

Proyecto Final

Modelo de segmentación de imágenes para la localización de lesiones asociadas al cáncer de piel

Angel David Piñeros Sierra¹, Camilo Andres Roncancio Toca², Kelly Johana Solano Calderón³

¹²³ Facultad de Ingeniería, Universidad Nacional de Colombia, Sede Bogotá
{apineross, croncancio, ksolanoc}@unal.edu.co

Diciembre 2025

Resumen:

Este documento expone los resultados obtenidos en la creación de una red neuronal convolucional para la detección de hongos en hojas, la cual sirve como base para un proceso posterior de mejoramiento mediante *transfer learning* orientado a la detección de lesiones cutáneas.

Abstract:

This document presents the results obtained during the development of a convolutional neural network for leaf fungus detection, which serves as a baseline for future improvements using *transfer learning* aimed at identifying skin lesions.

Keywords: Lesiones cutáneas, Cáncer de piel, Detección, Segmentación, Transfer Learning, Fine-Tuning, Deep Learning

1. Introducción

El cáncer de piel es uno de los tipos de cáncer más comunes a nivel global. Entre sus tipos está el melanoma originado en los melanocitos, que son las células encargadas de producir melanina, el pigmento que otorga color a la piel [Mayo Clinic ORG, 2023]. En las últimas décadas, su incidencia ha aumentado de manera significativa, convirtiéndose en uno de los cánceres cutáneos más comunes y con mayor tasa de mortalidad, alcanzando aproximadamente 8,000 casos fatales por año [American Cancer Society, 2019]. Este tipo de cáncer suele formarse en zonas del cuerpo expuestas al sol, como los brazos, la espalda, la cara y las piernas; además, factores como la presencia de lunares atípicos, la piel clara, antecedentes familiares y sistemas inmuno-lógicos comprometidos incrementan considerablemente el riesgo de desarrollar la enfermedad.

Sin embargo, el panorama clínico de las lesiones cu-

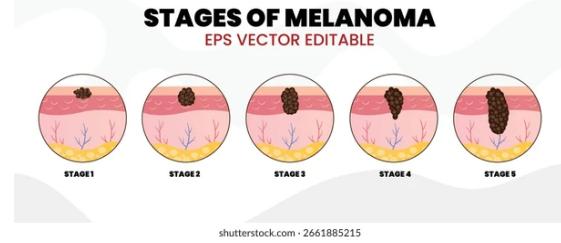


Figura 1: Etapas de formación del melanoma [Shutterstock, 2025].

táneas no se limita al melanoma. Existen múltiples afecciones dermatológicas que presentan variabilidad significativa en forma, color, textura y características estructurales. Entre ellas se incluyen queratosis actínicas, carcinomas basocelulares, lesiones vasculares, dermatofibromas y otras patologías pigmentadas. En la práctica clínica, la adecuada delimitación de estas lesiones es un paso fundamental tanto para la caracterización diag-

nóstica como para la selección del tratamiento. Debido a esta diversidad de casos, las soluciones computacionales actuales tienden a enfocarse no sólo en un tipo de cáncer, sino en la segmentación generalizada de lesiones cutáneas, lo que permite detectar patrones visuales relevantes independientemente de la patología específica.

Los avances recientes en Inteligencia Artificial han impulsado el desarrollo de modelos de visión por computadora capaces de apoyar el diagnóstico dermatológico mediante la segmentación automática de lesiones. Iniciativas como la International Skin Imaging Collaboration (ISIC) han permitido la creación de repositorios de más de 75.000 imágenes dermatoscópicas, facilitando el entrenamiento de modelos cada vez más precisos [Kopec, 2022]. La evidencia demuestra que las redes neuronales convolucionales (CNN) pueden alcanzar desempeños comparables con especialistas en tareas como clasificación y segmentación de lesiones pigmentadas [Melanoma Research Alliance, 2024]. Asimismo, soluciones comerciales como PathAI han reforzado la utilidad del aprendizaje profundo para la identificación de patrones visuales complejos en muestras clínicas [Path AI, 2025].

A pesar de su eficacia, el entrenamiento de estas arquitecturas requiere grandes cantidades de datos segmentados manualmente, lo cual es costoso y complejo. Para mitigar esta limitación, el aprendizaje por transferencia (*Transfer Learning*) se ha consolidado como una estrategia eficaz. Esta técnica permite reutilizar el conocimiento adquirido por un modelo entrenado en un dominio distinto, adaptándolo a tareas específicas mediante procesos de ajuste fino. En este sentido, el presente estudio no se limita exclusivamente al melanoma, sino que aborda la segmentación de múltiples clases de lesiones cutáneas, aprovechando la estructura del dataset HAM10000, que incluye siete categorías diagnósticas diferentes.

Tomando en cuenta lo anterior, este trabajo propone la adaptación de una arquitectura convolucional originalmente entrenada para la detección de hongos en hojas de café hacia la segmentación general de lesiones cutáneas. La hipótesis central es que los patrones texturales y estructurales aprendidos en el dominio agrícola pueden transferirse de manera efectiva al ámbito dermatológico, permitiendo que el modelo identifique los contornos y regiones de diferentes patologías cutáneas. El objetivo principal es la construcción de un modelo

apto para la correcta identificación de lesiones cutáneas a partir de un modelo preentrenado para un propósito similar pero con dominio diferente.

2. Metodología

La metodología implementada en este estudio se estructuró con el objetivo de evaluar la eficacia de un modelo de segmentación de lesiones cutáneas utilizando técnicas de aprendizaje profundo y, en particular, el enfoque de *Transfer Learning*. Para ello, se adaptó una arquitectura convolucional previamente entrenada para la identificación de hongos en hojas de café, reutilizando su capacidad de extracción de características texturales y estructurales. El procedimiento incluyó la selección y análisis del conjunto de datos, el preprocessamiento de las imágenes, la adaptación del modelo base, el entrenamiento supervisado y la evaluación cuantitativa y cualitativa del desempeño del modelo.

2.1. Descripción del conjunto de datos

El conjunto de datos utilizado corresponde al “*Skin Cancer: HAM10000*”, disponible en la plataforma Kaggle. Este dataset constituye una recopilación de 10,015 imágenes dermatoscópicas, acompañadas de sus respectivas máscaras de segmentación en formato PNG. Las imágenes fueron inicialmente publicadas en Harvard Dataverse con el fin de proporcionar una base de datos suficientemente amplia y diversa para tareas de diagnóstico automatizado de lesiones cutáneas pigmentadas [Propias, 2025]. Las carpetas incluidas en el dataset se dividen en *images* y *masks*, cada una con el mismo número de archivos, asegurando la correspondencia directa entre imágenes y máscaras.

El dataset incluye imágenes pertenecientes a siete diagnósticos distintos: queratosis actínicas, carcinoma intraepitelial, carcinoma basocelular, lesiones queratósicas, dermatofibroma, melanoma y lesiones vasculares. Esta diversidad permite la construcción de modelos de segmentación con mayor capacidad de generalización, debido a la variabilidad amplia en patrones visuales, texturas y estructuras anatómicas representadas. La disponibilidad de este conjunto de datos y su relevancia en competencias internacionales, como las organizadas por ISIC (International Skin Image Collaboration), motivan su selección para este estudio.

2.2. Preprocesamiento de los datos

Antes del entrenamiento del modelo, todas las imágenes y máscaras fueron sometidas a procesos estandarizados de preprocesamiento. Inicialmente, se realizó un redimensionamiento a un tamaño de 256×256 píxeles, siguiendo lineamientos metodológicos utilizados en estudios recientes de segmentación dermatoscópica [Dhar et al., 2024]. Posteriormente, se normalizaron los valores de los píxeles entre 0 y 1, con el objetivo de estabilizar el proceso de entrenamiento.

El conjunto de datos se dividió en subconjuntos de entrenamiento (70 %), validación (15 %) y prueba (15 %), manteniendo siempre la asociación entre cada imagen y su respectiva máscara. Con el fin de aumentar la capacidad de generalización del modelo y evitar el sobreajuste, se implementaron técnicas de aumento de datos, incluyendo rotaciones, traslaciones, escalamiento, inversiones horizontales y verticales, así como distorsiones no lineales, comunes en segmentación médica. Estas transformaciones permiten simular variaciones clínicas reales y mejorar la robustez del modelo.

2.3. Adaptación del modelo mediante Transfer Learning

El modelo utilizado como base fue una arquitectura convolucional entrenada previamente para la identificación de hongos en hojas de café. Aunque dicho dominio difiere del biomédico, la red original había aprendido patrones relevantes, tales como texturas complejas, irregularidades y bordes no uniformes, los cuales pueden ser reutilizados en la segmentación de lesiones cutáneas mediante técnicas de *Transfer Learning*.

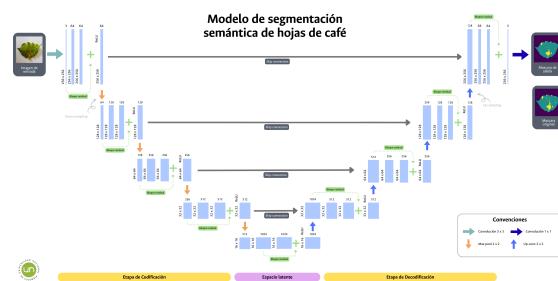


Figura 2: Arquitectura para la identificación de hongos en hojas de café (Autoría propia).

Para la adaptación, se conservaron los pesos del *encoder* y *decoder* original en todas sus capas, excepto la

final la cual cambia debido al número de clases consideradas. Este *decoder* incorpora conexiones de salto (*skip connections*) que permiten recuperar información espacial perdida, facilitando una segmentación precisa de los bordes de las lesiones. Finalmente, se habilitó el ajuste fino (*fine-tuning*) en todas las capas del modelo con el fin de optimizar el desempeño en el dominio dermatológico.

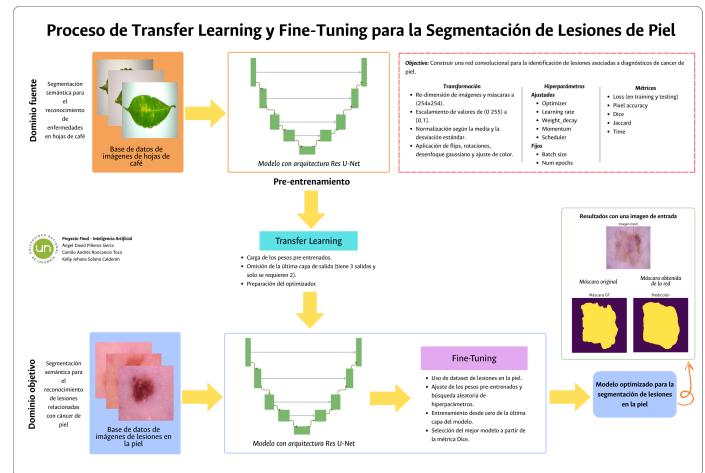


Figura 3: Esquema de aplicación de Transfer Learning y Fine-Tuning

2.4. Arquitectura del modelo propuesto

El modelo final adoptó una estructura de tipo *encoder-decoder* (ver figura 4). El diseño general se basa en arquitecturas exitosas en segmentación médica, como UNet y TransUNet, que combinan la capacidad de extracción local de características con la reconstrucción de mapas de segmentación de alta resolución [Dhar et al., 2024].

Aunque se revisaron opciones más avanzadas basadas en transformadores para segmentación (por ejemplo, TransUNet), el enfoque central de este estudio fue evaluar la capacidad del modelo transferido desde un dominio agrícola frente a modelos entrenados desde cero. Esta comparación permite establecer el valor real del *Transfer Learning* entre dominios visuales distintos.

2.5. Entrenamiento del modelo

El entrenamiento se realizó utilizando el optimizador Adam con tasas de aprendizaje iniciales evaluadas en el rango 10^{-3} – 10^{-5} . Se empleó la función de pérdida

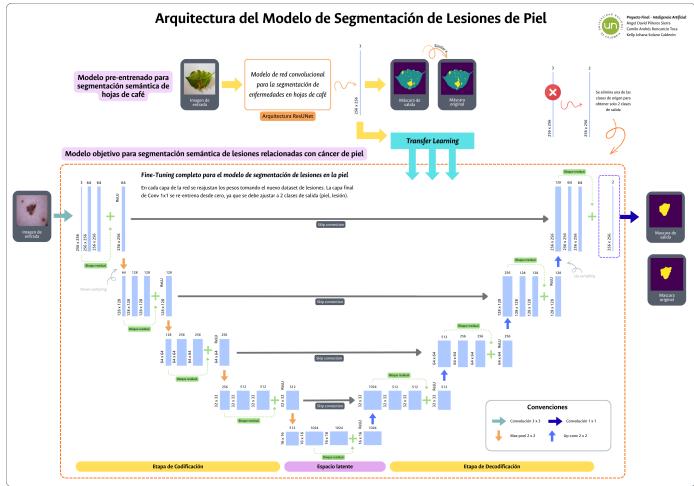


Figura 4: Arquitectura de red convolucional tipo ResU-Net para la segmentación de lesiones

da *Dice Loss*, ampliamente utilizada en segmentación biomédica debido a su capacidad para manejar clases desbalanceadas. El tamaño de lote utilizado fue de entre 8 y 16 imágenes, dependiendo de la configuración del modelo y las limitaciones de memoria.

Durante el entrenamiento, se monitorearon diversas métricas, entre ellas:

- **Coeficiente de Dice**, para evaluar la superposición entre predicciones y máscaras reales.
- **Intersección sobre Unión (IoU)**, una medida estándar en segmentación de imágenes.
- **Pérdida de entrenamiento y validación**, para identificar fenómenos de sobreajuste.

Estas métricas permiten comparar los resultados con estudios previos y evaluar la robustez del modelo.

2.6. Evaluación y validación

Finalmente, el modelo se evaluó utilizando el conjunto de prueba, aplicando tanto métricas cuantitativas como análisis cualitativo a través de visualizaciones de segmentación. Las predicciones generadas por el modelo fueron comparadas con sus correspondientes máscaras reales para analizar la precisión en la delimitación de bordes, la continuidad de las regiones segmentadas y la estabilidad del modelo en casos de alta complejidad clínica.

Con el fin de asegurar validez metodológica, se realizaron experimentos adicionales mediante variaciones en los hiperparámetros:

Este protocolo permitió establecer una evaluación robusta del desempeño general del modelo.

3. Resultados

Durante el proceso de evaluación del modelo se obtuvieron los siguientes resultados:

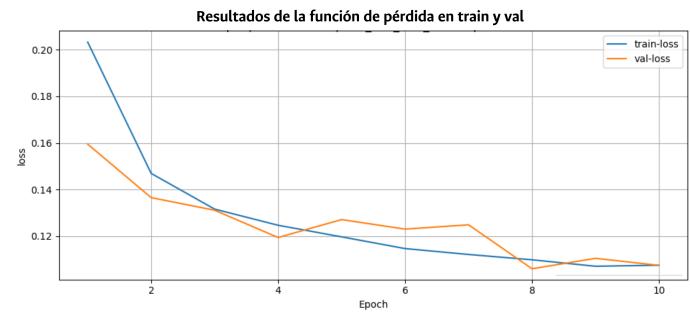


Figura 5: Resultado de la función de pérdida para Train y Val (10 épocas)

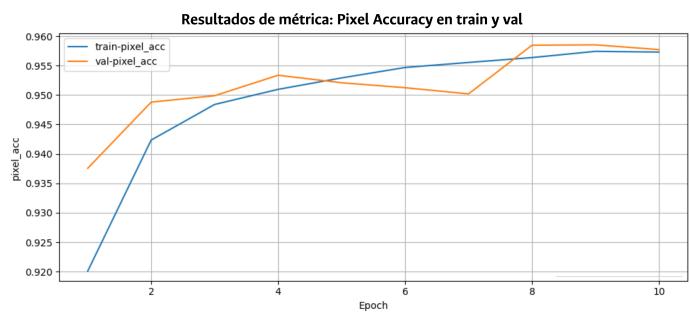


Figura 6: Resultado de la métrica de Pixel Accuracy para Train y Val (10 épocas)

Por un lado, el comportamiento de la función de pérdida (Fig. 5) muestra que tanto para el entrenamiento como en la validación, el descenso se realiza de forma estable, puesto que no se perciben oscilaciones o divergencias considerables a medida que aumentan las épocas. Como la diferencia entre la pérdida de train y val es pequeña ($\approx 0,0034$) se concluye que no se presentó sobreajuste y por tanto hubo una buena generalización. En el caso de Pixel Accuracy (Fig. 6) la clasificación de los pixeles es óptima, porque su convergencia en train y val se da en $\approx 0,9585$, no obstante, es importante comparar ese resultado con las métricas F1 para tener



Figura 7: Resultado de la métrica de Dice para Train y Val (10 épocas)

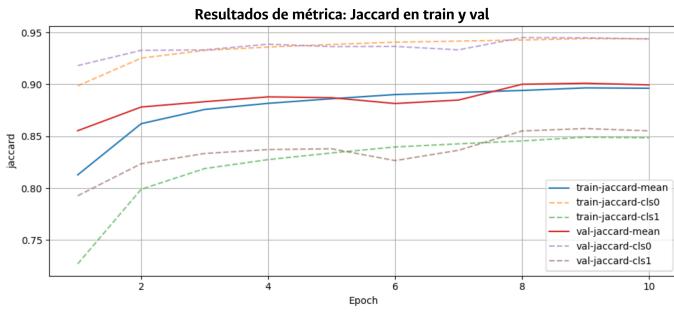


Figura 8: Resultado de la métrica de Jaccard para Train y Val (10 épocas)

un resultado más consistente.

En la métrica Dice (Fig. 7) se obtienen los resultados para la clase 0 (no lesión), la media y la clase 1 (lesión). La clase 0 es la que tiene valores más altos ya que las regiones de no lesión son más representativas, con esto, la segmentación de la piel es buena y estable. La clase 1 empieza con valores más bajos, subiendo progresivamente tanto en train como val. Como es la clase minoritaria se muestra un progreso de la segmentación, donde a mayores épocas se estabilizan los resultados.

En la métrica de Jaccard (Fig. 8) se utilizan esta misma división por clases y media, donde en la clase 0, los valores son altos pero muy similares (la diferencia es de $\approx 0,0045$) y en la clase 1 existe un progreso en el grado de solapamiento. Tanto Dice como Jaccard tienen una tendencia similar, rescatando el hecho de que los resultados de validación son mejores que los de entrenamiento. Esto puede deberse a que en la etapa de entrenamiento, la segmentación tuvo más dificultad al incluir imágenes transformadas (por el data augmentation) y al usar estadísticas por cada batch en el proceso de normalización del entrenamiento (en comparación

con las estadísticas acumuladas de validación).

Por otro lado, los resultados para la segmentación muestran qué las máscaras predichas (tercera columna) no representan de forma exacta la máscara original (segunda columna). Aunque la representación predicha difiere a lo esperado, sí está encerrando la mayoría de píxeles de la lesión, como un “vecindario” donde se encuentran los valores positivos verdaderos, en conjunto con otros píxeles que son falsos positivos. En este caso, el modelo sí tiene una capacidad de generalización, dado a qué se obtienen buenos resultados en validación como lo indican los valores de Dice y Jaccard, donde la superposición de las máscaras alcanza un valor de 0.9450 y 0.8969.

Métrica	Resultado
Test Loss	0.1108
Pixel Accuracy	0.9574
Dice	0.9450
Jaccard	0.8969
Dice Class 0	0.9712
Dice Class 1	0.9188
Jaccard Class 0	0.9439
Jaccard Class 1	0.8499

Cuadro 1: Resultados de las métricas del modelo con mejor índice de Dice

Algo a tener en cuenta es que, como las máscaras predichas toman valores como falsos positivos, el modelo está alineándose con una mayor sensibilidad para representar las lesiones. Este comportamiento, debe permitir a nivel médico tomar como punto de inicio de tratamiento el área identificada y descartar las zonas donde no exista el riesgo de lesión.

4. Conclusiones

El desarrollo de modelos de segmentación para apoyar el diagnóstico y tratamiento de enfermedades cutáneas, como son las lesiones relacionadas con cáncer de piel, favorece la identificación temprana de las áreas potencialmente comprometidas con la enfermedad. Los resultados obtenidos con el dataset de *HAM10000* demostraron que la aplicación de estrategias de tratamiento de datos pueden favorecer a una mejor generalización de los segmentos de las lesiones. Así mismo, que el uso de fine tuning completo permite que todas las carac-

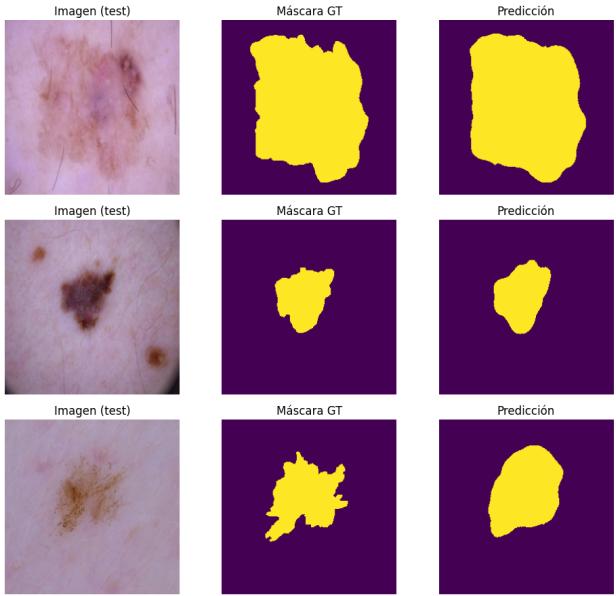


Figura 9: Evaluación sobre el conjunto de test y comparación entre la máscara real y la máscara predicha

terísticas del modelo pre-entrenado de hojas de café, puedan adaptarse mejor a la tarea específica de las lesiones, ahorrando recursos computacionales y tiempo de entrenamiento.

Cada una de las métricas obtuvo resultados aceptables tanto en entrenamiento como en validación, especialmente la función de pérdida, en donde se alcanzó un valor de $\approx 0,1108$, lo cual indica que no hubo sobreajuste en el modelo. El solapamiento de las máscaras, dado por Dice y Jaccard indica que, aunque los valores para la clase de lesión no fueron mayores al 92 %, la segmentación sí permitió rodear la lesión, al clasificar los píxeles de la zona afectada en conjunto con algunos píxeles en las fronteras, como clase 1. Es necesario tener en cuenta que, como la región identificada es más general a la esperada, a nivel médico deben realizarse las respectivas correcciones a las fronteras, para que la caracterización de la lesión (en etapas por ejemplo de clasificación) sea más específica y pueda orientar mejor en los procesos de tratamiento de la enfermedad.

5. Referencias

[American Cancer Society, 2019] American Cancer Society (2019). Factores de riesgo para el cáncer de piel tipo melanoma. <https://www.cancer.org/es/cancer/tipos/cancer-de-piel-tipo-melanoma/causas-riesgos-prevencion/factores-de-riesgo.html>. Recuperado: 2025-10-9.

[Dhar et al., 2024] Dhar, A., Prodhan, Y. A., Sikdar, D., and Joarder, S. (2024). Skin cancer semantic segmentation. Documento técnico utilizado como referencia metodológica.

[Kopec, 2022] Kopec, V. (2022). Tecnologías emergentes para los cánceres de piel de alto riesgo: herramientas para la toma de decisiones informada. <https://www.skincancer.org/es/blog/emerging-technologies-for-high-risk-skin-cancers-tools-for-informed-decision-making/>. Recuperado: 2025-10-9.

[Mayo Clinic ORG, 2023] Mayo Clinic ORG (2023). Melanoma. <https://www.mayoclinic.org/es/diseases-conditions/melanoma/symptoms-causes/syc-20374884>. Recuperado: 2025-10-9.

[Melanoma Research Alliance, 2024] Melanoma Research Alliance (2024). AI and melanoma: Innovation for the future. <https://www.curemelanoma.org/blog/ai-and-melanoma-innovation-for-the-future>. Recuperado: 2025-10-9.

[Path AI, 2025] Path AI (2025). About PathAI. <https://www.pathai.com/about-us>. Recuperado: 2025-10-9.

[Propias, 2025] Propias (2025). Notas internas del proyecto de segmentación de cáncer de piel. Documento interno.

[Shutterstock, 2025] Shutterstock (2025). Melanoma cancer infographic.