

# Título del Paper

Leismael Sosa Hernández

Claudia Puentes Hernández

David Cabrera García

Abdel Fregel Hernández

8 de julio de 2024

## Resumen

Este es el resumen del paper.

## 1. Introducción

El cáncer de piel representa uno de los problemas de salud pública más significativos a nivel mundial, con tasas de incidencia en aumento y consecuencias potencialmente devastadoras si no se detectan y tratan a tiempo. Este tipo de cáncer abarca diversas condiciones, desde formas relativamente benignas hasta tipos más agresivos como el melanoma y los carcinomas basocelular y espinocelular. La detección temprana juega un papel crucial en la mejora de las tasas de supervivencia y en la efectividad de los tratamientos disponibles. Identificar lesiones sospechosas en etapas iniciales no solo aumenta las opciones de tratamiento exitoso, sino que también reduce la carga emocional y económica asociada para los pacientes y el sistema de salud en general.

Solamente en los Estados Unidos, el cáncer de piel es la forma más común de cáncer, afectando a millones de personas cada año. Según el [Centro para el Control y Prevención de Enfermedades \(CDC\)](#)[1], se estima que más de 6.1 millones de adultos son diagnosticados anualmente con cáncer de células basales y de células escamosas, con un incremento constante en la incidencia debido a factores ambientales y cambios en el comportamiento social.

Entre los tipos más comunes de cáncer de piel se encuentran el melanoma (MEL), el carcinoma basocelular (BCC) y el carcinoma espinocelular. El melanoma, originado en los melanocitos, es uno de los cánceres de piel

más agresivos y potencialmente mortal. El carcinoma basocelular es el más común, aunque generalmente no se extiende más allá del sitio original. El carcinoma espinocelular, que aparece en las células escamosas de la epidermis, aunque menos común, presenta riesgos considerables debido a su capacidad de metastatizar si no se trata rápidamente.

Las técnicas tradicionales de diagnóstico del cáncer de piel, como las Redes Neuronales Convolucionales (CNNs) básicas, los modelos de Bag-of-Features y los basados en árboles de decisión, han sido utilizadas en la clasificación de imágenes dermatológicas [?, ?, ?]. Las CNNs básicas, ampliamente usadas por su capacidad para aprender características jerárquicas, pueden no capturar eficientemente todos los detalles relevantes de las lesiones cutáneas debido a su arquitectura relativamente simple. Los modelos de Bag-of-Features, que dependen de características locales extraídas manualmente y clasificadas con SVM o k-NN, pueden limitar la generalización en entornos complejos como la clasificación precisa de lesiones dermatológicas. Por otro lado, los modelos basados en árboles de decisión, aunque adaptados para clasificar imágenes dermatológicas mediante la extracción y reglas de clasificación, pueden tener dificultades con grandes volúmenes de datos y representaciones complejas [?].

En la actualidad, la clasificación de lesiones dermatológicas ha avanzado significativamente gracias al uso de técnicas avanzadas de aprendizaje automático. Entre los enfoques más destacados se encuentran el uso de redes neuronales convolucionales profundas, que han demostrado ser especialmente eficaces en la identificación y clasificación precisa de diversas condiciones cutáneas. Por ejemplo, estudios recientes han explorado el uso de modelos como ResNet para mejorar la precisión diagnóstica en el cáncer de piel [Gouda et al., 2020](#). La utilización de transfer learning se ha vuelto una dirección importante de investigación, permitiendo agilizar el proceso de entrenamiento [Kassem et al., 2020](#). Al igual que el uso de ensembles mejoran mucho la precisión del modelos llegando a un 98.6% [Kausar et al., 2021](#). Estos enfoques no solo mejoran la capacidad de diagnosticar melanomas y carcinomas, sino que también permiten una detección más temprana y precisa de condiciones menos comunes pero igualmente importantes.

El objetivo principal de este estudio es desarrollar un sistema avanzado para la clasificación de enfermedades dérmicas utilizando como conjunto de datos las imágenes proporcionadas por ISIC. También con este estudio buscamos analizar de forma experimental el impacto real que tiene el uso de modelos generativos en el entrenamiento de modelos de clasificación. Esto es

debido a que creemos que a medida que la distribución de datos aprendida por los modelos generativos se acerca más a la distribución real de los datos de entrenamiento esto debe tener un gran impacto en las métricas de los modelos de clasificación.

## 2. Estado del arte

El objetivo de esta sección es proporcionar una revisión exhaustiva de los avances recientes en el uso de técnicas de deep learning para la clasificación de enfermedades dérmicas, con un enfoque particular en el melanoma, carcinoma basocelular y carcinoma espinocelular. Esta revisión contextualiza el desarrollo del sistema propuesto dentro del marco actual de investigación y resalta las contribuciones más relevantes en el campo.

El diagnóstico de enfermedades dérmicas, como el melanoma y los carcinomas, se ha basado tradicionalmente en la evaluación clínica visual realizada por dermatólogos y la confirmación a través de biopsias. Aunque la experiencia del dermatólogo juega un papel crucial en la detección temprana y precisa, estos métodos no están exentos de limitaciones. La variabilidad en la experiencia y habilidad de los médicos puede llevar a diagnósticos inconsistentes, y las biopsias, aunque precisas, son procedimientos invasivos que pueden causar incomodidad al paciente. Además, el proceso de diagnóstico puede ser lento y costoso, lo que resalta la necesidad de métodos automáticos más eficientes y precisos.

El aprendizaje automático ha sido aplicado en la clasificación de cáncer de piel con el objetivo de superar las limitaciones de los métodos tradicionales. Técnicas como las Máquinas de Soporte Vectorial (SVM), los Árboles de Decisión y los Bosques Aleatorios han mostrado promesas en la clasificación de imágenes dérmicas. Sin embargo, estos métodos enfrentan desafíos significativos, como la necesidad de características manuales bien definidas y la dificultad para manejar grandes volúmenes de datos complejos. A pesar de sus logros iniciales, estos enfoques tradicionales a menudo no alcanzan la precisión y robustez necesarias para aplicaciones clínicas.

El uso de Redes Neuronales Convolucionales (CNNs) ha revolucionado la clasificación de imágenes médicas, incluyendo la detección de cáncer de piel. Las CNNs, capaces de aprender características directamente de los datos de imagen sin necesidad de intervención humana, han demostrado un rendimiento superior en múltiples estudios.

Por ejemplo, [Esteve et al. \(2017\)](#) desarrollaron un modelo basado en CNN que alcanzó un rendimiento comparable al de dermatólogos experimentados en la clasificación de melanoma. Otros estudios han explorado arquitecturas avanzadas como ResNet, Inception [Ahmet et al.\(2019\)](#), logrando mejoras significativas en la precisión y sensibilidad de los modelos. [Gouda et al. \(2020\)](#) utilizaron ResNet-34 para mejorar la precisión diagnóstica en el cáncer de piel, alcanzando una precisión del 83.31 % después de aplicar técnicas de reducción del valor RGB y redimensionamiento a 120x90 píxeles, junto con la segmentación y eliminación de ruido en las imágenes. La utilización de transfer learning se ha vuelto una dirección importante de investigación, permitiendo agilizar el entrenamiento, como lo demuestra el estudio de [Kassem et al. \(2020\)](#) que logró una precisión del 81 % usando un modelo pre-entrenado de GoogleNet con ImageNet realizando pequeños cambios en las capas finales del modelo.

El uso de ensembles de modelos de clasificación también ha mejorado significativamente la precisión de los modelos. [Kausar et al. \(2021\)](#) lograron una precisión del 98.6 %. Estos enfoques no solo mejoran la capacidad de diagnosticar melanomas y carcinomas, sino que también permiten una detección más temprana y precisa de condiciones menos comunes pero igualmente importantes.

Un estudio de [Tschandl et al. \(2018\)](#), utilizó EfficientNet en un enfoque de ensemble, logrando una precisión del 92.3 %. Además, el uso de U-Net para la segmentación de imágenes dermatológicas, como se muestra en el estudio de [Ronneberger et al. \(2015\)](#), ha demostrado ser altamente efectivo con métricas de segmentación como un IOU del 92.03 % y un DIC del 77.56 %.

En el estudio de [Goyal et al. \(2019\)](#), se utilizó el algoritmo Shades of Gray para la normalización de las imágenes y un modelo preentrenado llamado Faster-RCNN InceptionV2 combinado con DEXTR para la segmentación, logrando una precisión del 94.7 %, una sensibilidad del 91 % y una especificidad del 96.1 %. Estos resultados demuestran el potencial de los métodos avanzados de segmentación y ensemble en la mejora del diagnóstico del cáncer de piel.

## 2.1. Conjunto de datos

Para el desarrollo de este proyecto se seleccionó el dataset ISIC. Este conjunto de datos incluye una amplia variedad de imágenes de un total de 50,462 muestras de diferentes tipos de enfermedades de la piel. Cada imagen

está etiquetada con el tipo de enfermedad correspondiente. Las principales enfermedades a clasificar son: melanoma, con 6,987 imágenes; carcinoma basocelular, con 4,563 imágenes y carcinoma espinocelular, con 1,301 imágenes. Además, el dataset contiene imágenes de otras condiciones de la piel, tanto malignas como no malignas.

El dataset ISIC es fundamental utilizado en la investigación dermatológica y aplicaciones médicas, especialmente en el entrenamiento de modelos de clasificación de imágenes de la piel. Se utiliza para desarrollar sistemas de diagnóstico asistido por computadora, mejorar la detección temprana del cáncer de piel y evaluar algoritmos de segmentación de imágenes. Además, permite analizar y mitigar sesgos en los algoritmos de visión por computadora, generar imágenes sintéticas con GANs y capacitar a estudiantes y profesionales en el análisis de imágenes dermatológicas. Su uso facilita la comparación de modelos y algoritmos, mejorando el diagnóstico y tratamiento de enfermedades de la piel.

AQUI VAN LAS FOTOS

El desbalance de datos en el dataset ISIC presenta un desafío significativo para la clasificación precisa de cáncer de piel. Con la inmensa mayoría de imágenes en la categoría **Otros** y una representación mínima en clases como "Carcinoma Espinocelular", existe el riesgo de que el modelo se sesgue hacia las clases más frecuentes, afectando la precisión en la identificación de condiciones menos comunes. Para abordar esto, implementamos oversampling en las clases minoritarias, generando imágenes sintéticas con técnicas basadas en GANs. Este enfoque equilibra el dataset, mejorando la capacidad del modelo para detectar y clasificar con precisión una variedad de condiciones dermatológicas.

### 3. Metodología

Profesores, por cuestiones de tiempo nosotros no pudimos terminar nuestra implementación del proyecto pero en esta sección explicaremos la metodología que buscamos seguir.

En los proximos TODOs esta paso a paso lo que buscamos hacer.

**Preprocesamiento de las imágenes:** Las imágenes que componen nuestro Dataset tienen varias peculiaridades que dificulta su uso directo para el entrenamiento de modelos de Machine Learning. Algunos de los problemas que tenemos son los siguientes

- Nuestro dataset se encuentra bastante desbalanceado entre las clases a clasificar y por esto tenemos que utilizar técnicas de balanceo de dato. La que usamos es Oversampling.
- Algunas de nuestras imágenes tienen presentes grandes cantidades de pelo que según algunos artículos en la literatura provocan un detrimento del rendimiento de los modelos de clasificación. Es por esto que nosotros tenemos que aplicar técnicas de eliminación de pelos. Actualmente la que tenemos implementada es Dull Razor (1997). La elección de este algoritmo es primeramente su velocidad de procesamiento y segundo que era fácil de obtener acceso a implementaciones publicas. Sin embargo su rendimiento no es muy bueno en los casos en que las imágenes tienen mucho pelo o el pelo presente es muy grueso. De aquí que nosotros queremos emplear alguna técnica mas moderna para esta tarea, pero el tiempo no lo ha permitido.
- Otro problema presente es que a pesar de que todas las imágenes son de alta resolución (aproximadamente 4000x4000 pixeles), algunas de estas imágenes tienen la lesión en un pedazo relativamente pequeño, esto provoca que cuando se re-escale la imagen a las resoluciones requeridas por los modelos de clasificación (con la forma estándar de re-escalado), estas contengan poca información sobre la lesion. La forma en que nosotros queremos solventar este problema es entrenando primeramente un algoritmo de segmentación para luego con la información de la mascara que predice de la lesion, poder re-escalar la imagen de forma mas inteligente (haciendo un corte que bordeee la lesion). Nuevamente no llegamos a lograr esto por cuestiones de tiempo aunque el modelo lo tenemos implementado (UNET).
- Uno de los experimentos que tenemos pensado hacer desde hace poco tiempo para atacar el problema de re-escalado que se menciona en el punto anterior, es utilizar el algoritmo **Seam Carving** y comparar el rendimiento de los clasificadores usando este algoritmo como método para re-escalar la imagen vs los métodos usuales.
- Ultimo punto que también queríamos atacar era el de normalización del espacio de colores de las imágenes. Lo usual es tomar la media y la desviación estándar de la coloración de cada canal de la imagen y normalizar la imagen usando estos parámetros. Pero esto no lleva realmente todas las imágenes al mismo espacio de color y nosotros estábamos investigando sobre la posible utilización de técnicas de **Stain Normalization**. Pero todavía no estamos seguros si utilizarlo o no (por tiempo).

**Actualmente en la parte de pre-procesado lo que tenemos es:**

- Balanceo del dataset utilizando la técnica de **Oversampling**
- Luego aplicamos un re-escalado básico de la imagen a 1000x1000 pixeles y aplicamos el algoritmo de DullRazor
- También aplicamos técnicas de aumentaron automática de los datasets aplicando rotaciones, escalado y otras transformaciones de forma aleatoria.

### **Entrenamiento de Modelos Clasificadores**

Luego de que tenemos las imágenes pre-procesadas el siguiente paso en el pipeline es entrenar varios modelos de clasificación, en específico, ResNet, VGG y EfficientNet. El entrenamiento se haría utilizando Transfer learning con los pesos de estos modelos al ser entrenados en **ImageNet**. Los modelos se entrenaran durante 10 Epochs, guardándose por cada Epoch la perdida en el entrenamiento, validación y test. Luego se guardarán las métricas estándar de cada modelo al finalizar su entrenamiento. Estos resultados idealmente quisiéramos tenerlos al entrenar los modelos sin remover pelos y removiendo pelos para poder analizar el impacto de este preprocesamiento en los clasificadores.

### **Integración de Modelos Generativos**

Como uno de los objetivos centrales de nuestra investigación es analizar si realmente el uso de modelos generativos mejora las métricas en el entrenamiento de modelos de clasificación. Entonces en nuestro proyecto queríamos inicialmente tener mas de 1 modelo generativo implementado para poder hacer un estudio experimental más completo. Pero debido nuevamente a cuestiones de tiempo solamente tenemos Conditional GANs implementado.



**¿Cuál es la idea para entrenar los modelos generativos y evaluar su impacto en los modelos de clasificación?**

1. Entrenamiento del Conditional GAN: Se entrena un modelo generativo Conditional GAN (CGAN) para generar imágenes de lesiones cutáneas condicionadas a una categoría específica (e.g., Melanoma, Carcinoma Basocelular, etc.). El entrenamiento del CGAN se detiene inicialmente cuando la pérdida del generador baja de un umbral específico (como no hemos empezado el entrenamiento de este modelo no tenemos mucha idea de exactamente cual seria el umbral). La idea es que este umbral permita que las imágenes generadas tengan una calidad aceptable.

2. Luego de detenerse el entrenamiento del modelo generativo se conforma un nuevo dataset: Este nuevo dataset incluye tanto imágenes reales como imágenes generadas por el CGAN. La proporción de imágenes generadas se ajusta para mantener un balance adecuado entre datos reales y sintéticos. **Requiere experimentación**

3. Luego con este nuevo dataset se vuelven a entrenar los clasificadores con los pesos iniciales con los que se entrenaron sin el modelo generativo involucrado (usando los pesos de entrenamiento de ImageNet). Se vuelven a guardar las métricas de estos modelos y su historial de como fue variando su función de pérdida con el paso de los Epochs. (Con esto se harán las comparaciones).

4. **Iteración de mejora del CGAN:** Luego se continúa el entrenamiento del CGAN por 5 epochs y se vuelve al paso 2. Este ciclo de mejorar el GAN y reentrenar los clasificadores se repite 10 veces, con el objetivo de incrementar gradualmente la calidad de las imágenes generadas y evaluar su impacto en el rendimiento de los clasificadores. 5. Luego se comparan los resultados.

## 4. Resultados

Discusión de los resultados.

## 5. Conclusión

Conclusiones del estudio.

Asegurarse de que todos los autores han sido mencionados.

Revisar este resumen para añadir más detalles.

## Referencias

- [1] Nombre del Autor, *Título del libro o artículo*, Editorial, Año.
- [2] Nombre del Autor, *Título del libro o artículo*, Editorial, Año.