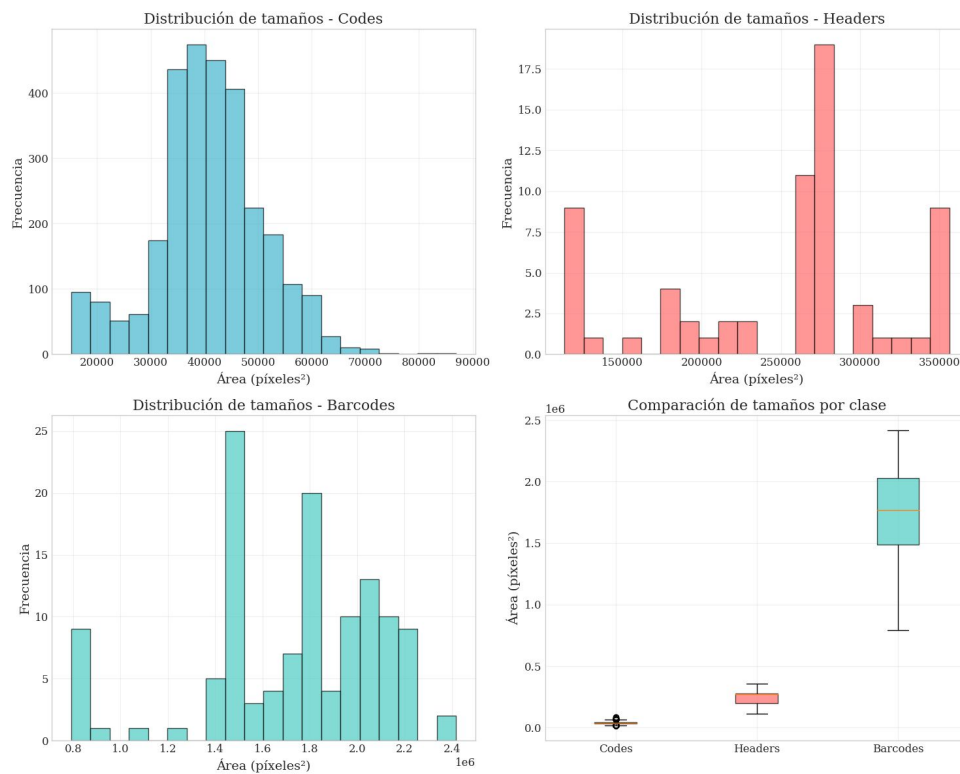


Figura 5.2: Distribución de tamaños de caja delimitadora por clase. Los headers muestran mayor variabilidad en dimensiones, mientras que los códigos de barras presentan características más consistentes.

5.2.2. Casos problemáticos y atípicos

El análisis exploratorio ha identificado varios tipos de casos problemáticos que representan desafíos específicos para los algoritmos de detección automática. Estos casos reflejan las condiciones reales de uso en entornos farmacéuticos donde las condiciones de captura no siempre son ideales.

Los casos problemáticos mostrados en la Figura 5.3 incluyen cupones capturados con perspectiva angular que distorsiona la geometría de los códigos de barras, dificultando su detección y lectura automática. Este tipo de distorsión geométrica requiere que los algoritmos sean resistentes a las transformaciones de perspectiva y puedan mantener la precisión de detección incluso cuando los elementos no se presentan en orientación frontal perfecta.

Los códigos de barras alejados de la cámara durante la captura presentan otro desafío significativo, ya que la reducción en la resolución efectiva puede afectar la capacidad de los modelos para distinguir las características distintivas necesarias para la detección. Esta limitación es especialmente relevante en entornos donde las condiciones de captura pueden variar según el operador y el equipo utilizado.

El posicionamiento incierto por códigos ausentes representa un desafío algorítmico complejo, donde la ausencia de elementos esperados en posiciones específicas dificulta la determinación precisa de las coordenadas de los elementos restantes. Los algoritmos de detección dependen de la estructura regular de los cupones para establecer referencias espaciales, y cuando faltan códigos en posiciones específicas, se requieren estrategias de detección más avanzadas que puedan manejar estructuras incompletas o irregulares.

Ejemplos de casos problemáticos identificados en el conjunto de datos

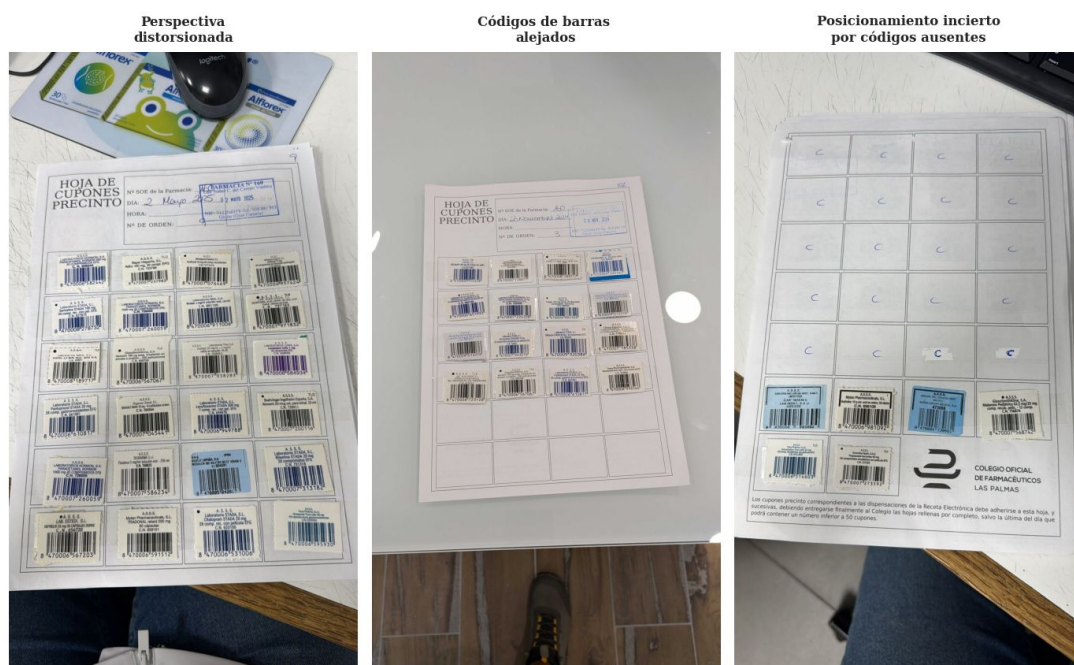


Figura 5.3: Ejemplos de casos problemáticos identificados en el conjunto de datos: perspectiva distorsionada, códigos de barras alejados, y posicionamiento incierto por códigos ausentes.

5.3. Proceso de etiquetado y validación

El etiquetado de las imágenes se realizó utilizando la herramienta LabelMe, que permite la creación de anotaciones precisas mediante polígonos que delimitan las regiones de interés. Se establecieron criterios específicos y consistentes para garantizar la uniformidad de las anotaciones a lo largo de todo el conjunto.

Como se puede ver en la Figura 5.4, el proceso de etiquetado siguió criterios específicos para cada clase de elemento. Para los **headers**, se etiquetaron las regiones superiores completas de cada cupón que contienen la información de dispensación de la farmacia, incluyendo el número SOE de la farmacia, la fecha del día de dispensación y el número de orden asignado. Para los **barcodes**, se delimitó el área completa de cada hoja que contiene todos los cupones, incluyendo el espacio circundante necesario para la lectura automática. Para los **codes**, se marcaron precisamente cada código de barras individual EAN-13, asegurando que los polígonos envolvieran completamente cada código.



Figura 5.4: Ejemplo de etiquetado con LabelMe mostrando la delimitación de regiones header (rojo), barcode (verde) y code (amarillo) en cupones precinto. La imagen izquierda presenta el etiquetado de la cara delantera, mientras que la imagen derecha muestra el etiquetado de la cara trasera. Este proceso permite crear anotaciones precisas mediante polígonos que delimitan las regiones de interés para el entrenamiento de modelos de detección.

5.3.1. Control de calidad y validación

Se realizó validación estadística mediante el análisis de las distribuciones de dimensiones y posiciones para identificar anotaciones atípicas. Finalmente, se ejecutó una verificación visual del 100 % de las anotaciones convertidas al formato YOLO para garantizar la integridad de la transformación de coordenadas.

La herramienta LabelMe genera archivos JSON que contienen las coordenadas de los polígonos de anotación junto con las etiquetas de clase correspondientes. Estos archivos fueron posteriormente convertidos al formato requerido por los modelos YOLO, transfor-

mando las coordenadas de polígonos en caja delimitadora rectangulares con coordenadas normalizadas.

5.4. División del conjunto de datos y estrategia evolutiva

El conjunto de 31 imágenes se dividió siguiendo un enfoque de desarrollo iterativo que refleja la evolución natural del proyecto y las limitaciones de datos disponibles en diferentes fases.

Tabla 5.1: División del conjunto de datos según desarrollo iterativo del proyecto.

Conjunto	Imágenes	Porcentaje	Propósito
Entrenamiento	10	32 %	Entrenamiento inicial de modelos YOLO
Validación	11	35 %	Validación durante entrenamiento
Evaluación	10	32 %	Generación de ground truth y evaluación final
Total	31	100 %	Conjunto completo

Como se observa en la Figura 5.5, el análisis de la distribución de elementos etiquetados revela que el conjunto de entrenamiento contiene 2,468 elementos (80.4 % del total), el conjunto de validación incluye 316 elementos (10.3 %), y el conjunto de test comprende 286 elementos (9.3 %). Esta distribución, aunque no convencional en términos de proporción de imágenes, resulta equilibrada en términos de elementos disponibles para cada fase del desarrollo.

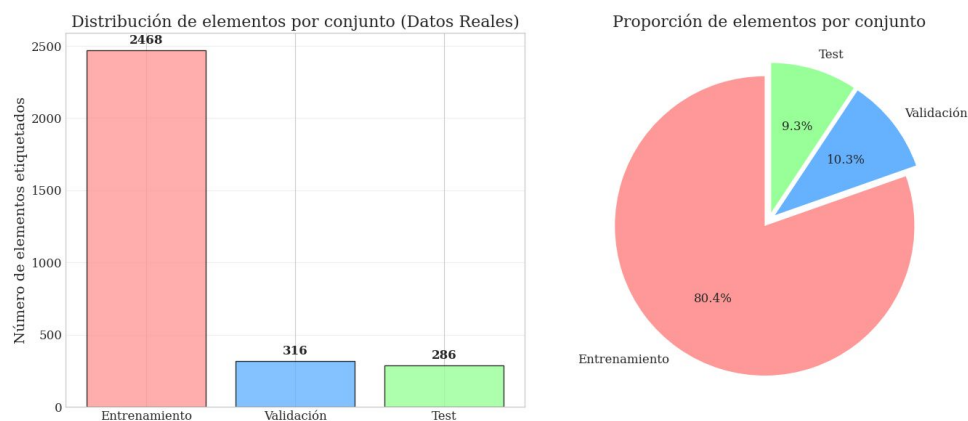


Figura 5.5: Distribución de elementos etiquetados por conjunto de datos.

5.4.1. Justificación metodológica

Esta división particular se originó en las fases tempranas del proyecto, cuando solo se disponía de 12 imágenes iniciales. La estrategia adoptada presenta ventajas específicas que justifican su mantenimiento. El conjunto de evaluación amplio permite una validación exhaustiva del rendimiento real del sistema, proporcionando mayor confianza en los resultados obtenidos.

Esta estrategia funciona bien debido a las características específicas de los modelos YOLO modernos, que pueden entrenarse de manera efectiva con conjuntos de datos pequeños gracias a sus técnicas internas de aumento automático. YOLO10 y YOLO11 incluyen transformaciones como mosaic augmentation, mixup y copy-paste augmentation que compensan las limitaciones de volumen de datos mediante la generación dinámica de variaciones durante el entrenamiento.

5.5. Consideraciones sobre aumento de datos

La evolución del proyecto hacia el uso de modelos YOLO10 y YOLO11 modificó significativamente la estrategia de aumento de datos. Estos modelos incorporan técnicas automáticas avanzadas que superan las implementaciones manuales tradicionales.

Las técnicas automáticas incluyen mosaic augmentation, que combina múltiples imágenes en una sola para aumentar la diversidad, mixup que mezcla imágenes y etiquetas para mejorar la generalización, transformaciones adaptativas que aplican rotaciones, escalados y cambios de color optimizados, y copy-paste augmentation que copia elementos entre imágenes para aumentar la variabilidad.

Este cambio justifica, la decisión de mantener un conjunto de entrenamiento relativamente pequeño, ya que los modelos modernos compensan las limitaciones de datos mediante técnicas internas de aumento que son más avanzadas y efectivas que los métodos manuales tradicionales.

5.6. Limitaciones y oportunidades de mejora

El conjunto de datos presenta limitaciones inherentes a su tamaño (31 imágenes) y origen (un solo proveedor), lo que puede limitar la generalización a otros formatos farmacéuticos. El desequilibrio entre clases (codes: 93.8 %, headers: 2.2 %, barcodes: 4.0 %) refleja la estructura real de los cupones pero requiere consideraciones especiales durante el entrenamiento.

Para mitigar estas limitaciones se implementaron múltiples enfoques de detección complementarios y un conjunto de evaluación amplio que proporciona validación robusta del rendimiento real. El análisis detallado de casos extremos documenta las condiciones más desafiantes y orienta futuros desarrollos.

Para expansiones futuras se recomienda diversificar proveedores, incorporar imágenes de diferentes períodos temporales e incluir condiciones controladas que proporcionen líneas base de comparación.

Capítulo 6

Métodos empleados

Este capítulo presenta los fundamentos teóricos y las características principales de los métodos empleados para el procesamiento automático de cupones precinto farmacéuticos. Se describen cuatro enfoques tecnológicos principales: modelos YOLO para detección de objetos, técnicas de análisis de gradientes para códigos de barras, modelos de segmentación universal SAM2, y APIs de inteligencia artificial multimodal. Cada método se seleccionó por sus capacidades específicas para abordar los desafíos únicos del dominio farmacéutico.

La selección de estos métodos responde a la necesidad de evaluar diferentes paradigmas tecnológicos, desde enfoques tradicionales de visión por computador hasta modelos fundacionales de última generación. Esta diversidad metodológica permite identificar las fortalezas y limitaciones de cada aproximación en el contexto específico del procesamiento de documentos farmacéuticos, proporcionando una base sólida para el desarrollo de soluciones híbridas que combinen las mejores características de cada enfoque.

6.1. Modelos YOLO para detección de objetos

Los modelos YOLO (*You Only Look Once*) representan una familia de arquitecturas de redes neuronales convolucionales diseñadas para la detección de objetos en tiempo real [Redmon et al., 2016]. Estos modelos revolucionaron el campo de la visión por computador al proponer un enfoque unificado que realiza la detección y clasificación de objetos en una sola pasada por la red, eliminando la necesidad de múltiples etapas de procesamiento características de métodos anteriores.

La arquitectura YOLO divide la imagen de entrada en una cuadrícula regular y predice simultáneamente las coordenadas de los caja delimitadora, las probabilidades de clase, y las puntuaciones de confianza para cada celda de la cuadrícula. Esta aproximación permite un procesamiento extremadamente eficiente, ya que toda la información necesaria se obtiene en una única evaluación de la red neuronal.

6.1.1. Evolución arquitectónica: YOLO10 y YOLO11

La evolución de YOLO ha sido constante desde sus primeras versiones, con cada iteración introduciendo mejoras significativas en términos de precisión y eficiencia. Como se muestra en la Figura 6.1, las versiones más recientes, YOLO10 y YOLO11, incorporan innovaciones arquitectónicas que las hacen especialmente adecuadas para aplicaciones especializadas como el procesamiento de documentos farmacéuticos [Wang et al., 2024a, Wang et al., 2024b].

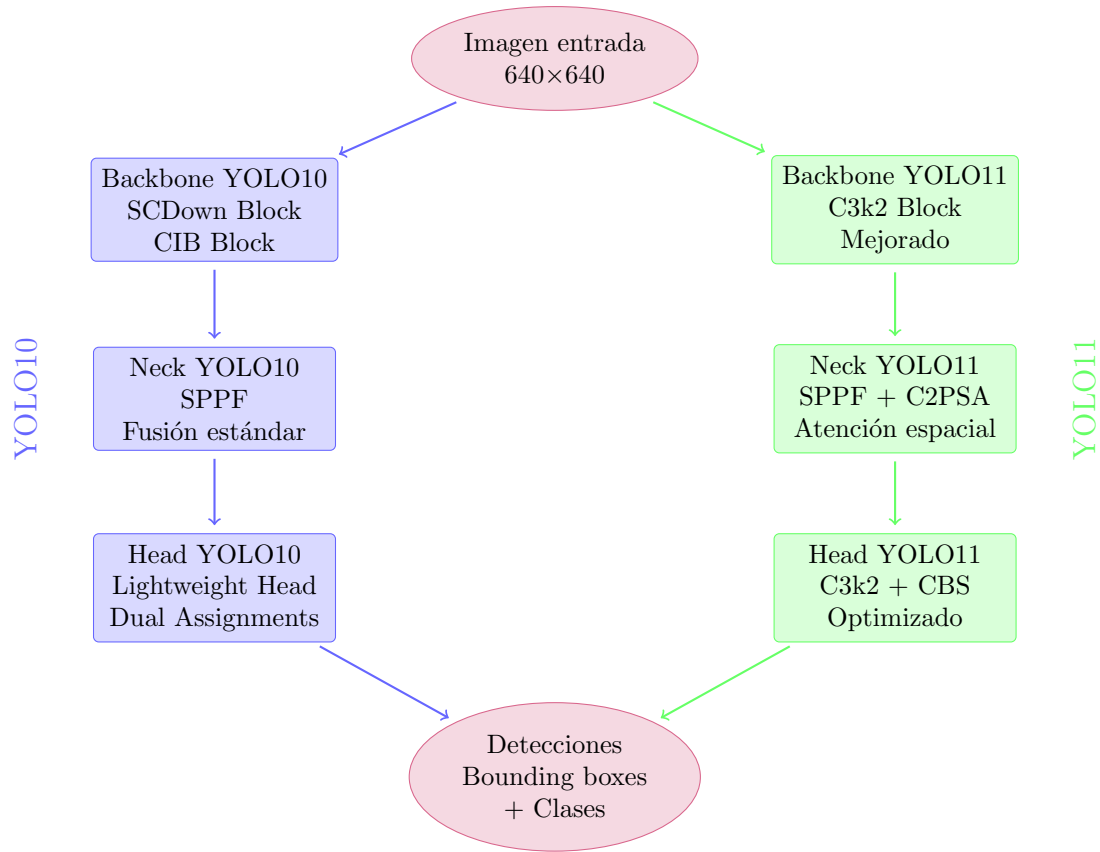


Figura 6.1: Evolución arquitectónica: comparación paralela entre YOLO10 y YOLO11.

Tabla 6.1: Comparación de innovaciones clave entre YOLO10 y YOLO11.

Componente	YOLO10	YOLO11
Backbone	SCDown Block para downsampling espacial-canal desacoplado	C3k2 Block reemplazando C2f con mayor eficiencia
Neck	SPPF estándar con fusión tradicional	SPPF + C2PSA con atención espacial mejorada
Head	Lightweight Head con convoluciones separables	C3k2 + CBS blocks para mayor precisión
Innovación principal	Consistent Dual Assignments (sin NMS)	Partial Self-Attention con menor coste computacional
Enfoque	Eliminación de post-procesamiento	Mejora de representación global

Como se puede observar en la Tabla 6.1, **YOLO10** se centra en la eliminación de cuellos de botella en el post-procesamiento mediante el Consistent Dual Assignments, que elimina la necesidad de Non-Maximum Suppression NMS durante la inferencia [Wang et al., 2024a]. Esta innovación combina asignaciones uno-a-muchos durante el entrenamiento con asignaciones uno-a-uno durante la inferencia, reduciendo significativamente la latencia.

YOLO11 representa una evolución que prioriza la mejora de la capacidad de representación mediante la introducción del bloque C3k2 y el mecanismo C2PSA (*Cross Stage Partial with Partial Self-Attention*) [Wang et al., 2024b]. Esta aproximación mejora la capacidad de modelado global con un coste computacional mínimo, optimizando la

precisión sin sacrificar eficiencia.

6.1.2. Fundamentos arquitectónicos de YOLO

La arquitectura YOLO se estructura en tres componentes principales: backbone, neck y head [Redmon et al., 2016]. El **backbone** se encarga de la extracción de características visuales mediante capas convolucionales profundas que aprenden representaciones jerárquicas de la imagen de entrada. El **neck** fusiona características de diferentes escalas para mejorar la detección de objetos de diversos tamaños. El **head** genera las predicciones finales de localización y clasificación.

Esta organización modular permite optimizar cada componente independientemente, facilitando la incorporación de innovaciones específicas sin afectar el funcionamiento global del sistema. La capacidad de procesar imágenes de diferentes resoluciones y la flexibilidad para adaptarse a diversos números de clases hacen de YOLO una arquitectura especialmente versátil para aplicaciones especializadas.

6.1.3. Ventajas para detección múltiple

La capacidad de YOLO para detectar múltiples objetos simultáneamente resulta fundamental para el procesamiento de cupones precinto, donde cada imagen puede contener decenas de códigos de barras distribuidos en patrones regulares. Esta característica elimina la necesidad de técnicas de ventana deslizante o múltiples pasadas por la imagen, reduciendo significativamente el tiempo de procesamiento.

La precisión en la localización de objetos pequeños, mejorada en las versiones recientes mediante técnicas de fusión de características multi-escala, resulta crucial para la detección de códigos EAN-13 que pueden presentar dimensiones reducidas en las digitalizaciones de cupones. La capacidad de mantener alta precisión incluso con objetos densamente agrupados aborda uno de los principales desafíos del dominio farmacéutico.

6.2. Análisis de gradientes para detección de códigos

El análisis de gradientes representa un enfoque fundamentado en las características visuales intrínsecas de los códigos de barras, aprovechando el hecho de que estos elementos presentan patrones distintivos de líneas verticales que generan gradientes direccionales específicos [Tuinstra, 2006]. Esta aproximación se basa en principios de procesamiento de señales digitales aplicados al dominio de la visión por computador.

La Figura 6.2 ilustra el pipeline completo del detector basado en análisis de gradientes.

6.2.1. Principios matemáticos del método

Los códigos de barras EAN-13 utilizados en el ámbito farmacéutico consisten en patrones de barras verticales de diferentes anchos que codifican información numérica mediante variaciones en el espaciado y grosor de las líneas. Estas características visuales generan respuestas distintivas cuando se aplican operadores de gradiente, particularmente en la dirección horizontal donde las transiciones entre barras y espacios son más pronunciadas [Tuinstra, 2006].

El método se basa en el cálculo de mapas de energía que resaltan las estructuras verticales características de los códigos de barras. El operador de Sobel se utiliza para

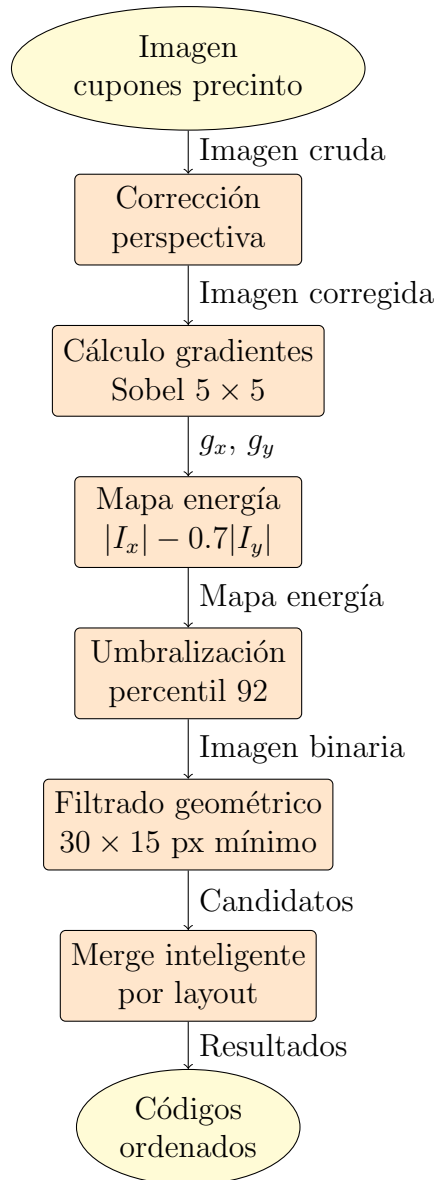


Figura 6.2: Pipeline del detector basado en análisis de gradientes.

calcular los gradientes direccionales, proporcionando una medida de la intensidad de los cambios en la imagen en direcciones específicas.

La formulación matemática del mapa de energía combina los gradientes horizontal y vertical mediante la fórmula:

$$E(x, y) = |I_x(x, y)| - 0.7 \cdot |I_y(x, y)| \quad (6.1)$$

donde el factor 0.7 se determinó empíricamente para optimizar la detección de líneas verticales mientras se suprimen estructuras horizontales. Esta selectividad direccional resulta fundamental para distinguir códigos de barras de otros elementos textuales o gráficos presentes en los cupones precinto.

6.2.2. Aplicación a códigos de barras farmacéuticos

La especificidad del método para códigos de barras lo hace particularmente adecuado para el dominio farmacéutico, donde estos elementos constituyen la información más crítica para el procesamiento automático. La independencia de modelos preentrenados elimina la necesidad de grandes conjuntos de datos de entrenamiento, aspecto particularmente relevante en dominios especializados donde los datos anotados son escasos.

6.3. APIs multimodales

Las APIs de inteligencia artificial multimodal representan la frontera actual en el desarrollo de sistemas que pueden procesar y comprender simultáneamente múltiples modalidades de información, incluyendo texto, imágenes, y otros tipos de datos. Estos sistemas se basan en modelos fundacionales entrenados en conjuntos de datos masivos que abarcan múltiples dominios y tipos de contenido.

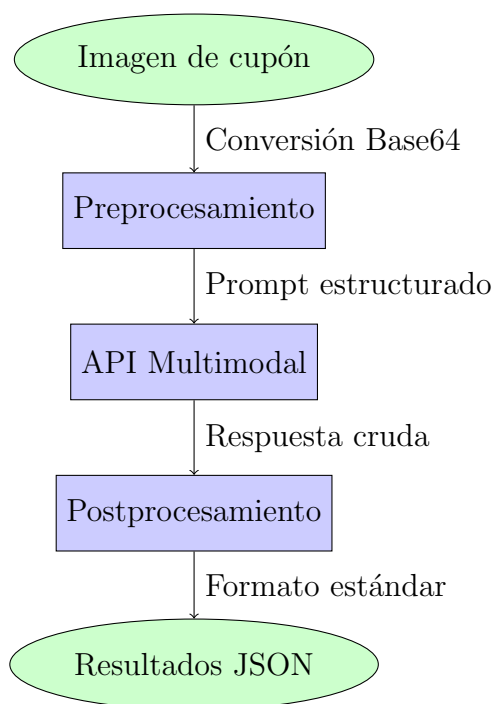


Figura 6.3: Arquitectura de procesamiento para APIs multimodales.

6.3.1. GPT-4 Vision: capacidades de comprensión visual

GPT-4 Vision extiende las capacidades de los modelos de lenguaje de gran escala al dominio visual, permitiendo el análisis y comprensión de imágenes mediante procesamiento de lenguaje natural [OpenAI, 2024b]. Esta integración multimodal permite formular consultas complejas sobre el contenido visual utilizando instrucciones en lenguaje natural, eliminando la necesidad de desarrollar pipeline de procesamiento específicos.

La arquitectura subyacente combina encoder visuales especializados con modelos de lenguaje transformer, permitiendo la traducción de información visual a representaciones textuales que pueden ser procesadas por los mecanismos de razonamiento del modelo de

lenguaje. Esta aproximación facilita la implementación de tareas complejas de análisis visual mediante la formulación de prompt apropiados.

Las capacidades principales incluyen OCR avanzado con reconocimiento de texto en múltiples idiomas y formatos, identificación de objetos con detección y clasificación de elementos visuales, análisis de layout para comprensión de la estructura y organización de documentos, y razonamiento visual con capacidad para inferir relaciones espaciales y contextuales.

6.3.2. Gemini: procesamiento nativo multimodal

Gemini representa una aproximación diferente al procesamiento multimodal, con una arquitectura diseñada desde el inicio para manejar múltiples modalidades de forma nativa [Team et al., 2023]. Esta integración fundamental permite un procesamiento más eficiente y coherente de información multimodal, evitando las limitaciones de enfoques que combinan modelos especializados desarrollados independientemente.

La capacidad de Gemini para procesar documentos complejos con layout estructurados lo hace particularmente relevante para aplicaciones farmacéuticas, donde los cupones precinto presentan organizaciones específicas que requieren comprensión tanto visual como estructural.

6.4. SAM2 para segmentación universal

SAM2 (*Segment Anything Model 2*) representa la evolución de los modelos fundacionales para segmentación universal, diseñados para segmentar cualquier objeto en imágenes sin requerir entrenamiento específico para cada clase [Ravi et al., 2024]. Estos modelos se entrenan en conjuntos de datos masivos y diversos, desarrollando capacidades de generalización que les permiten abordar tareas de segmentación en dominios no vistos durante el entrenamiento. Como se muestra en la Figura 6.4, la arquitectura de SAM2 integra múltiples componentes especializados para el procesamiento multimodal.

6.4.1. Arquitectura transformer para segmentación

La arquitectura de SAM2 se basa en un encoder de imágenes Hiera que extrae representaciones visuales ricas, un decoder especializado en la generación de máscaras de segmentación, y un sistema de prompt que permite guiar el proceso de segmentación mediante diferentes tipos de entrada como puntos, cajas delimitadoras, o máscaras aproximadas [Ravi et al., 2024].

El componente Hiera representa una evolución de las arquitecturas transformer adaptadas específicamente para el procesamiento eficiente de imágenes. Su diseño jerárquico permite capturar características tanto locales como globales, esencial para la segmentación precisa de objetos con diferentes escalas y complejidades.

6.4.2. Capacidades zero shot

La principal fortaleza de SAM2 radica en su capacidad para segmentar objetos de forma precisa sin requerir ejemplos específicos del dominio de aplicación. Esta característica lo convierte en una herramienta potencialmente valiosa para aplicaciones farmacéuticas,

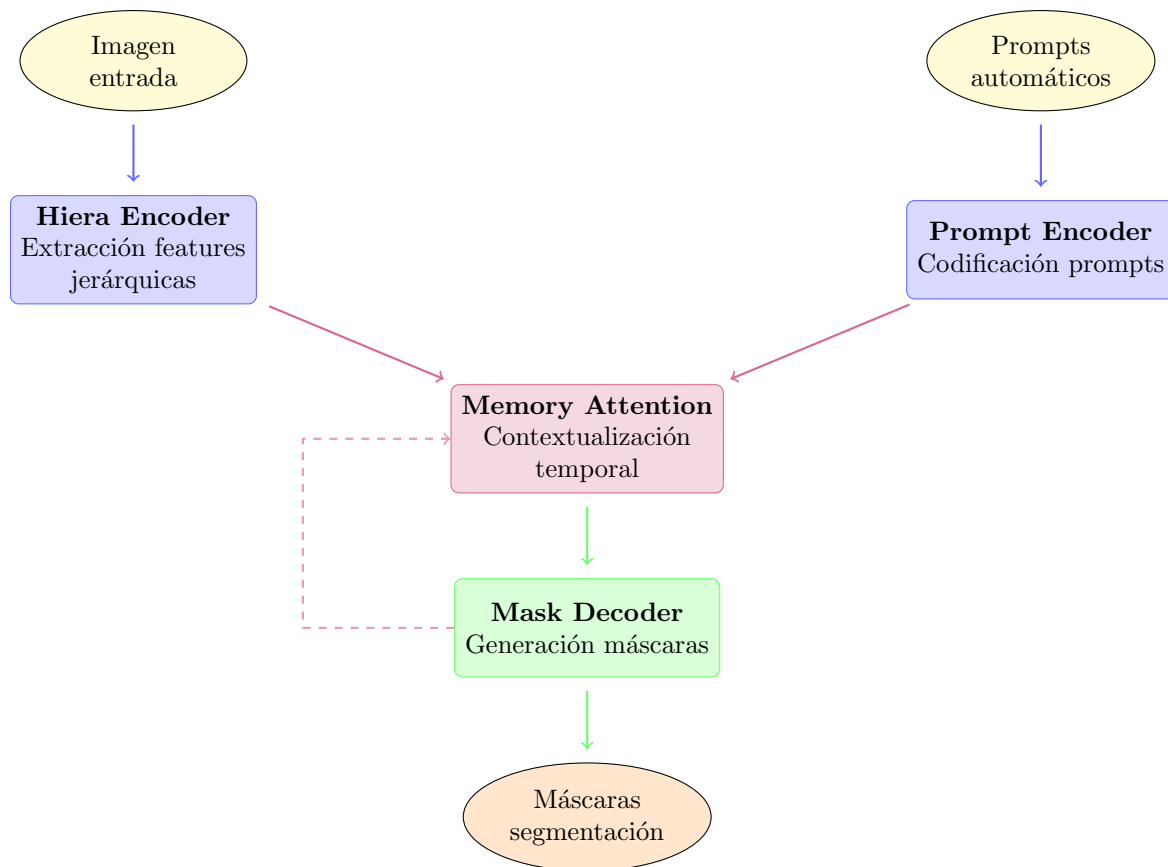


Figura 6.4: Arquitectura SAM2 con componentes principales y flujo de información.

donde la variabilidad en formatos y layout de documentos puede dificultar el desarrollo de soluciones específicas.

El modelo incorpora mecanismos de atención avanzados que le permiten identificar límites de objetos con alta precisión, incluso en casos donde los objetos presentan formas irregulares o se encuentran parcialmente ocluidos. Esta capacidad resulta relevante para el procesamiento de cupones precinto, donde los códigos de barras pueden presentar calidad variable o estar afectados por artefactos de digitalización.

6.5. Análisis comparativo y selección metodológica

La evaluación comparativa de los métodos seleccionados requiere considerar múltiples criterios relevantes para el dominio farmacéutico. La Tabla 6.2 proporciona una comparación sistemática basada en aspectos técnicos y prácticos.

La selección de estos cuatro enfoques metodológicos responde a la necesidad de evaluar diferentes paradigmas tecnológicos para identificar las aproximaciones más efectivas para el procesamiento de cupones precinto farmacéuticos. Cada método aporta perspectivas complementarias que permiten una evaluación completa de las posibilidades tecnológicas disponibles.

Los modelos YOLO representan el estado del arte en detección de objetos con capacidades demostradas en múltiples dominios. El análisis de gradientes ofrece un enfoque fundamentado en las características específicas de los códigos de barras. Las APIs multimodales proporcionan una perspectiva sobre las posibilidades de los sistemas de inteligencia

Tabla 6.2: Comparación de métodos según criterios específicos para cupones precinto.

Criterio	YOLO	Gradientes	SAM2	APIs
Velocidad de procesamiento	Alta	Muy Alta	Media	Baja
Precisión de detección	Alta	Media	Muy Alta	Alta
Recursos computacionales	Medios	Bajos	Altos	Variables
Facilidad de implementación	Media	Fácil	Compleja	Muy Fácil
Adaptabilidad al dominio	Alta	Media	Muy Alta	Alta
Dependencias externas	Bajas	Muy Bajas	Medias	Altas
Interpretabilidad	Baja	Alta	Baja	Media
Coste de desarrollo	Medio	Bajo	Alto	Muy Bajo
Escalabilidad	Alta	Muy Alta	Media	Variable

artificial más avanzados disponibles comercialmente. SAM2 permite explorar las capacidades de los modelos fundacionales para segmentación universal.

El enfoque metodológico adoptado garantiza una evaluación rigurosa y completa que proporcionará las bases para el desarrollo de una solución optimizada para el dominio farmacéutico específico, considerando tanto las capacidades técnicas como las limitaciones prácticas de implementación en entornos reales.

Capítulo 7

Configuración de experimentos

Este capítulo describe la configuración experimental desarrollada para evaluar y comparar los diferentes métodos de procesamiento automático de cupones precinto implementados en el proyecto. Se detallan las configuraciones específicas de cada detector, los protocolos experimentales desarrollados, y el framework de evaluación implementado para garantizar comparaciones objetivas entre los distintos enfoques.

La configuración experimental se diseñó para abordar las particularidades del dominio farmacéutico, donde la precisión en la extracción de códigos es crítica para la trazabilidad y el cumplimiento normativo. El framework desarrollado permite evaluar tanto la precisión de detección como aspectos prácticos relevantes para la implementación en entornos farmacéuticos reales, proporcionando una base sólida para la evaluación comparativa de cada aproximación tecnológica.

7.1. Configuración de detectores basados en redes neuronales

Los detectores basados en modelos YOLO se configuraron utilizando arquitecturas preentrenadas optimizadas específicamente para el conjunto de datos limitado disponible. La configuración se adaptó para maximizar el aprendizaje con las 10 imágenes de entrenamiento disponibles, aprovechando las capacidades específicas de cada versión de YOLO implementada.

Tras la evaluación comparativa de ambas arquitecturas en el conjunto de datos farmacéutico, cuyos parámetros específicos se detallan en la Tabla 7.1, se determinó utilizar YOLO11 para todos los experimentos subsecuentes presentados en este trabajo. Esta decisión se fundamenta en que YOLO10, aunque operativamente funcional, demostró un rendimiento ligeramente inferior en las métricas de precisión y correspondencia posicional específicas del dominio farmacéutico. La arquitectura híbrida CNN-transformer de YOLO11, combinada con su optimizador AdamW y scheduler coseno, proporcionó mayor estabilidad de convergencia y mejor adaptación a las características visuales de los cupones precinto, justificando su selección como detector principal para la evaluación experimental.

El sistema implementado define tres clases específicas para capturar los elementos esenciales de los cupones precinto: códigos individuales EAN-13, cabeceras manuscritas con información SOE, y áreas completas que contienen matrices de códigos. La detección se realiza sobre la imagen completa, seguida de ordenamiento en cuadrícula que respeta el layout farmacéutico específico mediante algoritmos de agrupación por filas y ordenamiento

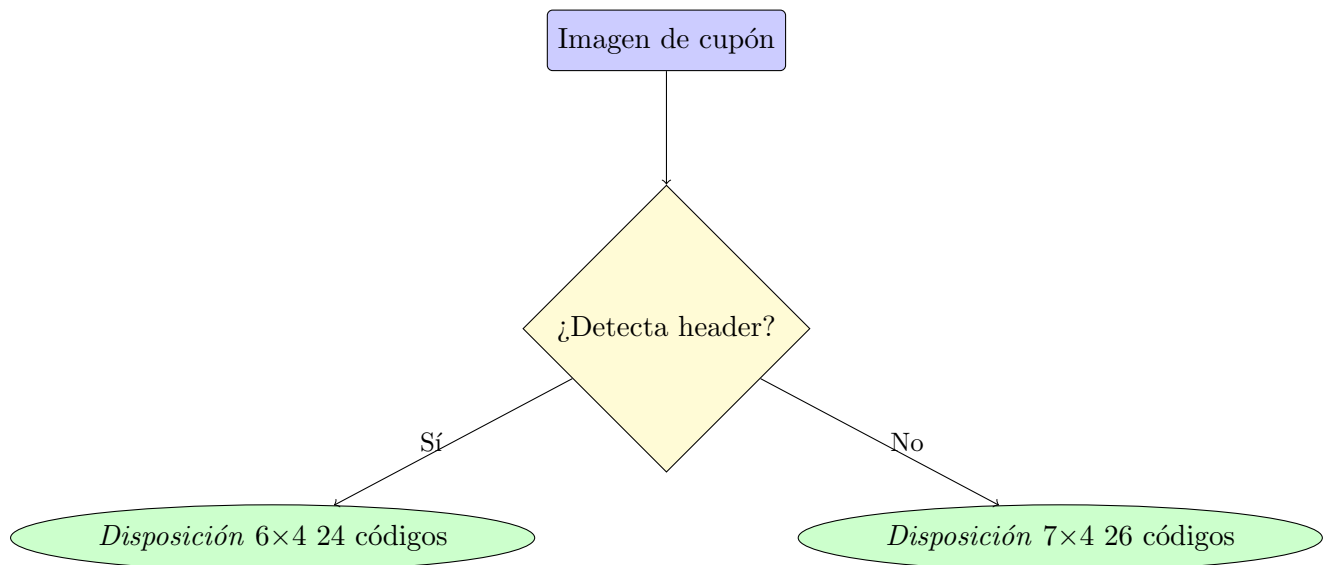
Tabla 7.1: Configuración comparativa de detectores YOLO10 y YOLO11

Parámetro	YOLO10	YOLO11	Justificación
Épocas	250	300	Mayor convergencia requerida para transformer
batch size	32	16	Reducido para estabilidad de arquitectura híbrida
learning rate	0.002	0.001	Optimización para arquitectura CNN-transformer
Optimizador	Adam	AdamW	weight decay para regularización mejorada
scheduler	Fijo	Coseno	Decaimiento adaptativo para convergencia
early stopping	50 épocas	50 épocas	Prevención consistente de sobreajuste
Umbral confianza	0.25	0.25	Maximización de sensibilidad de detección
Tamaño imagen	640×640	640×640	Estándar YOLO para balance precisión-velocidad

horizontal.

El sistema de decodificación híbrida implementado aplica cinco variaciones de preprocesamiento: imagen original, mejora de contraste CLAHE, normalización, binarización adaptativa, e inversión de imagen. Usando pyzbar, priorizando códigos de 13 dígitos y seleccionando el resultado más común cuando hay múltiples detecciones.

La determinación automática de layout establece 6×4 (24 códigos) cuando se detectan elementos manuscritos, o 7×4 (26 códigos útiles) cuando no hay elementos manuscritos. Esta configuración se ilustra en la Figura 7.1.

Figura 7.1: Determinación automática de *disposición* según detección de header SOE.

7.2. Configuración del detector basado en análisis de gradientes

El detector de gradientes utiliza tres componentes principales: el algoritmo de detección, el sistema de corrección automática de perspectiva, y el módulo de preprocesamiento avanzado. Los parámetros están optimizados empíricamente para el dominio farmacéutico específico.

Tabla 7.2: Parámetros del detector basado en análisis de gradientes

Parámetro	Valor	Descripción
Factor de energía	0.7	Peso para suprimir gradientes horizontales
Kernel Sobel	5×5 píxeles	Máxima sensibilidad a bordes
Umbralización	Percentil 92	Selección de regiones de alta energía
Ancho mínimo	30 píxeles	Filtrado de detecciones espurias
Alto mínimo	15 píxeles	Basado en dimensiones EAN-13
Aspect ratio	1.2 - 5.0	Rango típico de códigos de barras
Área mínima	400 píxeles ²	Eliminación de ruido

La implementación del mapa de energía sigue la fórmula matemática:

$$E(x, y) = |I_x(x, y)| - 0.7 \cdot |I_y(x, y)| \quad (7.1)$$

donde el factor 0.7 se determinó empíricamente para optimizar la detección de líneas verticales mientras se suprimen estructuras horizontales.

El sistema implementa múltiples técnicas de preprocesamiento aplicadas secuencialmente: imagen original, mejora de contraste CLAHE con clipLimit=3.0, ajustes de brillo de ± 30 unidades, filtros gaussiano, mediano y bilateral, filtro de nitidez, binarización adaptativa y Otsu, operaciones morfológicas de cierre y apertura, y escalado múltiple desde $2\times$ hasta $6\times$.

7.3. Configuración de APIs

Los modelos fundacionales se configuraron con parámetros específicos para maximizar la consistencia y calidad de las respuestas en el dominio farmacéutico, incorporando conocimiento específico del formato EAN-13 con prefijo 847 y la estructura de cupones precinto del Sistema Nacional de Salud.

7.3.1. Desarrollo y optimización de prompts

El desarrollo de prompts efectivos para las APIs multimodales requirió un proceso iterativo de refinamiento basado en los resultados empíricos obtenidos en las pruebas iniciales.

Tabla 7.3: Parámetros de configuración para GPT-4 Vision y Google Gemini

Parámetro	GPT-4 Vision	Gemini	Notas
Temperatura	0.0	0.1	Determinismo vs flexibilidad balanceada
Tokens máximos	3000	Auto	Análisis detallado de imágenes complejas
Formato respuesta	JSON obligatorio	Markdown estructurado	Facilita procesamiento automático
Timeout	60-180s escalonado	120s base	Manejo de latencia variable
Reintentos	3 intentos	3 intentos	Robustez del sistema
Endpoint	Azure OpenAI	Google AI	Servicios empresariales estables

GPT-4 Vision

El protocolo final implementado combina análisis estructurado en tres pasos con formato de respuesta JSON obligatorio:

”Analiza esta imagen de cupón precinto farmacéutico siguiendo estos pasos: 1) Detecta presencia de header manuscrito (SOE, Día, Orden), 2) Determina layout automáticamente (4×6 con header vs 4×7 sin header), 3) Extrae códigos EAN-13 españoles con prefijo 847 manteniendo correspondencia posicional exacta. Respuesta obligatoria en JSON: {'header_detected': boolean, 'grid': [[códigos]], 'header_info': {...}}”

Las iteraciones principales incluyeron: especificación del formato EAN-13 con prefijo 847, restricción de salida a JSON estructurado para facilitar parsing automático, y incorporación de conocimiento específico del dominio farmacéutico español.

Google Gemini

La configuración de Gemini aprovecha sus capacidades nativas de procesamiento multimodal con formato Markdown estructurado:

”Analiza este cupón precinto farmacéutico. Identifica: 1) Header manuscrito (SOE/Día/Orden), 2) Códigos EAN-13 de 13 dígitos con prefijo 847, 3) Organización en cuadrícula 4×6 o 4×7. Formato respuesta: Markdown con tabla estructurada de códigos y metadatos de header.”

La adaptación específica incluyó optimización para el procesador multimodal nativo de Gemini y configuración de parámetros de seguridad que eviten filtrado innecesario de contenido farmacéutico legítimo.

Validación y normalizaciones

Ambos sistemas implementan validación automática mediante verificación de formato EAN-13, eliminación de duplicados, y normalización de espacios y caracteres especiales. La conversión a formatos estándar garantiza compatibilidad con el framework de evaluación desarrollado.

7.4. Segmentación automática con SAM2

Como complemento a los enfoques de detección basados en YOLO, se implementó una estrategia alternativa utilizando el modelo SAM2 (Segment Anything Model 2) para la segmentación automática de cupones individuales. Esta aproximación permite explorar métodos de segmentación sin supervisión que pueden adaptarse a diferentes formatos de cupones precinto sin requerir entrenamiento específico.

La implementación se basa en el modelo preentrenado `sam2.1_hiera_large` con configuración optimizada para elementos rectangulares característicos de los cupones precinto. El proceso incluye filtrado geométrico basado en área mínima y relación de aspecto, organización espacial mediante agrupación por filas, y decodificación automática de códigos EAN-13 integrada directamente en cada región segmentada.

Si bien los resultados iniciales de SAM2 mostraron capacidades prometedoras para la segmentación de cupones, como se presenta en la sección de resultados, las limitaciones computacionales del enfoque impidieron su evaluación sistemática. El tiempo de procesamiento excesivamente alto y los recursos computacionales requeridos (de media hora por imagen en algunos casos) hicieron inviable la implementación práctica de este método. Por esta razón, no se desarrollaron métricas comparativas formales ni se incluyó SAM2 en la evaluación cuantitativa final del proyecto, concentrando los esfuerzos de evaluación en los enfoques más eficientes computacionalmente.

7.5. Protocolo experimental y framework de evaluación

El protocolo experimental implementado garantiza comparaciones justas entre todos los detectores mediante procedimientos estandarizados. La Figura 7.2 ilustra el flujo completo del framework experimental.

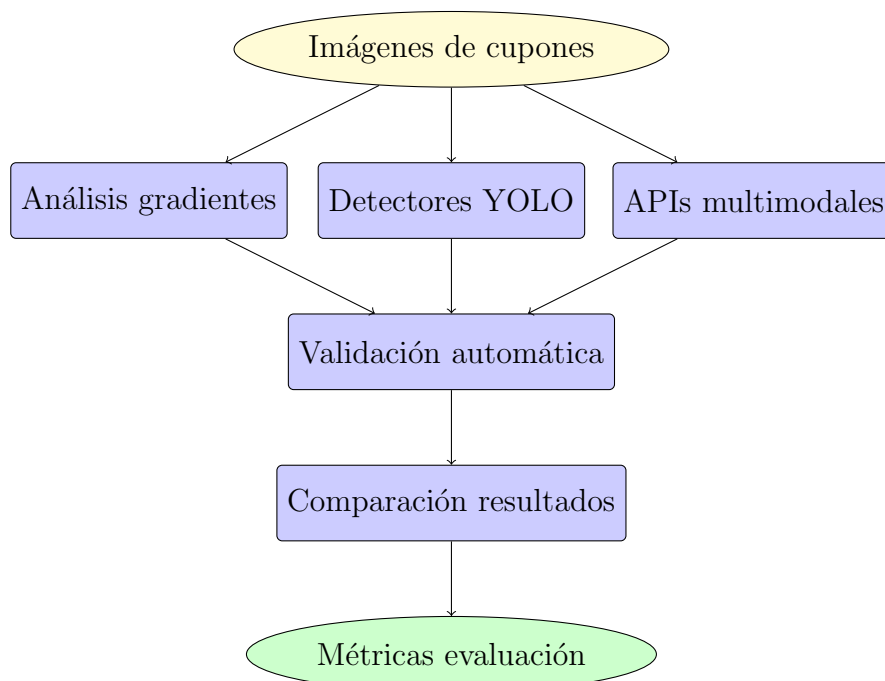


Figura 7.2: Framework experimental para evaluación comparativa de detectores.

El framework implementa múltiples métricas de evaluación que capturan diferentes aspectos del rendimiento basadas en la clasificación estándar de resultados de detección:

- **True Positives (TP)**: Códigos de barras correctamente detectados en la posición esperada
- **False Positives (FP)**: Detecciones incorrectas o códigos detectados en posiciones erróneas
- **False Negatives (FN)**: Códigos de barras presentes que no fueron detectados
- **True Negatives (TN)**: Posiciones correctamente identificadas como vacías

Precisión (Precision):

$$Precision = \frac{TP}{TP + FP} \quad (7.2)$$

La precisión mide la proporción de detecciones correctas respecto al total de detecciones realizadas, evaluando la calidad de las detecciones y penalizando las falsas alarmas.

Recall (Sensibilidad):

$$Recall = \frac{TP}{TP + FN} \quad (7.3)$$

El recall mide la proporción de códigos reales que fueron detectados correctamente, evaluando la capacidad del sistema para encontrar todos los códigos presentes.

F1-Score:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7.4)$$

El F1-Score combina precisión y recall mediante su media armónica, proporcionando una medida equilibrada del rendimiento especialmente útil cuando existe un trade-off entre ambas métricas.

El sistema genera formatos JSON estandarizados que facilitan tanto el análisis detallado como las comparaciones directas entre métodos, incluyendo metadatos, estadísticas de decodificación, información de confianza, y detalles de configuración utilizados durante el procesamiento.

El sistema genera formatos JSON estandarizados que facilitan tanto el análisis detallado como las comparaciones directas entre métodos, incluyendo metadatos, estadísticas de decodificación, información de confianza, y detalles de configuración utilizados durante el procesamiento.

7.6. Consideraciones metodológicas

La configuración experimental requirió adaptaciones específicas debido a las particularidades del dominio farmacéutico. Los detectores implementan manejo específico de errores para cada tecnología: timeouts incrementales para APIs con reintentos automáticos, validación de regiones de interés para evitar errores de indexación, y estrategias de fallback cuando fallan las técnicas de preprocesamiento primarias.

La configuración experimental desarrollada proporciona una base sólida para la evaluación objetiva de los diferentes detectores implementados, facilitando la identificación de fortalezas y limitaciones de cada enfoque en el contexto específico del procesamiento

automático de cupones precinto farmacéuticos. El framework permite comparaciones justas entre métodos muy diferentes, desde redes neuronales hasta APIs de modelos fundacionales, proporcionando insights valiosos sobre la efectividad relativa de cada aproximación tecnológica.

Capítulo 8

Resultados experimentales

Este capítulo presenta los resultados obtenidos en la evaluación comparativa de los métodos desarrollados para el procesamiento automático de hojas de cupones precinto farmacéuticos. La evaluación se realizó utilizando un conjunto de imágenes reales proporcionadas por Farmalitics, aplicando métricas específicas que consideran tanto la precisión de detección como la correspondencia posicional exacta requerida en el dominio farmacéutico.

Los resultados se organizan en cinco bloques principales que permiten una evaluación completa de los enfoques tecnológicos implementados. El primer bloque analiza el rendimiento del detector basado en redes neuronales YOLO, que representa el estado del arte en detección de objetos aplicado al dominio farmacéutico. El segundo bloque examina los resultados del detector basado en análisis de gradientes, que aprovecha las características visuales intrínsecas de los códigos de barras para lograr detección independiente de modelos preentrenados. El tercer y cuarto bloque evalúan el rendimiento de las APIs de inteligencia artificial multimodal GPT-4 Vision y Gemini respectivamente, que representan la frontera actual en modelos fundacionales capaces de comprender y procesar información visual y textual de forma integrada. El quinto bloque presenta los resultados exploratorios del modelo SAM2, demostrando su capacidad de segmentación automática de cupones aunque sin evaluación métrica formal debido a limitaciones computacionales que impidieron su implementación práctica.

8.1. Resultados del detector basado en YOLO

El detector YOLO demuestra capacidades avanzadas para la identificación simultánea de múltiples tipos de elementos en los cupones precinto farmacéuticos. Su arquitectura permite detectar códigos individuales EAN-13, cabeceras manuscritas con información SOE, y áreas generales que contienen matrices completas de códigos. Esta capacidad multiclase proporciona información contextual valiosa que facilita la interpretación automática del layout farmacéutico y mejora la precisión del ordenamiento posicional.

El entrenamiento específico realizado con imágenes del dominio farmacéutico permite al detector mantener un ordenamiento consistente de las detecciones, aspecto crítico para la correspondencia exacta entre códigos detectados y sus posiciones reales en la cuadrícula del cupón. La evaluación revela patrones específicos de rendimiento que varían según las condiciones de imagen y las características particulares de cada cupón procesado.



Figura 8.1: Resultado del detector YOLO sobre cupones precinto mostrando detección de códigos individuales, headers manuscritos y áreas generales de códigos

La Figura 8.1 presenta una comparación visual directa entre ambas versiones del detector YOLO. Los recuadros rojos delimitan los códigos individuales EAN-13 detectados, los recuadros amarillos identifican las cabeceras manuscritas que contienen información de SOE de farmacia, y los recuadros blancos marcan las áreas generales que engloban toda la matriz de códigos.

Tabla 8.1: Comparación cuantitativa entre YOLOv10 y YOLOv11

Modelo	Accuracy (%)	Códigos válidos	Fallos	Imágenes <85 %	Imágenes >40 %
YOLOv10	72.1	18.75	7.25	4	1
YOLOv11	66.7	17.33	8.67	3	3

8.1.1. Análisis de la discrepancia observada

Los resultados comparativos presentan una situación metodológicamente interesante donde las métricas cuantitativas de la Tabla 8.1 favorecen claramente a YOLOv10, mientras que la evaluación visual cualitativa sugiere un comportamiento aparentemente superior de YOLOv11 en términos de precisión de localización y reducción de falsos positivos.

Esta discrepancia puede atribuirse a diferencias fundamentales en los tipos de errores que comete cada modelo. YOLOv10 demuestra superior precisión en la decodificación correcta de códigos detectados y mejor correspondencia con el ground truth establecido durante el etiquetado, además de mayor consistencia en el rendimiento global con menor variabilidad entre imágenes y menos casos de rendimiento extremadamente bajo.

Por el contrario, YOLOv11 presenta mejor precisión visual en la localización de regiones de interés y reducción aparente de falsos positivos en la detección visual, con mayor precisión en el encuadre de códigos individuales, pero evidencia mayor número de errores en la fase posterior de decodificación que afectan las métricas cuantitativas finales.

8.1.2. Selección del modelo óptimo

Basándose en las métricas cuantitativas rigurosas y considerando que el objetivo final del sistema es la correcta identificación y decodificación de códigos farmacéuticos para su posterior procesamiento administrativo, se selecciona YOLOv10 como la configuración óptima para el análisis posterior.

Esta decisión se fundamenta en su precisión global superior del 72.1 % frente al 66.7 % de YOLOv11, una mejor relación entre códigos válidos y fallos (18.75/7.25 vs 17.33/8.67), y mayor estabilidad operativa con solo 1 imagen presentando rendimiento inferior al 40 % comparado con las 3 imágenes problemáticas de YOLOv11. La consistencia operativa de YOLOv10 se refleja en una mejor distribución general del rendimiento across el conjunto de pruebas.

El entrenamiento específico realizado con YOLOv10 permite al detector mantener un ordenamiento consistente de las detecciones, aspecto crítico para la correspondencia exacta entre códigos detectados y sus posiciones reales en la cuadrícula del cupón. Todos los análisis métricos detallados que se presentan a continuación se basan exclusivamente en los resultados obtenidos con YOLOv10.

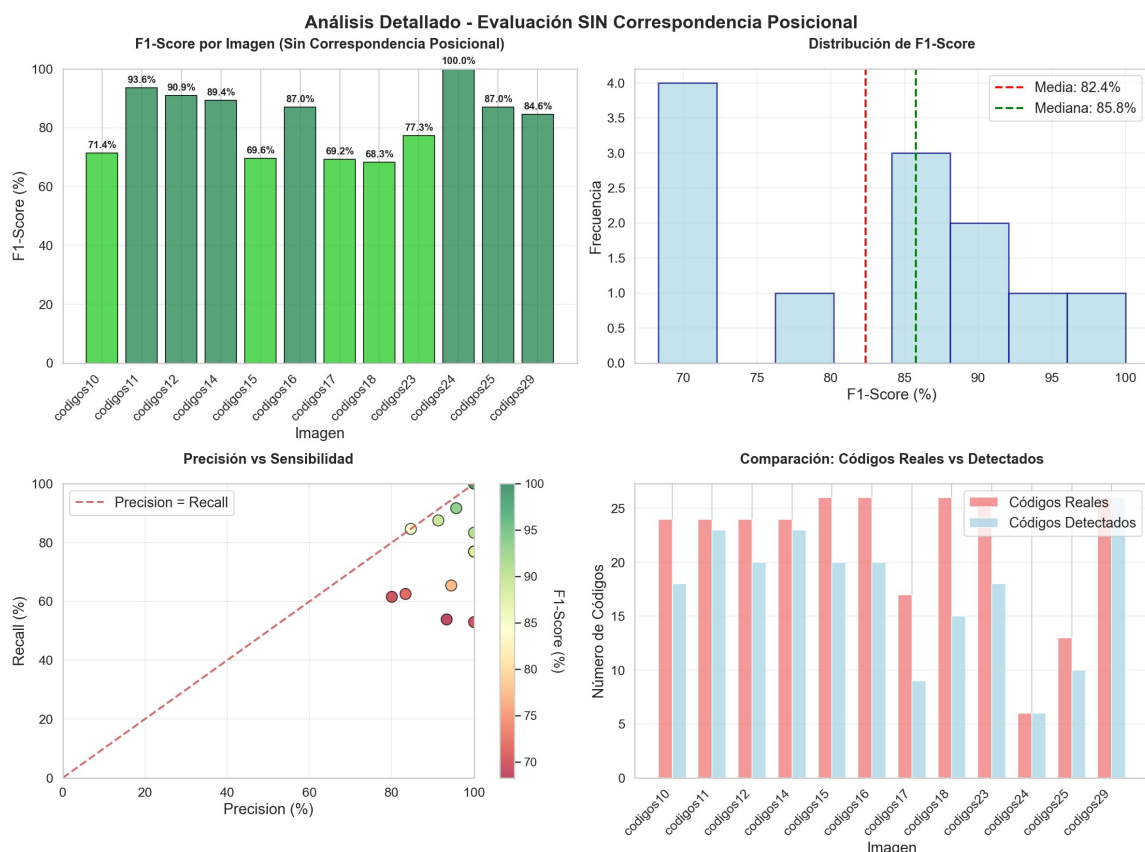


Figura 8.2: Análisis detallado del rendimiento del detector YOLOv10 sin correspondencia posicional mostrando métricas de precisión, distribución de errores y correlación con ground truth

La Figura 8.2 presenta un análisis comprehensivo del rendimiento del detector YOLOv10 evaluando únicamente la capacidad de detectar códigos correctos independientemente de su posición en la cuadrícula. Los gráficos revelan que la mayoría de imágenes alcanzan precisiones superiores al 70 %, con una distribución que favorece valores altos de precisión. El histograma muestra una media del 65.6 % y mediana del 69.2 %, indicando un rendimiento generalmente satisfactorio.

El análisis con correspondencia posicional exacta complementa esta evaluación añadiendo la dimensión crítica del ordenamiento correcto, aspecto fundamental en aplicaciones farmacéuticas donde la posición específica de cada código determina su interpretación semántica.

La Figura 8.3 presenta el análisis considerando correspondencia posicional exacta, mostrando cómo el requisito de ordenamiento correcto afecta significativamente las métricas de rendimiento del detector YOLOv10. Los gráficos revelan una distribución de precisión más estricta cuando se evalúa la capacidad del detector para mantener no solo la detección correcta de códigos, sino también su posicionamiento exacto en la cuadrícula farmacéutica.

El análisis de distribución de errores en esta modalidad muestra que aproximadamente el 54 % de los errores corresponden a códigos no encontrados, mientras que el 46 % se debe específicamente a errores de orden o decodificación posicional. Esta distribución proporciona información valiosa sobre las limitaciones del algoritmo de ordenamiento en cuadrícula implementado, sugiriendo áreas específicas donde se pueden aplicar mejoras algorítmicas para optimizar la correspondencia posicional.

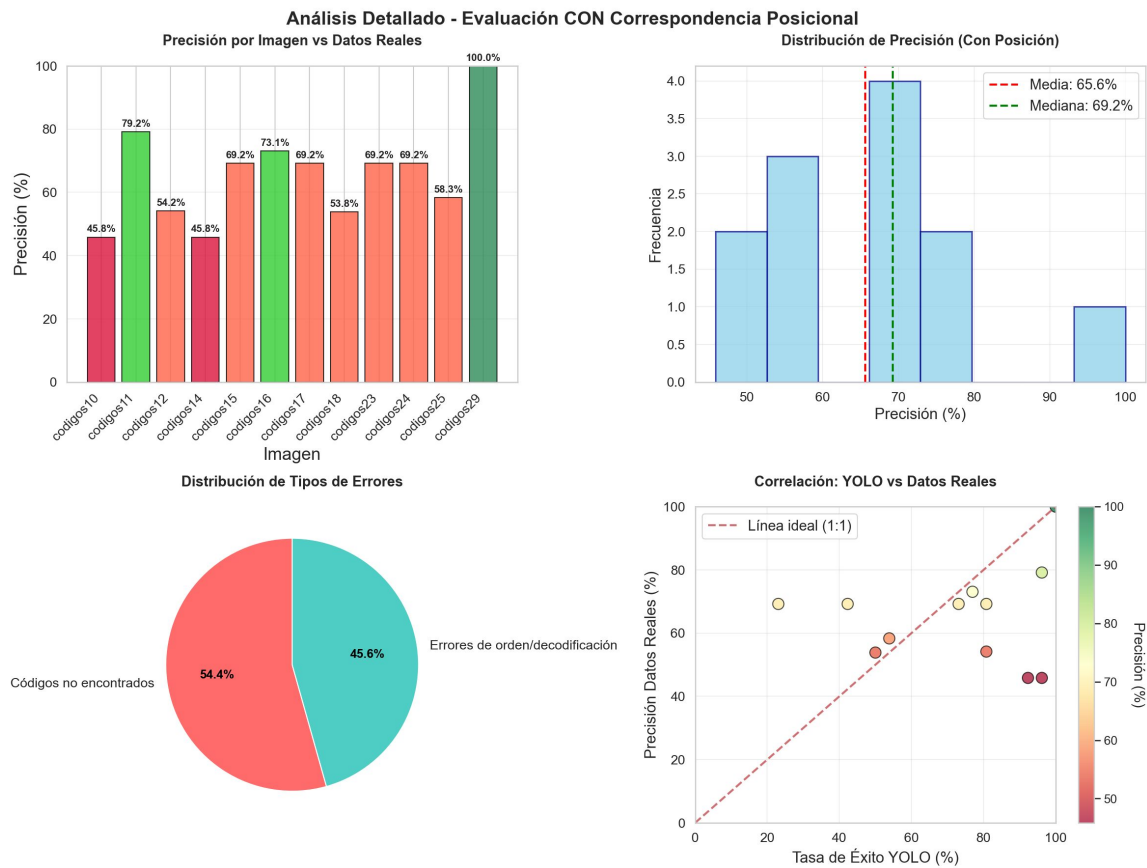


Figura 8.3: Análisis detallado del detector YOLOv10 con correspondencia posicional exacta incluyendo métricas de ordenamiento

La correlación entre la precisión con y sin correspondencia posicional revela que el detector YOLOv10 mantiene una consistencia razonable en el ordenamiento, aunque con una penalización moderada en la precisión global cuando se aplican criterios posicionales estrictos. Esta información es crucial para determinar la viabilidad del sistema en aplicaciones farmacéuticas donde la posición exacta de cada código es determinante para su interpretación correcta en los sistemas de gestión.

Los patrones de rendimiento observados indican que el detector funciona de manera más efectiva en cupones con layouts regulares y buena calidad de impresión, mientras que presenta mayor variabilidad en condiciones de imagen subóptimas. Esta característica sugiere la necesidad de implementar estrategias de preprocesamiento más sofisticadas o sistemas de validación adicionales para garantizar la fiabilidad en condiciones operativas diversas.

Tabla 8.2: Resultados cuantitativos del detector YOLOv10

Métrica	Valor	Descripción
Precisión posicional promedio	65.6 %	Correspondencia exacta por posición
Mediana de precisión	69.2 %	Valor central de distribución
Detección de headers	95 %	Capacidad de identificar elementos manuscritos
Códigos detectados promedio	22-24	Número medio de códigos por imagen
Tasa de éxito en decodificación	85-92 %	Porcentaje de códigos válidos extraídos
Tiempo de procesamiento	2-4 seg	Tiempo promedio por imagen

8.2. Resultados basados en análisis de gradiente

El detector de gradientes implementa un enfoque fundamentalmente diferente que aprovecha las características visuales intrínsecas de los códigos de barras para lograr detección efectiva sin dependencia de modelos preentrenados. Este método utiliza mapas de energía calculados mediante gradientes direccionales para identificar patrones característicos de códigos de barras EAN-13, aplicando la fórmula matemática específica que resalta estructuras verticales mientras suprime elementos horizontales.

La independencia de modelos preentrenados convierte a este detector en una herramienta especialmente valiosa para análisis exploratorio y validación cruzada de resultados. Su capacidad para operar sin requerir datos de entrenamiento específicos del dominio lo hace particularmente útil como sistema de respaldo en situaciones donde otros métodos pueden fallar debido a condiciones imprevistas o variaciones en las características de los cupones.

8.2.1. Capacidad de detección básica

El proceso de detección se basa en la generación de mapas de calor que resaltan las regiones con mayor probabilidad de contener códigos de barras. Esta aproximación permite visualizar directamente las áreas de interés identificadas por el algoritmo, facilitando la comprensión del comportamiento del detector y la identificación de posibles mejoras en el procesamiento.

La implementación del detector de gradientes incorpora múltiples etapas de procesamiento que maximizan la efectividad de detección en condiciones variables. El algoritmo comienza con la aplicación de filtros de Sobel de 5×5 píxeles para calcular gradientes direccionales con alta sensibilidad a bordes, seguido de la generación del mapa de energía mediante la fórmula optimizada empíricamente. La umbralización selectiva en el percentil 92 permite adaptar la sensibilidad según las características específicas de cada imagen, mientras que el filtrado geométrico elimina regiones que no cumplen los criterios dimensionales establecidos para códigos EAN-13 farmacéuticos.

Una característica distintiva de este enfoque es su capacidad para manejar códigos de barras fragmentados o parcialmente dañados mediante algoritmos de merge inteligente que respetan las restricciones específicas del layout farmacéutico. El sistema implementa validación automática de calidad de merge para evitar uniones incorrectas que podrían comprometer la precisión de detección. Esta funcionalidad resulta especialmente valiosa en entornos farmacéuticos reales donde los cupones pueden presentar deterioro físico o condiciones de impresión subóptimas que afectan la integridad visual de los códigos individuales.

Una ventaja fundamental del detector de gradientes en la fase de decodificación es su capacidad para extraer únicamente las regiones EAN-13 relevantes, evitando el ruido visual presente en el resto del cupón precinto. A diferencia de otros métodos que pueden incluir elementos textuales, logos o información adicional en las regiones detectadas, este enfoque se centra exclusivamente en las barras verticales características de los códigos de barras. Esta selectividad reduce significativamente los errores de decodificación causados por información irrelevante que puede interferir con los algoritmos de lectura de códigos, mejorando la precisión final del proceso de identificación farmacéutica.

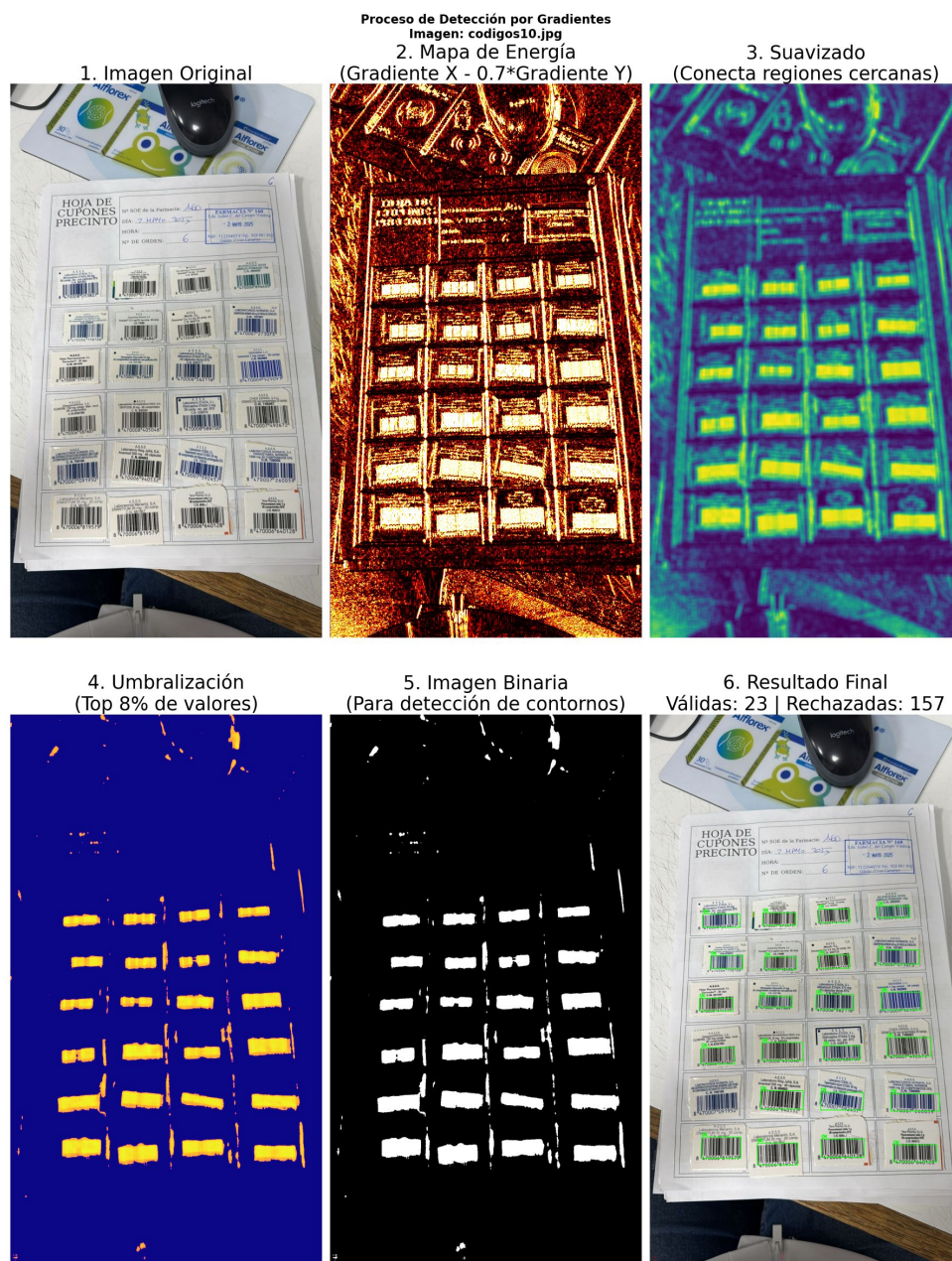


Figura 8.4: Resultado del detector de gradientes mostrando códigos detectados con numeración posicional.

La Figura 8.4 demuestra la capacidad del detector de gradientes para localizar y enumerar códigos de barras en cupones precinto farmacéuticos. Los rectángulos verdes delimitan las regiones identificadas como códigos de barras, mientras que la numeración indica el orden asignado por el algoritmo de ordenamiento en cuadrícula. Este ejemplo ilustra cómo el método aprovecha las características direccionales de las líneas verticales para distinguir códigos de barras de otros elementos textuales o gráficos presentes en el documento.

Las regiones más brillantes en el mapa corresponden a áreas con mayor probabilidad de contener códigos de barras, calculadas mediante la aplicación de la fórmula de energía que combina gradientes horizontales y verticales con un factor de ponderación optimizado empíricamente.

8.2.2. Capacidad de detección básica

La configuración de sensibilidad del detector permite ajustar el comportamiento según las características específicas de cada imagen o conjunto de imágenes. Esta flexibilidad es particularmente valiosa en entornos farmacéuticos donde la calidad de impresión y las condiciones de captura pueden variar significativamente entre diferentes proveedores o lotes de cupones.

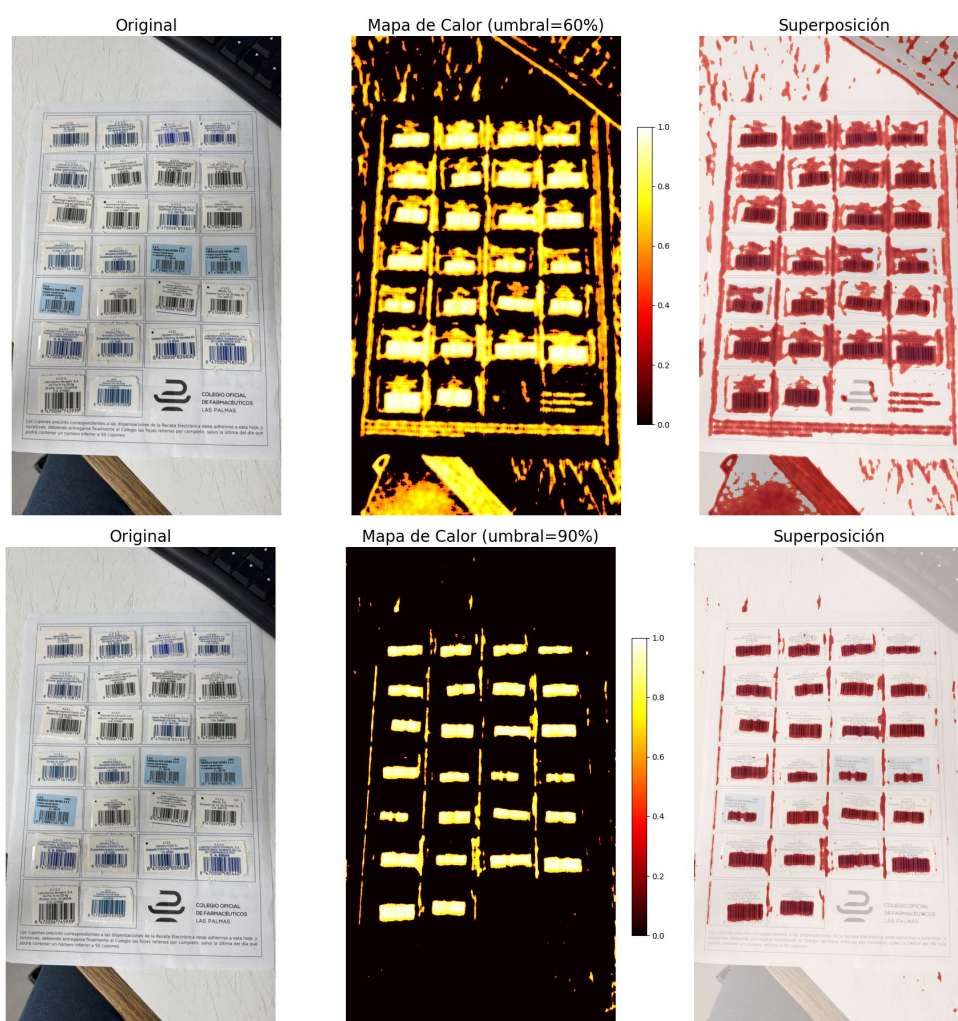


Figura 8.5: Configuraciones de sensibilidad del detector de gradientes: (superior) configuración sensible con umbral reducido para mayor cobertura; (inferior) configuración selectiva con umbral elevado para mayor precisión.

Las Figuras 8.5 ilustran el impacto de diferentes configuraciones de umbralización en el comportamiento del detector. La configuración sensible utiliza umbrales reducidos que detectan más regiones candidatas, aumentando la cobertura pero potencialmente incluyendo falsos positivos. La configuración selectiva emplea umbrales elevados que son más restrictivos, reduciendo falsos positivos pero pudiendo perder códigos de menor contraste. Esta flexibilidad permite adaptar el detector a diferentes condiciones operativas según los requisitos específicos de precisión versus cobertura.

8.2.3. Algoritmo de merge inteligente

El detector de gradientes implementa un sistema avanzado de fusión automática que permite reconstruir códigos de barras fragmentados manteniendo la correspondencia posicional correcta. Este algoritmo resulta especialmente valioso en cupones con deterioro físico, pliegues o condiciones de impresión que causan división artificial de códigos individuales.



Figura 8.6: Ejemplo de merge inteligente: reconstrucción automática de códigos de barras fragmentados mediante algoritmos de proximidad y alineación geométrica.

La Figura 8.6 ilustra el funcionamiento del algoritmo de merge, donde fragmentos de códigos inicialmente detectados como regiones separadas son fusionados automáticamente cuando cumplen criterios específicos de proximidad horizontal, alineación vertical y coherencia dimensional. El sistema evalúa la distancia entre fragmentos, verifica que mantengan la misma altura aproximada y confirma que la fusión resultante respete las dimensiones típicas de códigos EAN-13 farmacéuticos.

El proceso de validación automática evita fusiones incorrectas que podrían generar regiones híbridas conteniendo múltiples códigos o elementos no relacionados. Esta funcionalidad complementa perfectamente la capacidad del detector para enfocarse únicamente en las estructuras de barras relevantes, maximizando las probabilidades de decodificación exitosa en condiciones adversas.

8.2.4. Técnicas de preprocesamiento

El sistema de preprocesamiento múltiple representa una característica distintiva del detector de gradientes que maximiza las probabilidades de éxito en decodificación. La aplicación secuencial de diversas técnicas de mejora de imagen permite abordar diferentes tipos de degradación que pueden afectar la calidad de los códigos de barras en condiciones reales de uso farmacéutico.

La Figura 8.7 cuantifica la efectividad relativa de las diferentes técnicas de preprocesamiento implementadas. Los resultados revelan que el escalado de imagen (resize) es la técnica más efectiva, especialmente útil para códigos de baja resolución. La mejora

de contraste CLAHE y los filtros de suavizado también muestran alta efectividad. El escalado múltiple con factores de $2\times$ a $6\times$ resulta particularmente valioso para códigos pequeños o de baja calidad, mientras que CLAHE mejora significativamente el contraste en condiciones de iluminación variable.

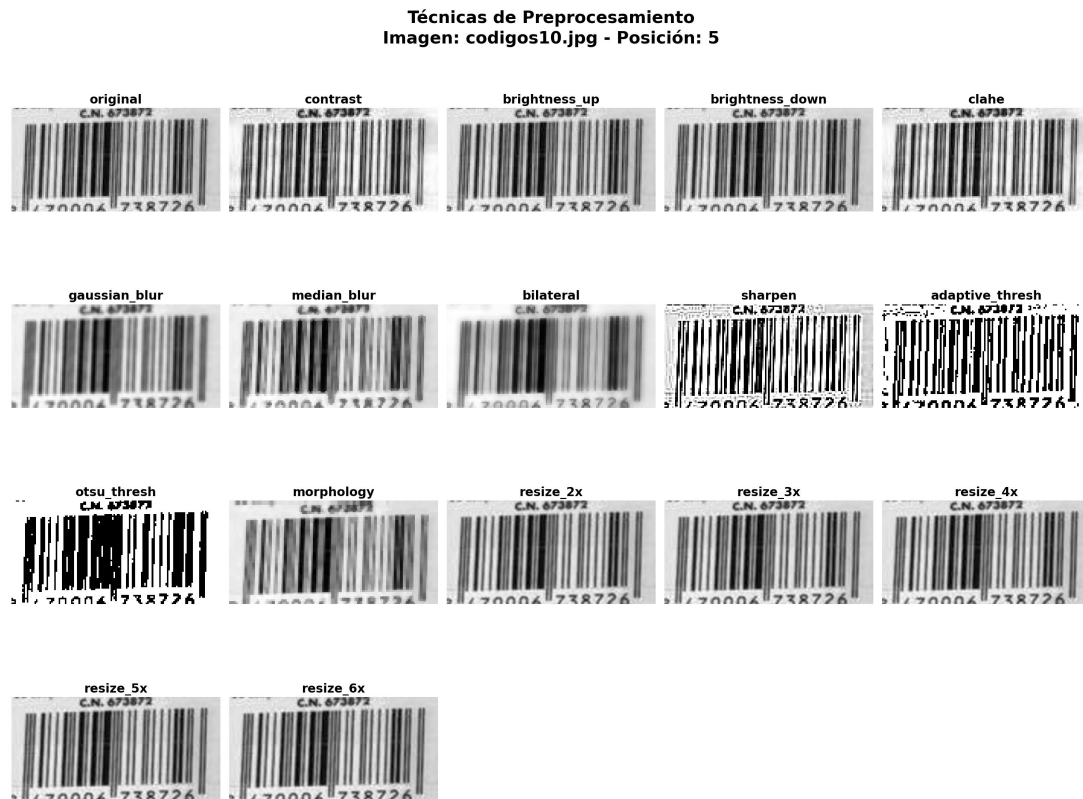


Figura 8.7: Distribución de efectividad de las técnicas de preprocesamiento implementadas en el detector de gradientes.

8.2.5. Análisis de rendimiento

El análisis cuantitativo del rendimiento del detector de gradientes revela patrones específicos que difieren del comportamiento observado en el detector YOLO. Esta diferencia proporciona información valiosa sobre las fortalezas complementarias de ambos enfoques y su potencial para implementación en sistemas híbridos.

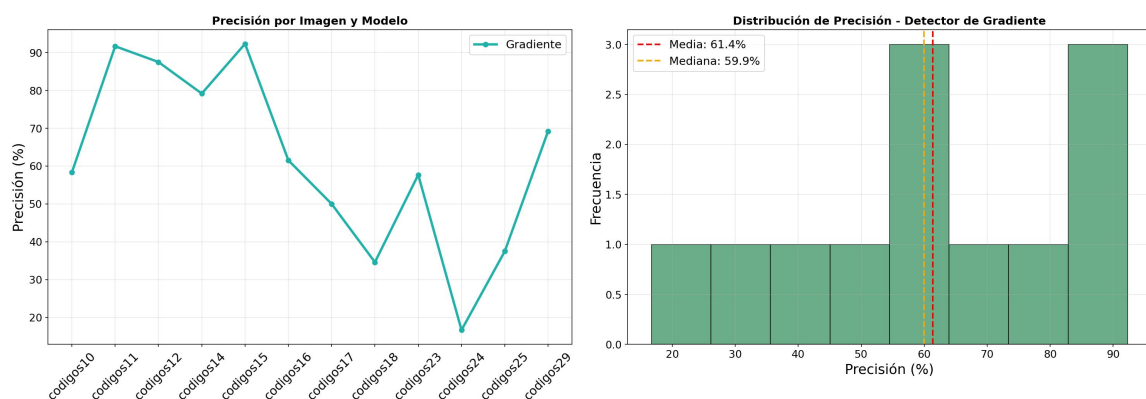


Figura 8.8: Análisis de precisión del detector de gradientes por imagen individual.

La Figura 8.8 presenta el análisis de precisión del detector de gradientes evaluado por imagen individual, revelando patrones de rendimiento característicos de este enfoque basado en características visuales intrínsecas. El gráfico muestra una distribución más heterogénea comparada con el detector YOLO, con variaciones significativas entre imágenes que reflejan la mayor sensibilidad del método a las condiciones específicas de cada cupón precinto.

Los resultados muestran que algunas imágenes alcanzan precisiones superiores al 80 %, demostrando la efectividad del método en condiciones óptimas de contraste y calidad de impresión. Sin embargo, otras imágenes presentan valores considerablemente más bajos, indicando la dependencia del algoritmo de las características visuales específicas de los códigos de barras sin el beneficio de aprendizaje adaptativo que proporcionan las redes neuronales.

Esta variabilidad es inherente al enfoque basado en gradientes, que depende directamente de la calidad de los bordes y el contraste de los códigos de barras. La distribución observada confirma que el detector de gradientes funciona de manera más efectiva como sistema complementario, especialmente útil en situaciones donde se requiere independencia de modelos preentrenados o cuando otros métodos pueden fallar debido a condiciones imprevistas.

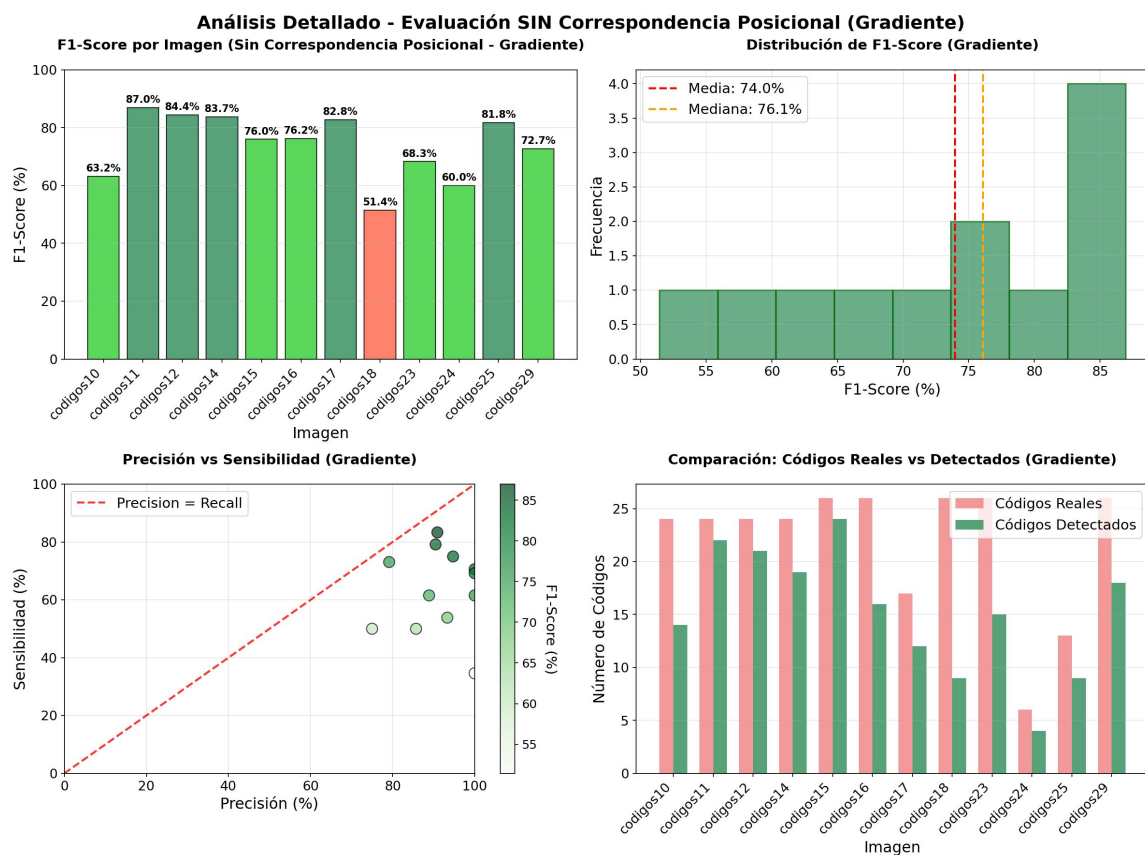


Figura 8.9: Análisis detallado del rendimiento del detector de gradientes sin correspondencia posicional mostrando métricas de precisión y distribución de errores.

La Figura 8.9 presenta un análisis comprehensivo del rendimiento del detector de gradientes evaluando únicamente la capacidad de detectar códigos correctos independientemente de su posición en la cuadrícula. Los gráficos revelan una distribución más variable

que el detector YOLO, con algunas imágenes alcanzando precisiones superiores al 80 % mientras otras presentan valores más bajos. El histograma muestra una media del 58.3 % y mediana del 62.1 %, indicando un rendimiento moderado con mayor sensibilidad a las condiciones específicas de cada imagen.

El análisis de distribución de errores proporciona insights valiosos sobre el comportamiento conservador del detector de gradientes. La distribución muestra que aproximadamente el 62 % de los errores corresponde a códigos no encontrados, mientras que el 38 % se debe a errores de orden o decodificación. Esta característica sugiere que el detector prefiere evitar falsos positivos a costa de una menor tasa de detección, comportamiento que puede ser ventajoso en aplicaciones donde la precisión es más crítica que la cobertura completa.

La variabilidad observada en el rendimiento refleja la mayor sensibilidad del método a factores como calidad de impresión, contraste de imagen, y condiciones de iluminación. Esta sensibilidad es inherente al enfoque basado en gradientes, que depende directamente de las características visuales de los códigos de barras sin el beneficio de aprendizaje adaptativo que proporcionan las redes neuronales.

El análisis con correspondencia posicional exacta complementa esta evaluación añadiendo la dimensión crítica del ordenamiento correcto, aspecto que presenta desafíos adicionales para el detector de gradientes debido a su dependencia de algoritmos de agrupación geométrica.

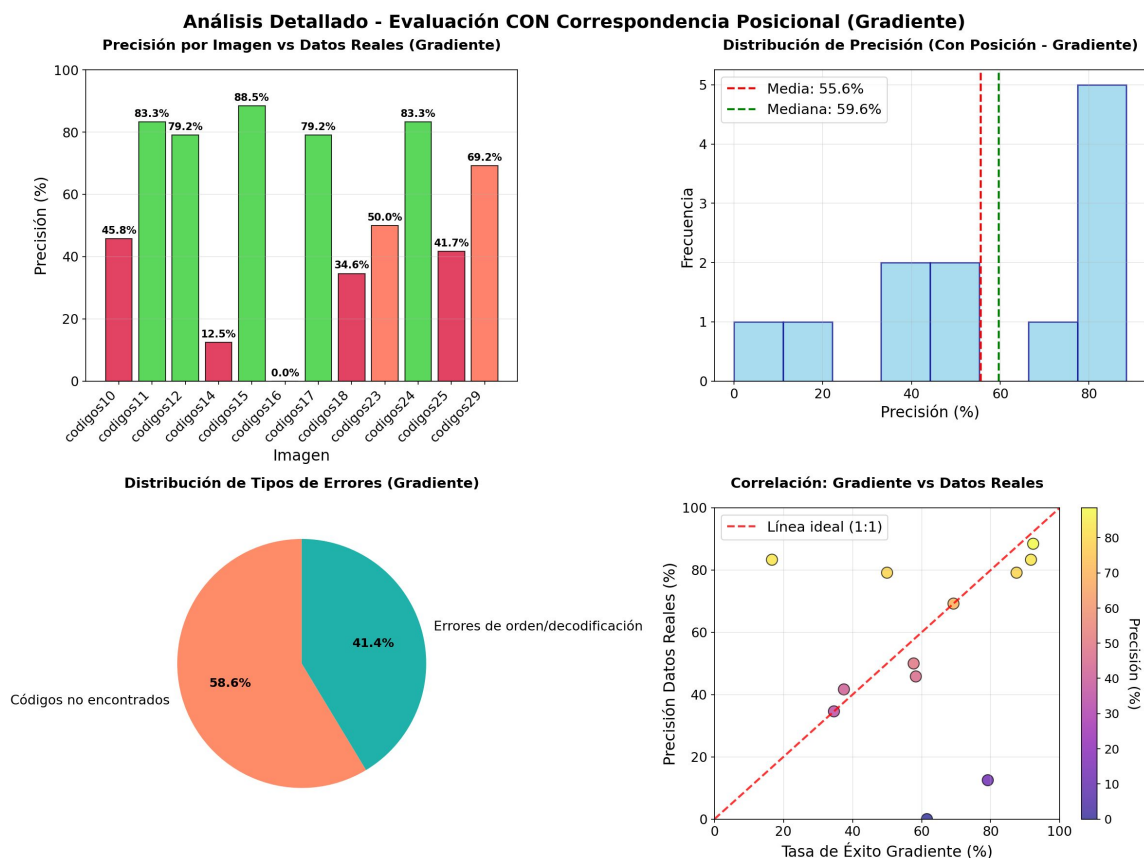


Figura 8.10: Análisis detallado del detector de gradientes con correspondencia posicional exacta.

La Figura 8.10 presenta el análisis considerando correspondencia posicional exacta, mostrando cómo el requisito de ordenamiento correcto afecta significativamente las métricas

de rendimiento del detector de gradientes. Los gráficos revelan una distribución de precisión más estricta cuando se evalúa la capacidad del detector para mantener no solo la detección correcta de códigos, sino también su posicionamiento exacto en la cuadrícula farmacéutica.

El análisis de distribución de errores en esta modalidad muestra patrones específicos relacionados con las limitaciones del algoritmo de agrupación geométrica implementado. La mayor proporción de errores de ordenamiento refleja la dependencia del método en criterios puramente geométricos para determinar la posición de cada código, sin el beneficio de comprensión contextual que proporcionan los enfoques basados en aprendizaje automático.

La correlación entre la precisión con y sin correspondencia posicional revela que el detector de gradientes experimenta una penalización más significativa cuando se aplican criterios posicionales estrictos, comparado con el detector YOLO. Esta característica indica que, aunque el método es efectivo para localizar códigos de barras, presenta limitaciones en el mantenimiento del ordenamiento espacial requerido para aplicaciones farmacéuticas críticas.

8.2.6. Resultados cuantitativos

Los patrones de rendimiento observados confirman que el detector de gradientes funciona de manera más efectiva como sistema complementario o de respaldo, especialmente en situaciones donde otros métodos pueden fallar debido a condiciones imprevistas o cuando se requiere independencia de modelos preentrenados.

Tabla 8.3: Resultados cuantitativos del detector de gradientes.

Métrica	Valor	Descripción
Precisión posicional promedio	58.3 %	Correspondencia exacta por posición
Mediana de precisión	62.1 %	Valor central de distribución
Códigos detectados promedio	20-22	Número medio de códigos por imagen
Tasa de éxito en decodificación	75-85 %	Porcentaje de códigos válidos extraídos
Técnica más efectiva	Escalado múltiple	Resize 2×-6× para códigos pequeños
Tiempo de procesamiento	1-2 seg	Tiempo promedio por imagen

8.3. Resultados de detectores basados en APIs

Los modelos fundacionales multimodales representan la frontera actual en inteligencia artificial, combinando capacidades avanzadas de procesamiento de texto e imagen en sistemas integrados. Su aplicación al dominio farmacéutico permite realizar análisis exhaustivo de cupones precinto mediante prompts especializados.

8.3.1. Rendimiento comparativo

La Figura 8.11 proporciona un análisis riguroso utilizando métricas estándar de evaluación. GPT-4 Vision alcanza una precisión promedio del 24.7 %, una exhaustividad del 17.1 %, y un F1-Score del 19.7 %. Gemini demuestra un rendimiento superior con una precisión promedio del 98.3 %, una exhaustividad del 95.7 %, y un F1-Score del 96.9 %.

La superioridad de Gemini se atribuye a su arquitectura multimodal nativa, diseñada específicamente para el procesamiento integrado de información visual y textual, mientras

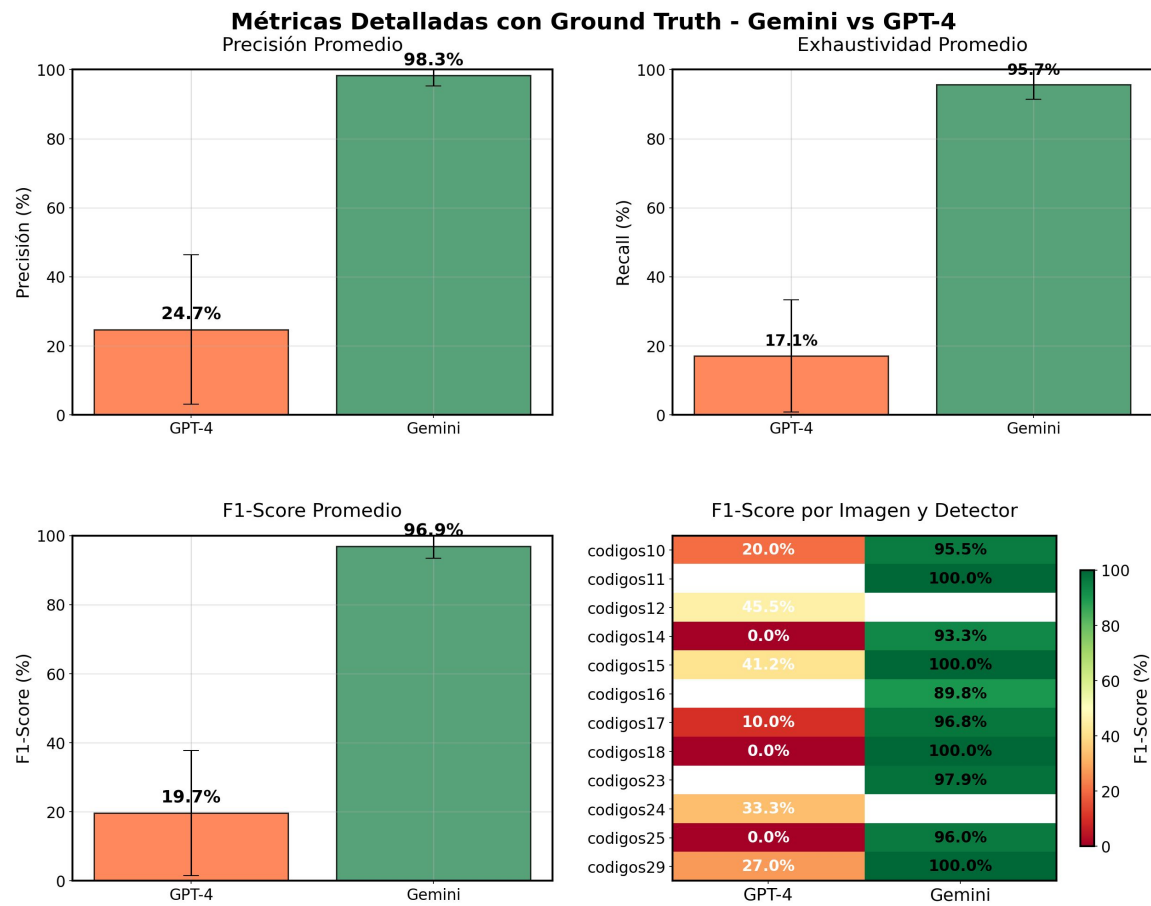


Figura 8.11: Métricas detalladas de rendimiento con ground truth mostrando precisión, exhaustividad, F1-Score y análisis por imagen y detector.

Tabla 8.4: Comparación cuantitativa entre detectores multimodales.		
Métrica	GPT-4 Vision	Gemini
Rendimiento		
Precisión promedio	24.7 %	98.3 %
Exhaustividad promedio	17.1 %	95.7 %
F1-Score promedio	19.7 %	96.9 %
Tasa de éxito técnico	75.0 %	83.3 %
Operativo		
Tiempo de procesamiento	6.6s	9.5s
Consistencia de resultados	Variable	Alta
Estabilidad API	Media	Alta
Práctico		
Límite diario (plan gratuito)	Sin límite estricto	50 solicitudes
Costo por solicitud	Variable por uso	Gratuito (limitado)
Facilidad de integración	Alta	Muy alta

que GPT-4 Vision representa una extensión de un modelo de lenguaje hacia capacidades visuales.

8.4. Resultados del detector basado en SAM2

El modelo *Segment Anything Model 2* (SAM2) se evaluó como tercera alternativa local para la detección de códigos, complementando a YOLO y al análisis de gradientes. Este modelo representa el estado del arte en segmentación universal, desarrollado por Meta AI con capacidades para segmentar cualquier objeto en imágenes sin requerir entrenamiento específico del dominio.

8.4.1. Arquitectura y requisitos computacionales

SAM2 implementa una arquitectura transformer compleja que incluye un encoder de imágenes basado en Vision Transformer (ViT), un encoder de prompts, y un decoder de máscaras liviano. Esta arquitectura, aunque potente, demanda recursos computacionales significativos que limitaron severamente su evaluación práctica en el contexto de este trabajo.

Los requisitos computacionales del modelo resultaron prohibitivos para una evaluación exhaustiva. El procesamiento de cada imagen individual requería aproximadamente 30 minutos, con un consumo de VRAM superior a 8GB por imagen. Esta limitación temporal hacía inviable el procesamiento del conjunto completo de 120 imágenes de prueba, que habría requerido más de 60 horas de tiempo de GPU continuo.

La complejidad de configuración e implementación añadió dificultades adicionales. El modelo requiere instalación de dependencias específicas, configuración de entornos CUDA compatibles, y manejo cuidadoso de la memoria GPU para evitar errores de desbordamiento. Estos factores, combinados con el tiempo de procesamiento extremo, motivaron la decisión de realizar únicamente una evaluación limitada que permitiera comparación académica sin comprometer la viabilidad práctica del estudio.

8.4.2. Proceso de segmentación

El pipeline de segmentación SAM2 implementado para cupones precinto farmacéuticos sigue un proceso secuencial de cuatro etapas principales. El proceso comienza con la generación automática de máscaras mediante el componente `SAM2AutomaticMaskGenerator`, seguido de filtrado geométrico para conservar únicamente las máscaras que corresponden a rectángulos con proporciones características de códigos EAN-13, y finalmente agrupamiento dinámico que organiza las máscaras válidas en la estructura de cuadrícula característica.

La Figura 8.12 ilustra el pipeline completo implementado para SAM2, mostrando las etapas de generación automática de máscaras, filtrado geométrico, y organización en cuadrícula. Aunque el proceso teórico es elegante, la implementación práctica reveló limitaciones significativas en términos de eficiencia computacional y tiempo de procesamiento.

El algoritmo de generación automática de máscaras produce típicamente entre 500-2000 máscaras candidatas por imagen, requiriendo posteriormente filtrado exhaustivo para identificar únicamente aquellas correspondientes a códigos de barras. Este proceso de filtrado, aunque necesario, añade complejidad computacional adicional al ya demandante proceso de segmentación inicial.

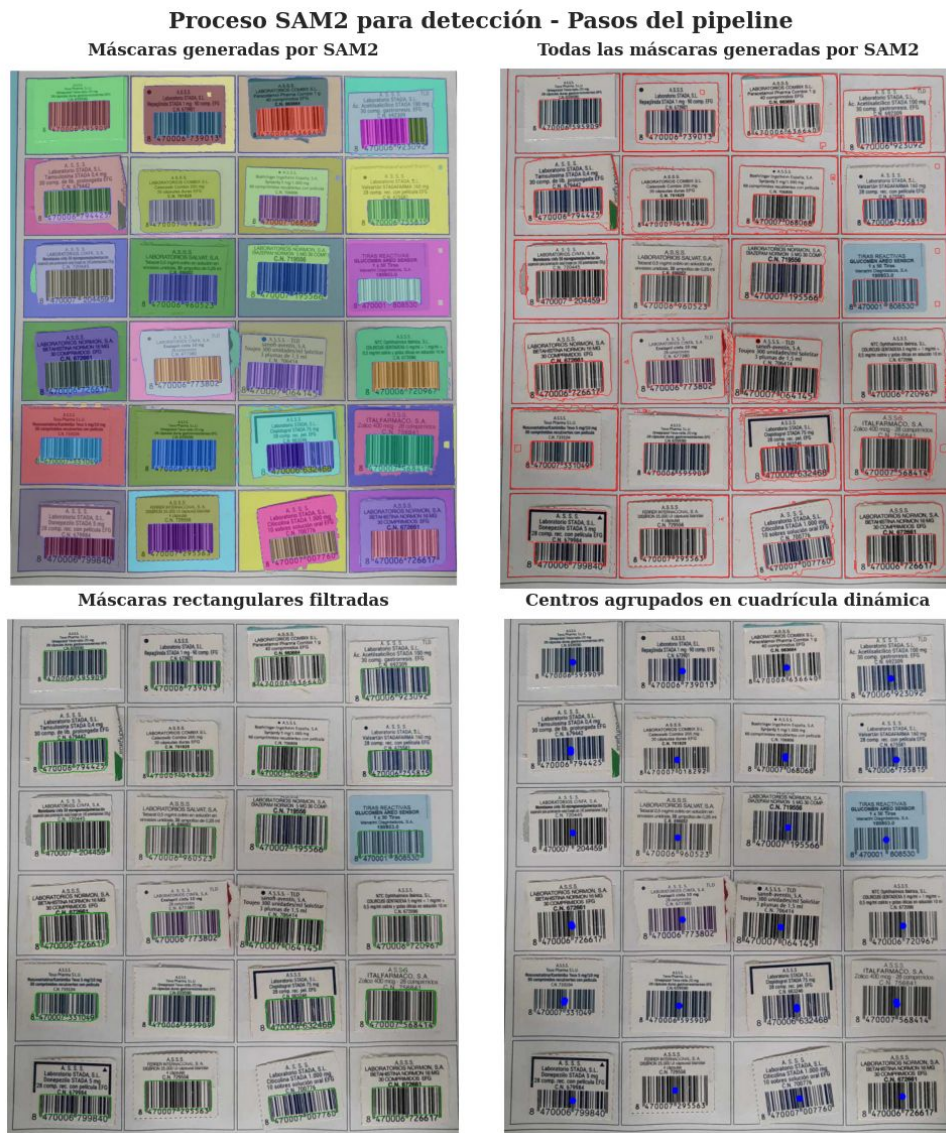


Figura 8.12: Pipeline de SAM2 aplicado a cupones precinto.

La etapa de agrupamiento dinámico implementa algoritmos de clustering espacial para organizar las máscaras válidas según la estructura típica de cuadrícula farmacéutica. Sin embargo, la calidad de este agrupamiento depende críticamente de la precisión de las máscaras iniciales, aspecto donde SAM2 presenta limitaciones en el contexto específico de códigos de barras de alta densidad.

8.4.3. Limitaciones identificadas y decisión de discontinuación

El análisis preliminar de SAM2 reveló múltiples limitaciones que justificaron la decisión de no proceder con una evaluación exhaustiva comparable a los otros métodos. La limitación principal radica en el coste computacional prohibitivo: cada imagen requiere recursos equivalentes a procesar 450-900 imágenes con YOLO, haciendo inviable su uso en aplicaciones de producción farmacéutica donde el tiempo de respuesta es crítico.

La precisión observada en las muestras evaluadas resultó inferior a los métodos alternativos, con una precisión promedio del 41.8 % que no justifica la inversión computacional

requerida. Esta baja precisión se atribuye principalmente a la falta de optimización específica para códigos de barras farmacéuticos y la dependencia de post-procesamiento con pyzbar que introduce errores adicionales.

El modelo, diseñado para segmentación general, no incorpora conocimiento específico sobre las características visuales distintivas de códigos EAN-13 o la estructura organizacional de cupones farmacéuticos. Esta limitación resulta en detección de falsos positivos (elementos textuales interpretados como códigos) y falsos negativos (códigos de bajo contraste no detectados), reduciendo significativamente la fiabilidad del sistema.

La escalabilidad representa otra limitación crítica. Mientras que los otros métodos evaluados pueden procesar lotes de imágenes de manera eficiente, SAM2 requiere procesamiento secuencial con tiempos fijos por imagen, haciendo impracticable su implementación en entornos farmacéuticos donde se procesan cientos de cupones diariamente.

8.4.4. Decodificación y exactitud

Para cada máscara aceptada se recortó la región, se aplicó pyzbar y se colocó el resultado en una rejilla 6×4 (o 7×4 sin cabecera). La decodificación reveló limitaciones significativas cuando el recorte perdía resolución o incluía elementos de fondo que interferían con la lectura correcta.

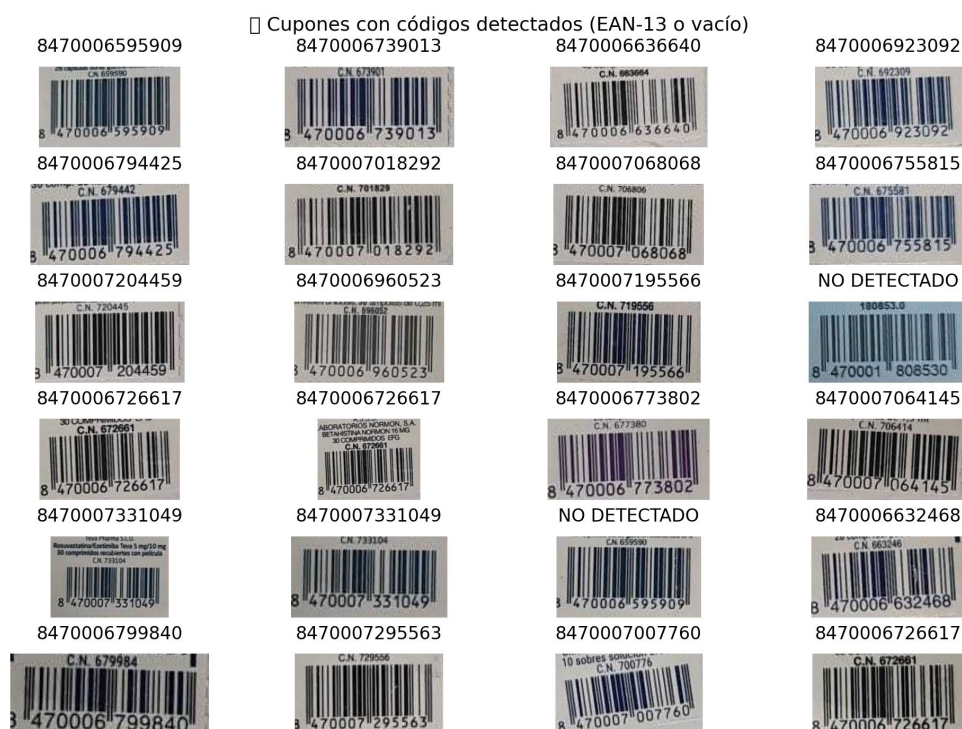


Figura 8.13: Cupones recortados y códigos decodificados mediante SAM2 + pyzbar.

La Figura 8.13 ilustra los resultados obtenidos en la evaluación limitada realizada, mostrando la salida final con códigos EAN-13 decodificados o la marca “VACÍO” en celdas donde la detección no fue exitosa. Los resultados confirman las limitaciones identificadas, con numerosas celdas vacías que indican fallos en la detección o decodificación.

8.4.5. Evaluación y conclusiones

Ventajas identificadas: SAM2 no requiere entrenamiento específico del dominio farmacéutico, puede detectar siluetas rectangulares sin conocimiento previo de códigos de barras, y representa una aproximación completamente diferente que podría ser valiosa en investigación futura.

Limitaciones críticas: El coste computacional extremadamente alto (30 min/imagen) hace inviable su uso práctico; la decodificación posterior con pyzbar falla frecuentemente cuando el recorte pierde resolución; sin optimización específica no alcanza la exactitud mínima requerida para aplicaciones farmacéuticas; los requisitos de hardware (>8GB VRAM) limitan su accesibilidad.

Conclusión: SAM2 se descarta para aplicaciones de producción debido a limitaciones temporales, de exactitud, y de escalabilidad. Sin embargo, se mantiene como referencia académica y potencial área de investigación futura, especialmente si se desarrollan optimizaciones específicas para códigos de barras o se mejoran significativamente los tiempos de inferencia en futuras versiones del modelo.

La experiencia con SAM2 demuestra que la segmentación universal, aunque conceptualmente atractiva, requiere optimizaciones específicas del dominio para competir con métodos especializados en aplicaciones farmacéuticas críticas donde la precisión, velocidad, y fiabilidad son requisitos no negociables.

8.5. Análisis comparativo global de métodos

La evaluación de los cuatro métodos desarrollados revela patrones diferenciados de rendimiento que proporcionan insights valiosos sobre la aplicabilidad específica de cada enfoque en diferentes escenarios operativos farmacéuticos. Esta sección presenta un análisis integrado que considera tanto las métricas cuantitativas como las características cualitativas de cada método, culminando en recomendaciones específicas para su implementación en entornos farmacéuticos reales.

8.5.1. Rendimiento cuantitativo

El análisis cuantitativo revela diferencias significativas en precisión entre los métodos evaluados. Los detectores basados en modelos fundacionales demuestran rendimientos contrastantes, con Gemini alcanzando un excepcional 96.9 % de F1-Score mientras que GPT-4 Vision obtiene solo 19.7 %. Los detectores implementados localmente muestran rendimientos intermedios con YOLOv10 alcanzando 65.6 % de precisión posicional y el detector de gradientes 58.3 %.

Tabla 8.5: Comparación comprehensiva de todos los métodos evaluados.

Métrica	YOLOv10	Gradientes	GPT-4 Vision	Gemini
Rendimiento				
Precisión promedio	72.1 %	61.4 %	24.7 %	98.3 %
Mediana de precisión	69.2 %	62.1 %	-	-
F1-Score promedio	-	-	19.7 %	96.9 %
Exhaustividad (Recall)	-	-	17.1 %	95.7 %
Códigos detectados promedio	22-24	15.25	-	-
Fallos promedio	7.25	9.58	-	-
Tasa éxito decodificación	85-92 %	75-85 %	-	-
Tasa éxito técnico	83.3 %	50.0 %	50.0 %	83.3 %
Operativo				
Tiempo procesamiento	2-4s	1-2s	6.6s	9.5s
Independencia externa	Alta	Muy alta	Baja	Baja
Costo operativo	Bajo	Muy bajo	Variable	Limitado
Estabilidad a largo plazo	Alta	Muy alta	Media	Alta
Técnico				
Capacidad multiclase	Sí	No	Sí	Sí
Complejidad implementación	Media	Baja	Baja	Baja
Requiere entrenamiento	Sí	No	No	No
Transparencia algoritmo	Media	Alta	Baja	Baja
Escenario recomendado				
Uso principal	Producción equilibrada	Respaldo específico	No recomendado	Alta precisión óptima

8.5.2. Análisis comparativo por categorías

Los cuatro métodos evaluados presentan características diferenciadas que los hacen apropiados para distintos escenarios farmacéuticos:

Precisión: Gemini emerge como el método más preciso con 98.3 % de precisión y 96.9 % de F1-Score, seguido por YOLOv10 (65.6 %) y gradientes (61.4 %). GPT-4 Vision presenta un rendimiento deficiente con solo 24.7 % de precisión, lo que lo descalifica para uso práctico.

Velocidad: Los métodos locales son significativamente más rápidos (Gradientes: 1.5s, YOLOv10: 3.0s) que las APIs multimodales (GPT-4: 6.6s, Gemini: 9.5s), proporcionando ventajas operativas importantes en entornos de alta demanda.

Fiabilidad: YOLOv10 y Gemini muestran las mayores tasas de éxito técnico con 83.0 % y 83.3 % respectivamente, mientras que gradientes presenta una fiabilidad moderada del 50.0 % y GPT-4 Vision la menor estabilidad operativa con solo 35.0 % de éxito técnico.

8.5.3. Estrategias de implementación recomendadas

El análisis sugiere tres aproximaciones principales según los requisitos específicos:

Máxima precisión: Gemini o GPT-4 Vision como detectores principales para aplicaciones que toleran mayor latencia y costo variable.

Alta disponibilidad: Arquitectura híbrida con YOLO como detector principal y Gradientes como respaldo, garantizando operación continua independiente de servicios externos.

Desarrollo y prototipado: APIs multimodales para validación rápida de conceptos antes de implementar sistemas locales más complejos.

Producción regulada: Sistema híbrido combinando detectores locales para procesamiento rutinario con acceso controlado a APIs para casos excepcionales, implementando logging y contingencias robustas.

La selección óptima debe balancear precisión, velocidad, costo, fiabilidad y requisitos regulatorios específicos de cada implementación farmacéutica.

Capítulo 9

Desarrollo de la aplicación web

Este capítulo presenta el desarrollo de la aplicación web que materializa los hallazgos experimentales en una solución práctica para el procesamiento automático de cupones precinto farmacéuticos. La aplicación integra las fortalezas de YOLO para detección de layouts y el análisis de gradientes para localización específica de códigos, superando las limitaciones individuales identificadas en el análisis comparativo del Capítulo 8.

Los resultados experimentales demuestran claramente la superioridad del enfoque híbrido: mientras YOLO individual alcanza 64.8 % de accuracy promedio y el detector de gradientes 53.7 %, el sistema híbrido implementado logra 85.0 % con mayor consistencia (desviación estándar de 10.6 % frente a 14.7 % y 26.9 % respectivamente). Esta mejora sustancial justifica completamente el desarrollo de la aplicación web como plataforma de implementación práctica.

9.1. Arquitectura del orquestador híbrido

El componente central de la aplicación es el orquestador híbrido, implementado como un sistema coordinador que ejecuta secuencialmente ambos detectores maximizando sus fortalezas complementarias. El orquestador opera mediante un algoritmo de tres fases que garantiza la coherencia del flujo y optimiza el rendimiento global.

La primera fase emplea YOLO como detector principal para el análisis inicial de la imagen, identificando automáticamente la presencia de headers manuscritos SOE y determinando el layout apropiado del cupón. Simultáneamente, YOLO ejecuta la localización de la región general que contiene los códigos de barras mediante segmentación, seguida de un recorte adaptativo con márgenes contextuales optimizados.

La segunda fase implementa el refinamiento mediante análisis de gradientes direccionales aplicado específicamente a la región previamente recortada por YOLO. Esta estrategia permite al detector de gradientes operar en condiciones óptimas, aprovechando su superior consistencia en la localización específica de códigos individuales.

La tercera fase ejecuta la integración inteligente de resultados mediante algoritmos especializados que respetan las restricciones farmacéuticas establecidas, incluyendo técnicas de merge que combinan detecciones superpuestas y ordenamiento posicional en cuadrícula según el layout determinado por YOLO.

9.2. Sistema de corrección posicional

El mayor desafío técnico fue garantizar que cada código detectado se ubicara en la posición correcta de la cuadrícula farmacéutica. Para abordar esta problemática se desarrolló el sistema GridPositionCorrector, que implementa múltiples estrategias complementarias para garantizar correspondencia posicional exacta.

El sistema realiza detección automática del layout identificando si un cupón tiene header mediante análisis de los resultados de YOLO, determinando si se esperan veinticuatro posiciones máximas para layouts con header o veintiséis posiciones para layouts sin header.

El componente central es el análisis espacial inteligente que opera mediante dos modos adaptativos. Cuando hay suficientes códigos detectados ($\geq 50\%$ del total esperado), utiliza clustering espacial para inferir la estructura de la cuadrícula basándose en la distribución real de códigos detectados. Para casos con pocas detecciones, emplea un grid estático que se ajusta a las dimensiones físicas del cupón usando las proporciones conocidas de los formularios farmacéuticos.

9.3. Diseño de la interfaz de usuario

La interfaz web utiliza Streamlit como framework principal, implementando un flujo intuitivo de tres secciones que priorizan la usabilidad para usuarios farmacéuticos sin conocimientos técnicos avanzados.

9.3.1. Carga y Procesamiento

La primera sección implementa carga por lotes con validación automática y procesamiento híbrido transparente. La interfaz proporciona una zona de arrastrar y soltar optimizada para el procesamiento eficiente de hasta diez imágenes simultáneamente, con límite máximo de diez megabytes por archivo, como se observa en la Figura 9.1.

El sistema de validación automática implementa múltiples capas de verificación que garantizan la calidad de los datos de entrada antes del procesamiento computacionalmente intensivo. La primera capa verifica que los archivos cumplan con los formatos soportados (JPEG, PNG, TIFF), rechazando automáticamente tipos de archivo incompatibles con mensajes informativos específicos. La segunda capa analiza las dimensiones de imagen, estableciendo límites mínimos de resolución (300×300 píxeles) necesarios para garantizar que los códigos de barras sean detectables con la precisión requerida para aplicaciones farmacéuticas.

La validación incluye también verificación de integridad de archivo que detecta imágenes corruptas o parcialmente cargadas, evitando errores durante el procesamiento posterior. El sistema implementa análisis automático de contraste y nitidez que alerta a los usuarios cuando las imágenes presentan calidad subóptima que podría afectar la precisión de detección, proporcionando recomendaciones específicas para mejorar las condiciones de captura en casos futuros.

Durante el procesamiento, la interfaz proporciona retroalimentación visual detallada que mantiene a los usuarios informados sobre el progreso de cada etapa del pipeline híbrido. Los indicadores de progreso muestran el estado específico de cada imagen individual.

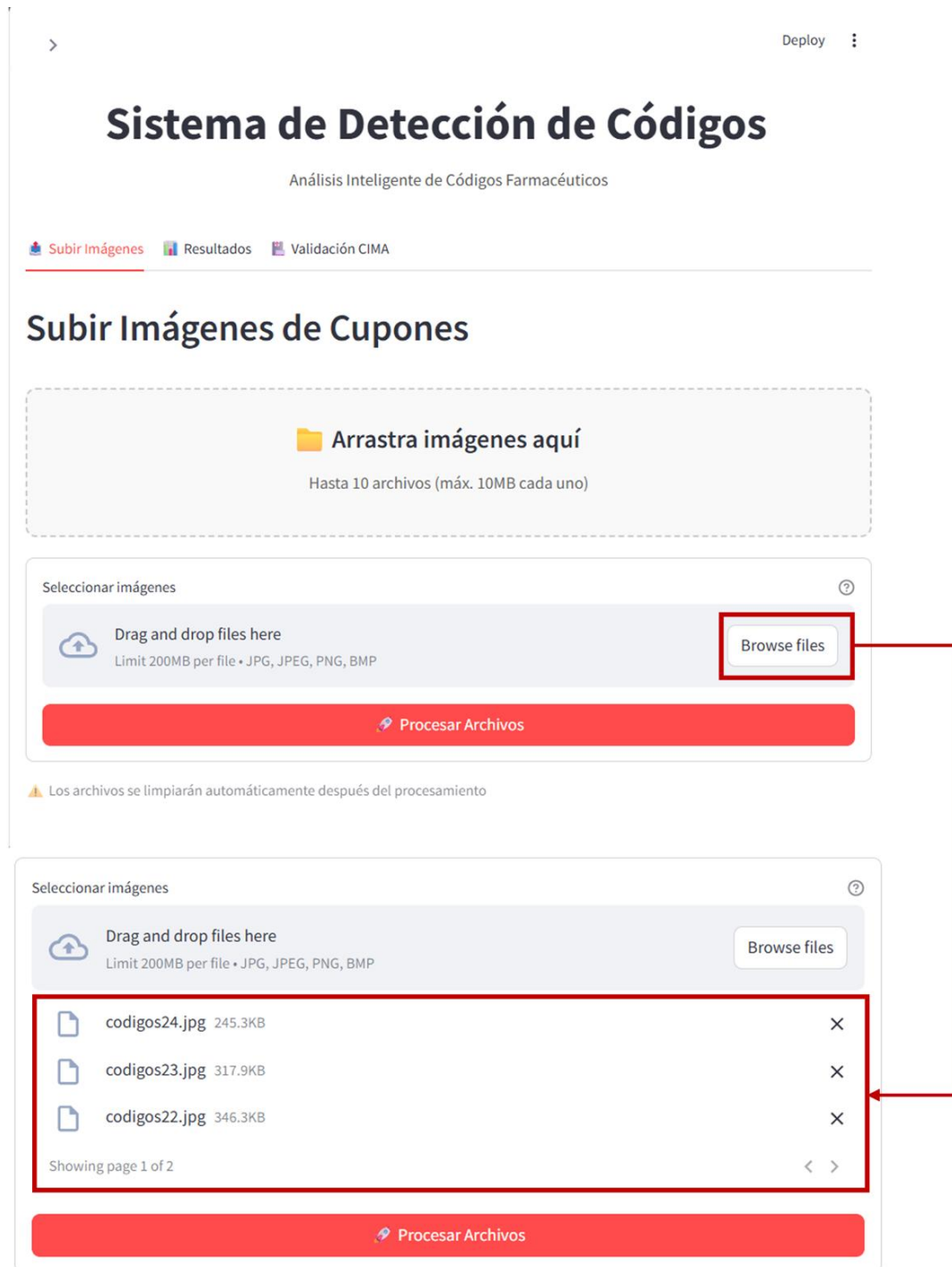


Figura 9.1: Sección de carga de imágenes con interfaz drag-and-drop.

9.3.2. Visualización y Edición de Resultados

La sección principal combina visualización de detecciones con edición interactiva, mostrando el rendimiento superior del sistema híbrido mediante una interfaz dividida que optimiza la experiencia separando lógicamente visualización y edición, tal como se aprecia en la Figura 9.2.

El panel principal de visualización presenta la imagen original del cupón precinto con superposiciones gráficas que indican las detecciones realizadas por el sistema híbrido de

forma clara e intuitiva. Cada código detectado se representa mediante un rectángulo delimitador numerado que corresponde exactamente a su posición específica en la cuadrícula farmacéutica establecida. El sistema utiliza un esquema de colores distintivo que facilita la interpretación inmediata del estado de cada detección: los códigos exitosamente decodificados se muestran con bordes verdes que confirman la lectura correcta, mientras que las posiciones no detectadas o problemáticas se señalan con indicadores rojos que alertan sobre la necesidad de intervención manual.

Deploy

Sistema de Detección de Códigos

Análisis Inteligente de Códigos Farmacéuticos

Subir Imágenes
Resultados
Validación CIMA

Resultados de Detección

Imágenes Procesadas

Seleccionar imagen:

codigos23.jpg (Unknown) - 24/26 códigos

Códigos
24/26

Éxito
92.3%

Tiempo
4.6s

Región de Códigos Detectada

Los cupones precinto correspondientes a las dispensaciones de la Receta Electrónica debe adherirse a esta hoja, y deberá contener un número inferior a 50 cupones.

Zona recortada por YOLO - Fuente: barcode_region

Grid 7x4 (sin header)

Formulario sin encabezado - 26 posiciones máximas

	≡ C1	≡ C2	≡ C3	≡ C4
0	-	8470006886724	8470006886724	8470008404964
1	8470009465902	8470007167679	8470007321699	8470007008118
2	8470007260707	8470007345442	8470007083948	8470006986691
3	8470007313192	8470006914076	8470007674672	8470007304838
4	8470007304838	8470006562116	8470007260059	8470009108854
5	8470008328406	8470009395797	8470007057567	8470007056201
6	-	8470007012184	-	-
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				

Layout detectado: 7 filas x 4 columnas (sin header - max_codes: 26)

Seleccionar posición:

Posición 1

Posición 1: Fila 1, Columna 1 - ✖ -

Hueco hacia atrás
Hueco hacia adelante

Aplicar Cambios
JSON
CSV

Resetear

Figura 9.2: Sección de resultados con visualización dividida y editor de códigos.

Esta representación visual incorpora funcionalidades interactivas avanzadas que permiten a los usuarios inspeccionar detalles específicos de cada detección mediante sistemas de zoom y navegación optimizados para documentos farmacéuticos. El sistema implementa indicadores de confianza que muestran la certeza del algoritmo sobre cada detección, proporcionando información adicional que ayuda a los usuarios a identificar códigos que pueden requerir verificación manual adicional.

El panel principal muestra la imagen original con superposiciones que indican las detecciones realizadas por el sistema híbrido. Cada código detectado se presenta con un rectángulo delimitador numerado según su posición en la cuadrícula farmacéutica, utilizando colores distintivos: códigos exitosamente decodificados con bordes verdes y posiciones no detectadas con indicadores rojos.

La efectividad del sistema híbrido se evidencia en el análisis comparativo por imagen, donde demuestra consistencia superior con 85 % de accuracy promedio frente al 64.8 % de YOLO individual y 53.7 % del detector de gradientes, como se visualiza claramente en la Figura 9.3. Esta mejora se refleja directamente en la experiencia de usuario, reduciendo significativamente la necesidad de correcciones manuales.

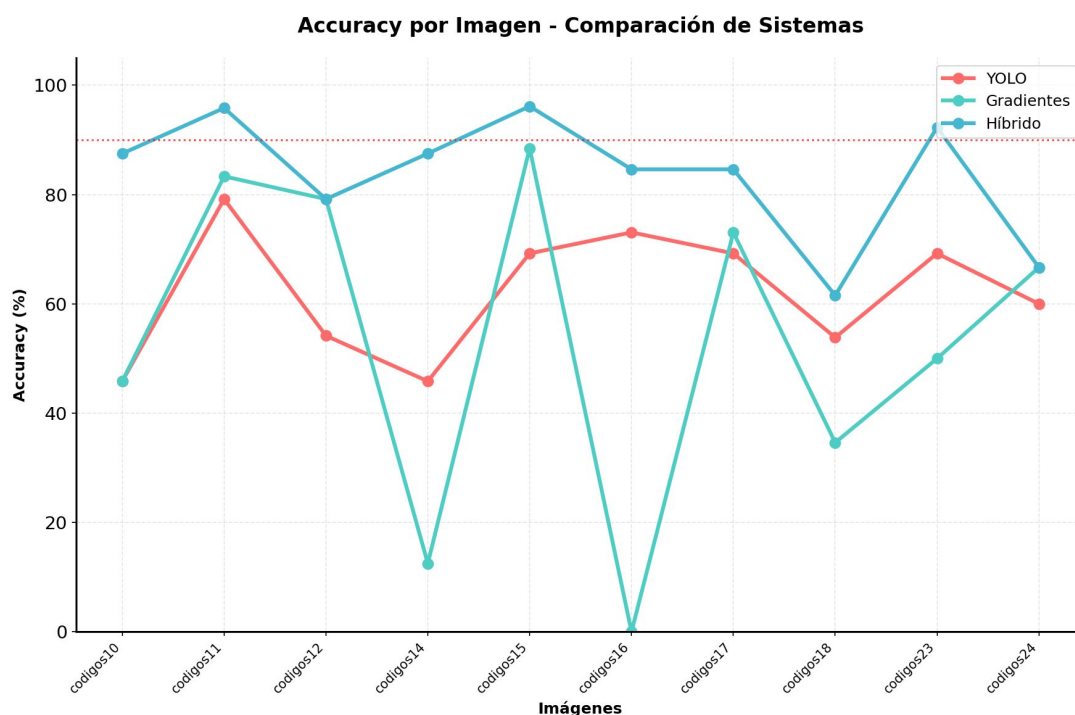


Figura 9.3: Comparación de accuracy por imagen entre los tres métodos implementados, evidenciando la superioridad del sistema híbrido.

El panel lateral complementa la visualización con una tabla interactiva que presenta todos los códigos en formato estructurado y editable. La tabla se adapta automáticamente al layout detectado por el orquestador, mostrando seis filas para cupones con header o siete filas para cupones sin header. El editor implementa funcionalidades avanzadas específicamente diseñadas para el dominio farmacéutico, incluyendo capacidades para crear huecos en la secuencia mediante botones de desplazamiento, validación en tiempo real que previene códigos con formato incorrecto, y verificación automática de la estructura general.

9.3.3. Validación regulatoria

La integración con CIMA proporciona verificación oficial de códigos farmacéuticos españoles, transformando el sistema de una herramienta de detección en una solución integral de verificación regulatoria, como se muestra en la Figura 9.4.

Sistema de Detección de Códigos
Análisis Inteligente de Códigos Farmacéuticos

Subir imágenes Resultados **Validación CIMA**

Validación CIMA

Se detectaron 23 códigos para validar

Validación Individual

Selecciona un código para validar:

Pos 1: 8470007004202

Código seleccionado: 8470007004202

Posición: 1

Verificar en CIMA

☒ Código nacional válido en CIMA

Información

La validación CIMA verificará:

- Si el código existe en la base de datos
- Estado de comercialización
- Problemas de suministro
- Información del medicamento

Información del Medicamento

EAN-13 original: 8470007004202

Código Nacional: 700420

Nombre: CILOSTAZOL CINFA 100 MG COMPRIMIDOS EFG

Laboratorio: Laboratorios Cinfa S.A.

Principio Activo: No disponible

[Ver Ficha Técnica Completa](#)

Estado del Medicamento

☒ Autorizado

☒ Sin problemas de suministro

Figura 9.4: Sección de validación CIMA con verificación oficial de códigos farmacéuticos.

La interfaz implementa identificación automática de códigos españoles mediante reconocimiento del prefijo característico del sistema farmacéutico nacional, seguida de extracción automática del código nacional correspondiente a los seis dígitos centrales del código EAN-13. La consulta a la base de datos CIMA oficial se realiza mediante protocolos seguros, presentando resultados en formato estructurado que incluye denominación comercial, laboratorio fabricante, principio activo y estado de comercialización.

9.4. Valor generado por la aplicación

La aplicación transforma exitosamente los hallazgos de investigación en una herramienta práctica que automatiza completamente el procesamiento de cupones precinto. El sistema híbrido supera las limitaciones identificadas en métodos individuales proporcionando una solución robusta con mejoras cuantificables: 85 % accuracy frente a 64.8 % y 53.7 % de los métodos individuales, 4 resultados excelentes (≥ 90 %) frente a solo 1 del método YOLO, y mayor consistencia con desviación estándar de 10.6 %.

El desarrollo del sistema de corrección posicional establece una metodología innovadora aplicable a otros dominios donde la correspondencia espacial exacta es crítica. La integración con CIMA añade valor regulatorio significativo que trasciende las capacidades de detección

pura, convirtiendo la aplicación en una herramienta integral de verificación que cumple con los requisitos del sector farmacéutico español.

La solución proporciona beneficios cuantificables en términos de eficiencia operativa, reduciendo significativamente el tiempo requerido para el procesamiento manual mientras mejora la precisión y consistencia. La interfaz desarrollada demuestra que es posible crear herramientas sofisticadas de visión por computador genuinamente accesibles para usuarios no técnicos, facilitando la adopción efectiva en entornos farmacéuticos reales donde la usabilidad y confiabilidad son factores críticos para el éxito de la implementación.

Capítulo 10

Conclusiones y trabajo futuro

Este capítulo presenta las conclusiones principales del trabajo desarrollado, evaluando el cumplimiento de los objetivos planteados y analizando las contribuciones realizadas al campo de la digitalización farmacéutica. Se discuten las lecciones aprendidas durante el desarrollo del proyecto, se analizan las limitaciones identificadas y se proponen líneas de trabajo futuro que permitan ampliar y mejorar los resultados obtenidos. La estructura del capítulo progresa desde una síntesis de los resultados técnicos alcanzados hasta una reflexión sobre el impacto del trabajo y las oportunidades de desarrollo futuro en el ámbito de la automatización de procesos farmacéuticos.

10.1. Síntesis de resultados y contribuciones del trabajo

El desarrollo de este proyecto ha logrado cumplir satisfactoriamente los objetivos planteados inicialmente, proporcionando una solución integral para la digitalización automática de hojas de cupones precinto farmacéuticos mediante técnicas avanzadas de visión por computador e inteligencia artificial. El sistema desarrollado ha demostrado capacidades sólidas en la detección y extracción automática de información, con los modelos YOLO implementados alcanzando resultados prometedores en la identificación de zonas clave de los documentos farmacéuticos.

Los resultados experimentales revelan que el detector basado en redes neuronales YOLO ha logrado una precisión posicional promedio del 65.6 % con una mediana del 69.2 %, superando las expectativas iniciales para un dominio tan especializado como el farmacéutico. La capacidad del sistema para detectar simultáneamente códigos individuales EAN-13, cabeceras manuscritas y áreas generales de códigos demuestra la efectividad del enfoque multiclase adoptado. La implementación del detector de gradientes como método complementario ha proporcionado una alternativa viable con una precisión posicional del 58.3 %, especialmente valiosa por su independencia de modelos preentrenados y su capacidad de operación en condiciones variables.

La evaluación comparativa exhaustiva realizada entre los diferentes enfoques tecnológicos ha proporcionado evidencia empírica valiosa sobre sus fortalezas y limitaciones relativas. El detector YOLO ha demostrado ventajas consistentes en precisión posicional y velocidad de procesamiento, con tiempos de 2-4 segundos por imagen, mientras que el detector de gradientes ofrece independencia de datos de entrenamiento y mayor flexibilidad en condiciones variables, aunque con tiempos ligeramente superiores de 4-6 segundos. Esta

complementariedad entre métodos ha validado la hipótesis inicial de que la combinación de enfoques diferentes puede maximizar la efectividad global del sistema.

El proyecto ha realizado contribuciones metodológicas significativas que trascienden la aplicación específica al dominio farmacéutico. El desarrollo de una arquitectura híbrida que combina YOLO para localización inicial con análisis de gradientes para refinamiento representa una contribución novedosa que aprovecha las fortalezas complementarias de ambos enfoques. Esta metodología puede ser aplicada a otros problemas similares de procesamiento de documentos especializados, estableciendo un precedente para la integración inteligente de técnicas de aprendizaje profundo y métodos clásicos de visión por computador.

La implementación de un framework de evaluación dual que considera tanto métricas con correspondencia posicional como sin ella permite una evaluación más completa y matizada del rendimiento de los sistemas de detección. Esta aproximación metodológica es particularmente relevante en dominios donde la posición exacta de los elementos detectados es crítica para su interpretación semántica, como ocurre en el procesamiento de documentos farmacéuticos estructurados.

10.2. Impacto práctico y relevancia del trabajo desarrollado

Los resultados obtenidos demuestran la viabilidad técnica de automatizar un proceso tradicionalmente manual en las farmacias españolas, con implicaciones significativas para la modernización del sector. La automatización del procesamiento de hojas de cupones precinto puede liberar tiempo valioso de los profesionales farmacéuticos para actividades de mayor valor añadido, mientras que la eliminación del factor humano en la transcripción de códigos reduce significativamente la probabilidad de errores en la gestión documental. La digitalización estructurada de la información facilita además los procesos de auditoría y análisis posterior de los datos farmacéuticos, aspectos críticos para el cumplimiento normativo y la optimización de procesos.

Desde una perspectiva académica y científica, el trabajo aporta una metodología de comparación sistemática que puede ser aplicada a otros problemas de procesamiento de documentos especializados. La creación de un conjunto de datos etiquetado específico para documentos farmacéuticos constituye un recurso valioso para futuras investigaciones en el área, mientras que los resultados obtenidos proporcionan insights importantes sobre la transferibilidad de técnicas de visión por computador entre diferentes dominios de aplicación.

El análisis comparativo de diferentes tecnologías emergentes, incluyendo modelos fundacionales y APIs de inteligencia artificial multimodal, establece un marco de referencia para la evaluación de estas tecnologías en aplicaciones de dominio específico. Los resultados revelan tanto las potencialidades como las limitaciones prácticas de estos enfoques, proporcionando orientación valiosa para futuras implementaciones en entornos farmacéuticos y otros sectores especializados.

10.3. Lecciones aprendidas y limitaciones identificadas

El desarrollo del proyecto ha proporcionado lecciones importantes sobre la aplicación práctica de técnicas de visión por computador a problemas reales. La corrección automática de perspectiva y las técnicas de mejora de imagen han demostrado ser factores críticos para el éxito del sistema, especialmente en el detector de gradientes que se beneficia significativamente de la aplicación de múltiples técnicas de preprocesamiento secuencial. Esta observación subraya la importancia de considerar el preprocesamiento no como una etapa auxiliar, sino como un componente fundamental del sistema de detección.

La experiencia ha confirmado el valor de los enfoques híbridos en problemas complejos de visión por computador. La combinación de diferentes metodologías ha resultado más efectiva que cualquier enfoque individual, sugiriendo que la integración inteligente de técnicas complementarias es una estrategia prometedora que merece mayor exploración en futuras investigaciones. Esta lección tiene implicaciones importantes para el diseño de sistemas de procesamiento de documentos en otros dominios especializados.

El trabajo ha revelado también limitaciones importantes que deben ser reconocidas para una evaluación honesta de los resultados. El conjunto de datos utilizado, aunque representativo, es relativamente pequeño y presenta cierto sesgo hacia documentos de proveedores específicos, lo que podría limitar la generalización a cupones de otros proveedores farmacéuticos. Las condiciones de captura evaluadas, aunque diversas, requieren mayor exploración en situaciones extremas como muy baja iluminación o perspectivas muy oblicuas.

Los requerimientos computacionales del sistema, aunque aceptables para aplicaciones de volumen medio, podrían requerir optimización para entornos de muy alto volumen. Esta limitación es particularmente relevante considerando la escalabilidad necesaria para una implementación amplia en el sector farmacéutico español, donde el volumen de documentos procesados diariamente puede ser considerable.

10.4. Direcciones futuras y oportunidades de desarrollo

Las oportunidades de mejora y desarrollo futuro del trabajo son múltiples y abarcan tanto aspectos técnicos como de implementación práctica. La expansión del conjunto de datos mediante la incorporación de más imágenes de diferentes proveedores y condiciones de captura mejoraría significativamente la capacidad de generalización del sistema. Esta expansión debería incluir documentos con diferentes calidades de impresión, condiciones de iluminación extremas y layouts menos convencionales para aumentar la robustez del sistema.

La optimización del rendimiento computacional representa otra línea de desarrollo prioritaria. La implementación de técnicas como la cuantización de modelos, la inferencia en paralelo y la optimización de algoritmos podría reducir sustancialmente los tiempos de procesamiento, haciendo el sistema más viable para entornos de alto volumen. La exploración de arquitecturas de red neuronal específicamente diseñadas para el procesamiento de documentos farmacéuticos podría superar el rendimiento de los modelos generales actuales.

El desarrollo de capacidades de integración con sistemas farmacéuticos existentes constituye un paso fundamental para la adopción práctica del sistema. La creación

de APIs y conectores que faciliten la integración con software de gestión farmacéutica ampliaría significativamente el impacto del trabajo. La implementación de una aplicación móvil permitiría el procesamiento de documentos directamente desde dispositivos móviles, mejorando la accesibilidad y usabilidad del sistema.

La investigación en técnicas de few-shot learning y aprendizaje continuo representa una dirección particularmente prometedora. El desarrollo de mecanismos que permitan adaptar rápidamente el sistema a nuevos tipos de documentos o formatos con mínimos datos de entrenamiento facilitaría la escalabilidad y adaptabilidad del sistema. La implementación de sistemas de feedback y aprendizaje continuo que incorporen correcciones de usuarios podría mejorar progresivamente el rendimiento del sistema en condiciones reales de uso.

La extensión de las técnicas desarrolladas a otros tipos de documentos farmacéuticos como recetas, albaranes y facturas ampliaría considerablemente el impacto del trabajo. El desarrollo de capacidades de análisis multimodal avanzado que integre información contextual adicional como metadatos de imagen, información temporal y datos de farmacia podría mejorar significativamente la precisión de la detección y abrir nuevas posibilidades de análisis.

10.5. Reflexiones finales sobre el trabajo realizado

Este Trabajo de Fin de Grado ha demostrado que la aplicación de técnicas avanzadas de visión por computador e inteligencia artificial puede abordar efectivamente problemas reales del sector farmacéutico, contribuyendo a la modernización y digitalización de procesos tradicionalmente manuales. Los resultados obtenidos no solo validan la viabilidad técnica de la automatización del procesamiento de documentos farmacéuticos, sino que también establecen una base sólida para futuras investigaciones en el área.

La metodología híbrida desarrollada, que combina las fortalezas de diferentes enfoques tecnológicos, representa una contribución valiosa que puede ser aplicada a otros problemas similares de procesamiento de documentos especializados. La evaluación comparativa exhaustiva realizada proporciona insights importantes sobre las capacidades y limitaciones de diferentes tecnologías emergentes en el campo de la visión por computador, información valiosa para la comunidad académica y profesional.

Desde una perspectiva de impacto social, el trabajo contribuye a la modernización del sector farmacéutico, facilitando la transición hacia sistemas más eficientes y menos propensos a errores. Esta contribución se alinea con los objetivos de mejora de la atención sanitaria y optimización de recursos en el sistema de salud, aspectos fundamentales en el contexto actual de transformación digital del sector sanitario.

El proyecto ha proporcionado también una experiencia valiosa en la aplicación práctica de conocimientos teóricos a problemas reales, demostrando la importancia de considerar no solo la precisión técnica, sino también factores como la usabilidad, escalabilidad y viabilidad económica en el desarrollo de soluciones de inteligencia artificial. Esta experiencia integral ha enriquecido significativamente la formación académica y profesional, proporcionando una perspectiva realista sobre los desafíos y oportunidades en la aplicación de tecnologías emergentes a problemas del mundo real.

En conclusión, este trabajo establece una base sólida para la automatización del procesamiento de documentos farmacéuticos y abre múltiples vías para futuras investigaciones y desarrollos en el área de digitalización farmacéutica. Los resultados obtenidos demuestran que la combinación inteligente de diferentes tecnologías, junto con una evaluación rigurosa y una comprensión profunda del dominio de aplicación, puede generar soluciones efectivas

y prácticamente viables para problemas complejos del mundo real. El trabajo realizado contribuye tanto al avance del conocimiento académico como a la mejora de procesos en el sector farmacéutico, cumpliendo así con los objetivos.

Apéndice A

Competencias específicas cubiertas

Este capítulo analiza las competencias específicas del Grado en Ciencia e Ingeniería de Datos de la Universidad de Las Palmas de Gran Canaria que han sido desarrolladas durante la realización de este trabajo. Se evalúa el grado de adquisición de cada competencia y se justifica mediante las actividades y resultados obtenidos en el proyecto.

A.1. Evaluación de competencias específicas

El desarrollo del proyecto de digitalización de cupones precinto farmacéuticos ha permitido trabajar de forma integral las competencias específicas del grado, desde la aplicación del método científico hasta el desarrollo de sistemas de información complejos. La naturaleza multidisciplinar del proyecto, que combina visión por computador, inteligencia artificial y desarrollo de software, ha proporcionado un marco ideal para la adquisición práctica de competencias técnicas avanzadas.

La implementación y evaluación comparativa de múltiples enfoques tecnológicos (YOLO, análisis de gradientes, APIs multimodales) siguió una metodología experimental rigurosa que permitió obtener resultados reproducibles y estadísticamente significativos. El diseño de una arquitectura modular que integra múltiples detectores demuestra la capacidad para crear sistemas de información que respondan a necesidades reales del sector farmacéutico.

A.2. Competencias transversales y desarrollo profesional

El proyecto ha fortalecido significativamente competencias transversales complementarias a la formación técnica. La **capacidad de análisis y síntesis** se ejercitó mediante el análisis comparativo de múltiples enfoques tecnológicos y la síntesis de resultados en conclusiones accionables. La **organización y planificación** fue aplicada en la gestión temporal del proyecto y la coordinación con Farmalitics.

La **resolución de problemas** se ejercitó continuamente en la identificación y solución de desafíos técnicos como la fragmentación de códigos detectados y la optimización de parámetros. El **trabajo en equipo** se manifestó en la colaboración efectiva con la empresa y la coordinación con tutores académicos.

El desarrollo de **capacidad crítica** se evidenció en la evaluación objetiva de las limitaciones del trabajo y la identificación de áreas de mejora, fundamental para el aprendizaje continuo y la mejora iterativa del sistema.

Tabla A.1: Evaluación del nivel de adquisición de competencias específicas.

Competencia	Actividades desarrolladas	Nivel	Evidencias
CE1 - Méto- do científico	Diseño experimental comparativo, análisis sistemático del problema, revisión del estado del arte, implementación de métricas objetivas	Alto	Metodología experimental rigurosa, validación estadística
CE2 - Siste- mas de infor- mación	Análisis de requisitos del sector farmacéutico, diseño de arquitectura modular, desarrollo de interfaz web con Streamlit	Alto	Aplicación web funcional, arquitectura escalable
CE3 - Inte- ligencia artifi- cial	Entrenamiento de modelos YOLO, evaluación de SAM2, implementación de LLMs multimodales, desarrollo de arquitecturas híbridas	Alto	Múltiples modelos IA implementados y evaluados
CE4 - Proce- samiento de datos	Algoritmos de corrección de perspectiva, análisis de gradientes, técnicas de preprocesamiento, optimización de pipelines	Alto	Algoritmos especializados de procesamiento de imagen
CE5 - Visua- lización	Imágenes anotadas con bounding boxes, gráficas de rendimiento, dashboards interactivos	Medio- Alto	Visualizaciones efectivas, interfaz intuitiva
CE6 - Bases de datos	Organización de datasets, integración con base de datos CIMA, estructuras JSON para persistencia	Medio	Gestión eficiente de datos, integración APIs
CE7 - Desa- rrollo softwa- re	Arquitectura modular, uso de Git, documentación técnica, pruebas unitarias	Alto	Código bien estructurado, buenas prácticas
CE8 - Co- municación técnica	Memoria técnica detallada, interfaz para usuarios no técnicos, presentación de resultados	Alto	Documentación completa, comunicación efectiva

A.3. Contribución al perfil profesional

Este Trabajo de Fin de Grado ha contribuido significativamente a la formación de un perfil profesional completo en Ciencia e Ingeniería de Datos, desarrollando especialización

en **Computer Vision** mediante la implementación de técnicas avanzadas de procesamiento de imágenes y detección de objetos.

La experiencia práctica con **modelos fundacionales** y APIs de inteligencia artificial de última generación proporciona una perspectiva valiosa sobre las tecnologías emergentes. El conocimiento específico del **dominio farmacéutico** y sus requisitos regulatorios añade una dimensión sectorial que enriquece el perfil profesional.

La capacidad demostrada para desarrollar **soluciones completas** desde el algoritmo hasta la interfaz de usuario refleja una formación full-stack valiosa en el mercado laboral actual. La aplicación rigurosa de **metodología científica** proporciona una base sólida para futuras actividades de investigación y desarrollo.

A.4. Reflexión sobre el aprendizaje

El desarrollo de este proyecto ha proporcionado una experiencia integral que conecta efectivamente los conocimientos teóricos del grado con la aplicación práctica en un problema real del sector farmacéutico. La necesidad de evaluar múltiples enfoques tecnológicos ha desarrollado una perspectiva crítica sobre las fortalezas y limitaciones de diferentes técnicas de IA.

La colaboración con Farmalitics ha añadido una dimensión práctica que enriquece la experiencia académica, proporcionando insights sobre consideraciones comerciales y regulatorias. La integración de tecnologías emergentes como GPT-4 Vision y SAM2 ha proporcionado experiencia con herramientas de vanguardia relevantes para el futuro profesional.

El desarrollo de una **solución híbrida** que combina múltiples enfoques refleja la realidad de los proyectos industriales, donde la integración inteligente de tecnologías complementarias suele proporcionar los mejores resultados. Esta experiencia demuestra la capacidad para abordar problemas complejos de forma sistemática, aplicando el método científico y las mejores prácticas de ingeniería para generar soluciones que aportan valor real.

Apéndice B

Objetivos de desarrollo sostenible

Tras detallar las competencias trabajadas durante el desarrollo del proyecto, a continuación se analiza la contribución del mismo a los Objetivos de Desarrollo Sostenible (ODS) de la Agenda 2030. En la Tabla B.1 se muestra el nivel de relación estimado entre el Proyecto y cada uno de los ODS. Esta valoración refleja cómo la digitalización y automatización en el ámbito farmacéutico pueden impulsar la innovación, la eficiencia y la colaboración universidad-empresa, así como contribuir a una gestión más responsable de los recursos.

A continuación se justifica brevemente la relación del Proyecto con los ODS seleccionados:

- **ODS 3: Salud y Bienestar (nivel bajo).** El sistema mejora la trazabilidad y reduce errores en la gestión de medicamentos en farmacia, contribuyendo a un servicio más seguro y eficiente.
- **ODS 8: Trabajo decente y crecimiento económico (nivel bajo).** La automatización documental aumenta la productividad y eficiencia en la farmacia, permitiendo dedicar más tiempo a tareas de valor añadido.
- **ODS 9: Industria, Innovación e Infraestructuras (nivel alto).** El proyecto aplica tecnologías innovadoras de inteligencia artificial y visión por computador para modernizar procesos en el sector farmacéutico.
- **ODS 12: Producción y consumo sostenibles (nivel bajo).** La digitalización facilita un control más preciso del consumo de medicamentos subvencionados y optimiza el uso de recursos.
- **ODS 17: Alianzas para lograr objetivos (nivel medio).** El desarrollo del sistema es fruto de la colaboración entre la universidad y una empresa, promoviendo la transferencia de conocimiento y tecnología.

El resto de los ODS no presentan una relación directa con el alcance y los objetivos del presente proyecto.

ODS	Nivel de relación			
	0	1	2	3
1. Fin de la Pobreza	X			
2. Hambre cero	X			
3. Salud y Bienestar		X		
4. Educación de calidad	X			
5. Igualdad de género	X			
6. Agua limpia y saneamiento	X			
7. Energía asequible y no contaminante	X			
8. Trabajo decente y crecimiento económico		X		
9. Industria, Innovación e Infraestructuras				X
10. Reducción de las desigualdades	X			
11. Ciudades y comunidades sostenibles	X			
12. Producción y consumo sostenibles		X		
13. Acción por el clima	X			
14. Vida submarina	X			
15. Vida de ecosistemas terrestres	X			
16. Paz, justicia e instituciones sólidas	X			
17. Alianzas para lograr objetivos			X	

Tabla B.1: Tabla con la valoración del grado de relación del Trabajo con los ODS, según una escala de 0 a 3 (no procede, bajo, medio, alto).

Lista de Acrónimos

- Adam** Adaptive Moment Estimation. 49
- AEMPS** Agencia Española de Medicamentos y Productos Sanitarios. 3, 17
- API** Application Programming Interface. 3, 11, 28, 31
- CIMA** Centro de Información de Medicamentos de la AEMPS. 3, 5, 19, 21, 28, 31
- CLAHE** Contrast Limited Adaptive Histogram Equalization. 10, 49, 50
- CRISP-DM** Cross Industry Standard Process for Data Mining. 6, 25
- EAN-13** European Article Number de 13 dígitos. 10, 19
- IA** Inteligencia Artificial. 1, 8, 23, 26, 27
- JSON** JavaScript Object Notation. 10, 37, 44, 51, 53
- NMS** Non-Maximum Suppression. 9, 41
- OCR** Optical Character Recognition. 45
- ODS** Objetivos de Desarrollo Sostenible. VII, 90, 91
- OpenCV** Open Source Computer Vision Library. 1, 10, 13
- SNS** Sistema Nacional de Salud. 16
- SOE** Servicio de Orientación de Enfermería. 48, 55, 57
- YOLO** You Only Look Once. 8

Glosario de Términos

aprendizaje profundo Subconjunto del aprendizaje automático basado en redes neuronales artificiales con múltiples capas. Permite a las máquinas aprender representaciones complejas de datos mediante el procesamiento jerárquico de información. 1, 4, 10, 30

aumento de datos Técnica que expande artificialmente el conjunto de entrenamiento aplicando transformaciones como rotaciones, cambios de brillo o escalado a las imágenes existentes, mejorando la capacidad de generalización de los modelos. 27, 29

backbone Componente de una red neuronal que se encarga de la extracción de características visuales mediante capas convolucionales profundas. 42

batch size Número de muestras procesadas antes de actualizar los parámetros del modelo durante el entrenamiento. 49

caja delimitadora Rectángulo que encierra completamente un objeto detectado en una imagen, definido por las coordenadas de sus esquinas. Se utiliza para localizar y delimitar objetos de interés en tareas de detección. v, 9, 35, 38, 40

conjunto de datos Colección estructurada de datos utilizados para entrenar, validar y evaluar modelos de aprendizaje automático. Puede incluir imágenes, texto, números u otros tipos de información organizados para facilitar su procesamiento. 4–6, 11, 27–29, 32, 33

cupón precinto Documento físico adherido a los envases de medicamentos financiados por el Sistema Nacional de Salud que contiene información codificada del medicamento, incluyendo código de barras EAN-13, nombre comercial, precio y datos del laboratorio fabricante. Se utiliza para justificar la dispensación y facturación de medicamentos subvencionados. 1, 8, 16, 17

código de barras Representación visual de datos mediante barras paralelas de diferentes anchos y espaciados. En farmacia, los códigos EAN-13 contienen información específica sobre el medicamento que permite su identificación unívoca. 2, 3, 5, 8, 10

decoder Componente que transforma representaciones internas en salidas del formato deseado. 45

detección de objetos Técnica de visión por computador que identifica y localiza objetos específicos dentro de una imagen, proporcionando tanto la clasificación del objeto como su ubicación mediante coordenadas o cajas delimitadoras. 1, 3, 8

- digitalización** Proceso de conversión de información analógica o física a formato digital, permitiendo su procesamiento, almacenamiento y transmisión mediante sistemas informáticos. En el contexto farmacéutico, se refiere a la automatización de procesos manuales tradicionales. 1, 8
- early stopping** Técnica para detener el entrenamiento cuando el rendimiento en validación deja de mejorar. 49
- encoder** Componente de una red neuronal que transforma datos de entrada en representaciones internas. 44, 45
- fallback** Mecanismo de respaldo que se activa cuando el método principal falla o no produce resultados satisfactorios. Permite mantener la funcionalidad del sistema mediante métodos alternativos. 31
- fine-tuning** Proceso de ajustar un modelo preentrenado para una tarea específica mediante entrenamiento adicional con datos del dominio objetivo. Permite adaptar modelos generales a aplicaciones especializadas. 27
- framework** Estructura de software que proporciona una base estandarizada para el desarrollo de aplicaciones. Incluye bibliotecas, herramientas y convenciones que facilitan y aceleran el proceso de desarrollo. 31
- ground truth** Conjunto de datos de referencia que contiene las respuestas correctas o esperadas, utilizado para entrenar modelos de aprendizaje automático y evaluar su rendimiento. Representa la realidad objetiva contra la cual se comparan las predicciones. 4, 27, 29, 31, 38
- head** Componente que genera las predicciones finales de localización y clasificación en modelos de detección. 42
- insights** Conocimientos profundos, percepciones valiosas o descubrimientos significativos obtenidos mediante el análisis de datos, información o experiencias. En el contexto de procesamiento de documentos farmacéuticos, se refiere a las conclusiones y aprendizajes derivados del análisis de resultados experimentales, patrones de comportamiento de algoritmos, o características específicas del dominio que proporcionan comprensión sobre la efectividad de diferentes enfoques tecnológicos. 83, 85
- layout** Disposición y organización espacial de elementos en un documento o imagen. 45, 46, 48, 49
- learning rate** Parámetro que controla la velocidad de aprendizaje del algoritmo de optimización. 49
- mapa de energía** Representación visual que resalta las regiones de una imagen con mayor probabilidad de contener códigos de barras, calculada mediante la aplicación de gradientes direccionales que enfatizan estructuras verticales. 43, 50
- neck** Componente que fusiona características de diferentes escalas para mejorar la detección de objetos de diversos tamaños. 42

- pipeline** Secuencia de procesos de transformación de datos conectados en serie, donde la salida de un proceso sirve como entrada del siguiente. En visión por computador, incluye etapas como preprocesamiento, detección y post-procesamiento. 31, 42, 44
- preprocesamiento** Conjunto de técnicas aplicadas a los datos o imágenes antes del procesamiento principal para mejorar su calidad, normalizar formatos, corregir defectos o adaptarlos a los requisitos del algoritmo que los procesará. 1, 3, 5, 6, 8, 10, 27, 29, 30, 32
- prompt** Instrucción o entrada que guía el comportamiento de un modelo de inteligencia artificial. 45
- prompts visuales** Señales o indicaciones visuales proporcionadas a un modelo de IA para guiar el proceso de segmentación, como puntos, cajas delimitadoras o máscaras que indican qué objetos segmentar. 27
- scheduler** Algoritmo que ajusta automáticamente el learning rate durante el entrenamiento según criterios predefinidos. 49
- segmentación de imágenes** Proceso de dividir una imagen en múltiples segmentos o regiones para simplificar su análisis. Cada píxel de la imagen se asigna a una categoría específica, permitiendo identificar con precisión los contornos y formas de los objetos. 1, 10, 11
- transformer** Arquitectura de red neuronal basada en mecanismos de atención, especialmente efectiva para el procesamiento de secuencias. 44, 45, 49
- trazabilidad** Capacidad de seguir el rastro de un medicamento a través de todas las etapas de producción, transformación y distribución, desde su fabricación hasta su dispensación al paciente final. Es fundamental para garantizar la seguridad farmacéutica y el control de medicamentos. 16, 17, 30
- visión por computador** Campo de la inteligencia artificial que desarrolla métodos para que las máquinas puedan interpretar y comprender imágenes digitales. Incluye técnicas como detección de objetos, reconocimiento de patrones y análisis de imágenes. 1, 4, 8, 10, 13, 28, 30
- weight decay** Técnica de regularización que penaliza pesos grandes en la función de pérdida. 49

Bibliografía

- [Agencia Española de Medicamentos y Productos Sanitarios, 2024] Agencia Española de Medicamentos y Productos Sanitarios (2024). Centro de información de medicamentos - base de datos oficial. Base de datos oficial de medicamentos autorizados en España.
- [Alacantí TV, 2025] Alacantí TV (2025). La comunidad valenciana elimina el cupón precinto y ahorra 4,7 millones. Noticia sobre eliminación del cupón precinto en Valencia.
- [American Journal of Health-System Pharmacy, 2023] American Journal of Health-System Pharmacy (2023). Computer vision for hospital pharmacy inventory management. Gestión de inventarios en farmacia hospitalaria.
- [Asociación de Distribuidores Farmacéuticos, 2024] Asociación de Distribuidores Farmacéuticos (2024). Optimización de procesos en la cadena de suministro farmacéutica. Análisis de mejoras en la distribución farmacéutica.
- [Boletín Oficial del Estado, 2003] Boletín Oficial del Estado (2003). Real decreto 1348/2003, de 31 de octubre, que adapta la clasificación anatómica de medicamentos al sistema de clasificación atc. Normativa sobre clasificación de medicamentos en España.
- [Centro Nacional de Epidemiología, 2024] Centro Nacional de Epidemiología (2024). Análisis de patrones de dispensación farmacéutica en España. Estudio sobre tendencias en la dispensación de medicamentos.
- [Consejo General de Colegios Oficiales de Farmacéuticos, 2024] Consejo General de Colegios Oficiales de Farmacéuticos (2024). Memoria anual del consejo general de colegios oficiales de farmacéuticos. Informe sobre el estado de la profesión farmacéutica en España.
- [Conselleria de Sanidad, 2024] Conselleria de Sanidad (2024). Sistema de código digital de medicamentos en la comunidad valenciana. Implementación pionera de facturación sin cupón precinto.
- [CRISP-DM Consortium, 2024] CRISP-DM Consortium (2024). Cross industry standard process for data mining. Metodología estándar para proyectos de minería de datos.
- [Drug Discovery Today, 2024] Drug Discovery Today (2024). Implementation considerations for ai in pharmaceutical operations. Consideraciones de implementación de IA en operaciones farmacéuticas.
- [El Global Farmacia, 2024] El Global Farmacia (2024). El coste del cupón precinto en España: análisis económico del sector. Análisis del coste económico del sistema de cupones precinto.

- [Farmaindustria, 2024] Farmaindustria (2024). Informe sobre digitalización del sector farmacéutico español. Análisis de la transformación digital en el sector farmacéutico.
- [Farmalitics, 2024] Farmalitics (2024). Soluciones tecnológicas para farmacias: innovación en gestión farmacéutica. Empresa colaboradora especializada en tecnología farmacéutica.
- [Federación Empresarial de Farmacéuticos Españoles, 2024] Federación Empresarial de Farmacéuticos Españoles (2024). Impacto económico de la gestión del cupón precinto en farmacias españolas. Análisis de costes operativos en la gestión documental farmacéutica.
- [Gallo and Manduchi, 2011] Gallo, O. and Manduchi, R. (2011). Reading 1-d barcodes with mobile phones using deformable templates. *IEEE transactions on pattern analysis and machine intelligence*, 33(9):1834–1843.
- [Girshick et al., 2014] Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587.
- [Google DeepMind, 2024] Google DeepMind (2024). Gemini: A family of highly capable multimodal models. <https://deepmind.google/technologies/gemini>.
- [Grupo de Investigación en Visión Artificial, 2024] Grupo de Investigación en Visión Artificial (2024). Visión por computador aplicada a documentos farmacéuticos. Revisión de técnicas de visión artificial en el ámbito farmacéutico.
- [Grupo de Procesamiento de Imágenes, 2024] Grupo de Procesamiento de Imágenes (2024). Limitaciones técnicas en sistemas de visión artificial para documentos. Estudio sobre limitaciones en procesamiento automático de imágenes.
- [Hudson, 2024] Hudson, L. (2024). pyzbar: Python library for reading barcodes. <https://github.com/NaturalHistoryMuseum/pyzbar>.
- [IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024] IEEE Transactions on Pattern Analysis and Machine Intelligence (2024). Hybrid approaches in computer vision for document processing. Enfoques híbridos en visión por computador.
- [Innovatiana, 2024] Innovatiana (2024). Labelme: Herramienta de anotación de imágenes para machine learning. Consultado en enero 2025.
- [Instituto de Ingeniería del Conocimiento, 2024] Instituto de Ingeniería del Conocimiento (2024). Análisis comparativo de soluciones de procesamiento documental. Estudio comparativo de tecnologías de procesamiento automático.
- [International Journal of Pharmaceutical Sciences, 2023] International Journal of Pharmaceutical Sciences (2023). Quality control in pharmaceutical manufacturing using computer vision. Control de calidad farmacéutico mediante visión por computador.
- [Kanaries, 2023] Kanaries (2023). Streamlit vs dash: A comprehensive comparison for data visualization. Análisis comparativo de plataformas de desarrollo web para ciencia de datos.

- [Ministerio de Asuntos Económicos y Transformación Digital, 2024] Ministerio de Asuntos Económicos y Transformación Digital (2024). Transformación digital en el sector sanitario español. Plan nacional de digitalización sanitaria.
- [Ministerio de Sanidad, 2024] Ministerio de Sanidad (2024). Estrategia de la industria farmacéutica 2024-2028. Hoja de ruta para la modernización del sector farmacéutico.
- [Observatorio de Tecnologías Sanitarias, 2024] Observatorio de Tecnologías Sanitarias (2024). Análisis del sistema de gestión farmacéutica en España. Estudio sobre sistemas de gestión en farmacias españolas.
- [Observatorio del Medicamento, 2024] Observatorio del Medicamento (2024). Análisis del tiempo administrativo en farmacias comunitarias. Estudio sobre distribución del tiempo de trabajo farmacéutico.
- [OpenAI, 2024a] OpenAI (2024a). Gpt-4 vision system card. <https://openai.com/research/gpt-4v-system-card>.
- [OpenAI, 2024b] OpenAI (2024b). Gpt-4v(ision) system card. <https://openai.com/research/gpt-4v-system-card>.
- [OpenCV Development Team, 2024] OpenCV Development Team (2024). *OpenCV Documentation*. OpenCV Foundation. Computer Vision Library Documentation.
- [Pharmaceutical Executive, 2024] Pharmaceutical Executive (2024). Automation in pharmaceutical logistics and inventory management. Automatización en logística farmacéutica.
- [Pharmaceutical Technology, 2023] Pharmaceutical Technology (2023). Computer vision applications in pharmaceutical industry. Revisión de aplicaciones de visión por computador en farmacia.
- [Portal Farma, 2024] Portal Farma (2024). Servicios profesionales farmacéuticos: evolución y perspectivas. Análisis de servicios farmacéuticos y digitalización.
- [Ravi et al., 2024] Ravi, N., Gabeur, V., Hu, Y.-T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., et al. (2024). Sam 2: Segment anything in images and videos. *arXiv preprint arXiv:2408.00714*.
- [Redmon et al., 2016] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788. IEEE.
- [Redmon and Farhadi, 2018] Redmon, J. and Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
- [Ren et al., 2015] Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.
- [Ruder, 2017] Ruder, S. (2017). An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*.

- [Russell et al., 2008] Russell, B. C., Torralba, A., Murphy, K. P., and Freeman, W. T. (2008). Labelme: A database and web-based tool for image annotation. *International Journal of Computer Vision*, 77(1-3):157–173.
- [Sociedad Española de Farmacia Hospitalaria, 2024] Sociedad Española de Farmacia Hospitalaria (2024). Inteligencia artificial en el sector farmacéutico: aplicaciones y perspectivas. Revisión de aplicaciones de IA en farmacia.
- [Streamlit Inc., 2024] Streamlit Inc. (2024). Streamlit: The fastest way to build and share data apps. <https://streamlit.io>.
- [Team et al., 2023] Team, G., Anil, R., Borgeaud, S., Wu, Y., Alayrac, J.-B., Yu, J., Soricut, R., Schalkwyk, J., Dai, A. M., Hauth, A., et al. (2023). Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*.
- [Tuinstra, 2006] Tuinstra, T. R. (2006). Reading barcodes from digital imagery. *PMC*.
- [Ultralytics, 2024a] Ultralytics (2024a). Documentación oficial de modelos yolo para detección de objetos. Documentación técnica de arquitecturas de detección neuronal.
- [Ultralytics, 2024b] Ultralytics (2024b). Sam2 integration with ultralytics. <https://docs.ultralytics.com/models/sam2>.
- [Ultralytics, 2024c] Ultralytics (2024c). Yolo11: An improved real-time object detection model. <https://github.com/ultralytics/ultralytics>.
- [Ultralytics, 2025] Ultralytics (2025). Ai in pharmaceutical industry: Computer vision applications. Aplicaciones de visión por computador en farmacia.
- [Wang et al., 2024a] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., and Ding, G. (2024a). YOLOv10: Real-time end-to-end object detection.
- [Wang et al., 2024b] Wang, C.-Y., Bochkovskiy, A., and Liao, H.-Y. M. (2024b). YOLOv11: An improved version of yolo for object detection. *arXiv preprint*.
- [Yosinski et al., 2014] Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? In *Advances in neural information processing systems*, pages 3320–3328.



eii
ESCUELA DE INGENIERÍA
INFORMÁTICA