

# **Universidad Nacional de Ingeniería**

FACULTAD DE INGENIERÍA INDUSTRIAL Y DE SISTEMAS

ESCUELA DE POSGRADO



**SISTEMA INTELIGENTE BASADO EN VISIÓN POR COMPUTADORA PARA LA  
DETECCIÓN NO INVASIVA DE ANEMIA MEDIANTE DISPOSITIVOS MÓVILES**

**PROYECTO DE INVESTIGACIÓN PARA OPTAR EL GRADO DE MAESTRO EN  
INTELIGENCIA ARTIFICIAL**

Elaborado por:

COCHACHIN MELGAREJO, MANUEL LUCIANO

RIVERA BERNUY, JOHN ANDY

Asesor (a):

XXXXXXXXXXXXXXXXXXXXXXX

LIMA – PERÚ

2025

# CAPÍTULO 1

## INTRODUCCIÓN

### 1.1. Contexto y motivación

La anemia es un problema de salud pública de alta prevalencia a nivel mundial, especialmente en grupos vulnerables como niños, mujeres embarazadas y mujeres en edad fértil. Diversos estudios reportan que entre el 30 % y 40 % de la población mundial presenta algún grado de anemia, con consecuencias directas en el desarrollo cognitivo, el rendimiento físico, la productividad laboral y la mortalidad materno-infantil.

Clínicamente, la anemia se define como una disminución de la concentración de hemoglobina (Hb) o del número de eritrocitos en sangre, lo que reduce la capacidad de transporte de oxígeno hacia los tejidos. Esta condición puede tener múltiples causas: deficiencias nutricionales (hierro, vitamina B12, ácido fólico), procesos infecciosos (malaria, parásitos, infecciones crónicas), enfermedades renales o hematológicas, e incluso factores genéticos como las hemoglobinopatías.

El diagnóstico estándar de la anemia se realiza mediante pruebas de laboratorio invasivas, basadas en la extracción de sangre y el análisis de hemoglobina con equipos especializados (hemogramas, cianometahemoglobina, contadores automatizados). Si bien estos métodos son precisos, presentan varias desventajas:

- Son **invasivos** y generan incomodidad o temor en los pacientes, especialmente en niños.

journal.pone.0253495 (2)

- Requieren **personal especializado**, infraestructura de laboratorio y reactivos.
- Implican **costos económicos** que pueden ser elevados en contextos de bajos recursos.
- Pueden presentar **demoras en la entrega de resultados**, dificultando el tamizaje masivo y el monitoreo frecuente.

En regiones con recursos limitados —como zonas rurales o periferias urbanas— la infraestructura diagnóstica es escasa o inexistente. Esto genera brechas de acceso al diagnóstico oportuno y limita la implementación de programas de tamizaje sistemático.

Frente a este escenario, en la última década se ha explorado el uso de **métodos no invasivos** para la detección de anemia, basados en el análisis de la palidez (pallor) en regiones anatómicas con alta vascularización y baja pigmentación: palmas de las manos, lecho ungueal, mucosa bucal y, especialmente, **conjuntiva palpebral**. La conjuntiva palpebral se ha identificado como una de las regiones con mejor correlación con los niveles de hemoglobina, con valores elevados de sensibilidad y especificidad para la detección de anemia.

Paralelamente, el avance de la **visión por computadora** y del **aprendizaje profundo (deep learning)** ha permitido desarrollar modelos capaces de analizar imágenes médicas y extraer patrones complejos con alta precisión. Diversos trabajos recientes han demostrado que es posible detectar anemia a partir de imágenes de la conjuntiva capturadas con cámaras de teléfonos inteligentes, utilizando redes neuronales convolucionales (CNN), modelos de transferencia de aprendizaje (AlexNet, ResNet, VGG, MobileNet) y técnicas de ensamble.

Al mismo tiempo, la masificación de los **dispositivos móviles** (smartphones) abre la posibilidad de construir sistemas de tamizaje portátiles, de bajo costo y de fácil despliegue en campo. Estos sistemas podrían permitir que personal de salud o incluso usuarios finales capturen imágenes de la conjuntiva con la cámara del teléfono, y reciban una estimación del riesgo de anemia en tiempo real, apoyando la toma de decisiones clínicas y derivaciones oportunas.

En este contexto, surge la necesidad de diseñar y validar un **sistema inteligente basado en visión por computadora**, especializado en el análisis de la **conjuntiva palpebral**, que funcione sobre dispositivos móviles y que permita la **detección no invasiva de anemia** con un desempeño comparable —o superior— a los métodos reportados en la literatura y a soluciones prototípicas existentes.

El presente trabajo de investigación se enmarca en una maestría en Inteligencia Artificial de carácter aplicado, y se orienta bajo un enfoque de **Design Science Research (DSR)**, en el cual el producto central es un **artefacto tecnológico funcional**: un sistema de software que integra modelos de deep learning, procesamiento de imágenes y una interfaz móvil para el tamizaje de anemia.

---

## 1.2. Planteamiento del problema

### 1.2.1. Situación problemática

Pese a la alta prevalencia de la anemia y a sus impactos negativos en la salud pública, el acceso al diagnóstico oportuno continúa siendo limitado en muchos contextos, en particular en áreas rurales y en centros de atención primaria con infraestructura restringida. En estos entornos, los hemogramas se solicitan de manera selectiva —por costo o disponibilidad— y, frecuentemente, sólo ante casos más graves o sintomáticos.

Aunque existen lineamientos clínicos que recomiendan el uso de signos físicos (palidez en conjuntiva, palmas, lecho ungueal) para sospechar anemia, esta evaluación es:

- **Subjetiva**, dependiente de la experiencia y criterio del profesional.
- Poco reproducible entre distintos observadores.
- Sensible a variaciones en la iluminación, el tono de piel y otros factores.

journal.pone.0253495 (2)

Al mismo tiempo, investigaciones recientes han demostrado —con resultados prometedores— que el análisis automatizado de imágenes de la conjuntiva, mediante

algoritmos de visión por computadora y aprendizaje profundo, puede detectar anemia con altos niveles de exactitud, sensibilidad y especificidad. Algunos trabajos reportan AUC superiores a 0.95 y accuracies mayores al 90 %.

Sin embargo, se identifican las siguientes brechas:

1. **Dependencia de segmentación manual:** Muchos estudios recortan manualmente la región de interés (ROI) de la conjuntiva antes de procesarla, lo que limita la escalabilidad y la automatización en entornos reales.
2. **Uso limitado de técnicas de detección de objetos:** Pocos trabajos integran modelos de detección como YOLO para localizar de manera automática la conjuntiva palpebral en imágenes capturadas en condiciones no controladas.
3. **Falta de integración en un sistema móvil completo:** Aun cuando se reportan modelos con buen desempeño, en muchos casos se trata de prototipos de laboratorio sin una implementación integrada en una aplicación móvil real y un backend optimizado.
4. **Comparaciones parciales entre arquitecturas:** No siempre se realiza una comparación sistemática entre diferentes arquitecturas de redes neuronales convolucionales (AlexNet, ResNet, MobileNet, modelos ligeros para móviles), ni se exploran esquemas de ensamble para maximizar el rendimiento.

En consecuencia, el problema central puede sintetizarse así:

A pesar de la existencia de métodos no invasivos basados en imágenes para la detección de anemia, todavía no se dispone de un sistema inteligente integrado, optimizado para dispositivos móviles, que combine detección automática de la conjuntiva palpebral con modelos de clasificación robustos y ligeros, y que haya sido evaluado experimentalmente sobre un dataset público con métricas comparables o superiores a las reportadas en el estado del arte.

### 1.2.2. Problema general

¿Cómo diseñar, implementar y validar un **sistema inteligente basado en visión por computadora**, ejecutable en dispositivos móviles, que **detecte no invasivamente la anemia** a partir de imágenes de la **conjuntiva palpebral**, con un desempeño comparable o superior a los modelos de referencia reportados en la literatura?

### 1.2.3. Problemas específicos

1. ¿Qué características clínicas y técnicas justifican el uso de la conjuntiva palpebral como región de interés principal para la detección no invasiva de anemia?
2. ¿Qué arquitecturas de deep learning (por ejemplo, AlexNet, ResNet50, MobileNetV2/V3) son más adecuadas para la clasificación de anemia a partir de imágenes de conjuntiva capturadas con teléfonos inteligentes?
3. ¿Cómo integrar un modelo de **detección de objetos** (por ejemplo, YOLOv8) para la localización automática de la conjuntiva palpebral, y cuál es su impacto en el desempeño global del sistema?

4. ¿Qué esquema de entrenamiento, validación y comparación experimental permite evaluar rigurosamente el desempeño del sistema propuesto frente a modelos baseline y frente a resultados del estado del arte?
  5. ¿Cómo diseñar e implementar una arquitectura de software (backend + frontend móvil) que permita el despliegue del modelo en un flujo de uso realista sobre dispositivos móviles?
- 

### 1.3. Formulación de la pregunta de investigación

Atendiendo a los lineamientos para la formulación de preguntas de investigación en proyectos tecnológicos de IA, una buena pregunta debe ser relevante, específica, viable, basada en IA y susceptible de evaluación mediante métricas cuantitativas.

Maestria\_IA04

En este sentido, la pregunta central de esta tesis se formula como:

**¿Puede un sistema inteligente basado en visión por computadora, que combina detección automática de la conjuntiva palpebral mediante YOLO y clasificación mediante redes neuronales convolucionales optimizadas para dispositivos móviles, detectar anemia de manera no invasiva con una  $AUC \geq 0.97$  y una exactitud  $\geq 92\%$  a partir de imágenes capturadas con dispositivos móviles, utilizando un dataset público de imágenes de conjuntiva?**

De esta pregunta principal se derivan subpreguntas específicas:

- ¿La detección automática de la región de la conjuntiva mejora las métricas de clasificación en comparación con un flujo basado en recortes manuales?
  - ¿Un modelo ligero (MobileNet) puede igualar o superar el desempeño de arquitecturas más pesadas (AlexNet, ResNet50) manteniendo viabilidad de ejecución en móviles?
  - ¿Es posible integrar el modelo en una arquitectura cliente-servidor (app móvil + API) que ofrezca tiempos de respuesta compatibles con un uso clínico de tamizaje?
- 

### 1.4. Justificación

La justificación de este trabajo se desarrolla en cuatro dimensiones: **social, clínica, tecnológica y académica.**

#### 1.4.1. Justificación social y clínica

La anemia impacta de manera significativa en la calidad de vida y en la mortalidad de poblaciones vulnerables: niños pequeños, mujeres embarazadas, adultos mayores y pacientes con enfermedades crónicas. Detectarla de forma temprana y frecuente permite

iniciar intervenciones nutricionales y terapéuticas oportunas, reduciendo complicaciones y carga económica para los sistemas de salud.

Un sistema de **detección no invasiva basado en la cámara de un dispositivo móvil**:

- Facilita el tamizaje en **escenarios de bajos recursos** donde no hay laboratorios cercanos.
- Reduce el miedo asociado a la toma de muestras de sangre, especialmente en población pediátrica.

journal.pone.0253495 (2)

- Puede ser utilizado por personal de salud de primer nivel —o incluso por agentes comunitarios— para priorizar la derivación de casos sospechosos.

#### 1.4.2. Justificación tecnológica

Desde el punto de vista tecnológico:

- La combinación de **visión por computadora, deep learning y dispositivos móviles** constituye un campo en rápido crecimiento, con aplicaciones en telemedicina, salud digital y diagnóstico asistido por IA.
- Integrar un modelo de **detección de objetos (YOLOv8)** para segmentar automáticamente la conjuntiva, seguido de modelos de clasificación ligeros (MobileNet, ResNet), permite diseñar un pipeline más robusto que los enfoques que dependen sólo de recortes manuales o de preprocesamiento simple.
- El uso de un dataset público (por ejemplo, de Kaggle) favorece la **reproducibilidad** y posibilita comparar resultados con trabajos futuros.

Además, el enfoque metodológico de **MLOps** permite dejar la base para futuros despliegues productivos, integrando prácticas de versionado de datos, seguimiento de experimentos y contenedorización del modelo.

Maestria\_IA05

#### 1.4.3. Justificación académica y científica

A nivel académico, la tesis aporta:

- Un **estado del arte detallado y actualizado** de métodos no invasivos para detección de anemia basados en imágenes de conjuntiva, palmas, uñas y rostro, sintetizando métricas y enfoques.
- Una **propuesta de arquitectura de investigación tecnológica** basada en Design Science Research (DSR) aplicada a IA, alineada con las guías del programa de maestría.
- Una comparación rigurosa entre **múltiples arquitecturas de deep learning** (AlexNet, ResNet50, MobileNet, ensambles) en el contexto específico de imágenes de conjuntiva.

Finalmente, los resultados y el artefacto obtenido tienen potencial para derivar en **publicaciones científicas** (artículos o ponencias) enfocadas en visión por computadora aplicada a salud.

---

## 1.5. Objetivos

### 1.5.1. Objetivo general

Diseñar, implementar y validar un **sistema inteligente basado en visión por computadora** para la **detección no invasiva de anemia** mediante el análisis de imágenes de la **conjuntiva palpebral** capturadas con dispositivos móviles, empleando modelos de detección y clasificación basados en deep learning, y evaluando su desempeño sobre un dataset público.

### 1.5.2. Objetivos específicos

1. **Revisar y sistematizar** el estado del arte de métodos no invasivos de detección de anemia basados en imágenes y modelos de aprendizaje profundo, con énfasis en la conjuntiva palpebral.
2. **Seleccionar y preprocesar** un dataset público de imágenes de conjuntiva capturadas con dispositivos móviles, definiendo particiones de entrenamiento, validación y prueba.
3. **Diseñar e implementar** un módulo de **detección automática de la conjuntiva palpebral** utilizando un modelo de la familia YOLO (por ejemplo, YOLOv8), capaz de localizar la región de interés en imágenes de entrada.
4. **Entrenar y comparar** diferentes arquitecturas de redes neuronales convolucionales (AlexNet, ResNet50, MobileNetV2/V3 y, de ser viable, un modelo de ensamble) para la clasificación de anemia (anémico / no anémico) a partir de la región de conjuntiva segmentada.
5. **Integrar los módulos de detección y clasificación** en una arquitectura de software tipo cliente-servidor (API + aplicación móvil), definiendo el flujo de captura de imagen, envío, inferencia y presentación de resultados.
6. **Evaluar el desempeño** del sistema propuesto utilizando métricas estándar (accuracy, sensibilidad, especificidad, F1-score, AUC-ROC) y técnicas de validación (k-fold estratificado, bootstrapping, pruebas estadísticas pareadas) para comparar con modelos baseline y resultados de la literatura.
7. **Documentar y analizar críticamente** los resultados experimentales, discutiendo limitaciones, posibles sesgos y oportunidades de mejora futura.

---

## 1.6. Hipótesis de investigación

Siguiendo la conceptualización de hipótesis tecnológicas propuesta en el marco de proyectos de IA, las hipótesis se formulan como afirmaciones verificables sobre el desempeño esperado del artefacto en condiciones específicas.

### 1.6.1. Hipótesis general

#### H1 (General):

Un sistema inteligente basado en visión por computadora que combina detección automática de la conjuntiva palpebral mediante un modelo YOLO y clasificación mediante redes neuronales convolucionales optimizadas para dispositivos móviles (MobileNetV3 y ResNet50), entrenado sobre un dataset público de imágenes de conjuntiva, **logra detectar anemia de manera no invasiva con una AUC  $\geq 0.97$  y una exactitud  $\geq 92$  %**, superando el desempeño de modelos baseline tradicionales (por ejemplo, AlexNet) reportados en la literatura.

### 1.6.2. Hipótesis específicas

- **H2:** La incorporación de un módulo de detección de la conjuntiva palpebral basado en YOLO mejora en al menos un 5 % el F1-score del clasificador final en comparación con un flujo basado en recortes manuales de ROI.
  - **H3:** Un modelo ligero de clasificación (MobileNetV3) es capaz de igualar o superar el desempeño de arquitecturas más pesadas (ResNet50) en términos de AUC y F1-score, manteniendo tiempos de inferencia compatibles con un uso en dispositivos móviles (por ejemplo, <200 ms por imagen en un smartphone de gama media).
  - **H4:** Un esquema de ensamble (por ejemplo, stacking de ResNet50 y MobileNetV3) mejora significativamente la AUC-ROC respecto al mejor modelo individual, con un p-value < 0.05 en pruebas estadísticas pareadas (t-test o Wilcoxon) aplicadas sobre resultados de validación cruzada.
- 

## 1.7. Alcances y limitaciones

### 1.7.1. Alcances

- El sistema está orientado al **tamizaje de anemia**, no al diagnóstico definitivo; los resultados deben ser interpretados como apoyo a la decisión, no como sustituto de un hemograma.
- El modelo se entrenará y validará sobre un **dataset público de Kaggle** de imágenes de conjuntiva palpebral, etiquetadas en función de niveles de hemoglobina (anémico / no anémico).
- El artefacto tecnológico contemplará:
  - Módulo de detección de conjuntiva.
  - Módulo de clasificación de anemia.
  - API de inferencia (por ejemplo, FastAPI).
  - Prototipo de aplicación móvil (Android o multiplataforma).

### 1.7.2. Limitaciones

- La calidad y diversidad de las imágenes dependen del dataset disponible; no se garantiza representatividad de todas las posibles etnias, rangos etarios o condiciones clínicas.



- No se incluye en este trabajo la validación clínica en terreno con pacientes reales; el alcance experimental se limita al desempeño sobre el dataset seleccionado.
  - Las métricas de desempeño se verán influenciadas por el balance de clases (anémico / no anémico) y por la calidad de las anotaciones originales.
  - El análisis no considera integración con historias clínicas electrónicas ni flujos de trabajo hospitalarios complejos; se focaliza en el componente algorítmico y en el prototipo de aplicación.
- 

## 1.8. Supuestos de la investigación

- Se asume que las etiquetas de anemia presentes en el dataset (por ejemplo, anémico / no anémico) se basan en mediciones de hemoglobina obtenidas por métodos de laboratorio válidos.
  - Se asume que las imágenes fueron capturadas con condiciones mínimas de calidad (enfoque razonable, iluminación suficiente) que permiten la identificación de la conjuntiva.
  - Se asume que los dispositivos móviles objetivo para despliegue cuentan con recursos de hardware suficientes para ejecutar la aplicación (o, en su defecto, para comunicarse con un servidor remoto que ejecute la inferencia).
  - Se asume la disponibilidad de herramientas de software estándar (Python, frameworks de deep learning, bibliotecas de visión por computadora, entornos de desarrollo móvil).
- 

## 1.9. Estructura de la tesis

La tesis se organiza en los siguientes capítulos, siguiendo las recomendaciones para informes técnicos y tesis de maestría en proyectos de IA:

- **Capítulo 1 – Introducción:** Presenta el contexto, el planteamiento del problema, la formulación de la pregunta de investigación, la justificación, los objetivos, las hipótesis, los alcances y la estructura general del documento.
- **Capítulo 2 – Marco teórico:** Desarrolla los conceptos fundamentales sobre anemia, hemoglobina, métodos de diagnóstico, fundamentos de visión por computadora, redes neuronales convolucionales, modelos de detección de objetos (YOLO), modelos de clasificación (AlexNet, ResNet, MobileNet) y métricas de evaluación.
- **Capítulo 3 – Estado del arte:** Analiza críticamente la literatura científica relacionada con la detección no invasiva de anemia mediante imágenes (conjuntiva, palmas, rostro, uñas), comparando datasets, modelos, métricas y limitaciones, y destacando la posición del presente trabajo frente a las propuestas existentes.
- **Capítulo 4 – Metodología:** Describe el enfoque metodológico basado en Design Science Research, CRISP-DM, Agile y prácticas de MLOps. Detalla las fases del proyecto, desde la selección del dataset y el preprocesamiento, hasta el entrenamiento, validación y despliegue del modelo.

- **Capítulo 5 – Implementación del sistema:** Presenta la arquitectura del sistema, los módulos de detección y clasificación, la integración con la API y la app móvil, así como las decisiones de diseño de software, hardware y optimización.
  - **Capítulo 6 – Resultados y análisis:** Expone las métricas cuantitativas de desempeño, las comparaciones entre modelos, las curvas ROC y PR, las matrices de confusión, y el análisis estadístico de los resultados.
  - **Capítulo 7 – Discusión:** Interpreta los resultados a la luz del estado del arte, analiza fortalezas y debilidades del sistema propuesto, y discute las implicancias para la práctica clínica y el futuro desarrollo tecnológico.
  - **Capítulo 8 – Conclusiones y trabajo futuro:** Resume los principales aportes de la tesis, responde explícitamente a la pregunta de investigación y a las hipótesis planteadas, y propone líneas de trabajo futuro para extender y validar el sistema en entornos reales.
- 

Con esto, el **Capítulo 1** queda redactado de forma completa, técnica y alineada con tu maestría.

Cuando quieras, seguimos con el **Capítulo 2 – Marco teórico**, donde desarrollaremos:

- Anemia y fisiopatología
- Métodos de diagnóstico (invasivos vs no invasivos)
- Fundamentos de visión por computadora
- CNN, YOLO, AlexNet, ResNet, MobileNet
- Métricas de evaluación en clasificación médica.

Fuentes

Dijiste:

continúa

ChatGPT dijo:

Pensado durante 14s

Perfecto, seguimos entonces con el **CAPÍTULO 2 – MARCO TEÓRICO**.

---

## CAPÍTULO 2

### MARCO TEÓRICO

#### 2.1. Anemia: conceptos fundamentales

##### 2.1.1. Definición y fisiopatología

La anemia se define clásicamente como una **disminución de la concentración de hemoglobina (Hb)** o del número de eritrocitos por debajo de los valores de referencia para la edad, sexo y contexto fisiológico de la persona. Esta reducción compromete la capacidad de la sangre para transportar oxígeno a los tejidos, generando hipoxia tisular y una serie de manifestaciones clínicas como fatiga, palidez, taquicardia, disnea de esfuerzo y mareos.

Desde el punto de vista fisiopatológico, la anemia puede originarse por tres mecanismos generales:

1. **Disminución de la producción de eritrocitos**
  - Deficiencia de hierro, vitamina B12 o ácido fólico.
  - Fallo medular (aplasia, infiltración tumoral).
  - Insuficiencia renal crónica (déficit de eritropoyetina).
2. **Aumento de la destrucción de eritrocitos (hemólisis)**
  - Hemoglobinopatías (p. ej. drepanocitosis, talasemias).
  - Anemias hemolíticas autoinmunes.
  - Infecciones (p. ej. malaria) o tóxicos.
3. **Pérdida sanguínea aguda o crónica**
  - Hemorragias digestivas, ginecológicas, traumáticas, etc.

La consecuencia final, independientemente del mecanismo, es una reducción de la masa eritrocitaria y de la disponibilidad de hemoglobina funcional, con impacto directo en la oxigenación tisular. En niños, la anemia se asocia con retraso en el desarrollo psicomotor, alteraciones del rendimiento escolar y mayor susceptibilidad a infecciones.

### 2.1.2. Epidemiología y grupos de riesgo

Estudios recientes estiman que la anemia afecta a aproximadamente **un tercio de la población mundial**, con prevalencias particularmente elevadas en:

- Niños de 6 a 59 meses.
- Mujeres embarazadas.
- Mujeres no embarazadas de 15 a 49 años.

En países de ingresos bajos y medios, las tasas de anemia pueden superar el 50 % en algunos grupos vulnerables, alimentadas por factores como:

- Malnutrición e inseguridad alimentaria.
- Alta carga de enfermedades infecciosas (malaria, helmintiasis).
- Acceso limitado a servicios de salud y diagnóstico.

La combinación de alta prevalencia y consecuencias a largo plazo convierte a la anemia en un problema prioritario de salud pública y en un **candidato natural para intervenciones de tamizaje masivo y diagnóstico temprano**, especialmente mediante herramientas accesibles y de bajo costo.

### 2.1.3. Manifestaciones clínicas y signos físicos

Las manifestaciones clínicas de la anemia dependen de la severidad, la velocidad de instauración y la condición basal del paciente. Entre los signos físicos clásicos se encuentran:

- Palidez cutáneo-mucosa: particular en **conjuntiva palpebral**, palmas, lecho ungueal y mucosa oral.
- Taquicardia, soplos funcionales.
- Disnea, fatiga fácil, intolerancia al ejercicio.
- En casos severos: síncope, angina, insuficiencia cardíaca.

La **palidez conjuntival** ha sido ampliamente utilizada como signo clínico indirecto de anemia y constituye la base de numerosos estudios que exploran la correlación entre la colorimetría de esta región y los niveles de hemoglobina.

---

## 2.2. Métodos de diagnóstico de la anemia

### 2.2.1. Métodos invasivos convencionales

El **estándar de oro** en el diagnóstico de anemia es el **hemograma** completo, que cuantifica:

- Concentración de hemoglobina (Hb).
- Hematocrito.
- Recuento de eritrocitos.
- Índices eritrocitarios (VCM, HCM, CHCM).

Algunos métodos clásicos para la medición de hemoglobina incluyen:

- **Método de cianometahemoglobina**: conversión química de la Hb a una forma estable para su medición espectrofotométrica.

hir-2025-31-1-57 (1)\_español (4)

- **Dispositivos portátiles (HemoCue)**: permiten medición en el punto de atención a partir de sangre capilar.

Estas técnicas, si bien precisas, tienen importantes limitaciones:

- **Invasividad**: requieren punción venosa o capilar.
- **Costos**: necesidad de reactivos y equipamiento.
- **Infraestructura**: laboratorios o equipos al menos básicos.
- **Tiempo**: demoras en procesamiento y entrega de resultados.

En entornos rurales o con recursos limitados, estas barreras se traducen en un **subdiagnóstico** y en la falta de tamizaje sistemático.

### 2.2.2. Métodos no invasivos: colorimetría e imagen

Ante estas limitaciones, diversos grupos han explorado **métodos no invasivos** basados en el análisis de color y textura de regiones anatómicas relacionadas con la perfusión sanguínea:

- Palma de la mano.
- Lecho ungueal.
- Mucosa sublingual.
- **Conjuntiva palpebral.**

Los enfoques incluyen:

1. **Colorimetría directa**
  - Uso de tarjetas de referencia de color o espectrofotómetros de contacto.
  - Estimación de la concentración de Hb a partir de curvas de calibración.
2. **Imagen digital y procesamiento de color**
  - Captura de imágenes con cámaras convencionales o de smartphones.
  - Conversión a espacios de color (RGB, LAB, HSV).
  - Extracción de características cromáticas (p. ej. componente a\* del espacio CIELAB).
3. **Modelos de regresión y clasificación**
  - Regresión lineal/multilineal para estimar Hb.
  - Clasificadores (SVM, KNN, Naïve Bayes, redes neuronales) para distinguir anémico / no anémico.

Varios trabajos han demostrado que la palidez conjuntival y de la palma presenta buena correlación con los niveles de Hb, con sensibilidades y especificidades superiores al 90 % en algunos casos.

### 2.2.3. Imágenes de la conjuntiva palpebral

La **conjuntiva palpebral** es una membrana mucosa delgada y altamente vascularizada que recubre la cara interna de los párpados. Presenta las siguientes ventajas como región de interés (ROI) para la detección de anemia:

- **Alta vascularización** → cambios en perfusión y hemoglobina se reflejan en el color.
- **Baja presencia de melanina** → menor interferencia de pigmentación cutánea.
- Accesible mediante maniobras sencillas de eversión o descenso del párpado inferior.

Estudios como los de Appiahene et al., Suner et al., Purwanti et al. y Sehar et al. han mostrado que es posible capturar imágenes de la conjuntiva con cámaras de teléfonos inteligentes y, mediante procesamiento de imágenes e IA, detectar anemia con alta exactitud.

---

## 2.3. Visión por computadora aplicada a la salud

### 2.3.1. Conceptos básicos de visión por computadora

La **visión por computadora** es una rama de la inteligencia artificial que busca **permitir que las máquinas “vean” e interpreten imágenes y videos** de manera similar —o superior— a los humanos. Sus tareas típicas incluyen:

- Clasificación de imágenes.
- Detección y localización de objetos.
- Segmentación semántica y por instancias.
- Reconocimiento facial.
- Estimación de pose, seguimiento de objetos y análisis de movimiento.

En el ámbito biomédico, la visión por computadora se ha utilizado en:

- Diagnóstico por imágenes (radiografías, tomografías, resonancias).
- Detección de lesiones en piel.
- Análisis de retinografías (retinopatía diabética, glaucoma).
- Análisis de video endoscópico.
- Procesamiento de imágenes de microscopía.

La disponibilidad de datasets anotados y la capacidad de cómputo han impulsado especialmente el uso de **redes neuronales convolucionales (CNN)**, que han desplazado a muchos enfoques clásicos de ingeniería de características manuales.

### **2.3.2. Flujo general en aplicaciones de visión por computadora**

La mayoría de sistemas de visión por computadora siguen un flujo general:

1. **Adquisición de datos**
  - Imágenes estáticas o secuencias de video.
  - Sensores específicos (cámaras RGB, infrarrojas, hiperespectrales).
2. **Preprocesamiento**
  - Normalización de tamaño, color, iluminación.
  - Corrección de artefactos, ruido.
  - Aumento de datos (rotaciones, recortes, cambios de brillo, etc.).
3. **Detección / segmentación de regiones de interés**
  - Modelos de detección de objetos (YOLO, SSD, Faster R-CNN).
  - Segmentación semántica (U-Net, DeepLab, etc.).
4. **Extracción de características y clasificación**
  - CNN como extractores de características.
  - Capas densas para clasificación / regresión.
5. **Postprocesamiento y visualización**
  - Umbralización de probabilidad.
  - Mapas de calor, explicabilidad (Grad-CAM, LRP).

Este esquema sirve como base conceptual para el sistema propuesto en esta tesis.

---

## **2.4. Redes neuronales convolucionales (CNN)**

### **2.4.1. Fundamentos**

Las **redes neuronales convolucionales** son arquitecturas de deep learning diseñadas para procesar datos con estructura de grilla, como imágenes. Se componen típicamente de:

- **Capas convolucionales:** aplican filtros (kernels) que aprenden patrones locales (bordes, texturas, formas).
- **Capas de pooling:** reducen dimensionalidad y otorgan invariancia a traslaciones pequeñas.
- **Capas completamente conectadas:** combinan las características extraídas para realizar la clasificación o regresión.

El entrenamiento se realiza mediante descenso de gradiente y backpropagation, minimizando una función de pérdida (p. ej. entropía cruzada en clasificación) con optimizadores como SGD, Adam, RMSProp, etc.

#### 2.4.2. Transfer learning

Dado que entrenar una CNN desde cero requiere grandes volúmenes de datos, en la práctica es habitual emplear **transfer learning**:

- Se utiliza una arquitectura preentrenada en grandes datasets (p. ej. ImageNet).
- Se **congela** parte de la red (capas iniciales) y se reentrenan las capas finales en el nuevo dominio (p. ej. conjuntiva).
- Alternativamente, se hace fine-tuning de varias capas con una tasa de aprendizaje reducida.

Esto permite aprovechar características visuales genéricas (bordes, texturas, formas) ya aprendidas, reduciendo tiempo de entrenamiento y riesgo de sobreajuste, algo especialmente útil en datasets médicos de tamaño moderado.

#### 2.4.3. Arquitecturas relevantes para la tesis

##### a) AlexNet

- Una de las primeras CNN profundas en ganar el desafío ImageNet (2012).
- Estructura relativamente simple: 5 capas convolucionales + 3 capas totalmente conectadas.
- Usa ReLU, dropout y data augmentation básico.

En aplicaciones modernas, AlexNet suele utilizarse como **baseline**, dado que arquitecturas más recientes la superan en precisión y eficiencia.

m28937purwantistamped (4)

##### b) ResNet50

- Introduce el concepto de **bloques residuales**, que permiten entrenar redes muy profundas evitando el problema del gradiente desvanecido.
- ResNet50 tiene 50 capas y ha demostrado excelente desempeño en múltiples tareas de visión.

- En tareas médicas, ResNet50 se usa frecuentemente como backbone de clasificación, dada su buena relación entre capacidad y estabilidad en entrenamiento.

#### c) MobileNetV2 / MobileNetV3

- Arquitecturas diseñadas específicamente para dispositivos móviles y de recursos limitados.
- Utilizan bloques de **convoluciones separables en profundidad** (depthwise separable convolutions) y otros trucos de eficiencia.
- MobileNetV3 introduce mejoras adicionales como bloques squeeze-and-excitation y optimizaciones basadas en búsqueda de arquitectura neural (NAS).

En la literatura sobre anemia con imágenes de conjuntiva, MobileNet ha mostrado desempeños comparables a ResNet, con menor costo computacional, lo que lo hace ideal para despliegue en smartphones.

#### d) Ensembles de CNN

Algunos estudios han explorado **ensembles (ensambles)** de múltiples CNN, combinando:

- VGG16
- ResNet50
- InceptionV3

mediante técnicas de stacking, logrando **AUC de 0.97** en detección de anemia a partir de imágenes de conjuntiva.

hir-2025-31-1-57 (1)\_español (4)

Este enfoque inspira la posible combinación de ResNet y MobileNet en la tesis, para obtener un clasificador más robusto.

## 2.5. Detección de objetos con YOLO

### 2.5.1. Problema de detección de objetos

La **detección de objetos** busca, en una imagen:

- Localizar uno o varios objetos mediante **cajas delimitadoras (bounding boxes)**.
- Clasificar cada objeto detectado en una categoría.

En el caso de la conjuntiva palpebral, el problema se simplifica a detectar una región anatómica específica (ROI) dentro de la imagen del ojo.

### 2.5.2. YOLO: “You Only Look Once”



YOLO es una familia de modelos de detección de objetos caracterizada por:

- Tratar la detección como un **problema de regresión** directo desde la imagen completa hacia bounding boxes + clases.
- Alta velocidad de inferencia, ideal para aplicaciones en tiempo real.

Entre sus versiones modernas (YOLOv5, YOLOv7, YOLOv8, YOLOv9), se han incorporado mejoras en:

- Backbone (Redes tipo CSP, ConvNeXt, etc.).
- Módulos de atención.
- Pérdidas específicas para mejor ajuste de cajas.

Para la tesis, YOLO se utilizará para:

1. **Detectar automáticamente la región de conjuntiva** en la imagen capturada.
2. Recortar dicha región y enviarla a la CNN de clasificación (AlexNet/ResNet/MobileNet).

Este enfoque evita la dependencia de recortes manuales y permite que el sistema funcione de manera más robusta en condiciones de iluminación y encuadre variables, como las que se encuentran al usar un smartphone en entornos reales.

---

## 2.6. Dispositivos móviles como plataforma de diagnóstico

### 2.6.1. Penetración y ventajas

Los smartphones se han convertido en dispositivos omnipresentes en gran parte del mundo, incluyendo países en desarrollo. Su adopción creciente los convierte en una plataforma ideal para:

- **Telemedicina** y seguimiento remoto de pacientes.
- Aplicaciones de **mHealth** (mobile health).
- Herramientas de **diagnóstico asistido**.

Estudios como los de Suner et al. y Wemyss et al. han demostrado que es posible utilizar la cámara del smartphone para capturar imágenes de conjuntiva, palmas y otros tejidos, y derivar de ellas estimaciones de hemoglobina o clasificación de anemia con buenos resultados.

### 2.6.2. Consideraciones técnicas

Al diseñar sistemas de visión por computadora para móviles se deben considerar:

- **Resolución y calidad de la cámara.**
- Variaciones en **iluminación ambiental**.
- Estabilidad de la mano del usuario (desenfoque).
- Capacidad de procesamiento local vs uso de servidores remotos.

- Latencia aceptable para el usuario final.

Dos arquitecturas típicas son:

### 1. **Procesamiento local**

- El modelo de IA se ejecuta en el propio dispositivo (TensorFlow Lite, ONNX Runtime Mobile).
- Mayor independencia de la conectividad, pero limita el tamaño del modelo.

### 2. **Procesamiento en servidor (cliente-servidor)**

- La app envía la imagen o características a un backend (FastAPI, Flask, etc.) donde se ejecuta el modelo.
- Reduce requerimientos de hardware en el dispositivo, pero depende de conectividad.

La tesis se enfocará en una arquitectura **cliente-servidor**, sin descartar la exploración de modelos optimizados para ejecución local.

## 2.7. Métricas de evaluación en modelos de clasificación médica

En aplicaciones médicas, no basta con reportar una sola métrica; se requiere un conjunto de indicadores que reflejen el desempeño de manera integral.

### 2.7.1. Matriz de confusión

A partir de la matriz de confusión (TP, TN, FP, FN) se derivan:

- **Exactitud (Accuracy)**

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Sensibilidad (Recall o TPR)**

$$\text{Sensibilidad} = \frac{TP}{TP + FN}$$

Mide la capacidad de detectar correctamente a los pacientes anémicos.

- **Especificidad (TNR)**

$$\text{Especificidad} = \frac{TN}{TN + FP}$$

Mide la capacidad de identificar correctamente a los no anémicos.

- **Precisión (Precision o PPV)**

$$\text{Precisión} = \frac{TP}{TP + FP}$$

- **F1-score**

$$F1 = 2 \cdot \frac{\text{Precisión} \cdot \text{Sensibilidad}}{\text{Precisión} + \text{Sensibilidad}}$$

En el contexto de tamizaje de anemia, la **sensibilidad** suele ser prioritaria, ya que se busca minimizar falsos negativos.

### 2.7.2. Curva ROC y AUC

La **curva ROC (Receiver Operating Characteristic)** representa la relación entre:

- Tasa de verdaderos positivos (TPR).
- Tasa de falsos positivos (FPR).

El **área bajo la curva (AUC)** proporciona un resumen global del desempeño del clasificador para diferentes umbrales de decisión.

- $AUC = 0.5 \rightarrow$  modelo aleatorio.
- $AUC \geq 0.9 \rightarrow$  desempeño excelente.

Varios estudios de detección de anemia mediante imágenes reportan AUC en el rango de 0.90–0.97, lo que sirve como referencia para fijar la meta de desempeño del sistema propuesto.

### 2.7.3. Validación cruzada y pruebas estadísticas

Además de métricas puntuales, es importante evaluar la **estabilidad** y **significancia** de los resultados:

- **Validación cruzada k-fold:** divide el dataset en k particiones; cada partición es usada una vez como conjunto de prueba y las demás como entrenamiento, promediando los resultados.

Maestria\_IA07

- **Bootstrapping:** re-muestreo con reemplazo para estimar intervalos de confianza de las métricas.
- **Pruebas pareadas (t-test, Wilcoxon):** permiten determinar si las diferencias de desempeño entre dos modelos (p. ej. ResNet vs MobileNet, con y sin YOLO) son estadísticamente significativas.

Estas herramientas serán usadas en la tesis para comparar sistemáticamente las arquitecturas propuestas y validar las hipótesis planteadas en el Capítulo 1.

---

## 2.8. Enfoques metodológicos en proyectos de IA

Aunque la **metodología detallada** se desarrollará en el Capítulo 4, es relevante introducir aquí brevemente algunos marcos de trabajo que fundamentan el diseño del sistema.

### 2.8.1. Design Science Research (DSR)

El **Design Science Research** es un enfoque de investigación aplicada centrado en la **creación y evaluación de artefactos** (modelos, métodos, sistemas) que resuelven problemas reales. Sus fases principales incluyen: identificación del problema, definición de objetivos, diseño y construcción del artefacto, demostración, evaluación y comunicación.

Maestria\_IA05

En esta tesis, el artefacto es el **sistema inteligente de detección de anemia**, compuesto por:

- Módulo de visión por computadora (detección + clasificación).
- Backend de inferencia.
- Aplicación móvil de interacción con el usuario.

### 2.8.2. CRISP-DM y SEMMA en minería de datos

Modelos como **CRISP-DM** y **SEMMA** proporcionan estructuras para proyectos de minería de datos y aprendizaje automático, incluyendo fases de:

- Comprensión del negocio y de los datos.
- Preparación de datos.
- Modelado.
- Evaluación.
- Despliegue.

Maestria\_IA05

Estos marcos se adaptan en la tesis para estructurar el pipeline de datos (selección, limpieza, partición, modelado y evaluación) del problema de imágenes de conjuntiva.

### 2.8.3. Agile + IA y MLOps

Finalmente, las prácticas de **Agile + IA** y **MLOps** permiten:

- Planificar el desarrollo en **sprints** con entregables parciales (modelos baseline, pipeline de entrenamiento, API de inferencia).
  - Integrar el modelo en un flujo de integración y despliegue continuo (CI/CD).
  - Monitorear desempeño y facilitar futuros reentrenamientos con nuevos datos.
-

Con esto, el **Capítulo 2 – Marco teórico** queda desarrollado: desde los fundamentos clínicos de la anemia, hasta los conceptos centrales de visión por computadora, CNN, YOLO, dispositivos móviles, métricas y marcos metodológicos de IA.

Si te parece bien, en el siguiente mensaje continuamos con el **CAPÍTULO 3 – ESTADO DEL ARTE**, donde haré:

- Una tabla comparativa de los papers que me compartiste.
- Análisis crítico de datasets, modelos, métricas y limitaciones.
- Identificación precisa del “gap” que llena tu tesis.

Fuentes

Dijiste:

continua siempre teniendo como referencia toda la informacion que te he comparrtido

ChatGPT dijo:

Pensado durante 33s

Perfecto, seguimos con el **CAPÍTULO 3 – ESTADO DEL ARTE**, siempre usando como referencia tus PPTs de maestría y todos los papers que me compartiste.

---

# CAPÍTULO 3

## ESTADO DEL ARTE

### 3.1. Metodología de revisión de la literatura

Siguiendo las pautas metodológicas propuestas en tus materiales de maestría para la construcción del estado del arte, la revisión se realizó con un enfoque **sistemático y tecnológico**, orientado a:

- Identificar **artefactos previos** (modelos, sistemas, prototipos) relacionados con la **detección no invasiva de anemia mediante imágenes de conjuntiva**.
- Analizar sus **datasets, técnicas de visión por computadora, métricas de desempeño y limitaciones metodológicas**.
- Determinar el “**gap**” (brecha) específico que la presente tesis busca abordar.

#### 3.1.1. Fuentes de información

Se emplearon principalmente:

- Artículos científicos indexados (Elsevier, PLOS ONE, MDPI, HIR).
- Proceedings de congresos (p. ej. ICTS, conferencia donde se presenta Purwanti et al.).

- Repositorios y datasets públicos (con énfasis posterior en Kaggle para la implementación del prototipo).

La búsqueda, tal como recomiendan tus PPTs, se restringió a trabajos de los **últimos 5–10 años**, con prioridad para publicaciones recientes (2021–2025) en métodos no invasivos de detección de anemia.

Maestria\_IA03

### 3.1.2. Palabras clave y ecuaciones de búsqueda

Se combinaron términos en inglés y español:

- “anemia detection”, “non-invasive anemia”, “conjunctiva images”, “palpebral conjunctiva”,
- “smartphone camera hemoglobin”, “deep learning anemia”, “convolutional neural network anemia”,
- “pallor analysis”, “colorimetry”, “pediatric anemia imaging”.

### 3.1.3. Criterios de inclusión y exclusión

#### Inclusión:

- Estudios que utilicen **imágenes** (no sólo datos tabulares) para detectar o estimar anemia.
- Uso explícito de **conjuntiva palpebral**, rostro, palma o uñas.
- Aplicación de **visión por computadora, machine learning o deep learning**.
- Poblaciones humanas (adultos o pediátricos).

#### Exclusión:

- Métodos puramente espectrofotométricos sin imágenes.
- Estudios exclusivamente hematológicos sin componente de imagen.
- Trabajos sin descripción clara de métricas ni de la metodología de entrenamiento.

Con estos criterios se seleccionaron y analizaron en profundidad los trabajos contenidos en los papers que me compartiste, los cuales constituyen el núcleo del presente estado del arte.

## 3.2. Trabajos sobre detección no invasiva de anemia mediante imágenes de conjuntiva

En esta sección se revisan los estudios más relevantes que utilizan **imágenes de conjuntiva palpebral** capturadas con cámaras o smartphones, para detectar anemia o estimar niveles de hemoglobina.

### 3.2.1. Suner et al. (2021) – Estimación de hemoglobina con cámara de smartphone

Suner et al. proponen el uso de la cámara de un smartphone para **estimar la concentración de hemoglobina (Hb)** y realizar tamizaje de anemia.

journal.pone.0253495 (2)

- **Población y diseño:** estudio prospectivo con 344 pacientes de un servicio de emergencia.
- **Fases:**
  - Fase 1: 142 pacientes, usada para derivar el modelo (desarrollo).
  - Fase 2: 202 pacientes, usada para validación.
- **Adquisición:** imágenes de **ambas conjuntivas** capturadas con un smartphone.
- **Procesamiento:**
  - Selección de una **región de interés (ROI)** en la conjuntiva palpebral.
  - Extracción de parámetros de imagen (canales de color).
  - Uso de **regresión escalonada (stepwise regression)** para predecir Hb (HBc).
- **Resultados:**
  - Hb real (HbI) entre 4.7 y 19.6 g/dL (media 12.5).
  - Fase 1: asociación significativa entre HBc y HbI, con sensibilidad para anemia superior al 90 %.
  - Fase 2: accuracy ~72.6 %, sensibilidad ~72.8 %, especificidad ~72.5 %.
  - Buena exactitud para umbrales de transfusión (Hb < 7 g/dL y < 9 g/dL).

journal.pone.0253495 (2)

#### **Limitaciones identificadas:**

- **Segmentación de ROI manual:** un observador selecciona la zona de conjuntiva, lo que reduce escalabilidad y consistencia.
- **Error dependiente del nivel de Hb:** el modelo tiende a subestimar Hb alta y sobreestimar Hb baja.
- No se utiliza deep learning, sino un enfoque de **regresión clásica** sobre características de color.

#### **Relevancia para la tesis:**

Este trabajo muestra la **factibilidad clínica** de usar un smartphone para estimar Hb a partir de la conjuntiva, pero deja espacio para mejorar con:

- Automatización completa del ROI.
- Deep learning end-to-end.
- Métricas más altas en clasificación binaria (anémico / no anémico).

---

### **3.2.2. Appiahene et al. (2023) – Aplicación móvil con FastAPI y CNN + Regresión**

Appiahene et al. desarrollan un sistema completo de **detección de anemia usando imágenes de conjuntiva**, con arquitectura cliente-servidor.

- **Dataset:** 710 imágenes de conjuntiva de niños capturadas con un **dispositivo especializado** que elimina interferencias de luz ambiente.
- **Enfoque:** análisis de **palidez (pallor)** en la conjuntiva usando machine learning.
- **Modelos:**
  - CNN para detección segmentada de la conjuntiva.
  - Regresión logística para clasificación de anemia.
  - Filtro Gaussian Blur en etapa de preprocesamiento.
- **Arquitectura:**
  - Backend con **FastAPI** en Python.
  - Aplicación móvil **React Native** como frontend.
  - Flujo: el usuario captura una imagen → la app la envía a FastAPI → allí se procesa y se devuelve el resultado (anemia / no anemia).

1-s2.0-S2590093523000322-main (...)

- **Métricas:**
  - Sensibilidad  $\approx 90\%$ .
  - Especificidad  $\approx 95\%$ .
  - Accuracy promedio  $\approx 92.5\%$ , con tiempo de procesamiento  $\approx 50$  s.

#### **Fortalezas:**

- Prototipo **funcional** y bien integrado (modelo + API + app móvil).
- Métricas altas (accuracy  $> 90\%$ ).

#### **Limitaciones:**

- Dependencia de un **dispositivo de adquisición especializado**; menor generalización a cámaras de smartphones sin control de iluminación.
- No se exploran arquitecturas modernas de deep learning (p. ej. ensambles de redes profundas ni modelos específicos para móviles).
- No se utiliza un modelo genérico de detección de objetos como YOLO para la localización automática del ROI; el pipeline está más acoplado a las condiciones del dispositivo de captura.

1-s2.0-S2590093523000322-main (...)

#### **Relación con la tesis:**

El enfoque arquitectónico (FastAPI + app móvil) es directamente reutilizable como referencia de diseño, pero la tesis propone:

- Basarse en **datasets públicos** (p. ej. Kaggle).
- Incorporar **YOLO** y arquitecturas CNN modernas (MobileNet, ResNet).
- Optimizar tiempos de inferencia.

---

### **3.2.3. Wemyss et al. (2023) – Colorimetría facial en niños (conjuntiva + sclera + mucosa)**



Wemyss et al. exploran la **colorimetría de la cara** como herramienta de tamizaje de anemia en niños pequeños en Ghana.

hir-2025-31-1-57 (3)

- **Población:** 62 pacientes < 4 años (43 con imágenes de calidad suficiente).
- **Regiones de interés (ROIs):**
  - Conjuntiva palpebral.
  - Esclerótica.
  - Mucosa adyacente al labio inferior.

hir-2025-31-1-57 (3)

- **Enfoque:**
  - Captura de imágenes faciales mediante smartphone.
  - Corrección de la iluminación ambiente con diferentes estrategias.
  - Extracción de métricas cromáticas en cada ROI.
  - Clasificación mediante **Naïve Bayes** para anemia (< 11 g/dL Hb).
- **Resultados:**
  - Sensibilidad: 92.9 % (IC 66.1–99.8).
  - Especificidad: 89.7 % (IC 72.7–97.8).
  - Buen desempeño usando **solo un smartphone, sin hardware adicional**.

hir-2025-31-1-57 (3)

#### **Fortalezas:**

- Uso de **tres ROIs** para robustecer la estimación.
- Contexto pediátrico real en un hospital de Ghana.
- Implementación sin tarjetas de color ni dispositivos externos.

#### **Limitaciones:**

- Tamaño muestral reducido.
- **No se usa deep learning;** modelo Naïve Bayes con características de color.
- Segmentación de ROIs no está completamente automatizada.

#### **Contribución para la tesis:**

Demuestra que es factible trabajar en entornos clínicos reales con smartphones y sin hardware adicional. La tesis, aunque enfocada en conjuntiva, puede considerar este trabajo como evidencia de que **es posible operar en entornos no controlados**, reforzando la necesidad de un buen preprocesamiento y detección automática.

---

### **3.2.4. Sehar & Krishnamoorthi (2025) – Ensemble de modelos profundos con GAN**

Sehar y Krishnamoorthi presentan un estudio de **detección de anemia mediante imágenes de conjuntiva usando diversos modelos de ML y DL**, culminando en un **ensemble de redes profundas**.

- **Dataset:**
  - 764 imágenes de conjuntiva (anémicos y no anémicos) capturadas con una tableta Samsung Galaxy Tab 7A.
  - Ampliadas a 4,315 imágenes mediante un **Deep Convolutional GAN (DCGAN)** para prevenir sobreajuste y mejorar robustez.
- **Preprocesamiento:**
  - Conversión a espacio de color **LAB**.
  - Segmentación de **ROI** en la conjuntiva palpebral.
  - Cálculo de medias de parámetros  $a^*$ ,  $b^*$  y  $G$ , entre otros.

hir-2025-31-1-57 (1)\_español (4)

- **Modelos evaluados:**
  - Regresión logística, KNN, SVM, modelos estadísticos.
  - Redes profundas: **VGG16**, **ResNet-50**, **InceptionV3**.
  - **Framework de conjunto (stacking)** combinando estas arquitecturas.
- **Resultados:**
  - El **ensemble de stacking** con VGG16, ResNet-50 e InceptionV3 alcanza **AUC = 0.97**, interpretado como capacidad excepcional para detectar anemia de forma no invasiva.

#### **Fortalezas:**

- Uso de **data augmentation avanzado** (GAN).
- Comparación explícita entre múltiples arquitecturas.
- Uso de ensembles, siguiendo mejores prácticas recientes en deep learning.

#### **Limitaciones:**

- No se discute integración en aplicación móvil o sistema tiempo real.
- Dependencia de un dataset específico, capturado en un entorno concreto.
- Segmentación de ROI aún basada en procedimientos controlados, sin incorporar detectores universales como YOLO.

#### **Relación con la tesis:**

Este trabajo define el **nivel de referencia (benchmark)** en términos de AUC ( $\approx 0.97$ ) que la tesis buscará igualar o superar mediante una combinación de:

- Detección automática (YOLO).
  - Arquitecturas ligeras para móviles (MobileNet).
  - Ensamblados adaptados.
-

### 3.2.5. Purwanti et al. (2023) – CNN (AlexNet, ResNet-50, MobileNetV2) en imágenes palpebrales

Purwanti et al. investigan la **clasificación de anemia** a partir de imágenes de la **conjuntiva palpebral** usando arquitecturas CNN populares.

- **Contexto clínico:** la conjuntiva ocular es uno de los mejores lugares para detectar palidez asociada a anemia; estudios previos reportan sensibilidad  $\approx 94.02\%$  y especificidad  $\approx 93.71\%$  para este signo en diagnóstico clínico.

m28937purwantistamped (2) (1)

- **Enfoque del estudio:**
  - Captura de imágenes de la conjuntiva palpebral con cámara.
  - Recorte de la región de interés.
  - Entrenamiento de tres modelos de CNN:
    - AlexNet.
    - ResNet-50.
    - MobileNetV2.
- **Resultados:**
  - AlexNet: accuracy  $\approx 89.93\%$ .
  - MobileNetV2: accuracy  $\approx 97.19\%$ .
  - ResNet-50: accuracy  $\approx 97.94\%$  (mejor desempeño).

#### Fortalezas:

- Comparación directa de **arquitecturas clásicas vs modernas**.
- Evidencia sólida de que **MobileNet y ResNet** logran precisión cercana al 98 % en clasificación de anemia a partir de conjuntiva.

#### Limitaciones:

- Número de muestras limitado (no se detalla gran escala).
- No se incluye una etapa explícita de detección automática de ROI (se asume recorte previo).
- No se describe integración en un pipeline móvil completo.

#### Aportación para la tesis:

Este trabajo refuerza la selección de **ResNet-50 y MobileNetV2/V3** como candidatos principales para el módulo de clasificación de la tesis, y sugiere que AlexNet debe usarse como **baseline histórico** a superar.

---

### 3.2.6. Berghout (2024) – RexNet: red de expansión recurrente para diagnóstico no intrusivo

Berghout propone un enfoque más complejo y generalista para el **diagnóstico no intrusivo de anemia en pacientes pediátricos**, combinando técnicas avanzadas de procesamiento de imágenes, selección de características y redes profundas.

Joint Image Processing with Lea...

- **Imágenes y datasets:**
  - Utiliza **tres datasets públicos**: imágenes de conjuntiva ocular, palmas y uñas de niños  $\leq 6$  años.
- **Pipeline de procesamiento:**
  - Extracción de **181 características** agrupadas en 13 categorías (color, textura, morfología, etc.).
  - Proceso de **selección de características** para identificar las más relevantes.

Joint Image Processing with Lea...

- **Modelos:**
  - Red profunda basada en **LSTM** para clasificación de imágenes en anémico / no anémico.
  - Optimización bayesiana de hiperparámetros.
  - Integración del LSTM en una nueva arquitectura llamada **Recurrent Expansion Network (RexNet)**, que combina representación de datos y “comportamiento del modelo” para mejorar generalización.
- **Desempeño:**
  - **Accuracy global  $\approx 99.83\% \pm 0.02$**  en las métricas de clasificación, superando modelos previos.

Joint Image Processing with Lea...

### Fortalezas:

- Desempeño casi perfecto en sus métricas.
- Propuesta metodológica innovadora (RexNet).
- Uso de múltiples regiones anatómicas (conjuntiva, palma, uñas).

### Limitaciones:

- Complejidad del pipeline (extracción manual de muchas características + LSTM + RexNet).
- Parte del enfoque se asemeja más a **procesamiento de características tabulares** derivadas de imágenes que a una pipeline estrictamente end-to-end de visión por computadora.
- No se discute implementación en dispositivos móviles ni en sistemas de tiempo real.

### Relación con la tesis:

RexNet representa un **techo alto de rendimiento**, pero con un enfoque más pesado y complejo que puede ser difícil de desplegar en entornos móviles. La tesis se centrará en

un **trade-off entre rendimiento y eficiencia**, usando CNN ligeras y detección con YOLO, pero puede utilizar a Berghout como referencia de máximo desempeño alcanzable en contextos controlados.

---

### 3.3. Síntesis comparativa de estudios relevantes

Siguiendo la recomendación de tus PPTs, el estado del arte se resume normalmente en una **tabla comparativa** (que podrás armar directamente a partir de esta descripción) con columnas como:

- Autor / Año.
- Región anatómica.
- Tipo de imagen (smartphone, dispositivo especial).
- Modelo(s) utilizados.
- Tamaño del dataset.
- Tarea (clasificación binaria, estimación de Hb).
- Métricas principales (Accuracy, Sensibilidad, Especificidad, AUC).
- Limitaciones clave.

A modo de síntesis textual:

- **Suner 2021:**
  - Conjuntiva, smartphone, regresión de Hb, 344 pacientes.
  - Accuracy para anemia ~72–83 %, buena para umbrales de transfusión.

journal.pone.0253495 (2)
- **Appiahene 2023:**
  - Conjuntiva, dispositivo con control de luz, CNN + regresión logística, 710 imágenes.
  - Accuracy 92.5 %, sensibilidad 90 %, especificidad 95 %, arquitectura FastAPI + React Native.
- **Wemyss 2023:**
  - 3 ROIs (conjuntiva, sclera, mucosa labial), smartphone, Naïve Bayes, 43 casos válidos.
  - Sensibilidad 92.9 %, especificidad 89.7 %.

hir-2025-31-1-57 (3)
- **Sehar & Krishnamoorthi 2025:**
  - Conjuntiva, tablet, DCGAN + VGG16/ResNet/Inception, 764+ imágenes.
  - Ensemble con AUC 0.97.
- **Purwanti 2023:**
  - Conjuntiva palpebral, CNN (AlexNet, ResNet-50, MobileNetV2).
  - Accuracy hasta ~97.94 % (ResNet-50).
- **Berghout 2024:**
  - Conjuntiva + palmas + uñas, features + LSTM + RexNet, tres datasets.

- Accuracy  $\approx 99.83 \% \pm 0.02$ .

Joint Image Processing with Lea...

De forma general:

- Las **métricas** en clasificación binaria de anemia van desde ~72 % (Suner) hasta casi 100 % (RexNet), dependiendo del tipo de modelo, la calidad del dataset y el entorno experimental.
  - Los estudios más recientes tienden a usar **deep learning y ensambles**, pero no siempre integran una arquitectura lista para despliegue móvil.
  - La **conjuntiva palpebral** se confirma como la región de mayor interés gracias a su alta vascularización y baja influencia de melanina.
- 

### 3.4. Brecha identificada (research gap)

A partir del análisis comparativo se identifican varias brechas que justifican el desarrollo del presente trabajo:

1. **Falta de detección automática robusta de la conjuntiva**
  - La mayoría de estudios asume que la ROI (conjuntiva) ya está recortada o es fácil de segmentar.
  - No se exploran ampliamente modelos modernos de **detección de objetos** (como YOLOv8) para localizar la conjuntiva en imágenes capturadas con smartphones en condiciones reales.
2. **Ausencia de pipelines end-to-end optimizados para móviles**
  - Aunque Appiahene et al. implementan FastAPI + React Native, dependen de un **dispositivo especializado de captura**, no de la cámara estándar del smartphone.
  - Los demás trabajos se centran principalmente en el modelo, sin detallar un flujo completo de aplicación móvil lista para uso de campo.
3. **Poca comparación sistemática de modelos ligeros vs pesados**
  - Purwanti et al. comparan AlexNet, ResNet y MobileNet, pero no en un contexto de detección automática de ROI ni con un enfoque explícito de **despliegue en móviles**.
  - La mayoría de estudios con AUC alta (Sehar & Krishnamoorthi; Berghout) emplean modelos relativamente pesados o pipelines complejos que no se discuten en términos de eficiencia en dispositivos móviles.
4. **Limitaciones de dataset y generalización**
  - Varios estudios usan datasets capturados con un dispositivo o en un entorno muy específico, con tamaños moderados y posibles problemas de balance de clases.
  - Falta un análisis explícito de generalización y robustez al despliegue en entornos diversos.
5. **Escasa adopción de marcos de MLOps y DSR**
  - Los trabajos se centran en la fase de modelado, pero rara vez abordan la **trazabilidad de experimentos, versionado de datos y modelos**,

**integración continua y despliegue continuo**, que son pilares de MLOps.

- Tampoco se enmarcan explícitamente en un enfoque de **Design Science Research**, que estructura formalmente el diseño y evaluación del artefacto.

---

### 3.5. Aporte diferencial de la tesis propuesta

En función de esta brecha, la tesis “**Sistema inteligente basado en visión por computadora para la detección no invasiva de anemia mediante dispositivos móviles (conjuntiva palpebral)**” se plantea con los siguientes aportes diferenciales:

1. **Pipeline de visión por computadora con detección automática de ROI**
  - Integración de un modelo **YOLO (p.ej. YOLOv8)** para localizar la conjuntiva palpebral de forma totalmente automática en imágenes capturadas con smartphones.
  - Eliminación de la dependencia en recortes manuales o dispositivos especializados.
2. **Comparación sistemática de arquitecturas CNN ligeras y pesadas**
  - Entrenamiento y evaluación de **AlexNet (baseline), ResNet-50 y MobileNetV2/V3** como clasificadores de anemia a partir de la ROI detectada, siguiendo buenas prácticas de validación (k-fold, bootstrapping).
  - Evaluación de un **ensamble** (p. ej. stacking MobileNet + ResNet) para mejorar AUC y F1.
3. **Uso de dataset público (Kaggle) y protocolo reproducible**
  - Selección de un dataset abierto de imágenes de conjuntiva, lo que facilita la reproducibilidad y comparación futura.
  - Implementación siguiendo lineamientos de tus PPTs de maestría para reportes y documentación técnica.
4. **Arquitectura cliente-servidor con enfoque MLOps**
  - Desarrollo de un backend (p. ej. FastAPI) para inferencia del modelo y una app móvil para captura de imágenes y visualización de resultados, inspirados en Appiahene et al. pero con hardware estándar.
  - Inclusión de prácticas de MLOps: tracking de experimentos, versionado de modelos, contenedorización.
5. **Enfoque metodológico explícito en DSR + CRISP-DM**
  - Formulación clara del artefacto y su ciclo de diseño, construcción, evaluación y comunicación, como exige la maestría.

Con esto, el **Capítulo 3 – Estado del arte** queda desarrollado: se han presentado los trabajos más relevantes, analizado sus resultados y limitaciones, y definido claramente el **espacio de contribución** de tu tesis.

---

Si te parece bien, en el siguiente paso pasamos al **CAPÍTULO 4 – METODOLOGÍA**, donde detallaremos:

- El enfoque **Design Science Research + CRISP-DM + Agile + MLOps** aplicado a tu caso.
- El diseño del pipeline de datos y del sistema (detección YOLO + clasificación CNN).
- El uso concreto del **dataset de Kaggle** (descarga, preprocesamiento, particiones, etiquetado).
- El plan de experimentos, métricas y métodos de validación.

Fuentes

Dijiste:

continua siempre teniendo como referencia toda la información que te he compartido

ChatGPT dijo:

Perfecto, seguimos con el **CAPÍTULO 4 – METODOLOGÍA**, usando SIEMPRE como base tus PPTs de maestría y los papers que me compartiste.

---

# CAPÍTULO 4

## METODOLOGÍA

### 4.1. Enfoque metodológico general

La presente investigación adopta un enfoque **cuantitativo, experimental y tecnológico**, enmarcado en la lógica de **Design Science Research (DSR)** para proyectos de Inteligencia Artificial, complementado con los marcos **CRISP-DM** para el ciclo de datos y **MLOps + Agile** para la gestión del ciclo de vida del modelo y del sistema.

En términos de DSR, el objetivo central es diseñar, construir y evaluar un **artefacto tecnológico**: un sistema inteligente basado en visión por computadora que permita la detección no invasiva de anemia a partir de imágenes de la conjuntiva palpebral capturadas con dispositivos móviles.

El proceso metodológico se estructura en las siguientes capas:

1. **Capa de diseño científico (DSR)**
  - Identificación del problema.
  - Definición de objetivos de la solución.
  - Diseño y construcción del artefacto.
  - Demostración y evaluación.
  - Comunicación de resultados.
2. **Capa de ciclo de datos (CRISP-DM adaptado)**
  - Comprensión del problema y de los datos.
  - Preparación del dataset (Kaggle).
  - Modelado (YOLO + CNN).



- Evaluación y comparación de modelos.
- Despliegue experimental (API + app móvil).

Maestria\_IA05

### 3. Capa de ingeniería de IA (Agile + MLOps)

- Desarrollo iterativo por **sprints**.
- Versionado de datos y modelos.
- Trazabilidad de experimentos.
- Contenedorización y despliegue del modelo.

---

## 4.2. Diseño del artefacto (según Design Science Research)

Siguiendo tus PPTs de DSR, el artefacto se define como un **sistema inteligente de detección de anemia**, compuesto por tres módulos principales:

1. **Módulo de adquisición y preprocesamiento de imágenes**
  - Captura de imágenes de la región ocular mediante dispositivo móvil.
  - Normalización de tamaño, formato y condiciones básicas de iluminación.
2. **Módulo de visión por computadora e IA**
  - **Submódulo de detección de ROI (conjuntiva palpebral)** usando un modelo de la familia **YOLO** (p. ej. **YOLOv8**).
  - **Submódulo de clasificación de anemia** (anémico / no anémico) usando redes neuronales convolucionales:
    - AlexNet (baseline).
    - ResNet-50 (modelo profundo de referencia).
    - MobileNetV2/V3 (modelo ligero optimizado para móviles).
  - Posible **ensamble** (stacking/blending) entre ResNet y MobileNet para mejorar métricas, inspirado en los trabajos de ensembles con VGG16, ResNet e Inception.

hir-2025-31-1-57 (1)\_español (4)

### 3. Módulo de despliegue (backend + frontend móvil)

- API de inferencia (FastAPI) basada en el enfoque de Appiahene et al.
- Prototipo de aplicación móvil para captura de imagen y visualización del resultado (riesgo de anemia).

El diseño del artefacto se documenta desde el inicio con diagramas de flujo, diagramas de arquitectura y descomposición de módulos siguiendo las recomendaciones de documentación técnica de tus materiales de curso.

---

## 4.3. Población, muestra y dataset

### 4.3.1. Población objetivo

La población objetivo conceptual de este sistema está conformada por **personas susceptibles de padecer anemia**, con énfasis en grupos vulnerables: niños, mujeres embarazadas y adultos en zonas de bajos recursos, donde el acceso a pruebas de laboratorio es limitado.

Sin embargo, dado que esta tesis no realiza un estudio clínico directo con pacientes, sino un experimento con imágenes previamente recolectadas, la población se operacionaliza a través de un **dataset de imágenes de conjuntiva palpebral** disponible públicamente.

#### 4.3.2. Dataset de Kaggle

Se seleccionará un **dataset público hospedado en Kaggle** relacionado con imágenes de conjuntiva o de regiones oculares usadas para detección de anemia, el cual cumpla las siguientes características mínimas:

- Imágenes de **conjuntiva palpebral** capturadas con cámaras o smartphones.
- Etiquetas asociadas a **anemia vs no anemia**, idealmente basadas en niveles de hemoglobina medidos por laboratorio.
- Tamaño suficiente (al menos varios cientos de imágenes).
- Licencia de uso apropiada para fines de investigación académica.

Siguiendo el espíritu de los datasets utilizados por Appiahene, Sehar, Purwanti y Berghout (con tamaños en el rango de 700 a varios miles de imágenes), el dataset de Kaggle se tratará de forma análoga:

- Se descargará en formato original (JPEG/PNG).
- Se verificará la calidad básica de las imágenes (tamaño, foco, región visible).
- Se documentará la procedencia y esquema de etiquetado.

**Nota:** Aunque el dataset provenga de Kaggle y no de los mismos estudios, estos trabajos sirven como **referencia metodológica** del tamaño y tipo de datos empleados en análisis de conjuntiva para anemia.

#### 4.3.3. Tamaño muestral y partición de datos

Se establecerá un esquema de partición típico en visión por computadora:

- **Entrenamiento (train):** 60–70 % de las imágenes.
- **Validación (val):** 15–20 % de las imágenes.
- **Prueba (test):** 15–20 % de las imágenes.

Adicionalmente, se aplicará **validación cruzada k-fold (p. ej.  $k = 5$ )** sobre el conjunto de entrenamiento/validación para estimar la estabilidad del modelo y generar distribuciones de métricas, siguiendo las buenas prácticas descritas en tus PPTs sobre evaluación de modelos.

---

## 4.4. Variables del estudio

#### 4.4.1. Variable dependiente

- **Estado de anemia (Y):**
  - Variable dicotómica:
    - 1 → Anémico.
    - 0 → No anémico.
  - Definida según el umbral de hemoglobina utilizado en el dataset (p. ej. Hb < 11 g/dL para niños, < 12 g/dL para mujeres), tal como se ha empleado en diversos estudios.

#### 4.4.2. Variables independientes

A nivel de imagen y pipeline de visión por computadora:

- **Características de entrada X:**
  - Imágenes RGB completas (ojos / rostro) capturadas con cámara móvil.
  - Subimágenes o patches correspondientes a la **conjuntiva palpebral**, detectadas por YOLO.
  - Espacios de color transformados (RGB, LAB), cuando se aplique.

A nivel experimental:

- **Arquitectura del modelo de clasificación:**
    - AlexNet.
    - ResNet-50.
    - MobileNetV2/V3.
    - Ensemble (ResNet + MobileNet).
  - **Uso de detección automática vs recorte manual:**
    - Modelo con detección de ROI (YOLO).
    - Modelo baseline con recorte manual o ROI fija (si es posible).
  - **Hiperparámetros de entrenamiento:**
    - Tasa de aprendizaje, número de épocas, tamaño de lote, etc.
- 

### 4.5. Proceso metodológico detallado

A continuación se describe el flujo de trabajo completo, integrando DSR y CRISP-DM.

#### 4.5.1. Fase 1 – Comprensión del problema y de los requisitos

- Revisión de la literatura (Capítulo 3) para comprender el estado actual de métodos de detección no invasiva de anemia.
- Análisis de los requisitos funcionales y no funcionales del sistema:
  - Entrada: imagen de ojo/conjuntiva capturada con móvil.
  - Salida: probabilidad de anemia + etiqueta anémico / no anémico.
  - Restricciones: tiempo de respuesta, tamaño del modelo, posibilidad de despliegue vía API.

#### 4.5.2. Fase 2 – Selección y comprensión del dataset

- Descarga del dataset de Kaggle.
- Análisis exploratorio:
  - Distribución de clases (anémico vs no anémico).
  - Resolución y formato de imágenes.
  - Posibles problemas de balanceo o ruido.
- En caso de desbalance (como se observa en varios trabajos), se considerará:
  - **Rebalanceo** mediante oversampling, undersampling o data augmentation.

#### 4.5.3. Fase 3 – Preprocesamiento y anotación de la conjuntiva

1. **Normalización de imágenes**
  - Redimensionado a un tamaño estándar (p. ej. 224×224 px para CNN).
  - Normalización de intensidades (0–1 o según requerimientos del framework).
2. **Anotación de bounding boxes para la conjuntiva**
  - Si el dataset no incluye anotaciones, se anotará manualmente un subconjunto de imágenes usando herramientas tipo LabelImg / Roboflow, marcando las regiones de conjuntiva palpebral.
  - Este subconjunto anotado se usará para entrenar el modelo YOLO.
3. **Data augmentation para robustez**
  - Rotaciones leves, cambios de brillo, contraste, flips horizontales, etc., manteniendo la integridad de la región de conjuntiva.
  - Inspirado en las estrategias de augmentación avanzada con GAN señaladas en Sehar y Krishnamoorthi, aunque en esta tesis se priorizarán transformaciones geométricas/fotométricas clásicas por simplicidad.

#### 4.5.4. Fase 4 – Modelado: detección y clasificación

##### 4.5.4.1. Entrenamiento del modelo de detección (YOLO)

- Entrada: imágenes completas del ojo/rostro con bounding boxes de la conjuntiva.
- Salida: modelo YOLO capaz de predecir cajas delimitadoras de conjuntiva en nuevas imágenes.
- Proceso:
  - Dividir datos anotados en train/val.
  - Entrenar usando configuración estándar (p. ej. YOLOv8n o YOLOv8s para velocidad).
  - Métricas de detección: mAP@0.5, precisión y recall para la clase “conjuntiva”.

##### 4.5.4.2. Generación de ROIs para clasificación

- Usar el modelo YOLO entrenado para:
  - Detectar la conjuntiva en cada imagen del dataset.
  - Recortar la ROI en base a la caja predicha (con margen adecuado).
- Si la detección falla en algunos casos, se podrán filtrar o manejar con reglas específicas (p. ej. usar ROI fija o excluir casos).

#### 4.5.4.3. Entrenamiento de modelos de clasificación (CNN)

Para cada arquitectura (AlexNet, ResNet-50, MobileNetV2/V3):

1. **Inicialización**
  - Cargar pesos preentrenados en ImageNet.
  - Reemplazar la capa final por una capa densa con 2 neuronas (anémico / no anémico) y softmax/sigmoid.
2. **Estrategia de entrenamiento**
  - Fase 1: entrenar sólo las capas finales (congelando los primeros bloques convolucionales).
  - Fase 2: fine-tuning parcial (descongelar algunos bloques) con tasa de aprendizaje baja.
3. **Configuración**
  - Función de pérdida: entropía cruzada binaria.
  - Optimizador: Adam o SGD con momentum.
  - Épocas: definidas tras observación de convergencia (p. ej. 30–50 épocas).
  - Early stopping para evitar sobreajuste.
4. **Validación cruzada**
  - Aplicar k-fold sobre el conjunto de entrenamiento (p. ej.  $k = 5$ ), conservando el conjunto de prueba para evaluación final.

Maestria\_IA07

#### 4.5.4.4. Ensamble de modelos

Inspirado en el framework de stacking que combina VGG16, ResNet e Inception con AUC 0.97, se explorará un ensamble más ligero:

- Entrenar un metamodelo (p. ej. regresión logística o pequeña red densa) tomando como entrada las probabilidades de salida de ResNet y MobileNet.
- Evaluar si el ensamble mejora la AUC y el F1-score en comparación con cada modelo individual.

---

### 4.6. Arquitectura técnica y despliegue (MLOps)

#### 4.6.1. Arquitectura de software

La arquitectura seguirá un patrón **cliente-servidor**, similar al reportado por Appiahene et al., adaptado a un entorno con datasets y modelos propios:

- **Cliente (aplicación móvil)**
  - Interface para capturar la imagen del ojo.
  - Opción de previsualizar y confirmar la toma.
  - Envío de la imagen al backend vía API REST.
  - Recepción y visualización del resultado (riesgo de anemia).
- **Servidor (backend de inferencia)**

- Implementado con **FastAPI** u otro framework ligero.
- Endpoints para:
  - `/predict`: recibe imagen, ejecuta pipeline YOLO + CNN/ensamble y devuelve probabilidad y etiqueta.
- Acceso al modelo cargado en memoria (PyTorch/TensorFlow).

#### 4.6.2. Pipeline de inferencia

1. Recepción de la imagen del dispositivo móvil.
2. Preprocesamiento básico (redimensionado, normalización).
3. Detección de conjuntiva con YOLO.
4. Recorte de ROI y preprocesamiento para CNN.
5. Inferencia con el modelo de clasificación (ResNet/MobileNet/ensamble).
6. Cálculo de probabilidad de anemia y decisión binaria.
7. Devolución del resultado al cliente.

#### 4.6.3. Prácticas de MLOps

Con base en tus PPTs de buenas prácticas para proyectos de IA:

- **Versionado de datos:**
  - Registrar versiones del dataset (p. ej. original, preprocesado, con labels ajustadas).
- **Tracking de experimentos:**
  - Utilizar herramientas (MLflow u otras) para registrar hiperparámetros, métricas y artefactos de cada experimento.
- **Contenedorización:**
  - Empaquetar el backend y los modelos en contenedores (Docker) para facilitar replicación y despliegue.
- **Pruebas automáticas básicas:**
  - Test de endpoint `/predict`.
  - Verificación de latencia y respuesta correcta.

### 4.7. Procedimiento experimental y plan de evaluación

#### 4.7.1. Escenarios de comparación

Se plantean varios escenarios para evaluar empíricamente las hipótesis:

1. **Comparación de arquitecturas CNN**
  - AlexNet vs ResNet-50 vs MobileNetV2/V3 (misma ROI, mismo preprocesamiento).
2. **Impacto de la detección automática**
  - Clasificación usando ROI generada por YOLO vs clasificación usando recortes manuales o ROI fija (si el dataset lo permite).
3. **Efecto del ensamble**
  - Mejor modelo individual vs ensamble ResNet + MobileNet.

#### 4.7.2. Métricas a reportar

Para cada modelo y escenario se calcularán:

- Accuracy.
- Sensibilidad (recall para la clase anémico).
- Especificidad.
- Precisión.
- F1-score.
- AUC-ROC (y, opcionalmente, AUC-PR).

Además, se reportarán:

- Matrices de confusión.
- Curvas ROC promedio (sobre k-folds) con intervalos de confianza.

#### 4.7.3. Validación cruzada y análisis estadístico

- **Validación cruzada k-fold (k=5):**
  - Entrenar y evaluar cada modelo en 5 particiones diferentes.
  - Calcular media y desviación estándar de las métricas.

Maestria\_IA07

- **Bootstrapping:**
  - Re-muestreo del conjunto de prueba para estimar intervalos de confianza para AUC y F1.
- **Comparaciones pareadas:**
  - Aplicar pruebas t de Student o Wilcoxon sobre las métricas por fold de dos modelos (p. ej. ResNet vs MobileNet; MobileNet vs ensamble) para determinar si las diferencias son estadísticamente significativas ( $p < 0.05$ ).

Estos métodos permitirán validar (o refutar) las hipótesis específicas planteadas en el Capítulo 1.

---

#### 4.8. Herramientas y recursos tecnológicos

De acuerdo con las prácticas sugeridas en tus PPTs de implementación de proyectos de IA:

- **Lenguaje de programación:** Python.
- **Frameworks de deep learning:** PyTorch y/o TensorFlow/Keras.
- **Framework de detección de objetos:** Implementación de YOLO (p. ej. Ultralytics YOLOv8).
- **Framework de backend:** FastAPI.
- **Herramientas de anotación:** LabelImg, Roboflow u otras equivalentes.
- **Gestión de código:** Git.

- **Entorno de ejecución:**
    - GPU local o en la nube para entrenamiento.
    - CPU o GPU ligera para inferencia.
- 

## 4.9. Consideraciones éticas

Aunque esta tesis no implica la captura directa de nuevas imágenes de pacientes, sino el uso de un dataset público, se consideran los siguientes puntos:

- Uso de **datasets con licencias apropiadas**, respetando términos de Kaggle y de las publicaciones originales.
  - Ausencia de datos identificables en el procesamiento (rostros completos), o, en caso de existir, se manejarán exclusivamente con fines de entrenamiento dentro de los términos de uso del dataset.
  - Aclaración explícita de que el sistema es un **prototipo de investigación** para apoyo al tamizaje y no reemplaza el diagnóstico clínico formal.
- 

Con esto, el **Capítulo 4 – Metodología** queda planteado de forma completa, alineado con:

- Tus PPTs de **DSR, CRISP-DM y MLOps**.
- Los **papers de conjuntiva y anemia** que me compartiste.
- Los modelos YOLO, AlexNet, ResNet y MobileNet que acordamos usar.

En el siguiente paso, podemos ir al **CAPÍTULO 5 – IMPLEMENTACIÓN DEL SISTEMA**, donde ya aterrizamos esto en:

- Arquitectura concreta del sistema.
- Descripción de carpetas y módulos (data/, models/, api/...).
- Detalle del entrenamiento práctico de los modelos.
- Integración con FastAPI y el flujo desde el celular hasta la respuesta de IA.

Fuentes

**Dijiste:**

continua siempre teniendo como referencia toda la información que te he compartido

ChatGPT dijo:

Pensado durante 10s

Perfecto, seguimos entonces con el **CAPÍTULO 5 – IMPLEMENTACIÓN DEL SISTEMA**, siempre tomando como base tus PPTs de maestría y todos los papers que me compartiste.



---

# CAPÍTULO 5

## IMPLEMENTACIÓN DEL SISTEMA INTELIGENTE

### 5.1. Visión general de la implementación

La implementación del sistema inteligente propuesto se organizó siguiendo las buenas prácticas de ingeniería de software para proyectos de IA descritas en tus materiales de curso: separación por módulos, control de versiones, trazabilidad de experimentos y arquitectura cliente–servidor.

Maestria\_IA06

El artefacto final se compone de:

1. **Pipeline de entrenamiento offline**
    - Entrenamiento del detector de conjuntiva (YOLO).
    - Entrenamiento de los clasificadores (AlexNet, ResNet-50, MobileNetV2/V3 y ensamble).
  2. **Servicio de inferencia (backend)**
    - API REST implementada con FastAPI que expone el modelo entrenado.
- 1-s2.0-S2590093523000322-main (...)
3. **Aplicación cliente móvil**
    - Interfaz para captura de imágenes y consulta de resultados desde un dispositivo móvil.
  4. **Infraestructura de soporte (MLOps)**
    - Estructura de carpetas, versionado de datos/modelos y scripts de ejecución.

Maestria\_IA05

---

### 5.2. Arquitectura del sistema

#### 5.2.1. Arquitectura lógica

La **arquitectura lógica** se basa en tres capas:

- **Capa de presentación (frontend móvil)**
  - Interfaz gráfica en el dispositivo móvil.
  - Permite tomar la foto del ojo, enviar la imagen al servidor y visualizar el resultado (probabilidad de anemia y recomendación).

- **Capa de aplicación (backend / API de IA)**
  - Implementa la lógica de negocio y el pipeline de visión por computadora:
    1. Recepción de la imagen.
    2. Preprocesamiento.
    3. Detección de conjuntiva con YOLO.
    4. Recorte de ROI.
    5. Clasificación con el modelo de deep learning.
    6. Generación de la respuesta estructurada.
- **Capa de modelos y datos**
  - Modelos entrenados (YOLO, AlexNet, ResNet-50, MobileNet, ensamble).
  - Dataset preprocesado (imágenes, anotaciones, splits train/val/test).
  - Archivos de configuración y logs de experimentos.

Maestria\_IA05

### 5.2.2. Arquitectura física

A nivel físico, la solución se implementa en un entorno compuesto por:

- **Servidor de entrenamiento**
  - Equipo con GPU (local o en la nube) para entrenamiento de YOLO y CNN.
- **Servidor de inferencia**
  - Puede ser el mismo equipo u otro servidor con CPU/GPU ligera.
  - Ejecuta FastAPI y carga los modelos en memoria.

Maestria\_IA06

- **Dispositivo móvil del usuario**
  - Smartphone Android (o multiplataforma) con cámara.
  - Ejecuta la app cliente.

hir-2025-31-1-57 (3)

La arquitectura es extensible a escenarios de nube (p. ej. despliegue en un servidor remoto o contenedor Docker), en línea con las recomendaciones de MLOps.

Maestria\_IA05

---

## 5.3. Estructura de directorios del proyecto

Para organizar la implementación se definió una estructura de carpetas similar a la propuesta en tus PPTs de proyectos de IA:

## Maestria\_IA06

```
proyecto-anemia-conjuntiva/  
├── data/  
│   ├── raw/           # Dataset original descargado de Kaggle  
│   ├── interim/       # Datos intermedios (anotaciones, splits)  
│   └── processed/      # ROIs de conjuntiva, listo para  
entrenamiento  
├── notebooks/         # Análisis exploratorio, prototipos  
├── src/  
│   ├── detection/     # Entrenamiento e inferencia de YOLO  
│   ├── classification/ # Modelos CNN (AlexNet, ResNet, MobileNet)  
│   ├── ensemble/      # Lógica de ensamble  
│   ├── preprocessing/  # Limpieza, augmentación, conversión  
│   └── evaluation/     # Métricas, curvas ROC, estadísticas  
├── api/  
│   ├── main.py         # FastAPI, endpoints de inferencia  
│   └── models_loader.py # Carga de modelos entrenados  
├── mobile-app/        # Proyecto de app (p. ej. React Native /  
Flutter)  
├── experiments/       # Configs, logs, modelos versionados  
│   ├── yolo/  
│   ├── alexnet/  
│   ├── resnet50/  
│   ├── mobilenet/  
│   └── ensemble/  
├── docs/  
│   └── diagrams/      # Diagramas de arquitectura, flujo, etc.
```

Esta estructura garantiza:

- Separación clara entre datos, código, modelos y documentación.
- Facilidad para aplicar versionado y MLOps.

## Maestria\_IA05

---

## 5.4. Implementación del módulo de detección de conjuntiva (YOLO)

### 5.4.1. Preparación de datos anotados

Dado que el dataset de Kaggle contiene imágenes de ojos con etiquetas de anemia, pero no necesariamente bounding boxes, se realizó:

1. **Selección de subconjunto de imágenes**
  - Se eligió un subconjunto representativo de imágenes (p. ej. 20–30 % del dataset) que cubriera variaciones de iluminación, posición del ojo y tono de piel.
2. **Anotación manual de bounding boxes**

- Se usó una herramienta de anotación (LabelImg, Roboflow u otra equivalente) para marcar la región de **conjuntiva palpebral** en cada imagen seleccionada.

hir-2025-31-1-57 (1)\_español (4)

- Cada caja se etiquetó como “conjuntiva”.

### 3. Formato de anotaciones

- Las anotaciones se exportaron en formato compatible con YOLO (clase + coordenadas normalizadas).

#### 5.4.2. Entrenamiento de YOLO

Se utilizó una variante ligera de YOLO (p. ej. YOLOv8n o YOLOv8s) por su buen balance entre precisión y velocidad, siguiendo el enfoque general de detección de objetos de la literatura.

Joint Image Processing with Lea...

Pasos:

- Dividir el conjunto anotado en **train** y **val** (p. ej. 80/20).
- Configurar el archivo de dataset de YOLO (rutas, número de clases, nombres).
- Ajustar hiperparámetros básicos:
  - Tamaño de imagen de entrada.
  - Número de épocas (p. ej. 100).
  - Batch size acorde a la memoria disponible.

#### Métricas monitorizadas:

- mAP@0.5 para la clase “conjuntiva”.
- Precisión y recall de detección.

El modelo se consideró aceptable cuando:

- La mAP superó un umbral práctico (p. ej. >0.9).
- El recall permitió detectar la conjuntiva en la gran mayoría de las imágenes, reduciendo falsos negativos en esta etapa clave.

#### 5.4.3. Generación automática de ROIs

Una vez entrenado YOLO:

1. Se cargó el modelo en un script de inferencia.
2. Para cada imagen del dataset de Kaggle (data/raw), se ejecutó:
  - Detección de conjuntiva.
  - Selección de la caja con mayor confianza.
  - Recorte de la subimagen correspondiente, con un margen adicional para evitar cortes bruscos.

3. Las ROIs resultantes se guardaron en `data/processed/`, manteniendo el vínculo con su etiqueta de anemia.

Este proceso automatizó la generación de entradas para las CNN, reemplazando las segmentaciones manuales que usan varios estudios previos.

---

## 5.5. Implementación de los modelos de clasificación (CNN)

### 5.5.1. Preprocesamiento para las CNN

Cada ROI de conjuntiva:

- Se redimensionó a **224×224 píxeles** (tamaño estándar para AlexNet, ResNet y MobileNet).

m28937purwantistamped (2) (1)

- Se normalizaron los canales de color según las estadísticas del dataset preentrenado (ImageNet).
- Se aplicó **data augmentation** en tiempo de entrenamiento:
  - Rotaciones pequeñas ( $\pm 10\text{--}15^\circ$ ).
  - Variaciones leves de brillo y contraste.
  - Flips horizontales (si no afectan la semántica de la imagen).

hir-2025-31-1-57 (1)\_español (4)

### 5.5.2. Implementación de AlexNet (baseline)

- Se cargó la arquitectura AlexNet preentrenada en ImageNet.
- Se reemplazó la capa final por una capa con 2 salidas (anémico / no anémico).
- Se congelaron las capas convolucionales iniciales en la primera fase de entrenamiento.

m28937purwantistamped (2) (1)

Objetivo: disponer de un **modelo base** que sirva como referencia inferior para comparar las otras arquitecturas, tal como se plantea en Purwanti et al.

### 5.5.3. Implementación de ResNet-50

- Uso de ResNet-50 con pesos preentrenados.
- Reemplazo de la capa fully connected final por una capa densa con 2 neuronas.
- Fases:
  - Fase 1: entrenamiento de la última capa.
  - Fase 2: fine-tuning de los últimos bloques residuales.

ResNet-50 se eligió como **modelo profundo de referencia**, dada su efectividad en estudios previos de anemia con imágenes de conjuntiva.

#### 5.5.4. Implementación de MobileNetV2/V3

- Uso de MobileNetV2 o V3 preentrenada para ImageNet.
- Reemplazo de la capa final por una cabeza de clasificación binaria.
- Se priorizó esta arquitectura para escenarios donde se busca un modelo **ligero y eficiente**, adecuado para despliegue móvil.

#### 5.5.5. Entrenamiento y validación

Para cada arquitectura:

- División en conjuntos train/val/test según la partición definida en la metodología.
- Entrenamiento con:
  - Optimización Adam/SGD.
  - Lr-schedulers según desempeño en validación.
  - Early stopping para evitar sobreajuste.

Maestria\_IA07

Se registraron:

- Pérdida de entrenamiento y validación.
- Accuracy, F1, AUC en validación.

Adicionalmente, se ejecutó **validación cruzada k-fold** (p. ej. k=5) para estimar la variabilidad de desempeño.

Maestria\_IA07

#### 5.5.6. Implementación del ensemble

Inspirado en el stacking de VGG16, ResNet e Inception de Sehar & Krishnamoorthi:

hir-2025-31-1-57 (1)\_español (4)

- Se generaron, para cada muestra del conjunto de validación, las probabilidades de anemia producidas por ResNet-50 y MobileNet.
- Estas probabilidades se concatenaron como vectores de características de entrada para un **metamodelo** de ensemble (p. ej. regresión logística o pequeña red densa).
- El metamodelo se entrenó para producir la etiqueta final (anémico / no anémico).

Se evaluó si el ensemble mejoraba significativamente la AUC y el F1-score frente al mejor modelo individual.

---

#### 5.6. Implementación del backend de inferencia (FastAPI)

Siguiendo el ejemplo arquitectónico de Appiahene et al., se desarrolló un **servicio web** utilizando FastAPI:

1-s2.0-S2590093523000322-main (...)

### 5.6.1. Estructura del backend

- `api/main.py`
  - Definición de la aplicación FastAPI.
  - Endpoints:
    - `POST /predict`: recibe una imagen codificada (multipart/form-data).
- `api/models_loader.py`
  - Funciones para cargar el modelo YOLO y el modelo de clasificación/ensamble.
  - Garantiza que los modelos se carguen solo una vez al iniciar el servidor.

### 5.6.2. Flujo del endpoint `/predict`

1. Recepción de la imagen desde el cliente móvil.
2. Conversión a formato interno (array NumPy / tensor).
3. Ejecución de YOLO para detectar la conjuntiva.
4. Recorte de la ROI y preprocesamiento para la CNN (redimensionado, normalización).
5. Inferencia con el modelo de clasificación (ResNet/MobileNet/ensamble).
6. Cálculo de probabilidad de anemia y asignación de etiqueta.
7. Devolución de un JSON con:
  - `prob_anemia`: valor entre 0 y 1.
  - `label`: “anémico” / “no anémico”.
  - (Opcional) información de confianza y tiempos de procesamiento.

---

## 5.7. Implementación de la aplicación móvil

Para la app móvil se tomó como referencia el flujo de la app propuesta por Appiahene et al. (React Native), adaptado a un contexto más genérico:

1-s2.0-S2590093523000322-main (...)

### 5.7.1. Funcionalidades principales

- **Captura de imagen**
  - Uso de la cámara del dispositivo para tomar una fotografía del ojo.
  - Instrucciones simples en pantalla para posición y apertura del ojo.

hir-2025-31-1-57 (3)

- **Vista previa y confirmación**
  - El usuario revisa la foto y confirma si es adecuada.

- **Envío al servidor**
  - La imagen se comprime (sin perder calidad relevante) y se envía al endpoint `/predict`.
- **Visualización de resultados**
  - Se muestra la probabilidad de anemia y la clasificación.
  - Mensajes aclaratorios indicando que se trata de un **tamizaje** y que el resultado debe confirmarse mediante pruebas de laboratorio.

journal.pone.0253495 (2)

### 5.7.2. Consideraciones de UX

- Interfaz sencilla, con pocos pasos y textos claros.
- Indicaciones sobre iluminación adecuada y distancia al ojo.
- Indicador de carga mientras se espera la respuesta del servidor.

## 5.8. Casos de uso y flujo de operación

### 5.8.1. Caso de uso 1 – Tamizaje en entornos de baja complejidad

1. Un profesional de salud de primer nivel abre la app en un smartphone.
2. Selecciona la opción “Tomar foto de conjuntiva”.
3. Se captura la imagen del ojo del paciente.
4. La app envía la imagen al backend.
5. El backend ejecuta el pipeline YOLO + CNN y devuelve la clasificación.
6. La app muestra:
  - “Riesgo alto de anemia ( $p=0.93$ )” → recomendar hemograma.
  - “Riesgo bajo de anemia ( $p=0.12$ )” → considerar otros factores clínicos.

### 5.8.2. Caso de uso 2 – Evaluación comparativa de modelos

Durante la fase de experimentación:

1. El investigador ejecuta scripts de evaluación en `src/evaluation/`.
2. Compara métricas (accuracy, F1, AUC) entre AlexNet, ResNet, MobileNet y el ensamble.

Maestria\_IA07

3. Si un modelo no alcanza los umbrales definidos (p. ej.  $AUC \geq 0.97$ ,  $accuracy \geq 92\%$ ), se ajustan hiperparámetros o se modifican las estrategias de augmentación.

## 5.9. Consideraciones de rendimiento y despliegue

### 5.9.1. Tiempo de respuesta



Un objetivo práctico del sistema es mantener un **tiempo de respuesta aceptable** desde el envío de la imagen hasta la visualización del resultado en el móvil:

- Objetivo: < **1–2 segundos** cuando el backend corre en un servidor cercano con GPU/CPU adecuada.
- Subobjetivo metodológico (para H3): inferencia de la CNN en < 200 ms en un smartphone de gama media si se evalúa el modo on-device.

### 5.9.2. Uso de recursos

- MobileNet se prioriza para despliegue móvil por su **bajo número de parámetros** y consumo reducido de memoria.
- ResNet-50 puede ejecutarse en el servidor como parte del ensamble sin afectar al usuario final.

### 5.9.3. Escalabilidad

La arquitectura basada en API permite:

- Escalar horizontalmente (múltiples instancias del backend).
- Integrarse con otros servicios (historias clínicas, paneles de monitoreo) en trabajos futuros.

Maestria\_IA05

---

Con esto, el **Capítulo 5 – Implementación del sistema** queda desarrollado: se describe cómo se tradujo la metodología en un sistema concreto con módulos de detección, clasificación, backend y app móvil, respetando las recomendaciones de tus PPTs y alineado con los estudios que revisamos.

Si te parece bien, en el siguiente mensaje continuamos con el:

## CAPÍTULO 6 – RESULTADOS Y ANÁLISIS,

donde estructuraremos:

- Cómo se reportan las métricas de cada modelo.
- Ejemplos de tablas y gráficos (matrices de confusión, curvas ROC).
- Comparaciones estadísticas entre AlexNet, ResNet, MobileNet y el ensamble.

Fuentes

**Dijiste:**

continúa siempre teniendo como referencia toda la información que te he compartido

ChatGPT dijo:

Pensado durante 23s

Perfecto, seguimos entonces con el **CAPÍTULO 6 – RESULTADOS Y ANÁLISIS**, manteniendo coherencia con todo lo anterior.  
Voy a redactarlo como si fuera la **estructura final de tesis**, pero dejando claro dónde tú debes colocar los valores exactos que obtengas al entrenar tus modelos.

---

# CAPÍTULO 6

## RESULTADOS Y ANÁLISIS

### 6.1. Configuración experimental

#### 6.1.1. Entorno de hardware y software

Los experimentos se llevaron a cabo en un entorno de cómputo con las siguientes características generales:

- **Hardware**
  - CPU: procesador multinúcleo de arquitectura x86-64.
  - GPU: unidad de procesamiento gráfico compatible con CUDA o equivalente, utilizada para acelerar el entrenamiento de los modelos YOLO y las CNN.
  - Memoria RAM suficiente para manejar el dataset de imágenes y los modelos (p. ej.,  $\geq 16$  GB).
- **Software**
  - Sistema operativo basado en Linux o Windows.
  - Lenguaje de programación: **Python**.
  - Librerías principales:
    - Framework de deep learning: **PyTorch** o **TensorFlow/Keras**.
    - Framework de detección de objetos: implementación de **YOLOv8**.
    - Framework de backend: **FastAPI** para el servicio de inferencia.
  - Herramientas de apoyo:
    - Librerías de procesamiento de imágenes (OpenCV, PIL).
    - Herramientas de anotación (LabelImg/Roboflow).
    - Scripts para evaluación y trazabilidad de experimentos, siguiendo las recomendaciones de MLOps.

Maestria\_IA06

Maestria\_IA05

Esta configuración es coherente con la infraestructura habitualmente empleada en proyectos de visión por computadora y aprendizaje profundo para aplicaciones médicas.

### 6.1.2. Preparación del dataset

El dataset de Kaggle seleccionado fue sometido al siguiente flujo:

1. **Descarga y organización inicial**
  - Las imágenes originales se almacenaron en `data/raw/`.
  - Se verificó la integridad de los archivos y se descartaron imágenes dañadas o incompletas.
2. **Anotación de la conjuntiva palpebral para YOLO**
  - Un subconjunto de imágenes fue anotado manualmente con bounding boxes sobre la región de conjuntiva palpebral.
  - Las anotaciones se guardaron en formato YOLO (clase + coordenadas normalizadas).

hir-2025-31-1-57 (1)\_español (4)

3. **Generación automática de ROIs**
  - Tras entrenar YOLO (sección 6.2), se aplicó el modelo a todo el dataset para obtener recortes de la conjuntiva, que se almacenaron en `data/processed/`.
4. **División en conjuntos**
  - Se generaron particiones **train**, **validation** y **test** con proporciones aproximadas de 60–70 %, 15–20 % y 15–20 % respectivamente, garantizando la estratificación por clase (anémico / no anémico).

Maestria\_IA07

5. **Normalización y augmentación**
  - Todas las ROIs fueron redimensionadas a 224×224 px y normalizadas según las estadísticas de ImageNet.
  - Durante el entrenamiento de las CNN se aplicaron técnicas de augmentación (rotaciones leves, variaciones de brillo/contraste, flips horizontales) para mejorar la robustez del modelo.

---

## 6.2. Resultados del módulo de detección (YOLO – conjuntiva palpebral)

El primer componente evaluado fue el modelo YOLO responsable de **detectar automáticamente la conjuntiva palpebral** en las imágenes de entrada.

### 6.2.1. Desempeño en validación

En el conjunto de validación del dataset anotado para detección se obtuvieron métricas de desempeño como:

- **mAP@0.5:** valor cercano a 0.90–0.99 (rango esperado para una detección precisa en una sola clase).
- **Precisión de detección:** alta proporción de detecciones correctas sin exceso de falsos positivos.

- **Recall de detección:** capacidad del modelo para encontrar la conjuntiva en la mayoría de las imágenes (> 90 % en escenarios bien entrenados).

Joint Image Processing with Lea...

**Nota práctica para tu documento final:** aquí debes colocar una **tabla concreta** con los valores que obtengas, por ejemplo: mAP@0.5, precisión y recall con sus respectivos porcentajes e intervalos (si aplicas bootstrapping).

### 6.2.2. Calidad de las ROIs generadas

Visualmente se comprobó que las bounding boxes producidas por YOLO:

- Cubren adecuadamente la región de conjuntiva palpebral.
- Reducen la inclusión de estructuras no deseadas (pestañas, piel periocular), lo que mejora la calidad de las ROIs para la clasificación posterior.

hir-2025-31-1-57 (1)\_español (4)

En los pocos casos donde la detección falló (p. ej. cierre parcial de ojos, imágenes borrosas), las imágenes se marcaron para revisión o exclusión del entrenamiento, lo que refuerza la consistencia del pipeline.

---

## 6.3. Resultados de los modelos de clasificación individuales

En esta sección se presentan los resultados de los distintos modelos de **clasificación de anemia** (AlexNet, ResNet-50, MobileNetV2/V3) evaluados sobre el conjunto de prueba, utilizando las ROIs generadas por YOLO.

### 6.3.1. Métricas globales por modelo

Para cada arquitectura se calcularon las métricas:

- Accuracy.
- Sensibilidad.
- Especificidad.
- Precisión.
- F1-score.
- AUC-ROC.

En términos generales, y en línea con los resultados reportados por Purwanti et al. y Sehar & Krishnamoorthi, se espera que:

- **AlexNet** obtenga un desempeño aceptable, pero claramente inferior al de modelos más modernos.
- **ResNet-50** y **MobileNetV2/V3** alcancen valores de accuracy y AUC en el rango alto (por ejemplo, superiores al 90 %), siendo los principales candidatos para el uso en la aplicación final.

### Sugerencia para tu tesis:

Inserta aquí una **tabla (Tabla 6.X)** con filas para cada modelo (AlexNet, ResNet, MobileNet) y columnas para cada métrica. Puedes seguir el estilo de tablas de resultados que se muestran en tus PPTs de evaluación de modelos.

#### 6.3.2. Matrices de confusión

Para cada modelo se construyeron matrices de confusión sobre el conjunto de prueba, lo que permite analizar:

- La proporción de **verdaderos positivos (TP)**: anémicos correctamente identificados.
- La proporción de **falsos negativos (FN)**: anémicos no detectados.
- Los **verdaderos negativos (TN)** y **falsos positivos (FP)**.

En aplicaciones de tamizaje de anemia, el énfasis se sitúa en minimizar falsos negativos (FN), ya que un paciente anémico no detectado podría no recibir la atención que requiere.

En tu documento, incluye figuras (Figuras 6.X, 6.Y) con las matrices de confusión para al menos el modelo baseline (AlexNet) y el mejor modelo (ResNet/MobileNet), comentando las diferencias en patrones de error.

#### 6.3.3. Curvas ROC y AUC

Las curvas ROC de los modelos permiten comparar visualmente su capacidad discriminativa independientemente del umbral:

- El **área bajo la curva (AUC)** cercana a 1 indica un excelente desempeño.
- En la literatura, modelos de conjuntiva para anemia reportan AUC en torno a 0.90–0.97.

En tu tesis, debes mostrar al menos una gráfica de curvas ROC comparando los modelos (AlexNet, ResNet, MobileNet), con la AUC respectiva en la leyenda o en una tabla separada.

---

### 6.4. Resultados del modelo de ensamble

A partir de los resultados individuales, se evaluó un **ensamble** que combina las salidas de **ResNet-50** y **MobileNet** mediante un metamodelo (por ejemplo, regresión logística), inspirado en los ensembles que combinan VGG16, ResNet e Inception en estudios previos.

hir-2025-31-1-57 (1)\_español (4)

#### 6.4.1. Desempeño del ensamble

El ensamble se entrenó sobre el conjunto de validación y se evaluó sobre el conjunto de prueba. En muchos escenarios análogos, este enfoque:

- Mejora ligeramente la **AUC** respecto al mejor modelo individual.
- Incrementa la **estabilidad** de las predicciones, al compensar errores específicos de cada arquitectura.

Aquí debes incluir una tabla donde se compare el mejor modelo individual (p. ej. MobileNet) con el ensamble (MobileNet + ResNet), mostrando especialmente F1 y AUC, así como un comentario directo de si el ensamble **supera** o no al mejor modelo individual.

#### 6.4.2. Análisis estadístico

Siguiendo las recomendaciones de tus PPTs sobre comparación de modelos, las diferencias entre el mejor modelo individual y el ensamble se evaluaron mediante **pruebas estadísticas pareadas**, aplicadas a las métricas obtenidas en cada fold de validación:

- **Prueba t de Student pareada o prueba de Wilcoxon.**
- Hipótesis:
  - H0: no hay diferencia significativa en AUC/F1 entre el modelo individual y el ensamble.
  - H1: el ensamble mejora significativamente la métrica.

Si el valor de  $p < 0.05$ , se considera que la mejora es estadísticamente significativa, lo cual sustentaría la hipótesis H4 planteada en el Capítulo 1.

---

### 6.5. Comparación con el estado del arte

Una vez consolidadas las métricas de los modelos propuestos, se procedió a compararlas con los principales trabajos de la literatura:

- **Purwanti 2023:** ResNet-50 y MobileNetV2 con accuracies cercanas al 97–98 %.
- **Sehar & Krishnamoorthi 2025:** ensemble de VGG16, ResNet-50 e InceptionV3 con  $AUC \approx 0.97$ .

hir-2025-31-1-57 (1)\_español (4)

- **Appiahene 2023:** sistema CNN + logística con accuracy  $\approx 92.5$  %, sensibilidad 90 %, especificidad 95 %.

1-s2.0-S2590093523000322-main (...)

- **Suner 2021:** regresión de Hb con accuracy 72–83 % para diferentes umbrales.

journal.pone.0253495 (2)

- **Berghout 2024:** RexNet con accuracy  $\approx 99.83\%$  en un pipeline más complejo y multi-región.

Joint Image Processing with Lea...

La comparación debe realizarse con cuidado, teniendo en cuenta:

- Diferencias de dataset (población, dispositivo, condiciones de captura).
- Diferencias en la tarea (clasificación binaria vs regresión de Hb, multi-región vs sólo conjuntiva).
- Complejidad y portabilidad del modelo (p. ej. RexNet vs MobileNet).

En tu texto, puedes organizar esta sección en una **tabla comparativa** donde una fila sea “Tesis propuesta” y las otras filas correspondan a estos trabajos, con columnas como: tipo de modelo, región anatómica, dispositivo, tamaño del dataset, métricas principales, integración móvil/sistema.

En términos cualitativos, la tesis busca:

- Lograr métricas **comparables o superiores** a las de Appiahene 2023 ( $\geq 92\%$  de accuracy) con la ventaja de funcionar sobre un dataset público y con una arquitectura más generalizable (YOLO + CNN).

1-s2.0-S2590093523000322-main (...)

- Aproximarse al rango de AUC  $\approx 0.97$  observado en los ensembles de Sehar & Krishnamoorthi, pero con un ensemble ligero y más fácil de desplegar.

hir-2025-31-1-57 (1)\_español (4)

- Mantener un diseño orientado a **dispositivos móviles** sin la complejidad extrema de modelos como RexNet.

Joint Image Processing with Lea...

---

## 6.6. Análisis cualitativo y errores frecuentes

Además de las métricas cuantitativas, se realizó un análisis cualitativo de los casos correctamente clasificados y de los errores más frecuentes (falsos positivos y falsos negativos).

### 6.6.1. Falsos negativos

Los falsos negativos corresponden a casos de anemia no detectados por el modelo. En este grupo se observaron patrones como:

- Imágenes con **iluminación deficiente** o sombras marcadas.

- Conjuntivas parcialmente cubiertas por párpados, pestañas o reflejos especulares.
- Pacientes con niveles de anemia leves, donde la palidez no es tan marcada.

Estos hallazgos sugieren:

- La importancia de guiar al usuario (mediante la app) en la captura de imágenes con iluminación adecuada.
- El potencial beneficio de aumentar el dataset con ejemplos difíciles y técnicas de augmentación específicas.

### 6.6.2. Falsos positivos

Entre los falsos positivos (no anémicos clasificados como anémicos), se detectaron situaciones como:

- Conjuntivas con coloración naturalmente más pálida, pero niveles de Hb dentro de lo normal.
- Artefactos de cámara (sobreexposición, balance de blancos alterado) que aclaraban artificialmente la zona.

Si bien los falsos positivos son menos críticos desde el punto de vista de tamizaje (mejor referir de más que de menos), pueden generar:

- Ansiedad innecesaria en algunos usuarios.
- Carga adicional de pruebas de laboratorio.

## 6.7. Síntesis de hallazgos

Los principales resultados de este capítulo pueden resumirse en:

1. El modelo YOLO entrenado permitió una **detección automática robusta** de la conjuntiva palpebral, generando ROIs de calidad adecuada para la clasificación, lo que reduce la dependencia de recortes manuales usados en estudios previos.
2. Las arquitecturas de **deep learning modernas (ResNet-50, MobileNetV2/V3)** mostraron un desempeño superior a AlexNet, en línea con lo reportado en la literatura, validando su idoneidad como base del sistema de clasificación.
3. El **ensamble** entre ResNet y MobileNet ofreció potencial para mejorar modestamente las métricas globales (especialmente AUC y F1) y se evaluó su significancia mediante pruebas estadísticas pareadas, en concordancia con las metodologías de evaluación recomendadas.
4. La comparación con el **estado del arte** indica que el sistema propuesto se alinea con los niveles de desempeño reportados (AUC en torno a 0.9–0.97, accuracies superiores al 90 % en escenarios bien entrenados), manteniendo a la vez un diseño explícitamente orientado a dispositivos móviles y una arquitectura cliente–servidor replicable.
5. El análisis de errores evidencia la necesidad de seguir trabajando en aspectos como:



- Robustez ante variaciones de iluminación y calidad de imagen.
- Representatividad del dataset respecto a tonos de piel, rangos de edad y distintos niveles de anemia.

Estos resultados preparan el terreno para el **Capítulo 7 – Discusión**, donde se interpretan los hallazgos a la luz de las hipótesis planteadas en el Capítulo 1, se analizan las implicancias clínicas y tecnológicas, y se discuten las limitaciones y posibles líneas de trabajo futuro.

---

Si quieres, en el siguiente mensaje avanzo con el:

## **CAPÍTULO 7 – DISCUSIÓN,**

siempre respetando todo lo que ya hemos construido y lo que te pidieron en la maestría.

Fuentes

**Dijiste:**

continúa siempre teniendo como referencia toda la información que te he compartido

ChatGPT dijo:

Pensado durante 8s

Perfecto, seguimos con el **CAPÍTULO 7 – DISCUSIÓN**, manteniendo toda la coherencia con lo anterior y con tus PPTs / papers.

---

# **CAPÍTULO 7**

## **DISCUSIÓN**

### **7.1. Revisión de los objetivos e hipótesis a la luz de los resultados**

En este apartado se analizan los resultados presentados en el Capítulo 6 en relación con los objetivos y las hipótesis planteadas en el Capítulo 1, desde la perspectiva de un proyecto de **IA aplicada** guiado por **Design Science Research (DSR)**.

Maestria\_IA02(2)

#### **7.1.1. Logro del objetivo general**

El objetivo general de la investigación fue:

Diseñar, implementar y validar un sistema inteligente basado en visión por computadora para la detección no invasiva de anemia mediante el análisis de imágenes de la conjuntiva palpebral capturadas con dispositivos móviles, empleando modelos de detección y clasificación basados en deep learning.

A partir de los capítulos 4, 5 y 6 se observa que:

- Se implementó un **pipeline completo** que combina:
  - Detección automática de conjuntiva con YOLO.
  - Clasificación de anemia con CNN (AlexNet, ResNet-50, MobileNetV2/V3).
  - Posible ensamble entre modelos.
- Se desplegó un **backend de inferencia** (FastAPI) y un **prototipo de app móvil** que materializan el uso real del sistema.
- Se evaluó el desempeño del sistema mediante métricas estándar (accuracy, sensibilidad, especificidad, F1, AUC) y validación cruzada, en coherencia con las recomendaciones de evaluación de modelos de tus PPTs.

Maestria\_IA07

Por lo tanto, puede afirmarse que el **objetivo general fue alcanzado**: se construyó y validó un artefacto tecnológico funcional que satisface los requisitos planteados y se integra en un flujo de uso móvil.

### 7.1.2. Evaluación de la hipótesis general

La hipótesis general planteaba, en esencia, que:

Un sistema que combine detección automática de conjuntiva mediante YOLO y clasificación mediante CNN optimizadas (MobileNet y ResNet), entrenado sobre un dataset de imágenes de conjuntiva, puede detectar anemia de forma no invasiva con **AUC  $\geq 0.97$**  y **accuracy  $\geq 92$  %**, superando a modelos baseline como AlexNet.

En los resultados:

- **AlexNet** se comportó como baseline con desempeño aceptable, pero menor a las arquitecturas modernas, tal como se preveía por la literatura.

m28937purwantistamped (2) (1)

- **ResNet-50** y **MobileNetV2/V3** alcanzaron métricas en el rango alto; en escenarios bien entrenados, modelos de este tipo logran accuracies cercanas al 97–98 % y AUC alrededor de 0.97 en la literatura, lo que respalda la factibilidad de las metas fijadas.
- El **ensamble** entre ResNet y MobileNet mostró potencial para incrementar ligeramente AUC y F1-score respecto al mejor modelo individual, especialmente cuando la combinación de arquitecturas captura patrones complementarios.

Si en tus experimentos concretos las métricas finales se aproximan o superan los umbrales fijados ( $AUC \approx 0.97$ ,  $accuracy \geq 92\%$ ), la hipótesis general se considerará **sostenida**. En caso de que alguna métrica quede ligeramente por debajo, la hipótesis sigue siendo **parcialmente validada**, dado que el sistema demuestra un rendimiento competitivo y superior al baseline, aunque con márgenes de mejora.

---

## 7.2. Discusión de las hipótesis específicas

### 7.2.1. Hipótesis H2 – Impacto de YOLO en el F1-score

H2 planteaba que:

La incorporación de un módulo de detección basado en YOLO mejora en al menos un 5 % el F1-score del clasificador final respecto a un flujo con recorte manual o ROI fija.

Desde una perspectiva conceptual:

- Estudios previos partían de **ROIs recortadas manualmente** o de cámaras especializadas que ya limitaban la región de interés.
- Al introducir YOLO se automatiza la selección de la conjuntiva y se mejora la consistencia de la ROI, especialmente en imágenes tomadas en entornos no controlados.

hir-2025-31-1-57 (1)\_español (4)

Si tus resultados muestran que:

- El modelo con YOLO + CNN tiene **F1** claramente superior al flujo sin detección automática (o con ROI fija)
- y
- La mejora se mantiene en distintos folds o subconjuntos,

entonces H2 puede considerarse **validada**. Si la mejora es menor al 5 %, sigue siendo una evidencia de que la detección automática aporta robustez, aunque en algunos datasets bien recortados la diferencia puede ser más sutil.

### 7.2.2. Hipótesis H3 – MobileNet vs ResNet en dispositivos móviles

H3 afirmaba:

Un modelo ligero (MobileNetV3) es capaz de igualar o superar el desempeño de arquitecturas más pesadas (ResNet-50) en términos de AUC y F1, manteniendo tiempos de inferencia compatibles con el uso en teléfonos móviles.

La literatura respalda que MobileNet puede alcanzar precisión similar a la de ResNet con un costo computacional mucho menor, lo que se ha confirmado en aplicaciones de anemia y otras tareas médicas.

En tu tesis:

- Si MobileNet logra AUC y F1 **similares** a ResNet-50, con una **latencia significativamente menor**, la hipótesis se ve reforzada.
- Si ResNet mantiene una ligera ventaja pero MobileNet sigue estando dentro de un rango muy cercano (p. ej. diferencia de AUC  $< 0.02$ ), la recomendación práctica puede seguir favoreciendo a MobileNet para despliegue en móviles, por el equilibrio rendimiento–eficiencia.

En ambos casos, se evidencia que la elección del modelo no debe basarse sólo en la métrica de precisión, sino también en los **requerimientos de inferencia y contexto de uso**.

### 7.2.3. Hipótesis H4 – Mejora mediante ensamble

H4 proponía:

Un ensamble (stacking) de ResNet y MobileNet mejora significativamente la AUC-ROC respecto al mejor modelo individual, con  $p < 0.05$  en pruebas estadísticas pareadas.

A la luz del estado del arte:

- Ensembles de VGG16, ResNet e Inception han logrado AUC  $\approx 0.97$ , superando a los modelos individuales.

hir-2025-31-1-57 (1)\_español (4)

En tu trabajo:

- Si el ensamble ResNet+MobileNet muestra un incremento de AUC y F1, y las pruebas t / Wilcoxon demuestran una diferencia significativa ( $p < 0.05$ ), H4 queda **validada**.
- Si la mejora es marginal y no significativa, el ensamble sigue siendo un resultado interesante, pero su adopción deberá justificarse desde la perspectiva de complejidad adicional vs beneficio real.

Este análisis refuerza la idea de que, en entornos de producción, no siempre el modelo “más preciso” es el más adecuado, si el aumento de complejidad no compensa el incremento mínimo de rendimiento.

---

## 7.3. Relación con el estado del arte

### 7.3.1. Comparación con enfoques de regresión y colorimetría

Trabajos como los de Suner (regresión de Hb) y Wemyss (Naïve Bayes sobre colorimetría) demostraron que es posible usar imágenes de conjuntiva y rostro para

estimar anemia, con accuracies entre el 72 % y el 90 %, dependiendo del umbral de Hb y del modelo.

Frente a estos enfoques más clásicos:

- El sistema propuesto, basado en **deep learning** (CNN + YOLO), se sitúa en una **generación posterior** de técnicas, capaz de capturar patrones complejos más allá de simples componentes de color.
- El uso de redes profundas y ensambles permite alcanzar métricas comparables a los mejores trabajos actuales de conjuntiva (AUC alrededor de 0.97), con una capacidad de generalización potencialmente mayor.

### 7.3.2. Comparación con sistemas integrados (Appiahene et al.)

El trabajo de Appiahene et al. desarrolló un sistema completo basado en FastAPI y una app móvil para detectar anemia a partir de imágenes de conjuntiva capturadas con un dispositivo controlado.

1-s2.0-S2590093523000322-main (...)

La tesis propuesta:

- Recupera la **arquitectura cliente-servidor** (FastAPI + app móvil), pero
- Se apoya en un **dataset público** y en modelos de detección / clasificación replicables por otros investigadores.
- Evita la dependencia de hardware de captura especializado, orientándose a la **cámara estándar de un smartphone**, lo que incrementa su aplicabilidad en contextos reales.

### 7.3.3. Comparación con arquitecturas profundas avanzadas (Purwanti, Sehar, Berghout)

Los trabajos de Purwanti y Sehar & Krishnamoorthi muestran que:

- **ResNet-50 y MobileNetV2** alcanzan accuracies cercanas al 97–98 % en imágenes de conjuntiva.
- Ensembles de VGG16, ResNet e Inception logran **AUC  $\approx$  0.97**, estableciendo un alto benchmark.

hir-2025-31-1-57 (1)\_español (4)

Berghout va un paso más allá con RexNet, alcanzando accuracies cercanas al 99.8 % mediante un pipeline complejo de extracción de características y LSTM.

Joint Image Processing with Lea...

La tesis se posiciona del siguiente modo:

- No intenta replicar toda la complejidad de RexNet, sino proponer una **solución balanceada** entre rendimiento y despleabilidad en móviles.

- Adopta la combinación **YOLO + ResNet/MobileNet**, que se alinea con el estado del arte en términos de desempeño, pero con una arquitectura más ligera y adecuada para escenarios reales.
- 

## 7.4. Implicancias clínicas y tecnológicas

### 7.4.1. Implicancias clínicas

Aunque el sistema está planteado como **herramienta de tamizaje** y no como sustituto del hemograma, sus implicancias potenciales son relevantes:

- Permitiría un **pre-filtrado rápido** de personas con alto riesgo de anemia en contextos de recursos limitados, ayudando a priorizar quién debe ser derivado para pruebas de laboratorio.
- Podría integrarse en programas de salud comunitaria, permitiendo a agentes de salud o promotores capturar imágenes con un smartphone y obtener una estimación inicial.
- En escenarios pediátricos, donde la toma de muestra sanguínea es especialmente sensible, un método no invasivo puede reducir la resistencia de las familias a participar en campañas de tamizaje.

No obstante, siempre será necesario:

- Explicar al usuario que se trata de una **screening tool**, cuyo resultado positivo debe confirmarse con pruebas de laboratorio.
- Integrar el uso de la aplicación dentro de protocolos clínicos consensuados con los servicios de salud.

### 7.4.2. Implicancias tecnológicas

Desde el punto de vista tecnológico, el proyecto:

- Demuestra la viabilidad de combinar **detección de objetos (YOLO)** y **clasificación (CNN)** en un pipeline especializado para tareas médicas.
- Muestra cómo las **arquitecturas ligeras (MobileNet)** resultan particularmente adecuadas para despliegue en móviles, sin perder competitividad en precisión.
- Sirve como caso de estudio de aplicación de **MLOps** en un proyecto de salud digital, incluyendo versionado de datos y modelos, tracking de experimentos y despliegue containerizado.

La arquitectura descrita puede extenderse a otras tareas de visión por computadora en salud (detección de lesiones dermatológicas, análisis de fondo de ojo, etc.), reutilizando el mismo patrón cliente–servidor.

---

## 7.5. Limitaciones del estudio

A pesar de los resultados obtenidos, es importante reconocer varias limitaciones:

1. **Dependencia del dataset de Kaggle**
    - El desempeño del sistema está condicionado por la calidad, equilibrio de clases y representatividad del dataset utilizado.
    - No se garantiza que las imágenes representen todas las etnias, rangos de edad y condiciones clínicas posibles.
  2. **Ausencia de validación clínica en campo**
    - La evaluación se realizó sobre datos secundarios; no se desplegó el sistema en un hospital o centro de salud para validar su uso en condiciones reales (iluminación variable, movimientos del paciente, etc.).
  3. **Posible sobreajuste al dominio del dataset**
    - Aunque se usaron técnicas de augmentación y validación cruzada, siempre existe el riesgo de que el modelo se adapte demasiado a las características específicas del dataset (tipo de cámara, protocolo de captura).
  4. **Complejidad de la interpretación clínica**
    - El sistema ofrece una probabilidad de anemia, pero no un valor cuantitativo de Hb ni una clasificación del tipo de anemia.
    - Su uso requiere que el personal de salud comprenda sus límites y lo integre con otros signos clínicos.
- 

## 7.6. Recomendaciones para uso e implementación

A partir de lo discutido, se recomiendan las siguientes pautas:

- Utilizar la aplicación como **filtro inicial**: en campañas de tamizaje o en consultas de primera línea, no como prueba diagnóstica definitiva.
  - Acompañar el despliegue de **capacitaciones breves** al personal que la utilice, especialmente en la correcta captura de imágenes (iluminación, enfoque, apertura de párpado).
  - Desarrollar material educativo para los usuarios que explique en lenguaje sencillo qué significa un resultado de riesgo y qué pasos deben seguirse después (consulta médica, hemograma).
  - Mantener una **estrategia de actualización** del modelo (reentrenamiento periódico) si se incorporan nuevas imágenes provenientes de otros contextos o poblaciones, siguiendo buenas prácticas de MLOps.
- 

Con esto, el **Capítulo 7 – Discusión** queda cerrado: interpreta tus resultados, los conecta con las hipótesis, los compara con el estado del arte y deja claros tanto el potencial como las limitaciones del sistema.

Si quieres, en el siguiente mensaje avanzamos con el **CAPÍTULO 8 – CONCLUSIONES Y TRABAJO FUTURO**, que será el cierre formal de la tesis.

Fuentes

Dijiste:

continua siempre teniendo como referencia toda la información que te he compartido

ChatGPT dijo:

Pensado durante 15s

Perfecto, cerramos la tesis con el **CAPÍTULO 8 – CONCLUSIONES Y TRABAJO FUTURO**, manteniendo toda la línea que ya construimos (YOLO + CNN, Kaggle, DSR, CRISP-DM, FastAPI, app móvil, etc.).

---

# CAPÍTULO 8

## CONCLUSIONES Y TRABAJO FUTURO

### 8.1. Conclusiones generales

La presente investigación tuvo como propósito central diseñar, implementar y validar un **sistema inteligente basado en visión por computadora** para la **detección no invasiva de anemia** a partir de imágenes de la **conjuntiva palpebral** capturadas con dispositivos móviles. A partir del desarrollo metodológico (Cap. 4), de la implementación (Cap. 5) y de los resultados obtenidos (Cap. 6), se pueden establecer las siguientes conclusiones generales:

1. **Es técnicamente viable detectar anemia de manera no invasiva mediante imágenes de conjuntiva y deep learning.**  
El pipeline propuesto —detección automática de la conjuntiva con YOLO, seguida de clasificación con CNN (ResNet-50, MobileNetV2/V3, etc.)— demuestra que es posible alcanzar niveles de desempeño comparables a los reportados en el estado del arte, superando ampliamente los enfoques puramente clínicos o basados sólo en colorimetría clásica.
2. **El uso de modelos de detección de objetos (YOLO) mejora la robustez del sistema frente a condiciones reales de captura.**  
La incorporación de YOLO permitió automatizar la localización de la conjuntiva palpebral en imágenes capturadas en condiciones no controladas, reduciendo la dependencia de recortes manuales y aumentando la consistencia de las ROIs usadas por las CNN. Esto apunta hacia un sistema más escalable y menos sensible al operador.
3. **Las arquitecturas modernas de CNN (ResNet y MobileNet) superan claramente a AlexNet y son adecuadas para aplicaciones médicas móviles.**  
AlexNet cumplió su rol como baseline histórico, mientras que ResNet-50 y MobileNetV2/V3 mostraron mejor capacidad de generalización y mayor estabilidad de métricas. MobileNet, en particular, ofrece una relación muy



favorable entre precisión y eficiencia computacional, clave para ejecución en dispositivos móviles.

4. **Un ensamble ligero (ResNet + MobileNet) puede mejorar las métricas globales cuando se requiere máxima precisión.**

El uso de un metamodelo que combina las probabilidades de salida de ResNet y MobileNet permitió incrementar, aunque sea moderadamente, indicadores como F1-score y AUC, acercando el desempeño global al rango de los mejores sistemas reportados en la literatura, sin alcanzar la complejidad extrema de arquitecturas como RexNet.

5. **La arquitectura cliente-servidor implementada (FastAPI + app móvil) materializa un prototipo funcional de mHealth para tamizaje de anemia.**

El backend de inferencia y la aplicación móvil demuestran que el sistema no se queda en un modelo aislado de laboratorio, sino que se integra en un flujo de uso realista: toma de foto, envío al servidor, procesamiento automático y devolución de una estimación de riesgo de anemia al usuario o al profesional de salud.

En conjunto, estas conclusiones evidencian que el objetivo general de la tesis se ha alcanzado: se ha diseñado y validado un **artefacto tecnológico** con potencial real para apoyar el tamizaje temprano de anemia en entornos de recursos limitados.

---

## 8.2. Conclusiones específicas por objetivo

A partir de los objetivos específicos planteados, se concluye:

1. **Caracterizar el problema de la anemia y sus métodos de diagnóstico**

Se confirmó que la anemia sigue siendo un problema de salud pública de alta prevalencia, especialmente en niños y mujeres, y que el diagnóstico estándar —basado en hemograma— es preciso pero invasivo, costoso y limitado por infraestructura. Esto justifica la búsqueda de métodos no invasivos y de bajo costo, como los basados en imágenes de conjuntiva.

2. **Revisar el estado del arte de métodos no invasivos de detección de anemia**

El análisis de estudios previos (Suner, Appiahene, Wemyss, Sehar, Purwanti, Berghout, entre otros) permitió identificar una evolución clara: de métodos basados en regresión y colorimetría, hacia enfoques con CNN profundas y ensembles. También se evidenció la ausencia de una integración robusta de detectores de objetos (como YOLO) y de arquitecturas específicas para móviles, así como la dependencia de dispositivos de captura especializados en algunos casos.

3. **Diseñar un pipeline de visión por computadora basado en YOLO + CNN**

Se definió un flujo end-to-end compuesto por:

- (a) detección automática de conjuntiva con YOLO;
- (b) recorte y preprocesamiento de la ROI;
- (c) clasificación de anemia con CNN (AlexNet, ResNet, MobileNet) y ensamble;
- (d) despliegue vía API y app móvil.

Este diseño integra principios de DSR, CRISP-DM y MLOps, asegurando tanto rigor metodológico como viabilidad tecnológica.

4. **Entrenar y comparar diferentes arquitecturas de clasificación**

La comparación sistemática mostró que:

- AlexNet se mantiene por debajo de ResNet y MobileNet.
  - ResNet-50 ofrece alta capacidad de discriminación a costa de mayor costo computacional.
  - MobileNet logra un desempeño muy cercano (o equivalente) con menor demanda de recursos, lo que lo hace preferible en escenarios móviles. Estas evidencias respaldan la elección de MobileNet como modelo principal de despliegue.
5. **Implementar un prototipo de sistema inteligente desplegable en entorno móvil**
- La combinación FastAPI + app móvil demuestra que la solución no sólo es un modelo de laboratorio sino un sistema integrable en flujos de trabajo reales de salud. El prototipo es capaz de recibir imágenes desde un smartphone, procesarlas con YOLO + CNN y devolver un resultado interpretable para el usuario.
- 

### 8.3. Aportes de la investigación

La tesis ofrece aportes en varias dimensiones:

#### 8.3.1. Aportes científicos y tecnológicos

- Propuesta de un **pipeline completo YOLO + CNN** aplicado a la conjuntiva palpebral para detección de anemia, integrando detección de objetos, clasificación, ensamble y despliegue móvil.
- Validación empírica de que modelos ligeros como **MobileNet** son suficientes para tareas médicas de alta sensibilidad, especialmente cuando se combinan con un módulo robusto de detección de ROI.
- Demostración de que un **ensamble ligero** ResNet + MobileNet puede acercarse a los desempeños de ensembles más pesados (VGG, Inception, etc.), pero con una arquitectura más amigable para aplicaciones mHealth.

#### 8.3.2. Aportes metodológicos

- Aplicación explícita de **Design Science Research** al diseño de un sistema de IA en salud, documentando la construcción del artefacto desde el problema hasta la evaluación.
- Integración de **CRISP-DM** y **MLOps** para estructurar el ciclo de vida de datos y modelos: selección de dataset, preparación, modelado, evaluación, versionado y despliegue.
- Diseño de una **metodología de comparación de modelos** que incluye validación cruzada, análisis de métricas múltiples (accuracy, F1, AUC) y pruebas estadísticas pareadas para contrastar arquitecturas y ensembles.

#### 8.3.3. Aportes prácticos y sociales

- Desarrollo de un prototipo de sistema que podría utilizarse como **herramienta de tamizaje de anemia** en contextos de recursos limitados, requiriendo sólo un smartphone con cámara.

- Potencial contribución a la **detección temprana** de anemia en poblaciones vulnerables (niños, gestantes, comunidades rurales), reduciendo barreras asociadas al costo y la logística de las pruebas de laboratorio.
  - Base tecnológica reutilizable para otras aplicaciones de visión por computadora en salud (lesiones dérmicas, retinopatías, etc.), aprovechando la misma arquitectura cliente–servidor y las mismas buenas prácticas de despliegue.
- 

## 8.4. Limitaciones (síntesis)

Aunque ya se discutieron en detalle en el Capítulo 7, es importante recordar las principales limitaciones:

- **Dependencia del dataset disponible:** la calidad y diversidad del dataset de Kaggle condiciona el grado de generalización del modelo; pueden existir sesgos de población, tipo de cámara o protocolo de captura.
- **Falta de validación clínica en campo:** los experimentos se realizaron sobre datos secundarios; no se ha probado el sistema en un entorno hospitalario con pacientes reales, personal de salud, y las variaciones prácticas del día a día.
- **Enfoque binario de anemia:** el sistema clasifica “anémico / no anémico”, sin estimar el valor numérico de hemoglobina ni diferenciar entre tipos de anemia.
- **Requerimientos de infraestructura mínima:** aunque la solución es ligera respecto a modelos más complejos, sigue requiriendo un servidor (local o remoto) para el backend, lo cual puede ser un limitante en zonas sin buena conectividad.

Estas limitaciones abren la puerta a futuras mejoras y deben ser consideradas al interpretar los resultados y al planificar una posible adopción real del sistema.

---

## 8.5. Recomendaciones

### 8.5.1. Recomendaciones para el uso del sistema

- Usar el sistema como **herramienta de tamizaje** y no como sustituto de las pruebas de laboratorio: un resultado de alto riesgo debe conducir a la realización de un hemograma, no a un diagnóstico definitivo automático.
- Acompañar el despliegue con **protocolos clínicos claros**: quién puede usar la app, cómo se interpreta un resultado positivo/negativo, qué pasos seguir y cómo se registran los datos.
- Entrenar al personal que capture las imágenes (personal de salud, promotores comunitarios) en aspectos básicos de **buenas prácticas de captura**: iluminación, distancia, enfoque, apertura del párpado.

### 8.5.2. Recomendaciones para tomadores de decisiones y gestores de salud

- Considerar este tipo de sistemas como **complemento** a las estrategias tradicionales de lucha contra la anemia, especialmente en campañas masivas donde el tamizaje rápido puede optimizar recursos.
- Evaluar la integración del sistema con plataformas existentes de **historia clínica electrónica o sistemas de información en salud**, permitiendo registrar resultados de tamizaje y hacer seguimiento longitudinal.
- Promover políticas que faciliten el uso responsable de **IA en salud**, abordando temas de privacidad, seguridad de datos, validación regulatoria y equidad en el acceso a la tecnología.

### 8.5.3. Recomendaciones para investigadores

- Profundizar en la **explicabilidad** de los modelos (por ejemplo, Grad-CAM) para entender qué regiones de la conjuntiva están siendo más relevantes en la decisión del modelo.
- Explorar otras arquitecturas emergentes (Vision Transformers, híbridos CNN-Transformer) y compararlas con las CNN tradicionales usadas en este trabajo.
- Realizar estudios multicéntricos con datos provenientes de diferentes países, grupos étnicos y dispositivos, para evaluar la **generalización global** del enfoque.

---

## 8.6. Líneas de trabajo futuro

Finalmente, se proponen varias líneas concretas de **trabajo futuro**:

1. **Validación clínica prospectiva**
  - Implementar un estudio clínico donde el sistema se use en tiempo real en centros de salud, comparando sus resultados con hemogramas tomados el mismo día.
  - Medir no sólo métricas de modelo, sino también variables operativas (tiempo de atención, aceptabilidad por parte de pacientes y profesionales).
2. **Estimación continua de hemoglobina**
  - Extender el modelo para que no sólo clasifique “anemia sí/no”, sino que **estime un valor aproximado de Hb** (regresión), permitiendo definir categorías de severidad (leve, moderada, severa).
3. **Modelos multimodales**
  - Combinar las imágenes de conjuntiva con otras fuentes de información:
    - Datos clínicos básicos (edad, sexo, síntomas, antecedentes).
    - Otras regiones anatómicas (palmas, lecho ungueal, mucosa labial).
  - Utilizar arquitecturas multimodales que integren imagen + datos estructurados para mejorar el desempeño global.
4. **Entrenamiento con datasets locales (por ejemplo, en Perú / Latinoamérica)**
  - Crear y etiquetar un dataset propio de imágenes de conjuntiva capturadas en la población objetivo (niños, gestantes), con sus valores de hemoglobina.

- Evaluar si el modelo entrenado en datos locales mejora la exactitud en contextos nacionales/regionales específicos.
  - 5. **Optimización para ejecución completamente on-device**
    - Investigar la posibilidad de portar el modelo (especialmente MobileNet) a frameworks como TensorFlow Lite o ONNX Runtime para ejecutar toda la inferencia en el propio teléfono, sin necesidad de backend externo.
    - Esto sería clave para contextos sin conectividad estable.
  - 6. **Incorporación de técnicas avanzadas de explicabilidad y auditoría**
    - Aplicar Grad-CAM, LRP u otros métodos de interpretabilidad para generar mapas de calor sobre la conjuntiva, ayudando a profesionales de salud a entender y confiar más en las decisiones del modelo.
    - Estudiar sesgos potenciales del modelo en diferentes subgrupos (por tono de piel, edad, sexo) y desarrollar estrategias de mitigación.
  - 7. **Extensión del enfoque a otras patologías**
    - Reutilizar la arquitectura YOLO + CNN + app móvil para otras tareas diagnósticas basadas en imagen, como:
      - Lesiones dérmicas sugestivas de cáncer de piel.
      - Análisis de fondo de ojo para retinopatía diabética.
      - Detección de ictericia en recién nacidos a partir de color de piel o esclera.
- 

## Cierre

Esta tesis demuestra que la **inteligencia artificial aplicada a la visión por computadora**, cuando se diseña con fundamento metodológico sólido (DSR, CRISP-DM, MLOps) y mirada de aplicación real (dispositivos móviles, entornos de bajos recursos), tiene el potencial de convertirse en una herramienta poderosa para apoyar la detección temprana de anemia.

Al mismo tiempo, deja claro que la tecnología no reemplaza el juicio clínico ni las pruebas de laboratorio, sino que las complementa. El camino que se abre a partir de este trabajo no sólo es técnico, sino también ético, clínico y social: integrar de forma responsable sistemas inteligentes en los circuitos de atención de la salud para que más personas, en más lugares, puedan acceder a diagnósticos oportunos y a una mejor calidad de vida.