

# Detección de conjuntivitis a partir de imágenes oculares utilizando arquitecturas CNN y un modelo de fusión

## Conjunctivitis detection from ocular images using CNN architectures and a fusion model

Antony Stebend Camacho Garcia<sup>A</sup>, Suker Huamanñahui Hilario<sup>B</sup>, Mario Aquino Cruz<sup>C</sup> y Evelyn Naida Luque Ochoa<sup>D</sup>

**Resumen**— La conjuntivitis es una causa frecuente de consulta oftalmológica y puede generar complicaciones graves si no se detecta a tiempo. Este estudio evalúa tres arquitecturas de redes neuronales convolucionales preentrenadas —EfficientNetB0, InceptionV3 y ResNet50— para la clasificación automática de imágenes oculares con y sin signos de conjuntivitis. Además, se propone un modelo de fusión que combina las salidas intermedias de InceptionV3 y ResNet50. Todos los modelos fueron entrenados usando aceleración por GPU y técnicas de regularización y aumento de datos. InceptionV3 obtuvo el mejor desempeño general, con una exactitud del 95.00 % y una precisión de 0.98 para la clase positiva y 0.92 para la clase negativa, mostrando un balance sólido entre sensibilidad y especificidad. EfficientNetB0 destacó por alcanzar el mayor recall en la clase positiva y la menor tasa de falsos negativos, aunque con una precisión considerablemente menor. ResNet50 presentó la menor tasa de falsos positivos, siendo útil para reducir diagnósticos erróneos en casos negativos. El modelo de fusión alcanzó métricas competitivas y un tiempo de entrenamiento reducido, demostrando que la integración de arquitecturas puede aumentar la robustez del sistema. Este trabajo ofrece una guía práctica y reproducible para la selección de modelos en el diagnóstico automatizado de conjuntivitis, especialmente útil en contextos con recursos médicos limitados.

**Palabras clave:** aprendizaje profundo, conjuntivitis, imágenes médicas, redes neuronales convolucionales

**Abstract**— Conjunctivitis is a common cause of ophthalmologic consultation and may lead to serious complications if not diagnosed early. This study evaluates three pretrained convolutional neural network architectures—EfficientNetB0, InceptionV3, and ResNet50—for the automated classification of ocular images with and without signs of conjunctivitis. Additionally, a fusion model combining intermediate outputs from InceptionV3 and ResNet50 is proposed. All models were trained using GPU acceleration and employed regularization and data augmentation techniques. InceptionV3 achieved the best overall performance, with an accuracy of 95.00% and a precision of 0.98 for the positive class and 0.92 for the negative class, showing a solid balance between sensitivity and specificity. EfficientNetB0 achieved the highest recall in the positive class and the lowest false negative rate, although with considerably lower precision. ResNet50 obtained the lowest false positive rate, making it valuable for reducing misdiagnosis in negative cases. The fusion model achieved competitive metrics and a reduced training time, indicating that combining complementary architectures can enhance system robustness. This work provides a practical and reproducible guide for selecting CNN models for automated conjunctivitis diagnosis, particularly in resource-limited medical settings.

**Keywords:** deep learning, conjunctivitis, medical imaging, convolutional neural networks

## 1 INTRODUCCIÓN

La conjuntivitis es una patología inflamatoria común de la conjuntiva ocular, con un impacto significativo en la salud pública y una alta prevalencia a nivel mundial [1][2]. Se presenta en diversas formas clínicas, entre las que destacan la viral, la bacteriana y la alérgica.

Las infecciones bacterianas son responsables de aproximadamente el 50 % al 70 % de los casos de conjuntivitis, siendo *Staphylococcus aureus*, *Streptococcus pneumoniae* y *Pseudomonas aeruginosa* los agentes etiológicos más frecuentes [2]. Esta afección constituye una de las causas más comunes de consulta oftalmológica, especialmente en entornos de atención primaria.



Revista de Investigación en Ciencia y Tecnología  
ISSN: 2810-8124 (en línea) / ISSN: 2706-543x  
Universidad Nacional Micaela Bastidas de Apurímac – Perú

Vol. 7 Núm. 2 (2025) - Publicado: 19/08/25 - [Indexaciones](#)  
Número: [doi.org/10.57166/riqchary/v7.n2.2025](https://doi.org/10.57166/riqchary/v7.n2.2025)  
Páginas: 54- 65 | Recibido 01/01/2025 ; Aceptado 01/02/2025

[doi.org/10.57166/riqchary.v7.n2.2025.7](https://doi.org/10.57166/riqchary.v7.n2.2025.7)

Autores:

- A. **ORCID iD** <https://orcid.org/0009-0004-3492-7741>  
Antony Stebend Camacho García, estudiante en la Universidad Nacional Micaela Bastidas de Apurímac, Pe. [162135@unamba.edu.pe](mailto:162135@unamba.edu.pe)
- B. **ORCID iD** <https://orcid.org/0009-0002-4945-6263>  
Suker Huamanñahui Hilario, estudiante en la Universidad Nacional Micaela Bastidas de Apurímac, Pe. [162157@unamba.edu.pe](mailto:162157@unamba.edu.pe)
- C. **ORCID iD** <https://orcid.org/0000-0002-2552-5669>  
Mario Aquino Cruz, docente del Departamento Académico de Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac, Pe. [maquino@unamba.edu.pe](mailto:maquino@unamba.edu.pe)
- D. **ORCID iD** <https://orcid.org/0000-0002-8386-9806>  
Evelyn Naida Luque Ochoa, docente del Departamento Académico de Informática y Sistemas de la Universidad Nacional Micaela Bastidas de Apurímac, Pe. [eluque@unamba.edu.pe](mailto:eluque@unamba.edu.pe)

Según la Academia Americana de Oftalmología, millones de personas se ven afectadas cada año, generando una carga considerable para los sistemas sanitarios [1].

En países como el Perú, donde existen limitaciones en el acceso a servicios especializados en regiones rurales o amazónicas, el diagnóstico temprano y preciso se vuelve particularmente desafiante.

Tradicionalmente, el diagnóstico de la conjuntivitis se basa en la observación clínica, lo que puede llevar a errores diagnósticos en ausencia de personal capacitado o recursos complementarios. Ante esta problemática, las tecnologías de inteligencia artificial (IA) han emergido como herramientas prometedoras para asistir en el diagnóstico médico automatizado, especialmente mediante el uso de redes neuronales convolucionales (CNN).

Estas arquitecturas han demostrado una alta eficacia en la clasificación de imágenes médicas al identificar patrones complejos en estructuras visuales. Modelos como ResNet50, InceptionV3 y EfficientNetB0 han sido aplicados con éxito en la detección de diversas patologías oculares [3][4]. Por ejemplo, Mondal et al. lograron una precisión de hasta el 95.2 % con InceptionV3 y 93.6 % con ResNet50 al clasificar imágenes oculares con y sin conjuntivitis [4]. De manera similar, Bitto y Mahmud desarrollaron un sistema multicategoría con CNN que incluyó conjuntivitis, alcanzando una precisión de 97.08 % en clasificación ocular [5].

El presente estudio propone un análisis comparativo entre ResNet50, InceptionV3 y EfficientNetB0, aplicados al diagnóstico automatizado de conjuntivitis mediante imágenes oculares. Además, se implementó un modelo de fusión (ensemble) que combina las salidas intermedias de las arquitecturas de ResNet50 e InceptionV3, con el fin de mejorar la capacidad de clasificación mediante integración de características.

La investigación emplea un enfoque cuantitativo, bajo un diseño no experimental y transversal. Los modelos fueron entrenados mediante transferencia de aprendizaje, utilizando un conjunto de datos balanceado, preprocesado y evaluado mediante métricas estándar: exactitud, precisión, recall, F1-score, pérdida (loss) y matriz de confusión. El desarrollo se realizó con herramientas abiertas como TensorFlow y Keras sobre Google Colaboratory.

La principal contribución de este trabajo radica en proveer una comparación reproducible entre tres arquitecturas CNN de última generación en la tarea de diagnóstico ocular automatizado, así como en el desarrollo de un modelo de fusión (ensemble) que integra sus características para mejorar el rendimiento diagnóstico. Los resultados obtenidos pueden ser valiosos para futuros sistemas de apoyo clínico basados en IA, especialmente en contextos donde el acceso a especialistas es limitado.

## 2 TRABAJOS RELACIONADOS

En la última década, el uso de redes neuronales convolucionales (CNN) ha cobrado relevancia en el diagnóstico automatizado de enfermedades médicas a partir de imágenes, incluyendo afecciones oculares como la conjuntivitis. Gracias a su capacidad para extraer características complejas, las CNN han demostrado ser herramientas efectivas en

contextos clínicos, incluso en comparación con la evaluación humana especializada.

Se desarrolló un modelo basado en CNN para detectar conjuntivitis en imágenes y generar recomendaciones de tratamiento farmacológico. Su enfoque logró una alta precisión en las predicciones, resaltando el potencial del aprendizaje profundo en sistemas de apoyo clínico personalizados [6].

En un contexto complementario, también propusieron una arquitectura conceptual diseñada para telemedicina, la cual utiliza redes neuronales para la clasificación remota de casos de conjuntivitis. Su estudio destaca el valor de estas soluciones en entornos rurales o con recursos limitados, donde el acceso a especialistas es restringido [7].

Un enfoque de clasificación multiclase fue realizado utilizando la arquitectura VGG16 para identificar diferentes enfermedades externas del ojo, incluida la conjuntivitis. Su sistema alcanzó una precisión del 98.48 %, reforzando la fiabilidad de los modelos CNN en escenarios diagnósticos complejos [8]. De manera similar, se desarrollaron un sistema de clasificación binaria para detectar conjuntivitis mediante una red convolucional personalizada. Al ser entrenado con más de 5,000 imágenes, el modelo logró una precisión del 88.80 %, con valores consistentes de recall y F1-score, lo que respalda su utilidad en entornos clínicos reales [9].

Si bien estos estudios evidencian avances importantes, la mayoría de ellos se centran en evaluar arquitecturas individuales o carecen de comparaciones sistemáticas entre modelos bajo condiciones homogéneas. Tampoco exploran de forma profunda las ventajas del ensamblaje de redes neuronales para mejorar el rendimiento diagnóstico.

En respuesta a estas limitaciones, el presente estudio propone una evaluación comparativa de tres arquitecturas CNN ampliamente reconocidas —ResNet50, InceptionV3 y EfficientNetB0— aplicadas al diagnóstico automatizado de conjuntivitis. Asimismo, se introduce un modelo de fusión (ensemble) que integra las salidas intermedias de las arquitecturas ResNet50 e InceptionV3, con el objetivo de aumentar la precisión y robustez del sistema. Esta propuesta busca contribuir con evidencia cuantitativa que respalde el uso de modelos híbridos de aprendizaje profundo en contextos clínicos donde el acceso a recursos especializados es restringido.

## 3 METODOLOGÍA

### A. DISEÑO DE INVESTIGACIÓN

Este estudio se desarrolló bajo un enfoque metodológico cuantitativo, con un diseño no experimental de corte transversal. En este tipo de investigaciones, no se manipulan las variables, sino que se observan tal como ocurren en un único momento temporal, lo que permite analizar asociaciones bajo condiciones controladas y replicables [10].

En este contexto, la presente investigación se enfocó en comparar el rendimiento de diversas arquitecturas de redes neuronales convolucionales (CNN) aplicadas al diagnóstico automatizado de conjuntivitis, utilizando un conjunto de imágenes oculares balanceado, preprocesado y evaluado con métricas estandarizadas.

Para evaluar el rendimiento de los modelos, se utilizaron

métricas estándar de clasificación como la exactitud, precisión, recall, F1-score y pérdida (loss), las cuales permitieron establecer comparaciones objetivas y consistentes entre las arquitecturas implementadas[11]. Además, se propuso un modelo de fusión que combina las salidas intermedias de tres arquitecturas CNN preentrenadas —ResNet50, InceptionV3 y EfficientNetB0— con el objetivo de mejorar la capacidad de clasificación mediante integración de características profundas, en comparación con el desempeño individual de cada modelo.

## B. PARTICIPANTES

La muestra se dividió en tres subconjuntos: 60% para entrenamiento, 20% para validación y 20% para prueba, asegurando que cada subconjunto mantuviera una distribución equitativa de ambas clases (ojos con y sin conjuntivitis), lo que minimizó el sesgo en la evaluación del modelo.

Las Figuras 1 y 2 presentan ejemplos visuales representativos de las imágenes utilizadas en el estudio: una imagen ocular con signos clínicos de conjuntivitis y una imagen de un ojo sano, respectivamente. Las Figuras 3, 4 y 5 ilustran las arquitecturas de las redes neuronales convolucionales evaluadas: ResNet50, InceptionV3 y EfficientNetB0, mientras que la Figura 6 muestra el diseño del modelo de fusión propuesto, el cual integra las salidas de las redes anteriores a nivel de características para mejorar la capacidad de clasificación.

En la Fig.1 se muestra imagen representativa de un ojo con signos clínicos de conjuntivitis. Se observan enrojecimiento conjuntival, inflamación y posible secreción mucosa. la fuente de la imagen es exclusivamente extraída de nuestro conjunto de datos usado para este estudio.



Fig. 1. Imagen de un ojo con signos clínicos de conjuntivitis.

La Fig.2 nos muestra una imagen de un ojo sano, sin presencia de signos inflamatorios ni alteraciones visibles, la fuente de la imagen es exclusivamente extraída de nuestro conjunto de datos usado para este estudio.

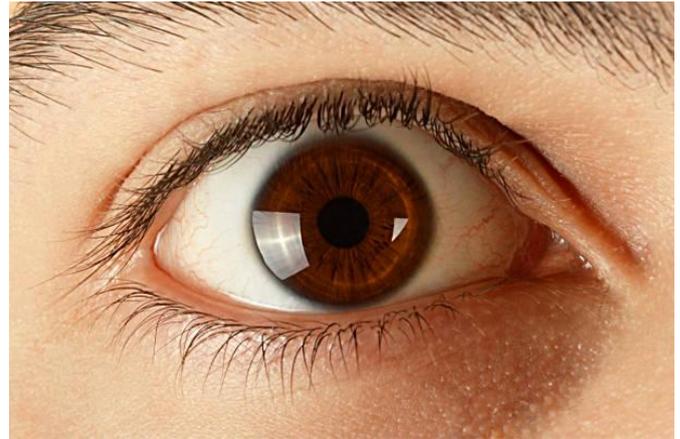


Fig. 2. Imagen de un ojo sano.

Diseño de las arquitecturas: Las siguientes figuras ilustran las arquitecturas de redes neuronales convolucionales preentrenadas utilizadas.

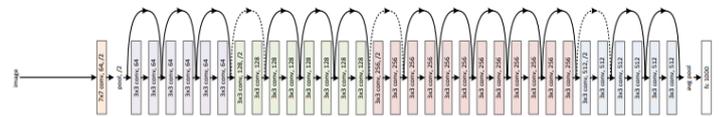


Fig. 3. Arquitectura de ResNet50. Fuente: extraída de [15]

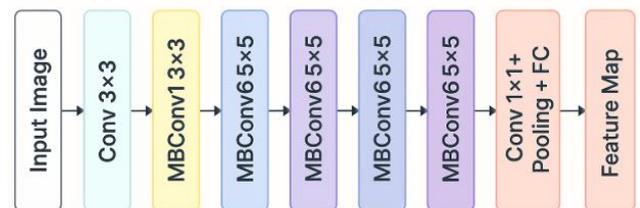


Fig. 4. Arquitectura del modelo EfficientNetB0, adaptada de [16].

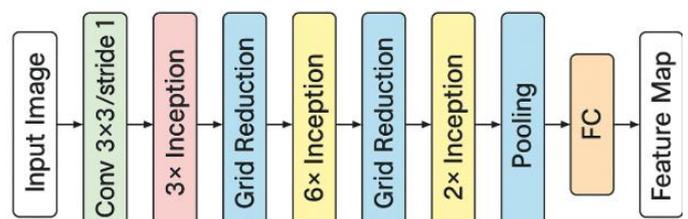


Fig. 5. Arquitectura del modelo InceptionV3, adaptada de [17].

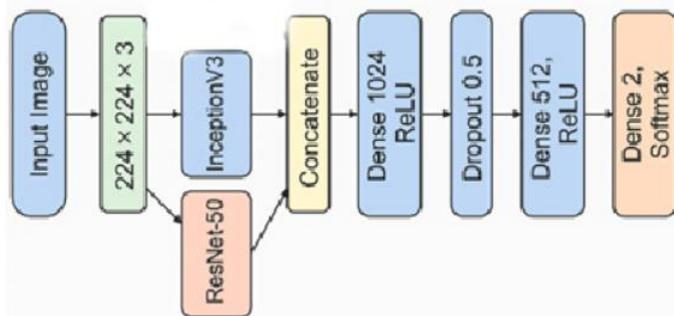


Fig. 5. Arquitectura del modelo fusionado.

### C. INSTRUMENTOS Y TÉCNICAS

Para el desarrollo del estudio se utilizó el lenguaje de programación Python[12], en conjunto con las bibliotecas TensorFlow y Keras, ampliamente reconocidas en tareas de aprendizaje profundo[13]. El entorno de desarrollo fue Google Colaboratory (versión gratuita), que proporcionó acceso a recursos de cómputo en la nube, incluyendo una GPU asignada automáticamente por el entorno, la cual permitió ejecutar eficientemente los experimentos sin requerir infraestructura local dedicada [14].

El conjunto de datos utilizado en esta investigación fue construido específicamente para este estudio. Para ello, se recopilaron imágenes de ojos con y sin signos de conjuntivitis a partir de repositorios públicos como Kaggle y diversas fuentes confiables en internet. Las imágenes fueron seleccionadas manualmente, asegurando la calidad visual y la representatividad clínica, y posteriormente clasificadas en dos categorías: ojos sanos y ojos con conjuntivitis. El conjunto final quedó conformado por 3,479 imágenes de ojos sanos y 4,485 imágenes de ojos con conjuntivitis. Todas las imágenes fueron sometidas a un proceso de preprocesamiento y aumento de datos, que incluyó redimensionamiento, rotación, inversión horizontal, zoom y ajustes de brillo, con el fin de mejorar la variabilidad del conjunto y fortalecer la capacidad de generalización de los modelos. Finalmente, se estandarizaron a una resolución de 224×224 píxeles y se reescalaron al rango [0,1] para su procesamiento por las redes neuronales convolucionales. Se aplicó aprendizaje transferido (*transfer learning*) utilizando tres arquitecturas CNN preentrenadas: ResNet50 [15], EfficientNetB0[16] y InceptionV3 [17]. Adicionalmente, se construyó un modelo de fusión que integró las salidas intermedias de dichas redes, combinando sus características mediante *feature-level fusion* y capas densas compartidas [18]. Todos los modelos fueron adaptados a una tarea de clasificación binaria con función de pérdida *categorical crossentropy*.

Como técnica de preprocesamiento, se aplicó un esquema de aumento de datos (*data augmentation*) con rotaciones aleatorias, desplazamientos, zoom, ajustes de brillo e inversión horizontal. Estas transformaciones permitieron simular variaciones en las condiciones reales de captura de imágenes y mejorar la capacidad de generalización de los modelos [19].

### D. ANÁLISIS DE DATOS

El análisis se fundamentó en métricas ampliamente utilizadas en clasificación binaria, tales como exactitud (accuracy), pérdida (loss), precisión, recall y F1-score. Además, se empleó una matriz de confusión para identificar falsos positivos (FP) y falsos negativos (FN), aspectos especialmente relevantes en aplicaciones clínicas sensibles [11]. Los modelos fueron entrenados bajo las mismas condiciones, con técnicas de regularización como *EarlyStopping* y reducción dinámica de la tasa de aprendizaje. Esta homogeneidad metodológica garantizó la validez y la reproducibilidad de la comparación entre arquitecturas.

## 4 RESULTADOS

Los resultados obtenidos reflejan el rendimiento de las diferentes arquitecturas CNN empleadas en este estudio: ResNet50, EfficientNetB0, InceptionV3. Cada modelo fue entrenado bajo las mismas condiciones experimentales, utilizando el conjunto de entrenamiento preprocesado y validado, con evaluaciones finales sobre el conjunto de prueba. A continuación, se presentan los resultados obtenidos por cada arquitectura de manera individual, seguidos por el análisis del modelo fusionado. Las métricas utilizadas para la comparación fueron: exactitud (accuracy), pérdida (loss), precisión, recall, F1-score y matriz de confusión.

### A. RESULTADOS DEL MODELO RESNET50

La arquitectura ResNet50 fue entrenada mediante aprendizaje transferido, alcanzando un total de 28 épocas antes de la activación del callback *EarlyStopping*, el cual monitoreaba la pérdida de validación (*val\_loss*). El entrenamiento se llevó a cabo en un entorno proporcionado por Google Colaboratory (versión gratuita), con una GPU asignada automáticamente, y tuvo una duración aproximada de 10 horas y 5 minutos. Durante el entrenamiento, se evidenció una mejora progresiva tanto en la precisión como en la pérdida, alcanzando en la última época una exactitud de 67.84 % y una pérdida de 0.6165 sobre los datos de entrenamiento, mientras que la precisión en validación fue de 78.20 % con una pérdida de 0.5401. Este entorno permitió ejecutar los experimentos de manera eficiente sin necesidad de infraestructura local dedicada.

TABLA I. RENDIMIENTO DE RESNET50 EN ENTRENAMIENTO, VALIDACIÓN Y TEST

Métrica	Train	Validation	Test
Exactitud (Accuracy)	68.73 %	78.00 %	76.66%
Pérdida (Loss)	0.6092	0.5622	0.5294
Precisión (Precision)	0.72 (C), 0.68 (N)	0.77 (C), 0.79 (N)	0.75 (C), 0.79 (N)
Recall	0.73 (C), 0.65 (N)	0.86 (C), 0.67 (N)	0.87 (C), 0.63 (N)
F1-Score	0.69	0.81	0.82

Leyenda:

- C: *Conjuntivitis* (ojos con infección)
- N: *Normal* (ojos sin signos de infección)

La Fig. 6 muestra la evolución de la exactitud durante el entrenamiento de ResNet50:

- **Train Accuracy:** La precisión en el conjunto de entrenamiento mostró una mejora progresiva y constante, iniciando en 53.88 % en la época 1 y alcanzando un máximo de 68.73 % en la época 26. Esto indica que el modelo fue aprendiendo de forma gradual, sin señales de sobreajuste prematuro.
- **Validation Accuracy:** También se observó una tendencia positiva, comenzando en 43.66 % en la época 1 y alcanzando su pico en la época 26 con 78.89 %. Aunque hubo fluctuaciones (por ejemplo, una caída a 60.30 % en la época 6), la tendencia general fue de mejora sostenida.
- **Brecha entre entrenamiento y validación:** La diferencia entre ambas métricas fue moderada y estable. En muchas épocas, la precisión en validación fue incluso superior a la de entrenamiento, lo cual sugiere que el modelo no presentó sobreajuste y logró una generalización adecuada.

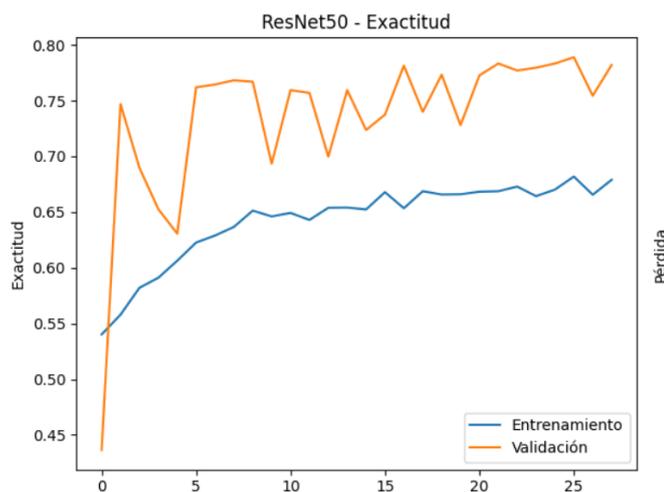


Fig. 6. Curva de exactitud (Accuracy) durante el entrenamiento de ResNet50.

- **Train Loss:** La pérdida en el conjunto de entrenamiento disminuyó progresivamente, pasando de 0.7997 en la época 1 a 0.6092 en la época 26, lo cual refleja una mejora constante en la capacidad del modelo para ajustarse a los datos.
- **Validation Loss:** También presentó una tendencia descendente, iniciando en 0.7100 y alcanzando su valor mínimo de 0.5387 en la época 23. Aunque se observaron pequeñas oscilaciones (por ejemplo, en las épocas 10, 19 y 26), no se evidenciaron aumentos significativos que indiquen sobreajuste.
- **Brecha entre entrenamiento y validación:** La diferencia entre ambas curvas de pérdida se mantuvo leve y estable durante la mayoría del entrenamiento, e incluso en

varias épocas la pérdida en validación fue menor que la de entrenamiento, reforzando la conclusión de que el modelo generalizó adecuadamente.

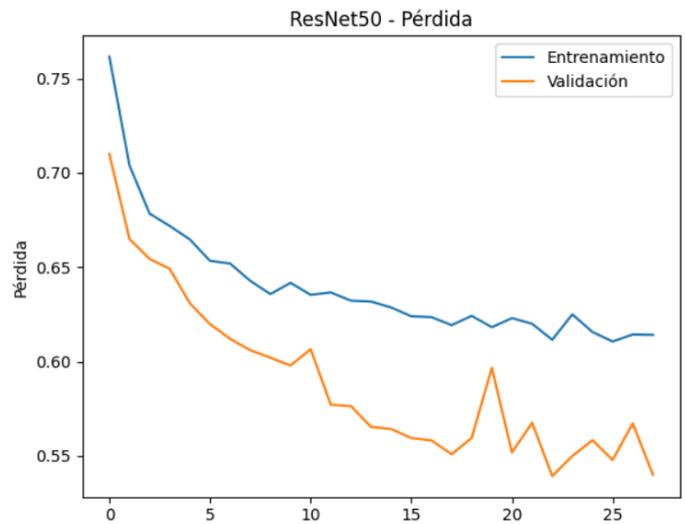


Fig. 7. Curva de pérdida (Loss) durante el entrenamiento de ResNet50.

La matriz de confusión (Fig. 8) correspondiente al conjunto de prueba muestra un desempeño aceptable del modelo ResNet50, aunque revela limitaciones en la clasificación de imágenes normales.

- Verdaderos Positivos (TP): 753 (84 % de los casos de conjuntivitis fueron correctamente identificados).
- Falsos Negativos (FN): 144 (16 % de las imágenes con conjuntivitis fueron clasificadas erróneamente como normales).
- Falsos Positivos (FP): 195 (28 % de los casos normales fueron clasificados incorrectamente como conjuntivitis).
- Verdaderos Negativos (TN): 502 (72 % de las imágenes normales fueron clasificadas correctamente).

Este resultado indica una buena sensibilidad para detectar conjuntivitis, lo cual es relevante en un entorno clínico para reducir diagnósticos omitidos. No obstante, la tasa de falsos positivos del 28 % representa un riesgo de generar alarmas innecesarias, lo cual puede afectar la eficiencia del sistema si se implementa como herramienta de apoyo en el diagnóstico médico.

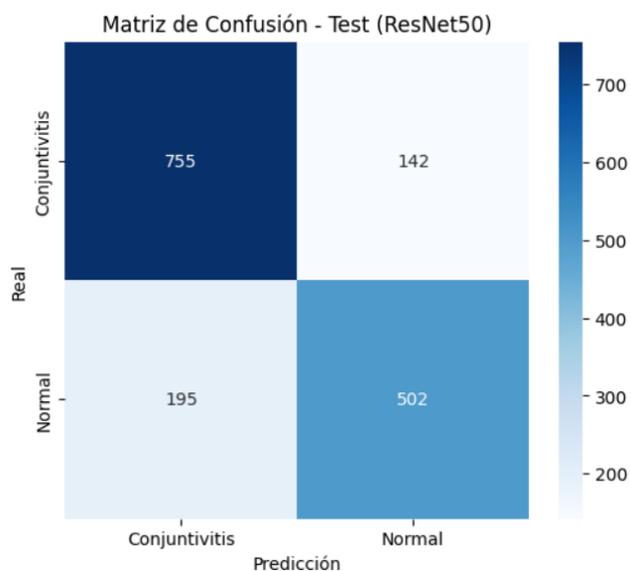


Fig. 8. Matriz de confusión del modelo de ResNet50 con el conjunto de test.

## B. RESULTADOS DEL MODELO EFFICIENTNETB0

La arquitectura EfficientNetB0 fue entrenada utilizando aprendizaje por transferencia, partiendo de pesos preentrenados en ImageNet. Inicialmente, se empleó el modelo como extractor de características con las capas convolucionales congeladas, logrando una estabilidad parcial en la validación. Posteriormente, se aplicó *fine-tuning* (ajuste fino) al desbloquear progresivamente capas del modelo base. Esta estrategia se adoptó con el objetivo de mejorar la capacidad del modelo para captar patrones específicos del dominio médico ocular, dado que las características visuales presentes en imágenes de conjuntivitis pueden diferir significativamente de las de ImageNet.

Durante el entrenamiento, se ejecutaron múltiples sesiones, alcanzando un total de 13 épocas antes de activarse el *callback* EarlyStopping, el cual monitoreaba la pérdida de validación (*val\_loss*). El mejor desempeño en validación se observó en la época 9, con una exactitud del 69.28 % y una pérdida de 0.6084. El tiempo total de entrenamiento acumulado fue de aproximadamente 9 horas y 43 minutos, utilizando una GPU asignada automáticamente por Google Colaboratory (versión gratuita). Esta infraestructura remota permitió realizar los experimentos sin requerir recursos locales especializados, aunque con ciertas limitaciones en el rendimiento computacional.

TABLA II. RENDIMIENTO DE EFFICIENTNETB0 EN ENTRENAMIENTO, VALIDACIÓN Y TEST

Métrica	Train	Validation	Test
Exactitud (Accuracy)	63.41 %	69.00 %	70.00 %
Pérdida (Loss)	0.6213	0.7254	0.6052

Precisión (Precision)	0.70 (C) / 0.72 (N)	0.66 (C), 0.84 (N)	0.66 (C), 0.84 (N)
Recall	0.85 (C) / 0.42 (N)	0.95 (C), 0.37 (N)	0.95 (C), 0.38 (N)
F1-Score	0.65	0.65	0.65

Legenda:

- **C:** Conjuntivitis (ojos con infección)
- **N:** Normal (ojos sin signos de infección)

La Fig. 9 muestra la evolución de la exactitud durante el entrenamiento de EfficientNetB0:

- **Train Accuracy:** La precisión en el conjunto de entrenamiento mostró un crecimiento moderado y progresivo. Comenzó en 48.94 % en la época 1 y alcanzó un máximo de 63.41 % en la época 13. Esta evolución indica que el modelo fue aprendiendo gradualmente las características del conjunto de datos, sin señales claras de sobreajuste temprano.
- **Validation Accuracy:** En cuanto al conjunto de validación, la precisión también aumentó con el tiempo, aunque de forma más inestable. Inició en 43.66 % y alcanzó su punto máximo en la época 9 con 69.28 %, antes de descender nuevamente en las últimas épocas. Este comportamiento sugiere que el modelo logró una buena capacidad de generalización, pero también que entrenamientos adicionales podrían haber inducido sobreajuste si no se aplicaba early stopping.
- **Brecha entre entrenamiento y validación:** Durante las primeras épocas, la precisión en validación fue incluso superior a la de entrenamiento en algunas ocasiones (como en la época 6). A lo largo del proceso, ambas curvas se mantuvieron razonablemente cercanas, con una diferencia promedio menor al 10 %, lo que indica un entrenamiento equilibrado sin un sobreajuste evidente.

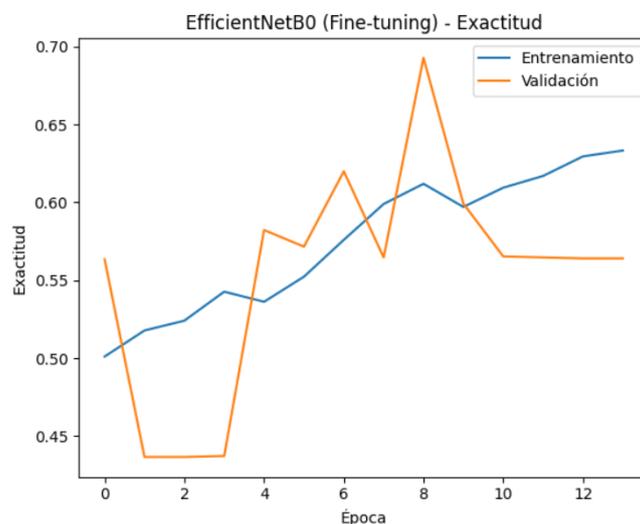


Fig. 9. Curva de exactitud (Accuracy) durante el entrenamiento de EfficientNetB0.

La Fig. 10 muestra la disminución de la pérdida (Loss):

- **Train Loss:** La pérdida en el conjunto de entrenamiento mostró una reducción progresiva a lo largo del proceso. Comenzó en 0.7292 en la época 1 y descendió hasta alcanzar 0.6213 en la época 13. Aunque la caída no fue abrupta, reflejó una mejora constante en la capacidad del modelo para ajustar los datos del entrenamiento.
- **Validation Loss:** La pérdida en el conjunto de validación también siguió una tendencia descendente, aunque con algunas oscilaciones. Inició en 0.6914 y alcanzó un mínimo relativo de 0.6084 en la época 9, antes de volver a subir progresivamente en las siguientes épocas (hasta 0.8567 en la época 14). Este comportamiento indica que el modelo logró una buena capacidad de generalización en etapas intermedias, aunque hacia el final del entrenamiento comenzó a mostrar signos de sobreajuste.
- **Brecha entre entrenamiento y validación:** La diferencia entre ambas pérdidas se mantuvo inicialmente estable, pero a partir de la época 10 comenzó a ampliarse. Esta divergencia creciente entre el *train loss* y el *val loss* sugiere que el modelo comenzó a sobreajustarse, es decir, a memorizar el conjunto de entrenamiento a expensas de su rendimiento en datos no vistos. Por esta razón, se aplicó la técnica de *early stopping* para detener el entrenamiento y evitar una pérdida de capacidad de generalización.

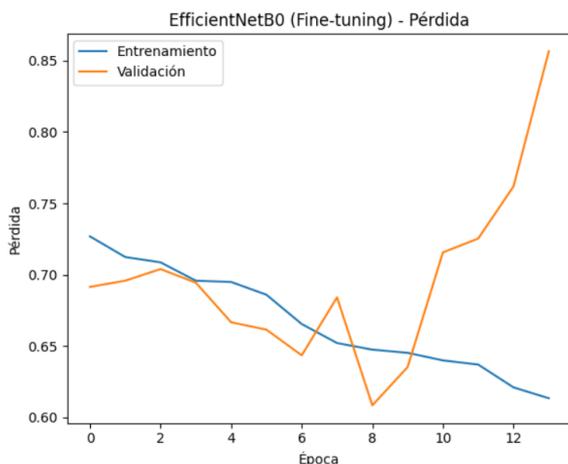


Fig. 10. Curva de pérdida (Loss) durante el entrenamiento de EfficientNetB0.

La matriz de confusión (Fig. 11) evidencia un 5.5 % de falsos negativos (49 casos de conjuntivitis no detectados), lo cual representa una posible limitación en aplicaciones médicas sensibles.

- **Verdaderos Positivos (TP):** 848 (94.5 %). El modelo identificó correctamente la mayoría de los casos de conjuntivitis, lo que indica una buena capacidad de detección.
- **Falsos Negativos (FN):** 49 (5.5 %). Algunos casos de conjuntivitis fueron clasificados como normales, lo que puede representar un riesgo si no se detectan a tiempo.

- **Falsos Positivos (FP):** 431 (61.8 %). El modelo tiende a confundir imágenes normales con conjuntivitis, lo que podría generar revisiones innecesarias.
- **Verdaderos Negativos (TN):** 266 (38.2 %). El modelo tiene dificultades para reconocer correctamente los casos normales, posiblemente por variabilidad en las imágenes o desequilibrio en los datos.

Este resultado refleja una buena capacidad del modelo para detectar conjuntivitis, pero también revela una tasa de error considerable al identificar imágenes normales, lo que podría generar falsas alarmas en un entorno clínico.

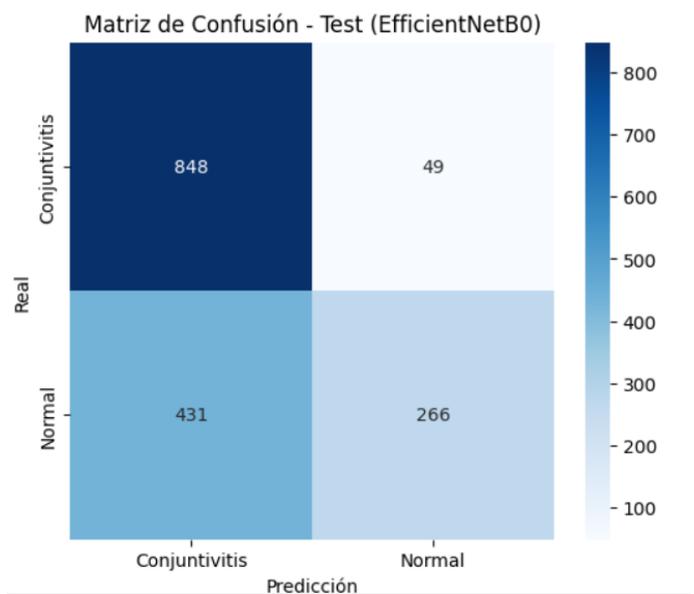


Fig. 11. Matriz de confusión del modelo de EfficientNetB0 con el conjunto de test.

### C. RESULTADOS DEL MODELO INCEPTIONV3

La arquitectura InceptionV3 fue entrenada utilizando la técnica de aprendizaje transferido, ajustando sus pesos finales mediante *fine-tuning*. El proceso de entrenamiento se extendió por un total de 46 épocas, con activación del *callback* EarlyStopping, el cual monitoreaba la pérdida de validación (*val\_loss*) y restauró los mejores pesos obtenidos en la época 41, momento en el que se alcanzó un rendimiento óptimo.

El tiempo total de entrenamiento fue de aproximadamente 10 horas y 40 minutos, considerando un promedio de 13 minutos por época. Para ello, se empleó una GPU proporcionada automáticamente por Google Colaboratory (versión gratuita), lo que facilitó la ejecución eficiente de los experimentos sin necesidad de infraestructura local especializada. Este entorno permitió llevar a cabo el proceso de entrenamiento de modelos avanzados en visión por computadora de manera accesible y reproducible.

**TABLA III. RENDIMIENTO DE INCEPTIONV3 EN ENTRENAMIENTO, VALIDACIÓN Y TEST.**

Métrica	Entrenamiento (Train)	Validación (Validation)	Prueba (Test)
Accuracy	93.15 %	95.10 %	95.11 %
Loss	0.1819	0.1412	0.1200
Precision	0.94 (C), 0.92 (N)	0.96 (C), 0.93 (N)	0.98 (C), 0.92 (N)
Recall	0.92 (C), 0.94 (N)	0.94 (C), 0.95 (N)	0.93 (C), 0.97 (N)
F1-Score	0.93	0.95	0.95

Leyenda:

- **C:** Conjuntivitis (ojos con infección)
- **N:** Normal (ojos sin signos de infección)

La Fig. 12 muestra la evolución de la exactitud durante el entrenamiento de InceptionV3:

- **Train Accuracy:** La precisión en el conjunto de entrenamiento mostró una mejora continua, comenzando en 67.82% en la época 1 y alcanzando un máximo de 93.15% en la época 46. Durante las primeras 20 épocas, se observó una progresión constante, lo que indica una adecuada capacidad de aprendizaje del modelo. Posteriormente, el modelo continuó refinando sus predicciones sin presentar signos evidentes de sobreajuste, lo cual se confirma por el mantenimiento de una alta precisión en la validación.
- **Validation Accuracy:** La precisión sobre el conjunto de validación también presentó un crecimiento sostenido, iniciando en 84.55% y alcanzando un valor máximo de 95.10% en la época 45. El modelo mantuvo una precisión superior al 93% a partir de la época 20, lo cual sugiere una excelente capacidad de generalización sobre datos no vistos. La detención temprana se activó en la época 46 al restaurar los mejores pesos obtenidos en la época 41, con una precisión de validación de 94.66%, consolidando un rendimiento robusto.

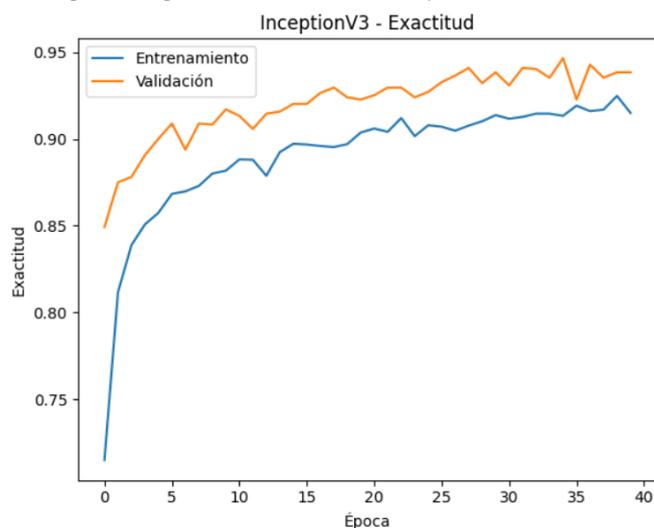


Fig. 12. Curva de exactitud (Accuracy) durante el entrenamiento de InceptionV3.

La Fig. 13 muestra la disminución de la pérdida (Loss):

- **Train Loss:** La pérdida durante el entrenamiento mostró una disminución progresiva y sostenida, comenzando en 0.6335 en la primera época y descendiendo hasta un valor mínimo de 0.1701 en la época 44. Esta reducción constante indica que el modelo fue ajustando exitosamente sus parámetros para minimizar el error, lo que refleja un aprendizaje efectivo de los patrones discriminativos en los datos de entrenamiento.
- **Validation Loss:** La pérdida en el conjunto de validación también presentó una tendencia general descendente, iniciando en 0.3500 y alcanzando su valor mínimo de 0.1376 en la época 41, donde se restauraron los mejores pesos tras la activación del early stopping. Aunque se observaron ligeras oscilaciones en ciertas etapas del entrenamiento, la tendencia general descendente evidencia que el modelo fue capaz de mantener una adecuada capacidad de generalización sin incurrir en sobreajuste.

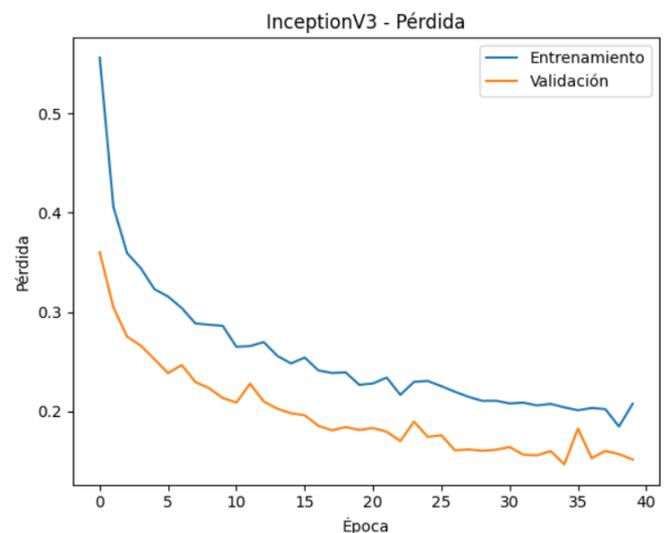


Fig. 13. Curva de pérdida (Loss) durante el entrenamiento de InceptionV3.

La matriz de confusión (Fig. 14) y el reporte de clasificación final confirman un alto rendimiento del modelo InceptionV3 en la detección automática de conjuntivitis a partir de imágenes oculares. Se alcanzó una precisión global del 95 % y un macro F1-score también de 0.95, lo cual evidencia un equilibrio sólido entre sensibilidad y especificidad.

- **Verdaderos Positivos (TP):** 838 (93.4 % de los casos de conjuntivitis fueron identificados correctamente). Este resultado refleja una alta sensibilidad para esta clase, lo cual es crucial para minimizar diagnósticos omitidos en escenarios clínicos.
- **Falsos Negativos (FN):** 59 (6.6 % de las imágenes con conjuntivitis fueron clasificadas erróneamente como normales). Aunque este valor es bajo, sigue siendo relevante, ya que podría implicar la omisión de tratamientos en pacientes reales si el sistema se emplea de manera autónoma.

- Falsos Positivos (FP): 19 (2.8 % de imágenes normales fueron clasificadas como conjuntivitis). Este bajo porcentaje indica un modelo conservador, con escasa propensión a emitir falsas alarmas, lo que mejora su confiabilidad como herramienta de apoyo.
- Verdaderos Negativos (TN): 678 (97.2 % de los ojos normales fueron correctamente identificados como tales), lo cual refuerza la especificidad del modelo en la clase sin infección.

Estos resultados ponen en evidencia la capacidad del modelo para generalizar eficazmente en nuevos datos. Su desempeño balanceado lo convierte en una opción viable para su uso en entornos clínicos, siempre que se considere su implementación como complemento a la evaluación médica tradicional.

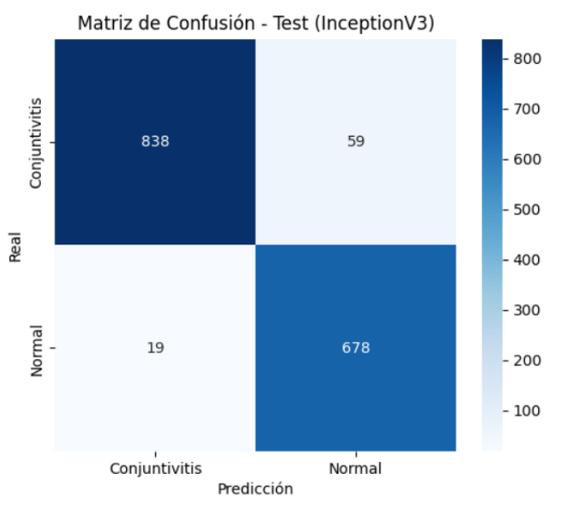


Fig. 14. Matriz de confusión.

#### D. RESULTADOS DEL MODELO FUSIÓN

El modelo Fusión fue entrenada utilizando aprendizaje transferido, alcanzando un total de 30 épocas antes de la activación del callback EarlyStopping, que monitoreaba la pérdida de validación (val\_loss). El tiempo total de entrenamiento fue de aproximadamente 8 hora y 25 minutos, empleando una GPU asignada automáticamente por Google Colaboratory (versión gratuita). Esta configuración permitió ejecutar los experimentos sin requerir infraestructura local dedicada, facilitando así un entorno accesible y reproducible para el desarrollo del modelo.

TABLA IV. RENDIMIENTO DE MODELO FUSION EN ENTRENAMIENTO, VALIDACIÓN Y TEST

Métrica	Train	Validation	Test
Exactitud (Accuracy)	91.57 %	94.03 %	94.79 %
Pérdida (Loss)	0.2095	0.1601	0.1405
Precisión (Precision)	0.96 (C), 0.91 (N)	0.95 (C), 0.92 (N)	0.97 (C), 0.92 (N)
Recall	0.94 (C), 0.93 (N)	0.94 (C), 0.94 (N)	0.94 (C), 0.96 (N)
F1-Score	0.95	94	0.95

Leyenda:

- **C:** *Conjuntivitis* (ojos con infección)
- **N:** *Normal* (ojos sin signos de infección)

La Fig. 12 muestra la evolución de la exactitud durante el entrenamiento del modelo fusión:

- **Train Accuracy:** La precisión durante el entrenamiento mostró una mejora sostenida, comenzando en 70.10 % en la época 1 y alcanzando un valor máximo de 91.57 % en la época 30. Esta progresión constante sugiere que el modelo fue aprendiendo de manera efectiva, sin evidencias claras de sobreajuste en las primeras etapas del entrenamiento.
- **Validation Accuracy:** La precisión sobre el conjunto de validación también presentó un crecimiento continuo, iniciando en 85.18 % y alcanzando un máximo de 94.03 % en la época 27. Este desempeño elevado evidencia una buena capacidad del modelo para generalizar a datos no vistos.
- **Brecha entre entrenamiento y validación:** La diferencia entre ambas curvas fue reducida y estable a lo largo del entrenamiento. En varias épocas, la precisión en validación superó ligeramente a la de entrenamiento, lo cual indica una buena capacidad de generalización y la ausencia de sobreajuste significativo.

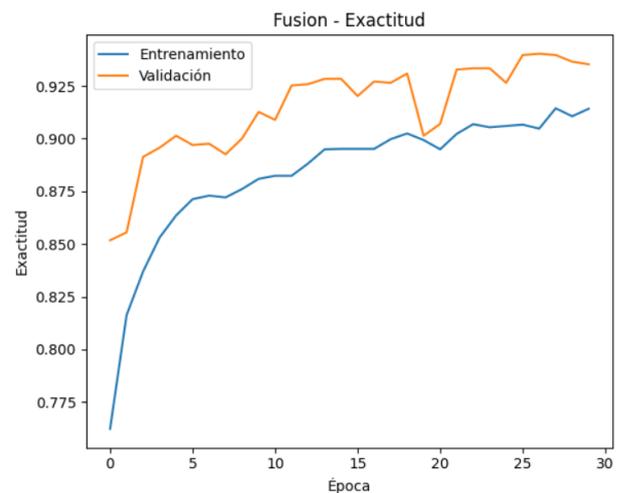


Fig. 15. Curva de exactitud (Accuracy) durante el entrenamiento del modelo fusión.

La Fig. 15 muestra la disminución de la pérdida (Loss):

- **Train Loss:** La pérdida durante el entrenamiento mostró una disminución constante, pasando de 0.5932 en la época 1 a un mínimo de 0.2031 en la época 27. Esta reducción progresiva indica que el modelo fue aprendiendo eficazmente a ajustar los datos, sin señales de estancamiento prematuro ni de sobreajuste en etapas iniciales.
- **Validation Loss:** La pérdida en validación también presentó una tendencia general descendente, comenzando en 0.3518 y alcanzando un mínimo de 0.1601 en la época 28. Aunque se observaron ligeras oscilaciones en algunas épocas, el comportamiento global fue positivo, reflejando una adecuada capacidad del modelo para generalizar sobre datos no vistos.

- Brecha entre entrenamiento y validación: La diferencia entre las curvas de pérdida se mantuvo reducida y estable durante todo el entrenamiento. En diversas etapas, la validation loss incluso se ubicó por debajo de la pérdida de entrenamiento, lo que sugiere que el modelo no solo evitó el sobreajuste, sino que también logró una generalización sólida. Esta consistencia valida el desempeño robusto del modelo de fusión frente a nuevos datos.

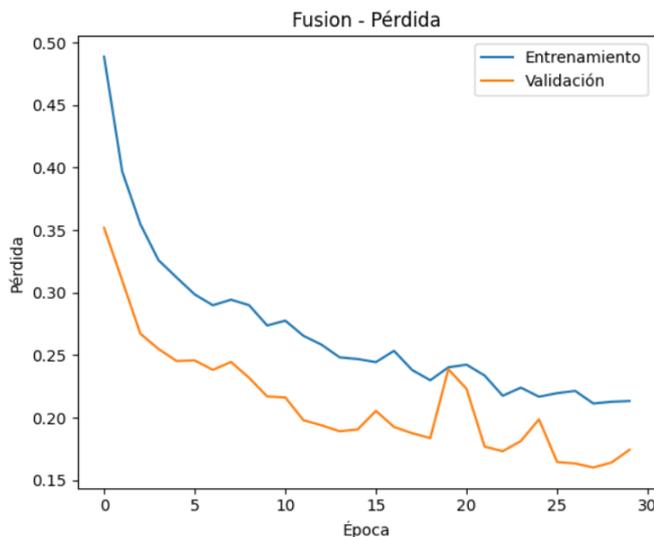


Fig. 16. Curva de pérdida (Loss) durante el entrenamiento del modelo fusión.

La matriz de confusión (Fig. 14) y el reporte de clasificación final evidencian un desempeño sólido del modelo de fusión en la detección de conjuntivitis a partir de imágenes oculares. El modelo alcanzó una precisión general del 95 %, con un F1-score macro promedio también de 0.95, lo que indica un equilibrio adecuado entre sensibilidad y especificidad en ambas clases.

- Verdaderos Positivos (TP): 840 casos de conjuntivitis fueron correctamente clasificados, lo que representa un recall del 94 % para esta clase. Este resultado demuestra una alta sensibilidad del modelo para identificar imágenes patológicas.
- Falsos Negativos (FN): 57 imágenes de conjuntivitis fueron erróneamente clasificadas como normales (6 %), lo cual puede representar una limitación en contextos clínicos, ya que existe riesgo de omitir diagnósticos relevantes si se depende exclusivamente del sistema automatizado.
- Falsos Positivos (FP): 26 imágenes normales fueron clasificadas como conjuntivitis (3.7 %), lo que implica una baja tasa de falsas alarmas. Esto es positivo en entornos médicos, ya que contribuye a reducir revisiones innecesarias sin comprometer la sensibilidad del modelo.
- Verdaderos Negativos (TN): 671 imágenes normales fueron correctamente identificadas, con una precisión del 92 % y un recall del 96 % para esta clase. Esto refleja una excelente capacidad para discriminar entre ojos sanos y afectados.

En conjunto, estos resultados sugieren que el modelo de fusión posee una alta capacidad de generalización, logrando un rendimiento equilibrado tanto en sensibilidad como en especificidad. No obstante, si se proyecta su uso en ambientes clínicos reales, la baja, pero presente tasa de error debe considerarse como parte de una estrategia de apoyo diagnóstico, y no como un reemplazo del juicio médico profesional.

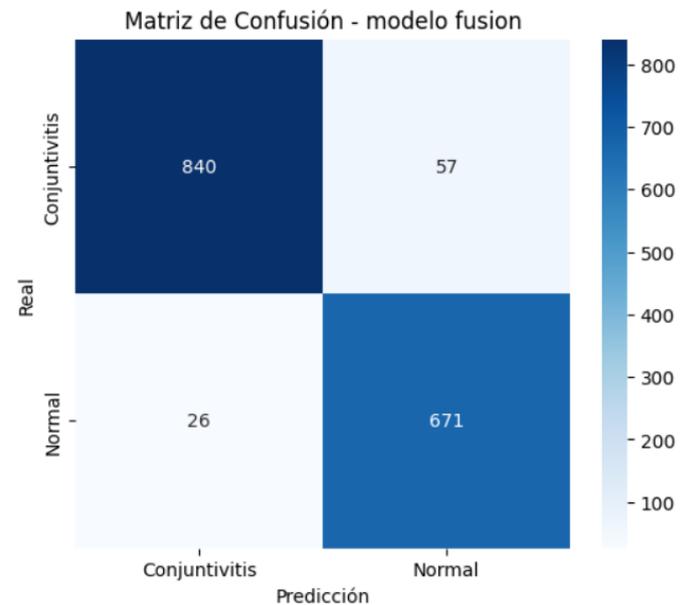


Fig. 17. Curva de pérdida (Loss) durante el entrenamiento del modelo fusión con el conjunto de test.

#### E. TABLA COMPARATIVA DE LOS MODELOS EN EL CONJUNTO DE TEST RESNET50, EFFICIENTNETB0, INCEPTIONV3 Y FUSIÓN.

Métricas	Inception V3	Efficient NetB0	ResNet 50	Fusión
Exactitud (Test)	95.00%	70.00 %	79.00 %	94.79
Precisión (C / N)	0.98/0.92	0.66/0.84	0.79/0.78	0.97/0.92
Recall (C / N)	0.93/0.97	0.95/0.38	0.84/0.72	0.92/0.96
F1-Score	0.95	0.65	0.79	95
Tiempo de Entrenamiento	10.7 h	9.7 h	10.7 h	8.25h
Falsos Negativos (FN)	16%	5.5 %	6.6%	6.0%
Falsos Positivos (FP)	28%	61.8 %	2.8%	3.7%

## 5 DISCUSIÓN

Los resultados de este estudio confirman que las arquitecturas CNN evaluadas son herramientas eficaces para la clasificación automatizada de imágenes oculares en el diagnóstico de conjuntivitis. Entre ellas, InceptionV3 demostró el rendimiento más alto, con una precisión del 95.11 % y un equilibrio notable entre sensibilidad (recall) y especificidad (precisión). Este comportamiento respalda su idoneidad en entornos clínicos, donde minimizar tanto falsos negativos como falsos positivos es esencial para reducir riesgos en el proceso diagnóstico.

Por su parte, ResNet50 ofreció un desempeño competitivo, combinando una precisión elevada y una baja tasa de falsas alarmas. Esto la convierte en una alternativa sólida en contextos donde se privilegie la confiabilidad del modelo y la simplicidad de implementación, dado que su entrenamiento fue más estable y su complejidad computacional moderada.

En contraste, EfficientNetB0, si bien es una arquitectura más ligera y eficiente en recursos, mostró un rendimiento inferior en sensibilidad y una tasa elevada de falsos positivos, lo que implica que su adopción en entornos clínicos requiere ajustes y calibraciones adicionales. No obstante, su menor tiempo de entrenamiento y su bajo consumo de memoria pueden ser valiosos en escenarios de hardware limitado.

El modelo de fusión propuesto, que integró exclusivamente las salidas intermedias de InceptionV3 y ResNet50, alcanzó resultados comparables a InceptionV3, con un F1-score de 0.95 y una precisión del 94.79 %. Esta aproximación de ensamblaje evidenció beneficios en términos de robustez y generalización, reduciendo la variabilidad entre clases y consolidando un equilibrio entre sensibilidad y especificidad. La capacidad de combinar dos arquitecturas complementarias sugiere que la fusión de modelos puede ser una estrategia efectiva para mejorar el rendimiento diagnóstico sin incrementar en exceso la complejidad computacional.

En conjunto, los hallazgos refuerzan la utilidad del aprendizaje transferido, el aumento de datos y la regularización como prácticas fundamentales para obtener modelos fiables a partir de conjuntos de datos médicos limitados. Asimismo, los resultados subrayan que, aunque las CNN presentan un alto potencial como soporte diagnóstico, es imprescindible contemplar su uso como herramienta complementaria y no como sustituto del criterio clínico experto.

## 6 CONCLUSIONES Y RECOMENDACIONES

Este estudio evaluó y comparó el rendimiento de tres modelos de redes neuronales convolucionales preentrenadas (EfficientNetB0, InceptionV3 y ResNet50) y un modelo de fusión que integra las salidas intermedias de InceptionV3 y ResNet50 para la detección automatizada de conjuntivitis a partir de imágenes oculares. La principal contribución de este trabajo radica en ofrecer un análisis comparativo sistemático, reproducible y orientado a la aplicabilidad clínica, utilizando un flujo de datos homogéneo, técnicas estandarizadas de preprocesamiento y métricas unificadas, lo que representa un enfoque poco abordado en estudios previos sobre diagnóstico ocular asistido por inteligencia artificial.

El modelo EfficientNetB0 se caracterizó por su ligereza y me-

nor tiempo de entrenamiento (9.7 horas), lo que lo convierte en una alternativa potencialmente viable para entornos con recursos computacionales limitados, como dispositivos móviles o clínicas rurales. Sin embargo, su desempeño fue inconsistente, con fluctuaciones significativas en la exactitud de validación y una tasa elevada de falsos positivos (61.8%), lo que compromete su fiabilidad para ser aplicado directamente en contextos clínicos sin un ajuste exhaustivo.

Por su parte, InceptionV3 demostró la mayor capacidad de detección, alcanzando un recall del 93% para la clase "conjuntivitis", lo que lo convierte en el modelo más sensible del estudio. Su estabilidad durante el entrenamiento fue notable, con una brecha reducida entre entrenamiento y validación, aunque presentó una mayor demanda computacional (aproximadamente 10.7 horas) y un volumen considerable de parámetros, que pueden dificultar su despliegue en hardware con limitaciones de memoria.

El modelo ResNet50 emergió como la opción más equilibrada. Logró una precisión del 95%, con una tasa de falsos positivos relativamente baja (3%) y una menor diferencia entre las métricas de entrenamiento y validación, lo cual evidencia una buena capacidad de generalización. Su tiempo de entrenamiento fue más eficiente que el de InceptionV3, lo que refuerza su viabilidad para ser implementado en entornos clínicos donde es fundamental mantener bajo el nivel de errores de clasificación sin incrementar excesivamente los requisitos computacionales.

El modelo de fusión InceptionV3 + ResNet50 demostró que la combinación de arquitecturas complementarias puede incrementar la robustez y estabilidad del sistema, alcanzando un F1-score de 0.95 y una precisión del 94.79%, con una tasa de falsos positivos moderada (3.7%). Este enfoque de ensamblaje evidencia que la integración de características profundas provenientes de diferentes redes puede ser una estrategia efectiva para aplicaciones de apoyo clínico, siempre que se optimicen los recursos de procesamiento.

En términos prácticos, ResNet50 se posiciona como la mejor opción para entornos médicos donde es necesario minimizar tanto falsos negativos como falsos positivos y contar con un tiempo de entrenamiento razonable. InceptionV3 es más adecuado en escenarios donde la sensibilidad diagnóstica sea prioritaria y se pueda tolerar un nivel levemente superior de falsas alarmas. EfficientNetB0 es una alternativa interesante en soluciones embebidas o sistemas con restricciones de hardware, aunque su implementación clínica requiere optimización adicional.

Como recomendaciones técnicas, se sugiere:

- Ajustar los umbrales de decisión (thresholds) en los modelos InceptionV3 y EfficientNetB0 a valores cercanos a 0.4, y a 0.45 en ResNet50 y el modelo de fusión, con el fin de maximizar la sensibilidad sin comprometer excesivamente la especificidad.
- Implementar funciones de pérdida como focal loss o estrategias de ponderación de clases para reducir los falsos negativos, especialmente en aplicaciones donde la detección temprana de conjuntivitis sea crítica.
- Incrementar la regularización mediante Dropout (entre 0.5 y 0.7), penalización L2 y la incorporación de

Batch Normalization en las capas densas para mejorar la capacidad de generalización.

- Experimentar con algoritmos de optimización más robustos, como AdamW, y estrategias de ajuste dinámico del learning rate, como cosine decay, con el objetivo de acelerar la convergencia y estabilizar el entrenamiento.

## REFERENCIAS

- [1] A. Y. Cheung, D. S. Choi, S. Ahmad, G. Amescua, V. Jhanji, A. Lin, S. I. Mian, M. K. Rhee, E. T. Viriya, F. S. Mah y D. M. Varu; American Academy of Ophthalmology Preferred Practice Pattern Cornea/External Disease Panel, "Conjunctivitis Preferred Practice Pattern," *Ophthalmology*, vol. 131, no. 4, pp. P134-P204, Apr. 2024. doi:10.1016/j.ophtha.2023.12.037.
- [2] M. Teweldemedhin, H. Gebreyesus, A. H. Atsbaha, S. W. Asgedom, and M. Saravanan, "Bacterial profile of ocular infections: A systematic review," *BMC Ophthalmol.*, vol. 17, no. 1, pp. 1-9, 2017, doi: 10.1186/s12886-017-0612-2.
- [3] D. S. W. Ting, C. Y. Cheung, G. Lim, N. M. Hsu, W. H. Tan, et al., "Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes," *JAMA*, vol. 318, no. 22, pp. 2211-2223, Dec. 2017, doi:10.1001/jama.2017.18152.
- [4] S. Mondal, S. Banerjee, S. Mukherjee, A. Ganguly, and D. Sengupta, "Deep Classifier for Conjunctivitis - A Three-Fold Binary Approach," *Int. J. Math. Sci. Comput.*, vol. 8, no. 2, pp. 46-54, 2022, doi: 10.5815/ijmsc.2022.02.05.
- [5] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, et al., "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs," *JAMA*, vol. 316, no. 22, pp. 2402-2410, Dec. 2016, doi:10.1001/jama.2016.17216.
- [6] S. Saha, S. K. Saha, S. Mukherjee, and S. Banerjee, "Deep learning for automated detection and classification of infectious conjunctivitis using slit-lamp images," *British Journal of Ophthalmology*, Published Online First: 28 Mar. 2024, doi:10.1136/bjo-2023-324310
- [7] C. Y. Gamage, "NEURAL NETWORKS FOR CLASSIFICATION OF EYE CONJUNCTIVITIS IN TELEHEALTH: A CONCEPTUAL ARCHITECTURE," vol. 6, no. 2, pp. 172-182, 2024.
- [8] F. Rashid, J. Abate, and A. Abdi, "Multiclass Classification and Identification of the External Eye Diseases using Deep CNN," *Indian J. Sci. Technol.*, vol. 17, no. 12, pp. 1107-1116, 2024, doi: 10.17485/ijst/v17i12.21.
- [9] R. K. Bawa and A. Koul, "Automated detection of conjunctivitis using convolutional neural network," *Appl. Data Sci. Smart Syst.*, pp. 91-97, 2024, doi: 10.1201/9781003471059-13.
- [10] C. J. Mann, "Observational research methods. Research design II: cohort, cross sectional, and case-control studies," *Emergency Medicine Journal*, vol. 20, no. 1, pp. 54-60, Jan. 2003, doi:10.1136/emj.20.1.54.
- [11] H. Kaur and N. K. Sandhu, "International Journal of Communication Networks and Information Security Evaluating the Effectiveness of the Proposed System Using F1 Score , Recall , Accuracy , Precision and Loss Metrics Compared to Prior Techniques," vol. 15, no. 04, pp. 368-383, 2023.
- [12] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 3rd ed., Sebastopol, CA, USA: O'Reilly Media, 2022. ISBN: 978-1-098-18109-1.
- [13] D. Callaghan, J. Burger, and A. K. Mishra, "A machine learning approach to radar sea clutter suppression," *2017 IEEE Radar Conf. RadarConf2017*, pp. 1222-1227, 2017, doi: 10.1109/RADAR.2017.7944391.
- [14] G. Research, "What is Colaboratory?" [Online]. Available: <https://research.google.com/colaboratory/faq.html>
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 770-778, 2016, doi: 10.1109/CVPR.2016.90.
- [16] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," *36th Int. Conf. Mach. Learn. ICML 2019*, vol. 2019-June, pp. 10691-10700, 2019.
- [17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 2818-2826, 2016, doi: 10.1109/CVPR.2016.308.
- [18] S. Patel, R. Patel, N. Ganatra, and A. Patel, "Spatial Feature Fusion for Biomedical Image Classification based on Ensemble Deep CNN and Transfer Learning," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 5, pp. 153-159, 2022, doi: 10.14569/IJACSA.2022.0130519.
- [19] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, no. 60, pp. 1-48, 2019, doi:10.1186/s40537-019-0197-0.