



UNIVERSIDAD ESAN
FACULTAD DE INGENIERÍA
INGENIERÍA DE TECNOLOGÍAS DE INFORMACIÓN Y SISTEMAS

Segmentación Avanzada de Características Morfológicas en la Piel del Rostro usando Redes
Neuronales Convolucionales: Enfoque en Arrugas y Manchas

Tesis para optar por el Bachiller de Ingeniería de Tecnologías de Información y Sistemas

Victor Manuel Ormeño Salazar
Jeferson Joseph Sandoval Díaz
Asesor: Junior John Fabián Arteaga

Lima, 10 de abril de 2025

Esta tesis denominada:

Segmentación Avanzada de Características Morfológicas en la Piel del Rostro usando Redes
Neuronales Convolucionales: Enfoque en Arrugas y Manchas

.....
Jurado 1

.....
Jurado 2

Universidad ESAN
2025

Segmentación Avanzada de Características Morfológicas en la Piel del Rostro usando Redes
Neuronales Convolucionales: Enfoque en Arrugas y Manchas

Dedicatoria

Thx.

Índice general

Resumen	1
Introducción	3
Capítulo I: Planteamiento del Problema	4
1.1 Descripción de la Realidad Problemática	4
1.2 Formulación del Problema	8
1.2.1 Problema General	8
1.2.2 Problemas Específicos	8
1.3 Objetivos de la Investigación	8
1.3.1 Objetivo General	8
1.3.2 Objetivos Específicos	8
1.4 Justificación de la Investigación	9
1.4.1 Teórica	9
1.4.2 Práctica	10
1.4.3 Metodológica	11
1.5 Delimitación del Estudio	11
1.5.1 Espacial	12
1.5.2 Temporal	12
1.5.3 Conceptual	13
Capítulo II: Marco Teórico	14
2.1 Antecedentes de la investigación	14
2.1.1 «Deep-Learning-Based Morphological Feature Segmentation for Facial Skin Image Analysis»	14
2.1.2 «High Performing Facial Skin Problem Diagnosis with Enhanced Mask R-CNN and Super Resolution GAN»	15
2.1.3 «Facial Wrinkle Detection Based on DeepLabV3+ and Semi-Automatic Labelling Strategy»	17
2.1.4 «Improved U-Net: Fully Convolutional Network Model for Skin-Lesion Segmentation»	20
2.1.5 «A New Approach to Facial Skin Analyzer»	22
2.1.6 «Dermatology: Pretraining with Texture Map-Based Weak Supervision»	23
2.2 Bases Teóricas	27
2.2.1 Inteligencia Artificial	27
2.2.2 Aprendizaje Automático	27

2.2.3	Aprendizaje Profundo	30
2.2.4	Segmentación de Imágenes	31
2.2.5	Redes Neuronales Convolucionales (CNN)	34
2.2.6	Redes de Atención	36
2.2.7	Modelos Avanzados de Segmentación en Imágenes Médicas	39
2.3	Marco Conceptual	41
2.3.1	Redes Neuronales Convolucionales (CNN)	41
2.3.2	Segmentación de Imágenes	41
2.3.3	Características Morfológicas del rostro	41
2.3.4	Modelos de Segmentación Utilizados	43
2.4	Hipótesis	44
2.4.1	Hipótesis General	44
2.4.2	Hipótesis Específicas	44
Capítulo III:	Metodología de la Investigación	45
3.1	Diseño de la investigación	45
3.1.1	Alcance de la investigación	45
3.1.2	Enfoque de la investigación	45
3.1.3	Población	46
3.1.4	Muestra	46
3.2	Metodología de Implementación de la Solución	46
3.2.1	Adquisición y Preparación de los Datos	46
3.3	Diagrama de la Metodología	47
3.3.1	Desarrollo de los Modelos de Segmentación	48
3.3.2	Evaluación y Validación del Sistema	49
3.3.3	Implementación y Pruebas Finales	49
3.4	Metodología para la Medición de Resultados de la Implementación	49
3.5	Cronograma de actividades y presupuesto	51
Capítulo IV:	Desarrollo de la Solución	53
4.1	Determinación y evaluación de alternativas de solución	53
4.2	Propuesta solución	53
4.2.1	Planeamiento y descripción de Actividades	53
4.2.2	Desarrollo de actividades	53
4.3	Medición de la solución	53
4.3.1	Análisis de Indicadores cuantitativo	53
Capítulo V:	Conclusiones y Recomendaciones	54

Referencias	57
Anexos	62
A Árbol de Problemas	63
B Árbol de Objetivos	64
C Matriz de Consistencia	65

Índice de Figuras

Figura 1.	Mercado mundial de productos para el cuidado de la piel: tendencias de la industria y pronóstico hasta 2031	4
Figura 2.	Ingresos del mercado mundial del cuidado de la piel de 2020 a 2030 (en millones de dólares estadounidenses)	5
Figura 3.	Ingresos del mercado del cuidado de la piel en Estados Unidos de 2019 a 2030 (en millones de dólares estadounidenses)	6
Figura 4.	Infografía de la Proyección del Mercado Cosmético e Higiene Personal en Perú	7
Figura 5.	Comparación de la segmentación de arrugas en (a) la frente y (b) la región de los ojos	16
Figura 6.	Comparación del modelo propuesto con otros 6 modelos de redes neuronales	18
Figura 7.	Interfaz de usuario del sistema de diagnóstico de problemas de la piel facial	19
Figura 8.	Estructura de red de DeepLabV3+	20
Figura 9.	El proceso de convolución separable en profundidad	20
Figura 10.	Diagrama de flujo del sistema propuesto	21
Figura 11.	Modelo totalmente convolucional propuesto para la segmentación de lesiones cutáneas	22
Figura 12.	Detección de edad y género	24
Figura 13.	Análisis del tono de piel	24
Figura 14.	Análisis de arrugas	25
Figura 15.	Entrenamiento en dos etapas para la segmentación de arrugas faciales . . .	26
Figura 16.	El algoritmo de K medias	29
Figura 17.	Relación entre IA, ML y DL	30
Figura 18.	Modelo de aprendizaje profundo	31
Figura 19.	Arquitectura de un modelo CNN	35
Figura 20.	Diagrama de la Metodología	47
Figura 21.	Cronograma de actividades	51

Índice de Tablas

Tabla 1.	Métricas de rendimiento de diferentes modelos	15
Tabla 2.	Presupuesto	52

Índice de Ecuaciones

Índice de Algoritmos

Resumen

La industria cosmética y de cuidado de la piel ha experimentado un crecimiento acelerado en las últimas décadas. Sin embargo, a pesar de la diversidad de productos y tratamientos disponibles, persisten desafíos importantes en la evaluación precisa y personalizada de problemas estéticos, como arrugas, poros dilatados y manchas faciales. Uno de los mayores obstáculos es la falta de herramientas tecnológicas avanzadas que permitan un análisis profundo y objetivo de las características morfológicas de la piel.

Este trabajo de investigación busca desarrollar un sistema de segmentación basado en redes neuronales convolucionales (CNN) para detectar y evaluar características morfológicas clave de la piel facial, como arrugas, poros y manchas. El sistema propuesto mejorará la precisión en la evaluación estética y la personalización de tratamientos cosméticos. Se utilizarán arquitecturas de CNN adaptadas para segmentar con precisión las imágenes de la piel facial, basándose en un conjunto de datos representativo que permita entrenar y validar el modelo. Las métricas de desempeño, como la precisión, el recall y el F1-score, serán empleadas para evaluar la eficiencia del sistema.

Este avance contribuirá tanto al ámbito de la dermatología computacional como al de la cosmética, optimizando los procesos de diagnóstico y recomendación de productos personalizados. La implementación de este sistema permitirá mejorar la precisión de los diagnósticos y personalizar tratamientos de manera más efectiva, beneficiando a consumidores que buscan soluciones específicas para sus necesidades estéticas.

Palabras clave: Segmentación, redes neuronales convolucionales, arrugas, poros, manchas, imagen de la piel, cosmética, tratamiento personalizado.

Abstract

The cosmetics and skincare industry has seen rapid growth in recent decades. However, despite the variety of products and treatments available, there remain significant challenges in the precise and personalized evaluation of aesthetic issues, such as wrinkles, enlarged pores, and skin spots. One of the major obstacles is the lack of advanced technological tools that allow for a deep and objective analysis of the skin's morphological features.

This research aims to develop a segmentation system based on Convolutional Neural Networks (CNN) to detect and evaluate key morphological features of facial skin, such as wrinkles, pores, and spots. The proposed system will enhance the accuracy of aesthetic evaluations and the personalization of cosmetic treatments. CNN architectures will be used to segment facial skin images accurately, using a representative dataset to train and validate the model. Performance metrics such as accuracy, recall, and F1-score will be used to assess the system's efficiency.

This advancement will contribute to both computational dermatology and cosmetics by optimizing diagnostic processes and personalized product recommendations. The implementation of this system will improve diagnostic precision and allow for more effective treatment personalization, benefiting consumers seeking solutions tailored to their specific aesthetic needs.

Keywords: Segmentation, convolutional neural networks, wrinkles, pores, spots, skin image, cosmetics, personalized treatment.

Introducción

En resumen, el principal objetivo de la presente investigación es el siguiente: desarrollar un sistema avanzado de segmentación de características morfológicas de la piel facial, es decir, las arrugas, los poros y las manchas, utilizando redes neuronales convolucionales (CNN). En otras palabras, este sistema tiene como objetivo mejorar la precisión de la valoración estética y, por lo tanto, fortalecer la personalización del tratamiento cosmético y favorecer el desarrollo de la industria de la piel. El sector cosmético se ha desarrollado de forma sostenida en las últimas décadas, a medida de una creciente preocupación en torno a problemas estéticos y el interés de mantener la piel sana y joven. No obstante, la mayor parte del diagnóstico de enfermedades cutáneas se basa en la observación manual o utilizando herramientas tradicionales, que son a menudo inexactas y no objetivas. Por lo tanto, se necesita urgentemente el desarrollo de nuevas tecnologías para evaluar con mayor detalle y precisión las características de la piel. En este contexto, las CNN han demostrado ser altamente efectivas para tareas de segmentación y análisis de imágenes, permitiendo identificar patrones complejos en datos visuales. Este estudio se enfoca en aplicar estas técnicas para desarrollar un modelo capaz de detectar y segmentar características clave de la piel, como arrugas, poros dilatados y manchas faciales, con el fin de brindar un análisis objetivo y detallado. La investigación sigue un enfoque metodológico que incluye la recopilación de datos, el desarrollo y prueba de modelos basados en CNN, y la evaluación de su desempeño mediante métricas como precisión, recall, F1-score y AUC-ROC. Este enfoque asegura la validación y confiabilidad del sistema desarrollado, destacando su aplicabilidad en el sector cosmético.

Capítulo I: Planteamiento del Problema

1.1 Descripción de la Realidad Problemática

La industria global de cuidado de la piel ha experimentado un crecimiento sostenido en los últimos años, como se puede ver en la Figura 1. En 2023, se estimó que el mercado mundial generó ingresos de aproximadamente 181.2 mil millones de dólares estadounidenses, con proyecciones que indican un aumento a más de 210 mil millones para 2028 . Este crecimiento refleja una tasa compuesta anual significativa, impulsada por la creciente conciencia sobre la importancia del cuidado de la piel y la demanda de productos innovadores. (Data Bridge Market Research, 2024)



Figura 1. Mercado mundial de productos para el cuidado de la piel: tendencias de la industria y pronóstico hasta 2031.

Fuente: Data Bridge Market Research (2024). «Mercado mundial de productos para el cuidado de la piel: tendencias de la industria y pronóstico hasta 2031».

Dentro de este mercado, el segmento de cuidado facial representa la mayor proporción de ingresos, como se puede ver en la Figura 2 alcanzando aproximadamente 108.94 mil millones de dólares en 2023 . Este segmento incluye productos como cremas hidratantes, sueros antiarrugas y tratamientos despigmentantes, dirigidos a abordar preocupaciones estéticas como arrugas dilatadas y manchas faciales. (Statista, 2023a)

Estados Unidos se posiciona como el mercado más lucrativo en la industria del cuidado

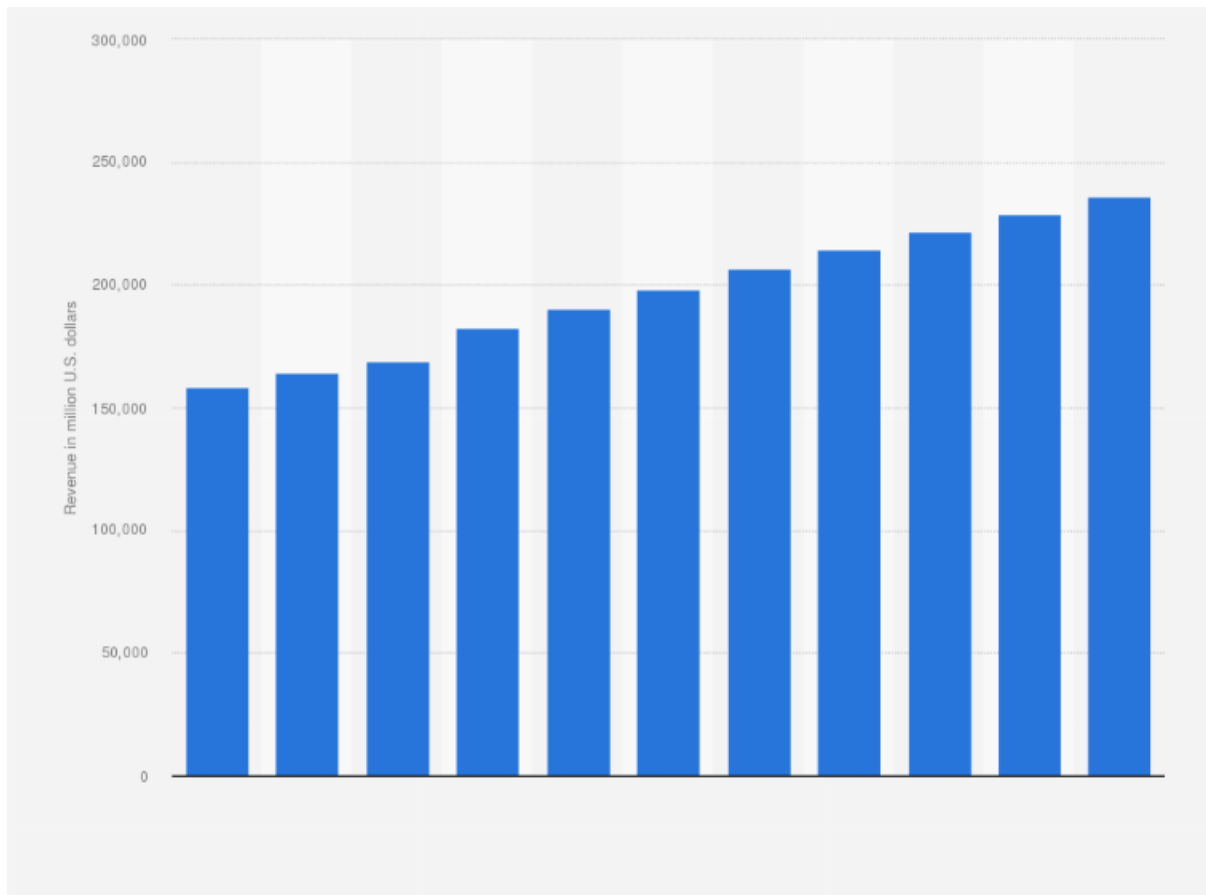


Figura 2. Ingresos del mercado mundial del cuidado de la piel de 2020 a 2030 (en millones de dólares estadounidenses).

Fuente: Statista (2023e). «Revenue in the Skin Care market worldwide».

de la piel. En 2023, el mercado estadounidense generó ingresos cercanos a 24 mil millones de dólares, superando a otros países como Japón y China . Se proyecta que para 2025, los ingresos alcancen los 26.01 mil millones de dólares, con una tasa de crecimiento anual compuesta del 4.53 % entre 2025 y 2029. (Statista, 2023f)

El segmento de cuidado facial domina el mercado estadounidense, lo podemos ver en la Figura 3, representando la mayor parte de las ventas. En 2023, las ventas de productos para el cuidado de la piel en Estados Unidos superaron los 23.5 mil millones de dólares, con el segmento facial contribuyendo significativamente a esta cifra. (Statista, 2023d)

El mercado peruano de belleza y cuidado personal también ha mostrado un crecimiento notable. En 2023, se estimó que el mercado generó ingresos de aproximadamente 2.94 mil millones de dólares, con una tasa de crecimiento anual compuesta proyectada del 4.60 % . Dentro de este mercado, el segmento de cuidado personal, que incluye productos para el cuidado de la

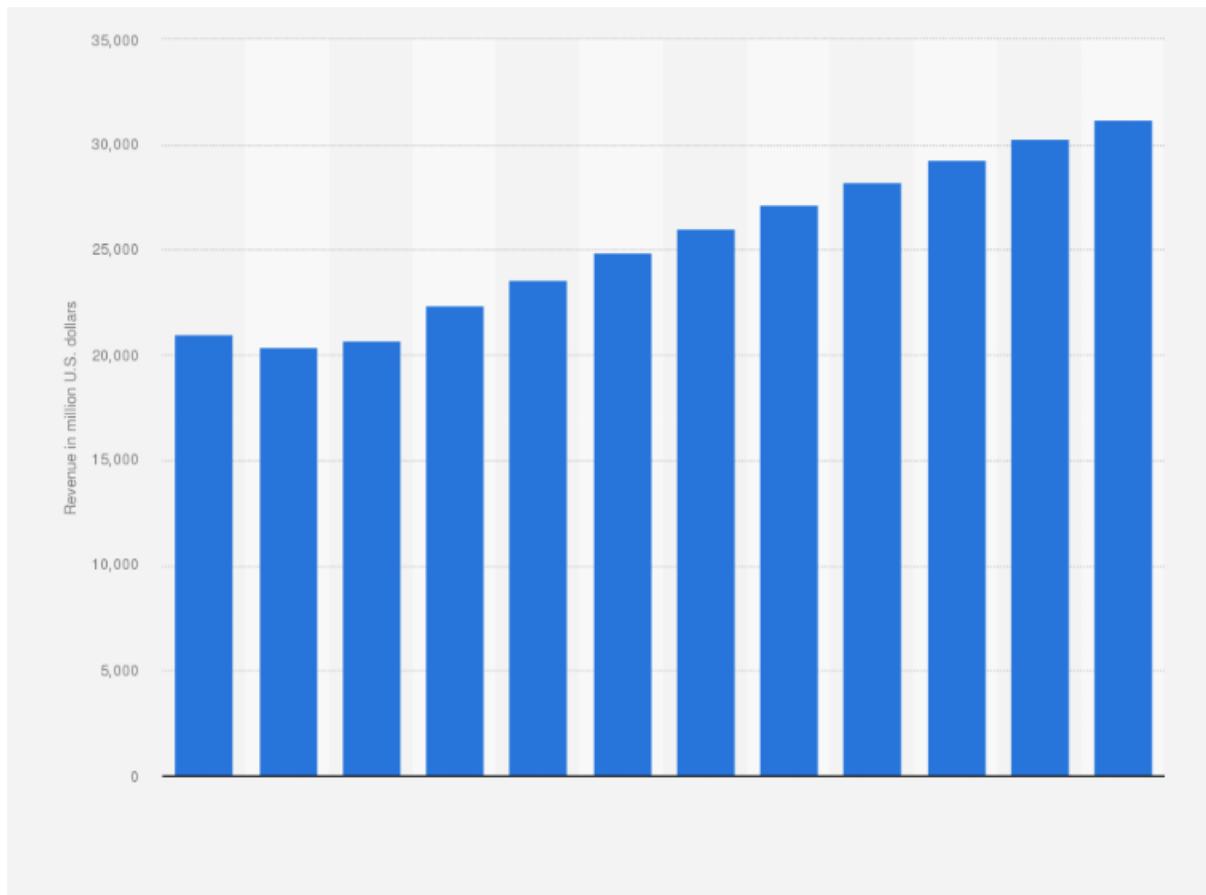


Figura 3. Ingresos del mercado del cuidado de la piel en Estados Unidos de 2019 a 2030 (en millones de dólares estadounidenses).

Fuente: Statista (2023f). «U.S. skin care market - Statistics & Facts».

piel, representa una proporción significativa, con un volumen estimado de 1.72 mil millones de dólares en 2025. (Statista, 2023c)

En el Perú, el mercado de belleza y cuidado personal ha mostrado un crecimiento sostenido, como se ve en la Infografía de la Figura 4, alcanzando los 2,240 millones de soles en ventas durante 2023, con proyecciones favorables para los próximos años, según datos proporcionados por la Cámara de Comercio de Lima (CCL). Este comportamiento refleja una creciente preocupación de los consumidores por el cuidado estético, especialmente en segmentos como el cuidado facial y productos anti-edad. (Peruano, 2025)

A pesar de la amplia disponibilidad de productos y tratamientos, persisten desafíos en la evaluación precisa y personalizada de problemas cutáneos. La falta de herramientas tecnológicas avanzadas dificulta el análisis cuantitativo de características morfológicas de la piel, como arrugas y manchas. Actualmente, muchos diagnósticos dependen de observaciones manuales,



Figura 4. Infografía de la Proyección del Mercado Cosmético e Higiene Personal en Perú.

Fuente: Peruano (2025). «Belleza y cuidado personal: una industria que se afianza».

lo que puede llevar a evaluaciones subjetivas y tratamientos menos efectivos. (Esteva et al., 2017)

Investigaciones recientes han demostrado el potencial de las redes neuronales convolucionales (CNN) en la clasificación y análisis de imágenes de la piel. Un estudio destacado por Andre Esteva y colaboradores utilizó una base de datos de 129,450 imágenes clínicas para entrenar una CNN capaz de clasificar lesiones cutáneas con una precisión comparable a la de dermatólogos certificados. Este avance sugiere que las CNN pueden ser herramientas valiosas para mejorar la precisión en la evaluación de características cutáneas y el desarrollo de tratamientos personalizados. (Cleveland Clinic, 2023)

El envejecimiento de la piel es una preocupación común entre adultos jóvenes. Según la Clínica Cleveland, las líneas finas pueden comenzar a aparecer después de los 25 años, y las arrugas se vuelven más prominentes entre los 40 y 55 años. Además, un estudio en Corea del Sur encontró que el 33 % de los adultos jóvenes presentaban algún grado de arrugas faciales, porcentaje que aumentaba al 87.8 % en personas de mediana edad y al 100 % en adultos mayores. (C. H. Lee et al., 2008)

1.2 Formulación del Problema

1.2.1 Problema General

PG: ¿De qué manera pueden las Redes Neuronales Convolucionales segmentar de forma eficiente las características morfológicas de la piel del rostro?

1.2.2 Problemas Específicos

-
- PE1: ¿De qué fuentes se obtendrán los datos necesarios para el entrenamiento y la validación del sistema de segmentación?
 - PE2: ¿Cuáles serán las etapas y métodos para el desarrollo del sistema de segmentación basado en Redes Neuronales Convolucionales?
 - PE3: ¿Qué métricas y procedimientos se utilizarán para evaluar la eficiencia del sistema en la segmentación de arrugas y manchas en la piel del rostro?
 - PE4: ¿Qué criterios se emplearán para seleccionar el modelo de segmentación más adecuado en la detección de características morfológicas del rostro?
 - PE5: ¿Cómo se implementará un sistema de segmentación morfológica facial que procese imágenes en tiempo real a partir de video en vivo, utilizando modelos de redes neuronales convolucionales integrados con mecanismos de atención?

1.3 Objetivos de la Investigación

A continuación, se presentan el objetivo general y los objetivos específicos.

1.3.1 Objetivo General

OG: Desarrollar un sistema inteligente de segmentación de características morfológicas en imágenes faciales, enfocado en la detección de arrugas y manchas, mediante el uso de Redes Neuronales Convolucionales, con el propósito de optimizar la precisión en la evaluación estética y facilitar la personalización de tratamientos cosméticos.

1.3.2 Objetivos Específicos

- OE1: Recopilar y preprocesar un conjunto representativo de imágenes faciales, que incluya diversidad en tipos de piel y afecciones cutáneas, para su utilización en el entrenamiento y validación del sistema de segmentación.

- OE2: Diseñar, desarrollar e implementar un sistema de segmentación basado en Redes Neuronales Convolucionales, orientado a la detección precisa de características morfológicas faciales, específicamente arrugas y manchas.
- OE3: Definir e implementar métricas de evaluación cuantitativas, con el fin de medir la efectividad del sistema de segmentación en la identificación de patrones morfológicos de la piel facial.
- OE4: Comparar el desempeño de diversas arquitecturas de Redes Neuronales Convolucionales y técnicas de aprendizaje profundo, con el objetivo de identificar el modelo que ofrezca el mejor compromiso entre exactitud en la segmentación y eficiencia computacional.
- OE5: Implementar un sistema de segmentación morfológica facial en tiempo real, capaz de procesar imágenes capturadas desde video en vivo, mediante el uso de Redes Neuronales Convolucionales potenciadas con mecanismos de atención para la detección eficiente de arrugas y manchas.

1.4 Justificación de la Investigación

1.4.1 Teórica

Este estudio plantea el diseño y desarrollo de un modelo basado en redes neuronales convolucionales (CNN), las cuales han demostrado un rendimiento sobresaliente en tareas de clasificación, detección y segmentación de imágenes, especialmente en contextos médicos y biomédicos (Esteva et al., 2017). Las CNN son particularmente eficaces al capturar patrones espaciales jerárquicos dentro de imágenes, lo que las convierte en herramientas idóneas para analizar características morfológicas complejas como arrugas, poros dilatados y manchas cutáneas.

La aplicación de estas arquitecturas de aprendizaje profundo al campo de la dermatología computacional responde a una creciente necesidad de automatización y precisión diagnóstica en el análisis de imágenes de piel, tanto en contextos clínicos como cosméticos. En la actualidad, gran parte del diagnóstico dermatológico depende de la experiencia visual subjetiva de los profesionales de la salud o de tecnologías que no permiten una segmentación detallada de las estructuras faciales. Esta falta de precisión dificulta el monitoreo objetivo de tratamientos dermatológicos o estéticos y limita la personalización de los mismos (J. Lee & Kim, 2020).

Mediante la incorporación de técnicas de deep learning en el procesamiento de imágenes faciales, este estudio busca contribuir no solo al perfeccionamiento de herramientas tecnológicas aplicadas al cuidado personal, sino también al cuerpo de conocimiento que existe

en la intersección entre inteligencia artificial y ciencias de la salud. El uso de CNN permite abordar tareas de segmentación con una granularidad que supera los métodos tradicionales, facilitando la detección temprana de signos de envejecimiento, hiperpigmentaciones o alteraciones cutáneas superficiales que podrían ser indicativas de patologías subyacentes o simplemente influir en la percepción estética del rostro (Gao et al., 2018; Phillips et al., 2020).

Además, este tipo de investigación fortalece el desarrollo de sistemas inteligentes con capacidad de autoaprendizaje, lo cual tiene implicancias directas en la industria cosmética, ya que permitiría la creación de aplicaciones móviles o plataformas web capaces de ofrecer diagnósticos preliminares, seguimiento de tratamientos, o recomendaciones personalizadas basadas en el análisis visual de la piel del usuario. Así, el presente estudio no solo aporta al desarrollo académico en el ámbito de la inteligencia artificial aplicada, sino también al diseño de soluciones tecnológicas que puedan ser integradas en productos de uso cotidiano.

1.4.2 Práctica

Desde un enfoque aplicado, el sistema de segmentación morfológica propuesto presenta un alto potencial de impacto en diversos sectores, particularmente en la industria cosmética, dermatológica y médica. En el ámbito cosmético, esta herramienta puede emplearse para ofrecer evaluaciones automáticas de la piel que guíen la selección de productos según las condiciones faciales específicas del usuario, como arrugas finas, hiperpigmentaciones o poros dilatados. Esto representa una oportunidad significativa para la personalización masiva de productos, mejorando la experiencia del consumidor y fortaleciendo la fidelidad hacia las marcas que integren inteligencia artificial en sus soluciones.

En dermatología, el sistema puede complementar la labor médica mediante la detección temprana y precisa de imperfecciones cutáneas, contribuyendo al diagnóstico preventivo de condiciones dérmicas y al monitoreo del progreso de tratamientos tópicos o procedimientos estéticos no invasivos. Esta automatización reduce la carga de trabajo en clínicas y consultorios, optimiza los tiempos de atención, y proporciona una evaluación objetiva, reproducible y cuantificable del estado de la piel, mitigando la variabilidad que suele existir entre profesionales humanos (Huang et al., 2020).

Adicionalmente, en el sector médico y tecnológico, esta herramienta puede integrarse en aplicaciones móviles, cabinas inteligentes o espejos digitales, abriendo nuevas posibilidades para el desarrollo de interfaces humano-máquina que permitan realizar escaneos faciales sin contacto y de manera remota. Este enfoque puede ser especialmente valioso en contextos rurales o con acceso limitado a dermatólogos especializados.

En suma, el sistema no solo aporta valor tecnológico al automatizar una tarea compleja

como la segmentación de estructuras faciales, sino que también genera ventajas competitivas en la industria de la belleza y la salud, al permitir ofrecer servicios personalizados, más eficientes y de mayor calidad.

1.4.3 Metodológica

Desde el punto de vista metodológico, este proyecto se sustenta en la utilización de redes neuronales convolucionales (CNN), una arquitectura de deep learning ampliamente validada en tareas de segmentación y clasificación de imágenes biomédicas (Ronneberger et al., 2015a). Su capacidad para identificar patrones espaciales jerárquicos y aprender representaciones discriminativas las convierte en herramientas particularmente efectivas para el análisis morfológico de imágenes faciales.

La metodología propuesta se apoya en la selección de arquitecturas CNN previamente documentadas, como U-Net, ResNet o EfficientNet, que ofrecen robustez y flexibilidad para abordar problemas de segmentación pixel a pixel. Este enfoque no solo mejora la precisión y reproducibilidad de los resultados, sino que también permite comparar el rendimiento entre distintas configuraciones, hiperparámetros y funciones de pérdida, facilitando así una exploración científica profunda sobre cuál es el modelo más eficiente para detectar y segmentar características específicas como arrugas, manchas y poros.

El uso de datasets etiquetados y técnicas de data augmentation permitirá entrenar modelos capaces de generalizar a diferentes condiciones de iluminación, tonalidades de piel y tipos de imperfecciones. Asimismo, el análisis de métricas cuantitativas como IoU (Intersection over Union), Dice Coefficient, precisión y recall, aportará un marco sistemático para validar y comparar las aproximaciones implementadas, asegurando una evaluación rigurosa del desempeño del sistema.

De este modo, la metodología del estudio no solo es replicable y escalable, sino también adaptable a otras áreas del análisis médico o estético, promoviendo así la expansión del conocimiento científico en la intersección entre visión computacional e inteligencia artificial aplicada a la salud y la cosmética.

1.5 Delimitación del Estudio

Este estudio se delimita al desarrollo y evaluación de un sistema de segmentación automática enfocado exclusivamente en arrugas y manchas presentes en imágenes de piel facial humana. No se considerarán otras imperfecciones cutáneas como acné, cicatrices, rosácea, poros dilatados u otras lesiones dérmicas. Asimismo, el estudio no tiene fines diagnósticos clínicos ni médicos, por lo que no se pretende reemplazar la evaluación profesional de dermatólogos,

sino más bien aportar herramientas tecnológicas para fines cosméticos, estéticos y de cuidado personal.

El enfoque se restringe al análisis de imágenes en dos dimensiones (2D) capturadas bajo condiciones controladas de iluminación y resolución estándar, por lo que no se incluirán modelos tridimensionales (3D), ni reconstrucción volumétrica de la piel facial. Además, las imágenes utilizadas corresponderán a rostros humanos adultos, descartando el análisis en otras zonas del cuerpo o en poblaciones pediátricas.

Finalmente, este trabajo se sitúa dentro del ámbito de la inteligencia artificial aplicada al sector cosmético y de belleza, y no en el contexto de herramientas clínicas o terapias dermatológicas. Por tanto, las conclusiones derivadas deberán interpretarse dentro del marco de la estética facial y el cuidado personal, sin extenderse a aplicaciones médicas o farmacológicas.

1.5.1 Espacial

Desde el punto de vista espacial, este estudio se desarrollará utilizando imágenes faciales obtenidas de bases de datos públicas internacionales, las cuales contienen fotografías de alta resolución capturadas bajo condiciones controladas. Las imágenes seleccionadas incluirán muestras representativas de diversos fototipos cutáneos (según la clasificación de Fitzpatrick) y procedencias geográficas variadas, con el objetivo de asegurar la diversidad étnica y regional en el análisis morfológico de la piel.

Cabe señalar que no se realizarán capturas de imágenes locales ni se recopilarán datos de sujetos en campo, ya que todo el trabajo se fundamentará en datasets disponibles de manera pública y ética para fines de investigación científica. Las bases de datos a emplearse incluyen, por ejemplo, UTKFace, CelebA y otras similares que contienen metadatos relevantes como edad, género y origen étnico, lo cual facilitará el análisis estratificado y la validación del modelo.

1.5.2 Temporal

Desde una perspectiva temporal, el desarrollo completo del presente estudio se proyecta en un período estimado de seis a doce meses, estructurado en varias fases consecutivas: recolección y selección de datos, preprocesamiento de imágenes, diseño e implementación del modelo de redes neuronales convolucionales, entrenamiento, validación y análisis de resultados. Este marco temporal incluye además iteraciones necesarias para la mejora del rendimiento del sistema y pruebas de robustez.

Las imágenes utilizadas para el entrenamiento y prueba de los modelos provendrán de datasets públicos recolectados o actualizados en los últimos cinco años. Esta selección se rea-

liza con el fin de asegurar que las características cutáneas analizadas como arrugas y manchas reflejen condiciones actuales de la población, considerando los cambios recientes en tendencias estéticas, estilos de vida, y exposición ambiental (como la contaminación o la radiación UV). Asimismo, se busca que las imágenes estén alineadas con los estándares contemporáneos de calidad visual y anotación, lo cual es fundamental para garantizar la validez y vigencia de los resultados obtenidos.

1.5.3 Conceptual

Este estudio se basa en conceptos fundamentales relacionados con la segmentación de imágenes, un área crucial en visión por computadora, y con el uso de redes neuronales convolucionales (CNN) para el procesamiento avanzado de imágenes faciales. A lo largo del trabajo se definirá con precisión el término segmentación, que hace referencia al proceso de dividir una imagen en regiones de interés para facilitar el análisis de características específicas.

El estudio también abordará las características morfológicas de la piel, tales como arrugas y manchas, explicando cómo estas se presentan en imágenes de alta resolución y qué técnicas de procesamiento de imágenes son necesarias para identificarlas de manera precisa. Se profundizará en la técnica de análisis morfológico, que examina la forma, tamaño y disposición de las imperfecciones cutáneas.

Además, el concepto de detección automática se explorará en su relación con el cuidado de la piel, destacando cómo la automatización del análisis de imágenes puede facilitar una evaluación precisa y objetiva de la piel, mejorando la personalización de productos cosméticos y los diagnósticos preventivos. Con todo, el trabajo construirá un marco teórico y práctico robusto que permita contextualizar la investigación en el área de la inteligencia artificial aplicada al cuidado de la piel.

Capítulo II: Marco Teórico

2.1 Antecedentes de la investigación

En esta parte de la investigación se presentan algunos antecedentes relacionados a la detección y pre-diagnóstico de nódulos en distintos órganos y a través de diversas metodologías. Estos ayudarán a entender el enfoque y obtener bases para un correcto desarrollo del proyecto en cuestión.

2.1.1 «Deep-Learning-Based Morphological Feature Segmentation for Facial Skin Image Analysis»

Introducción: El artículo «Deep-Learning-Based Morphological Feature Segmentation for Facial Skin Image Analysis», donde Yoon et al. (2023) presentan un enfoque innovador para el análisis de imágenes faciales con el objetivo de segmentar características morfológicas de la piel, como arrugas y poros, utilizando una red neuronal profunda basada en U-Net. A diferencia de los métodos tradicionales que se enfocan en el color de la piel, este método utiliza estructuras morfológicas y se integra con mecanismos de atención y codificación posicional. Esto permite una detección simultánea más precisa de arrugas y poros, un avance significativo en el campo de la dermatología estética y la industria cosmética.

Técnicas utilizadas: El método propuesto utiliza una versión mejorada de la arquitectura U-Net, que incorpora esquemas de atención para enfocarse en áreas relevantes de la piel y mejorar la transmisión de características al decodificador. Se desarrollaron mapas de textura para generar la "ground truth"(GT), utilizando filtros que resaltan detalles de alta frecuencia, como bordes y coherencia para las arrugas y pirámides Laplacianas para los poros. Además, se integró información posicional para mejorar la localización de estas características y se emplearon técnicas ligeras como el "zero-padding" para reducir los falsos positivos y optimizar el rendimiento computacional.

Metodología: La metodología empleada incluyó la creación de mapas de textura como "ground truth" para las arrugas y poros, lo cual refinó las anotaciones manuales mediante el uso de filtros y técnicas de análisis de imágenes. La red neuronal fue entrenada con imágenes faciales anotadas manualmente y optimizada mediante mecanismos de atención y codificación posicional. Los resultados obtenidos fueron comparados con enfoques tradicionales de procesamiento de imágenes y modelos recientes basados en aprendizaje profundo, lo que permitió evaluar la efectividad del método propuesto.

Base de datos utilizada: Aunque el artículo no menciona una base de datos pública

específica, se utilizaron imágenes faciales procesadas y anotadas manualmente para generar los datos de entrenamiento necesarios. Estas imágenes fueron empleadas para entrenar la red neuronal y crear los mapas de textura utilizados en la generación de la "ground truth" para las características faciales de interés, como arrugas y poros.

Resultados: El enfoque propuesto mostró una segmentación precisa de arrugas y poros, como se puede ver en la Tabla 1 y en la Figura 5, superando a los métodos tradicionales y otros modelos recientes de aprendizaje profundo en términos de precisión. La integración de mecanismos de atención mejoró significativamente la localización de las características faciales y redujo los falsos positivos. Además, la generación de la "ground truth" mediante mapas de textura demostró ser eficaz en la captura de las características morfológicas específicas de la piel facial.

Tabla 1
Métricas de rendimiento de diferentes modelos

Models	#Params	Loss	IoU of Wrinkle	IoU of Pore
U-Net	17.3 M	1.243	0.2078	0.3601
Reduced U-Net	4.3 M	1.250	0.2147	0.3646
Reduced U-Net, Attentions	5.2 M	1.242	0.2250	0.3714
Reduced U-Net, Attentions, Zero-padding (Proposed)	5.2 M	1.145	0.2341	0.4032

Conclusiones: Este estudio propone un avance significativo en la segmentación simultánea de arrugas y poros mediante redes neuronales profundas. El método mejora la precisión en comparación con enfoques anteriores y tiene aplicaciones potenciales en áreas como la estimación de la edad y la predicción de enfermedades cutáneas. Se sugiere que futuras investigaciones puedan ampliar esta metodología para abordar otras características faciales y problemas relacionados.

2.1.2 «High Performing Facial Skin Problem Diagnosis with Enhanced Mask R-CNN and Super Resolution GAN»

Introducción: El artículo de M. Kim y Song (2023) titulado «High Performing Facial Skin Problem Diagnosis with Enhanced Mask R-CNN and Super Resolution GAN», aborda la importancia de un diagnóstico preciso de problemas faciales de la piel, que afectan aspectos como la percepción de edad, belleza y salud. Tradicionalmente, este diagnóstico requiere visitas a clínicas especializadas, lo cual es costoso y poco práctico. Como alternativa, se propone el uso de algoritmos de aprendizaje automático para realizar diagnósticos más accesibles. Sin embargo, los enfoques basados en redes neuronales convolucionales (CNN) enfrentan desafíos técnicos, como la dificultad para detectar problemas pequeños, la identificación de hasta 20

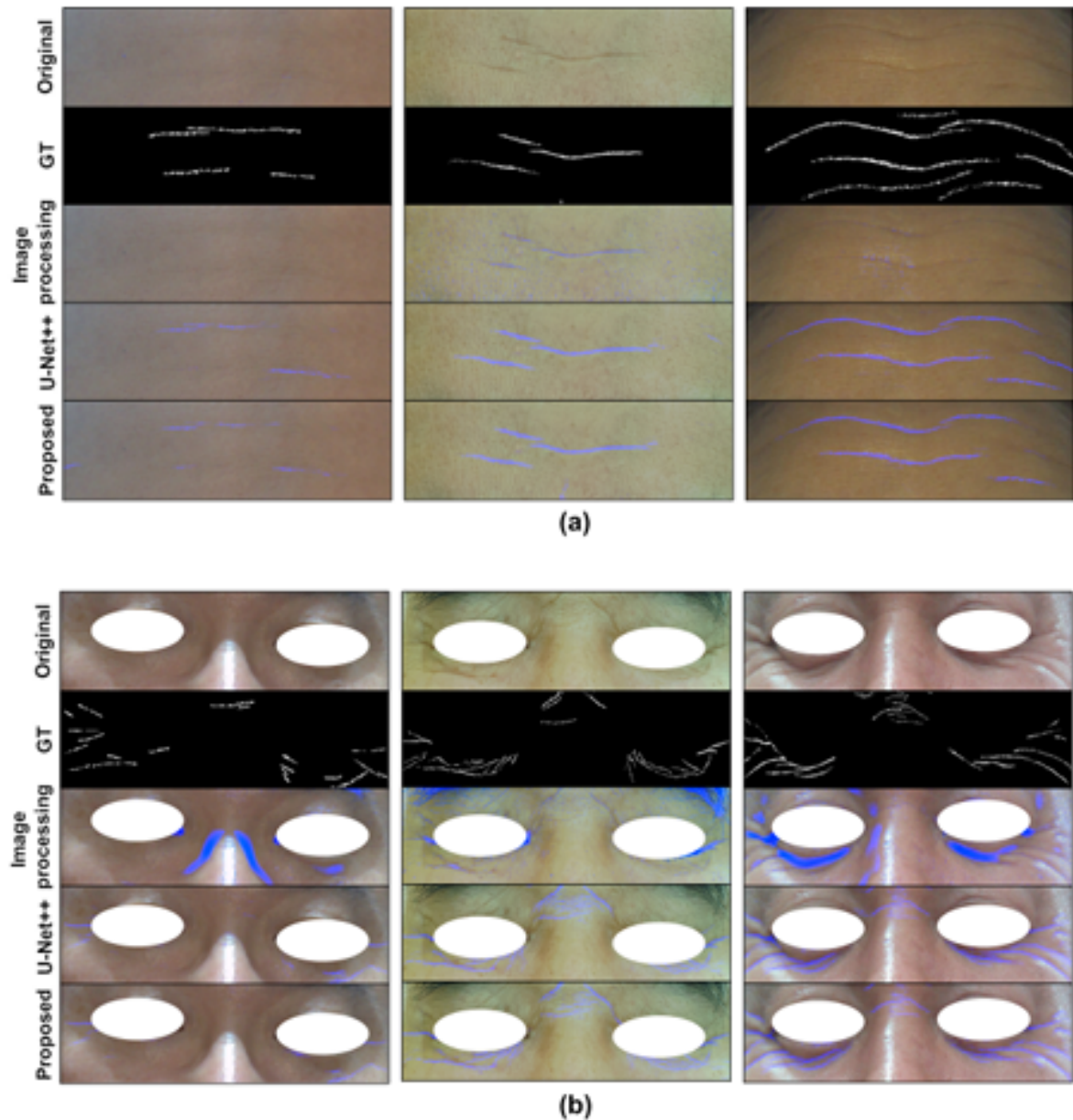


Figura 5. Comparación de la segmentación de arrugas en (a) la frente y (b) la región de los ojos.

Fuente: Yoon et al. (2023). «Deep-Learning-Based Morphological Feature Segmentation for Facial Skin Image Analysis». (p. 10)

tipos diferentes de problemas cutáneos, y la segmentación incorrecta en áreas no faciales. El objetivo del estudio es superar estas limitaciones mediante cinco tácticas específicas.

Técnicas utilizadas: El estudio propone un enfoque combinado que incluye el uso de Mask R-CNN mejorado para segmentar y detectar problemas cutáneos con mayor precisión,

así como Generative Adversarial Networks (GAN) para mejorar la resolución de las imágenes faciales. Además, se realizan ajustes en las redes neuronales para manejar objetos pequeños y mejorar la segmentación, así como el empleo de métodos para reducir el ruido y resaltar características relevantes en las imágenes. Se implementan técnicas para diferenciar visualmente problemas similares y mejorar la precisión en el diagnóstico.

Metodología: Se entrenaron 31 modelos de segmentación utilizando las tácticas propuestas, utilizando redes preexistentes como ResNet y MobileNetV2 como base. Los modelos fueron evaluados mediante métricas como el coeficiente de similitud Dice (DSC) y se compararon con otros enfoques tradicionales basados en CNN. Esta metodología permitió evaluar la efectividad de las técnicas implementadas para superar los desafíos técnicos mencionados y mejorar el rendimiento en el diagnóstico de problemas cutáneos.

Base de datos utilizada: Se emplearon 2225 fotos faciales con una resolución de 576×576 píxeles, que contenían diversos tipos de problemas cutáneos, como poros, lunares y acné, para entrenar y evaluar los modelos de segmentación. Estas imágenes proporcionaron un conjunto de datos diverso y representativo para evaluar la precisión del sistema en el diagnóstico de distintos problemas faciales.

Resultados: El enfoque propuesto alcanzó un rendimiento del 83.38 % en el diagnóstico, superando en un 32.58 % a los métodos tradicionales basados en CNN. Además, superó a modelos como MobileNetV2, Xception y VGG19, como se ve en la Figura 6, por un promedio del 17.89%. Uno de los logros más importantes fue la mejora significativa en la detección de objetos pequeños, como lunares, que tradicionalmente presentan baja precisión en los métodos tradicionales.

Conclusiones: El sistema desarrollado, como se ve en la Figura 7 demuestra ser una herramienta efectiva para diagnosticar problemas faciales de la piel con un nivel clínico de precisión, ofreciendo una alternativa viable y económica frente a las visitas a clínicas dermatológicas. Las tácticas propuestas tienen el potencial de revolucionar el cuidado facial mediante inteligencia artificial, con aplicaciones prácticas que pueden mejorar significativamente la accesibilidad y eficiencia de los diagnósticos dermatológicos.

2.1.3 «Facial Wrinkle Detection Based on DeepLabV3+ and Semi-Automatic Labelling Strategy»

Introducción: El artículo titulado «Facial Wrinkle Detection Based on DeepLabV3+ and Semi-Automatic Labelling Strategy» de Zhong et al. (2024) presenta un enfoque innovador para la detección de arrugas faciales, abordando las limitaciones de métodos tradicionales que se ven afectados por interferencias con otras características faciales. Para mejorar la precisión,

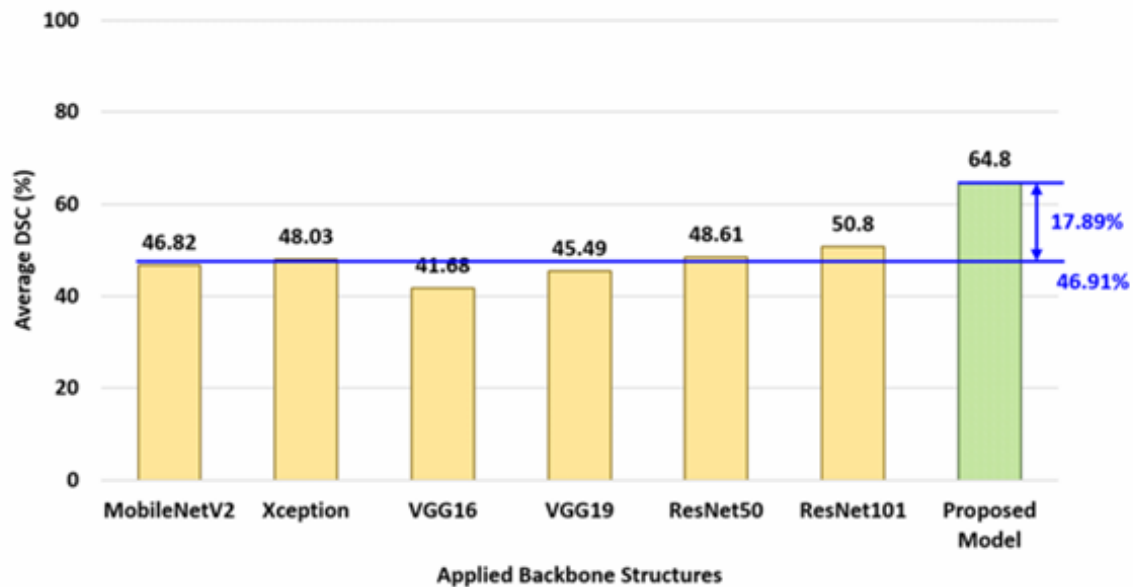


Figura 6. Comparación del modelo propuesto con otros 6 modelos de redes neuronales.

Fuente: M. Kim y Song (2023). «High Performing Facial Skin Problem Diagnosis with Enhanced Mask R-CNN and Super Resolution GAN». (p. 8)

los autores proponen el uso del modelo DeepLabV3+ junto con una estrategia de etiquetado semi-automática, lo que permite generar datos de entrenamiento más representativos y mejorar la segmentación de arrugas.

Técnicas utilizadas: La investigación, como podemos ver en las Figuras 8 y 9, emplea técnicas avanzadas como DeepLabV3+, una red optimizada para segmentación de imágenes, y MobileNetV2 para reducir la carga computacional. Además, se desarrolla una estrategia de etiquetado semi-automática combinando anotaciones dermatológicas con mapas de textura generados mediante filtros Wiener y umbralización adaptativa. La precisión del modelo se evalúa mediante el Índice de Similitud Jaccard (JSI).

Metodología: Se seleccionaron 300 imágenes faciales del conjunto Flickr-Face-HQ, con una resolución de 1024x1024 píxeles, para crear el conjunto de datos. Se generaron etiquetas ground truth utilizando la estrategia semi-automática propuesta. El conjunto de datos se dividió en 225 imágenes para entrenamiento, 25 para validación y 50 para pruebas. La configuración experimental incluyó un hardware con una CPU Intel Core i9-12900K, GPU RTX 3070, y 32 GB de RAM, además de utilizar el framework PyTorch para el desarrollo del modelo.

Base de datos utilizada: El conjunto de datos utilizado fue el Flickr-Face-HQ, que contiene imágenes faciales de alta resolución. Este conjunto de imágenes se procesó para ge-

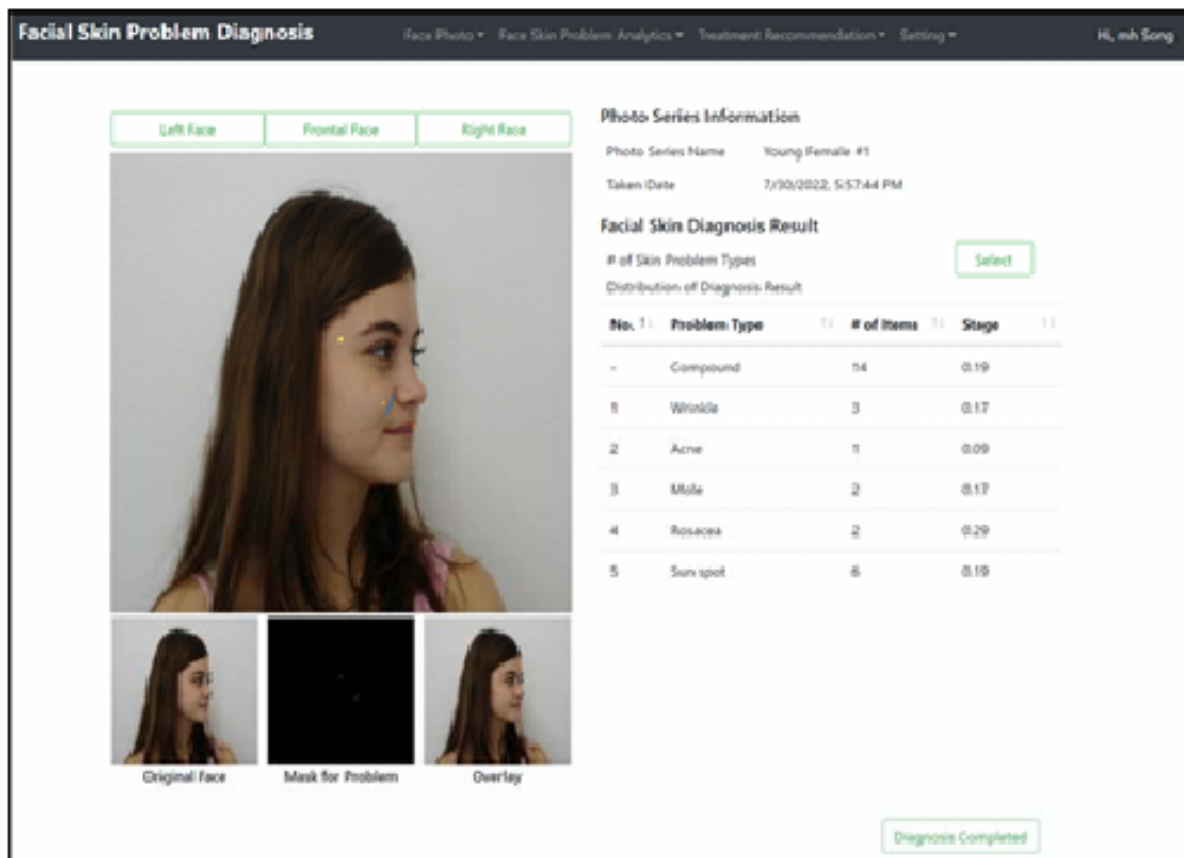


Figura 7. Interfaz de usuario del sistema de diagnóstico de problemas de la piel facial.

Fuente: M. Kim y Song (2023). «High Performing Facial Skin Problem Diagnosis with Enhanced Mask R-CNN and Super Resolution GAN». (p. 15)

nerar etiquetas ground truth específicas para la detección de arrugas utilizando la estrategia semi-automática propuesta.

Resultados: Los resultados demostraron que el método propuesto supera a enfoques tradicionales como el filtro Hessiano y el modelo U-Net, obteniendo mayores valores de JSI en la frente (0.62) y área ocular (0.64). Además, redujo la cantidad de falsos positivos y mejoró la segmentación de bordes, logrando un mejor desempeño en la detección de arrugas finas.

Conclusiones: En conclusión, el modelo DeepLabV3+ con etiquetado semi-automático se mostró más efectivo en la detección de arrugas faciales. Sin embargo, aún enfrenta desafíos en la diferenciación de arrugas muy finas y cabellos, lo que sugiere la necesidad de mejoras futuras para incrementar la precisión del sistema.

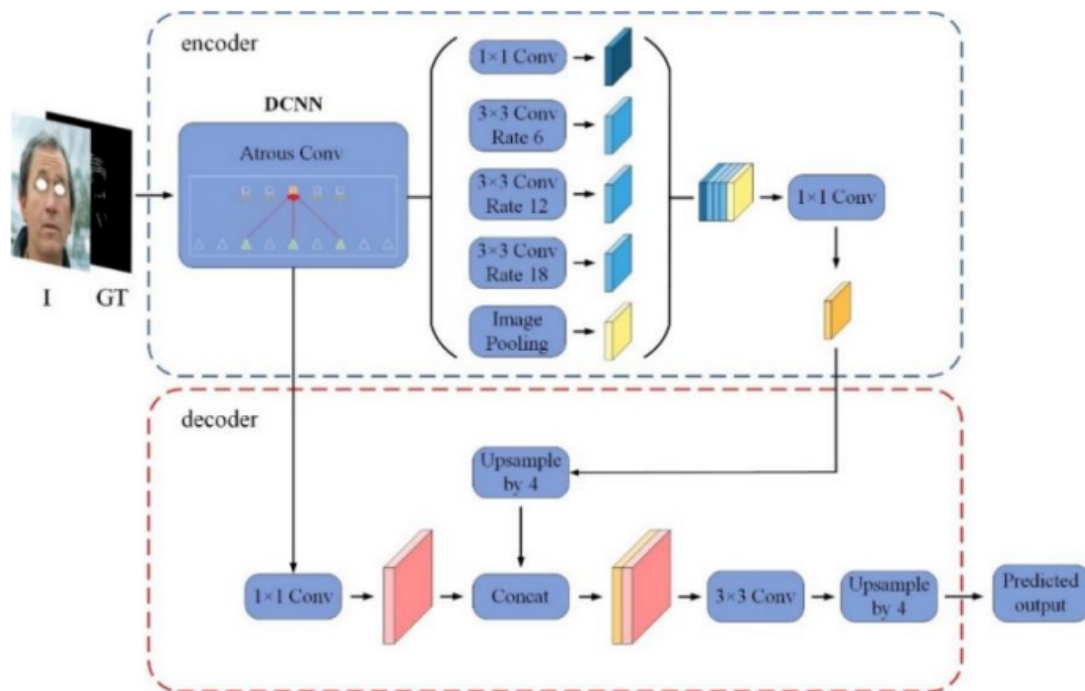


Figura 8. Estructura de red de DeepLabV3+.

Fuente: Zhong et al. (2024). «Facial Wrinkle Detection Based on DeepLabV3+ and Semi-Automatic Labelling Strategy». (p. 4)

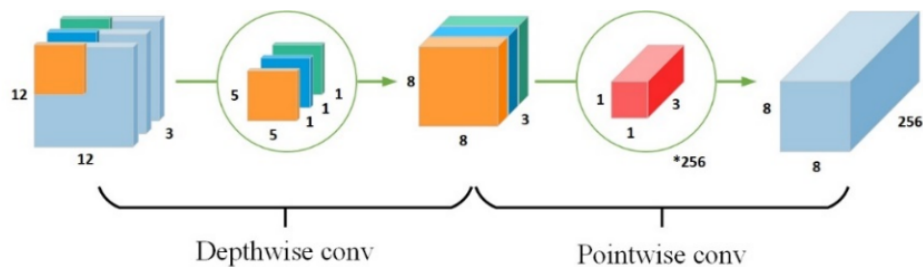


Figura 9. El proceso de convolución separable en profundidad.

Fuente: Zhong et al. (2024). «Facial Wrinkle Detection Based on DeepLabV3+ and Semi-Automatic Labelling Strategy». (p. 4)

2.1.4 «Improved U-Net: Fully Convolutional Network Model for Skin-Lesion Segmentation»

Introducción: En el artículo de Karshiev et al. (2020) llamado «Improved U-Net: Fully Convolutional Network Model for Skin-Lesion Segmentation» analiza el problema de la seg-

mentación de lesiones cutáneas, una tarea fundamental para el diagnóstico temprano del melanoma. Aunque el modelo U-Net ha sido ampliamente utilizado en segmentación médica, presenta limitaciones como ralentización en el entrenamiento y problemas con la función de activación ReLU. Para mejorar su desempeño, los autores proponen una versión optimizada que incorpora interpolación bilineal para el upsampling y la función de activación PReLU, lo que mejora la precisión y evita problemas como el sobreajuste.

Técnicas utilizadas: La investigación emplea un diagrama el cual se puede ver en la Figura 10 y a su vez varias técnicas clave: la interpolación bilineal sustituye la deconvolución tradicional para mejorar la segmentación de bordes, la función PReLU reemplaza ReLU para prevenir "neuronas muertas" optimizar la convergencia, y el dropout se utiliza después de cada bloque convolucional para reducir el sobreajuste. Estas modificaciones permiten mejorar la estabilidad y eficiencia del entrenamiento.

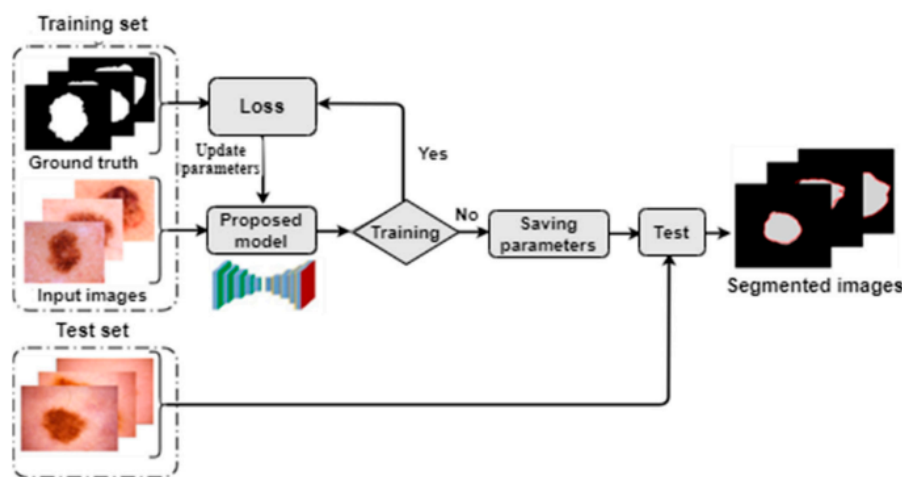


Figura 10. Diagrama de flujo del sistema propuesto.

Fuente: Karshiev et al. (2020). «Improved U-Net: Fully Convolutional Network Model for Skin-Lesion Segmentation». (p. 4)

Metodología: El modelo propuesto como se ve en la Figura 11, se basa en una arquitectura U-Net modificada con capas convolucionales optimizadas, PReLU, dropout y upsampling mediante interpolación bilineal. Fue entrenado en un sistema con un procesador Intel Core i7-9700K, 32 GB de RAM y una GPU NVIDIA GeForce RTX 2060 SUPER, garantizando un entorno adecuado para el procesamiento intensivo de imágenes médicas.

Base de datos utilizada: Para el entrenamiento y prueba del modelo se utilizó un conjunto de datos de imágenes dermoscópicas, con 2594 imágenes etiquetadas para entrenamiento y 1000 para prueba. Todas las imágenes fueron preprocesadas, redimensionadas a 256×256

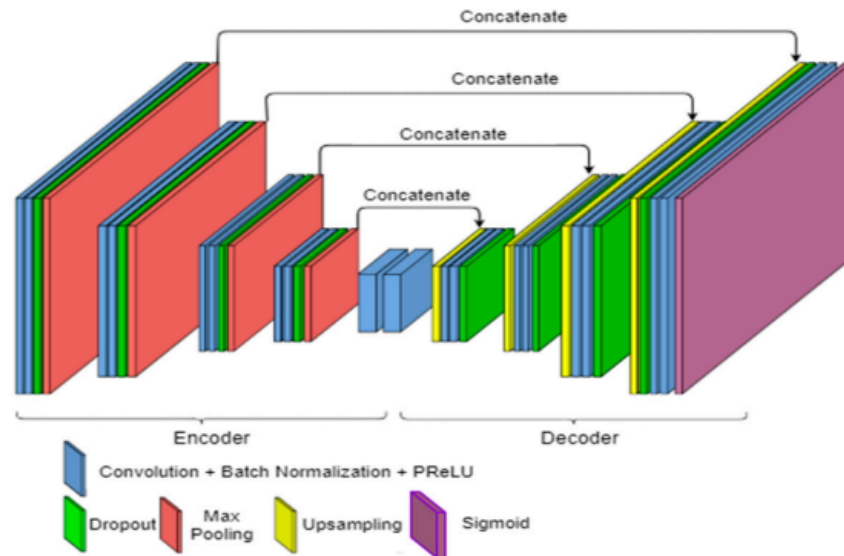


Figura 11. Modelo totalmente convolucional propuesto para la segmentación de lesiones cutáneas.

Fuente: Karshiev et al. (2020). «Improved U-Net: Fully Convolutional Network Model for Skin-Lesion Segmentation». (p. 4)

píxeles y convertidas a escala de grises, lo que facilitó la normalización y mejoró la eficiencia del modelo.

Resultados: Los resultados muestran un alto rendimiento del modelo mejorado, alcanzando una precisión por píxel del 94 % y un coeficiente Dice del 88 %. Estas mejoras permiten reducir artefactos en la segmentación y aumentar la eficiencia computacional en comparación con la U-Net estándar, consolidándose como una alternativa más precisa y robusta para la segmentación de lesiones cutáneas.

Conclusiones: En conclusión, la versión mejorada de U-Net supera las limitaciones del modelo original al abordar problemas como gradientes débiles y artefactos en la segmentación. La integración de interpolación bilineal, PReLU y dropout permite lograr una mayor precisión y eficiencia, posicionando este enfoque como una herramienta prometedora para la segmentación de imágenes médicas en el ámbito dermatológico.

2.1.5 «A New Approach to Facial Skin Analyzer»

Introducción: Tamilkodi et al. (2024) elaboraron el artículo titulado «A New Approach to Facial Skin Analyzer», que presenta un software avanzado para el análisis de la piel facial que permite identificar características clave como arrugas, poros, manchas y textura, además

de predecir edad y género. El sistema compara los resultados con datos de personas del mismo grupo etario, clasifica el tono de piel, genera máscaras visuales para resaltar zonas específicas del rostro y ofrece recomendaciones personalizadas. El objetivo central es desarrollar un modelo robusto capaz de realizar un análisis facial completo en diversos contextos.

Técnicas utilizadas: Para lograrlo, el software utiliza redes neuronales convolucionales profundas (D-CNN), apoyadas por modelos preentrenados para la detección de edad y género. También emplea técnicas como KMeans para clasificar el tono de piel y gradientes de Sobel para identificar arrugas. Estas herramientas se integran para ofrecer un análisis visual y cuantitativo de cada característica facial.

Metodología: La metodología del sistema incluye la captura de imágenes faciales desde múltiples ángulos, el preprocesamiento mediante técnicas como desenfoque gaussiano y conversión de formatos, seguido del análisis específico de textura, tono y arrugas. Finalmente, los resultados se visualizan en gráficos y máscaras y se utilizan para generar recomendaciones personalizadas para el usuario.

Base de datos utilizada: Como base de datos, se empleó la Indian Face Age Database (IFAD), compuesta por 3,296 imágenes de 55 personas con diversidad de edad, poses, expresiones e iluminación. Esta variedad permite entrenar modelos con mayor capacidad de generalización y robustez en escenarios reales de análisis facial.

Resultados: Los resultados como se pueden ver en las Figuras 12, 13 y 14 ,muestran una alta precisión en la detección de características faciales y predicciones fiables de edad y género. Además, el sistema logra una visualización efectiva mediante gráficos y máscaras que resaltan áreas clave del rostro, lo que mejora la comprensión de los resultados.

Conclusiones: En conclusión, el enfoque propuesto demuestra ser eficaz para detectar estructuras faciales importantes incluso en condiciones variables. Aunque aún no alcanza una precisión perfecta, se reconoce el valor de expandir el conjunto de datos para mejorar el rendimiento. Este modelo tiene un gran potencial para aplicaciones automatizadas en análisis dermatológico y cuidado personalizado de la piel.

2.1.6 «Dermatology: Pretraining with Texture Map-Based Weak Supervision»

Introducción: El artículo de Moon et al. (2024) llamado «Dermatology: Pretraining with Texture Map-Based Weak Supervision» se enfoca en la detección automática de arrugas faciales, un tema relevante en dermatología estética, debido a que los métodos manuales de segmentación son subjetivos, costosos y laboriosos. Para abordar este problema, los auto-

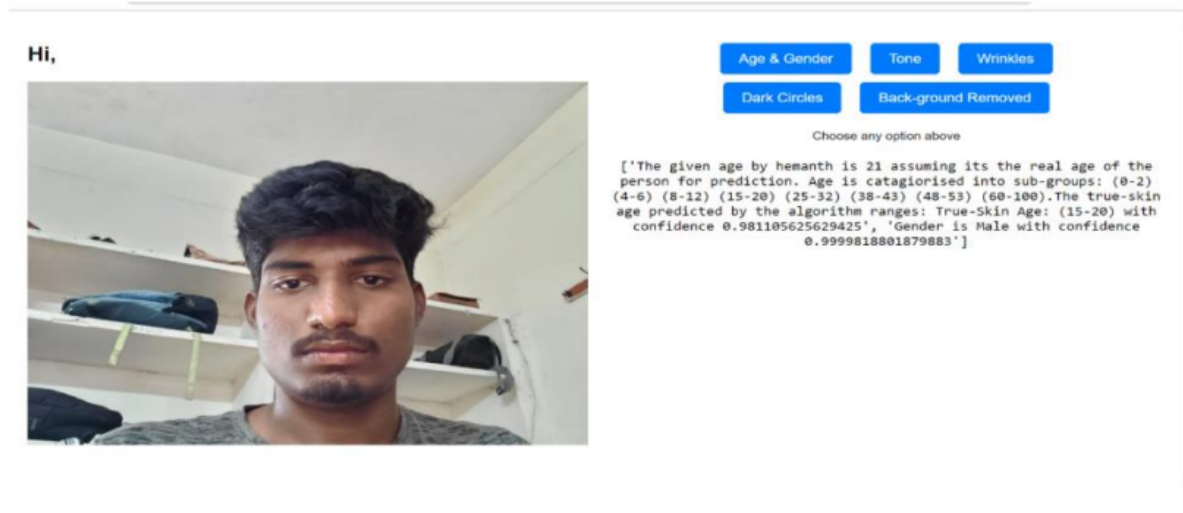


Figura 12. Detección de edad y género.

Fuente: Tamilkodi et al. (2024). «A New Approach to Facial Skin Analyzer». (p. 6)

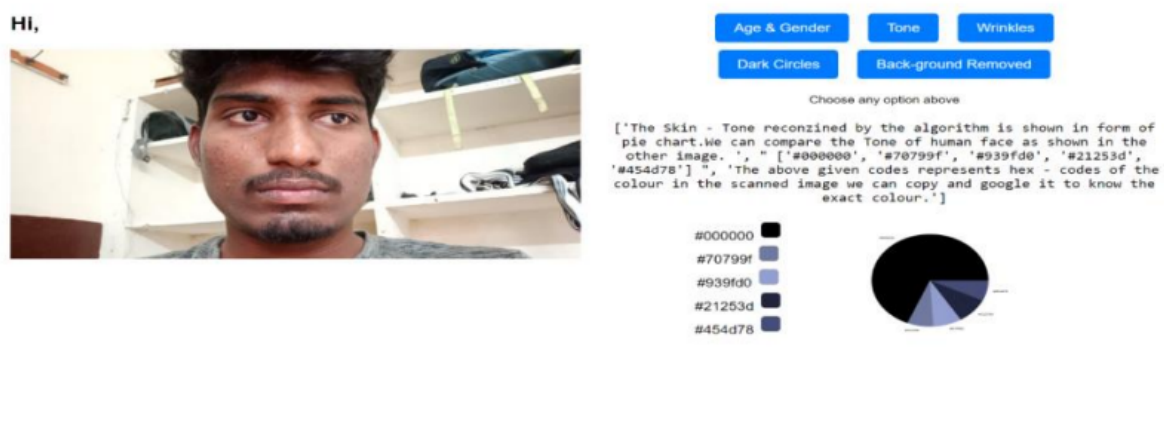


Figura 13. Análisis del tono de piel.

Fuente: Tamilkodi et al. (2024). «A New Approach to Facial Skin Analyzer». (p. 6)

res desarrollan un enfoque basado en aprendizaje profundo que incluye la creación del primer conjunto de datos público especializado en segmentación de arrugas (FFHQ-Wrinkle), y una estrategia de entrenamiento en dos etapas que combina un preentrenamiento débilmente supervisado con un ajuste fino supervisado, optimizando así el rendimiento del modelo con un uso eficiente de los recursos de etiquetado.

Técnicas utilizadas: Se aplican dos enfoques complementarios: primero, un preentrenamiento débilmente supervisado que genera automáticamente mapas de textura mediante

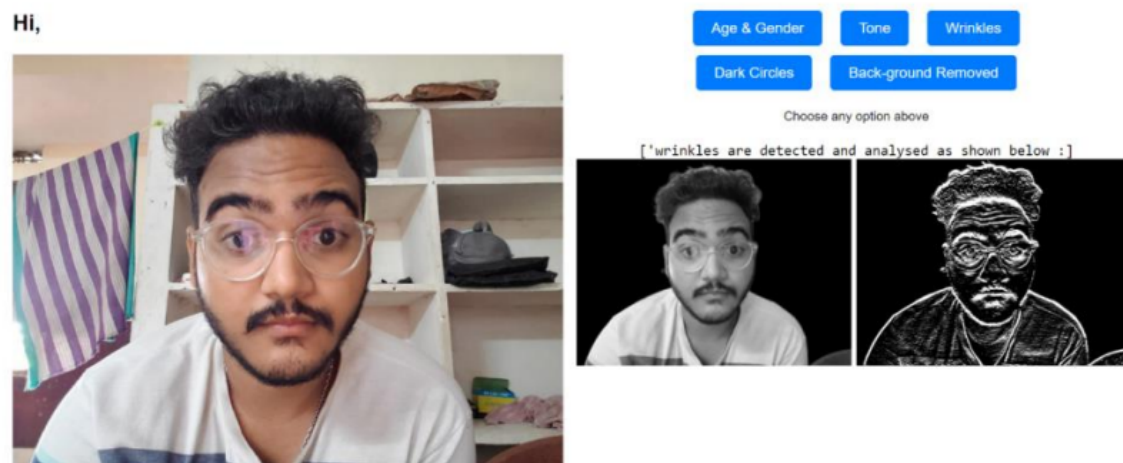


Figura 14. Análisis de arrugas.

Fuente: Tamilkodi et al. (2024). «A New Approach to Facial Skin Analyzer». (p. 6)

filtros Gaussianos sobre 50,000 imágenes no etiquetadas; y segundo, un ajuste fino supervisado que utiliza 1,000 imágenes con máscaras de arrugas anotadas manualmente por expertos. Para la segmentación se emplean arquitecturas de redes neuronales como U-Net y Swin UNETR, que permiten capturar tanto características locales como globales de las arrugas faciales.

Metodología: La metodología, como podemos ver en la Figura 15, consiste en generar etiquetas débiles mediante mapas de textura derivados de imágenes faciales, eliminando regiones no relevantes con un modelo de segmentación facial (BiSeNet). Luego, tres expertos generan manualmente máscaras de arrugas en regiones específicas como la frente y las líneas nasolabiales. El modelo se entrena en dos etapas: primero aprende a predecir texturas con pérdida MSE y luego realiza un ajuste fino con datos manualmente etiquetados y los mapas de textura como entrada adicional, fortaleciendo así su capacidad de generalización.

Base de datos utilizada: El estudio introduce el conjunto FFHQ-Wrinkle, derivado del conjunto de datos FFHQ, compuesto por 50,000 imágenes con etiquetas débiles generadas automáticamente y 1,000 imágenes con anotaciones manuales precisas. Este conjunto incluye una amplia variedad de edades, géneros y etnias, lo que contribuye a la robustez del modelo entrenado y favorece su aplicación generalizada en diversos contextos.

Resultados: Los modelos entrenados mediante la estrategia propuesta lograron un rendimiento superior en comparación con métodos anteriores, tanto en métricas cuantitativas como cualitativas. La combinación de datos etiquetados automáticamente y manualmente permitió

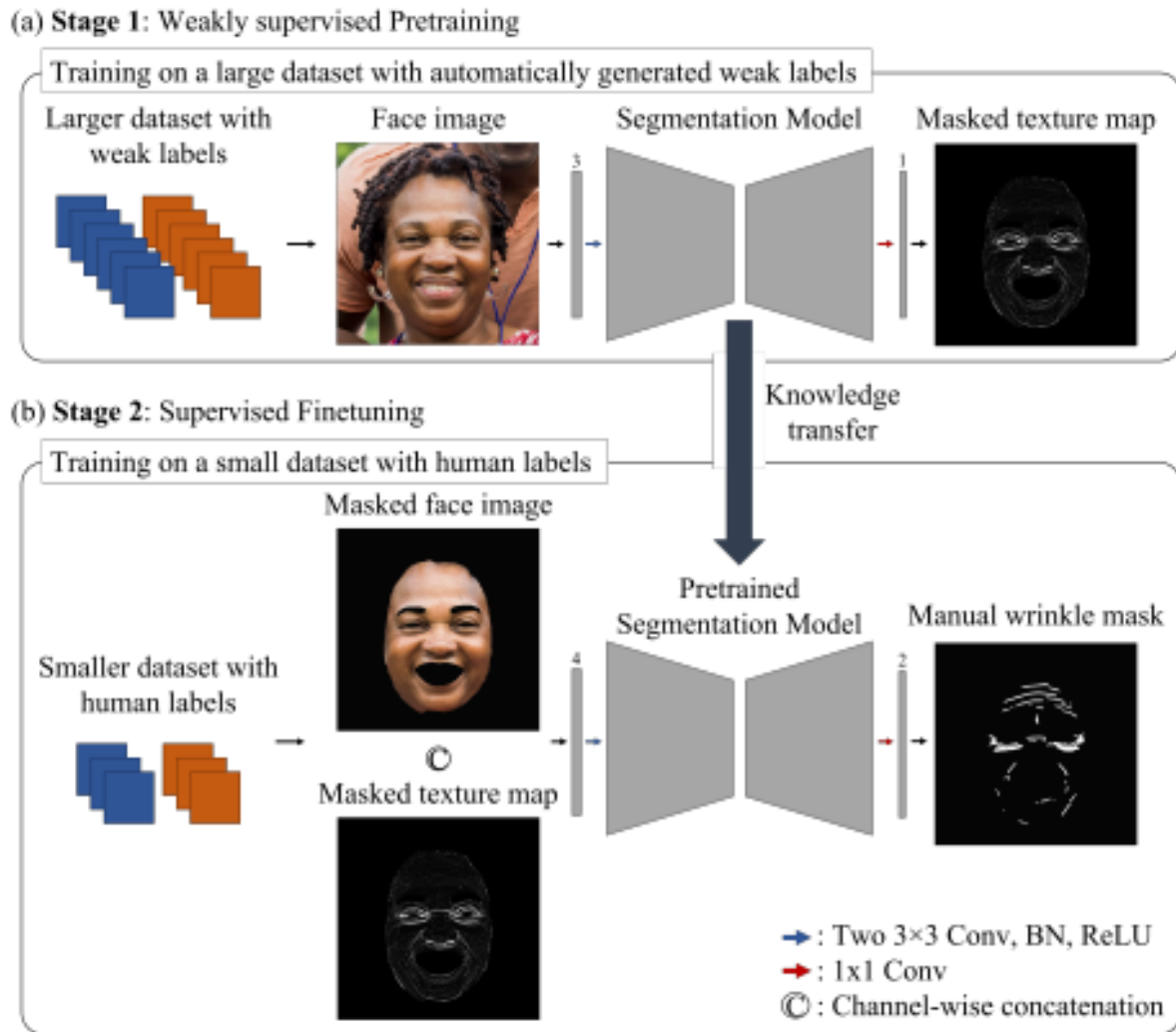


Figura 15. Entrenamiento en dos etapas para la segmentación de arrugas faciales.

Fuente: Moon et al. (2024). «Dermatology: Pretraining with Texture Map-Based Weak Supervision». (p. 3)

reducir significativamente los recursos necesarios para el etiquetado, manteniendo una alta precisión en la segmentación de arrugas faciales, especialmente en regiones clave del rostro.

Conclusiones: La investigación demuestra que la combinación de aprendizaje débilmente supervisado con ajuste fino supervisado mejora sustancialmente la segmentación automática de arrugas faciales. Además, el nuevo conjunto de datos FFHQ-Wrinkle representa un aporte valioso al campo, ya que ofrece una base estandarizada y diversa para futuras investigaciones sobre análisis facial automatizado y aplicaciones en dermatología estética.

2.2 Bases Teóricas

2.2.1 Inteligencia Artificial

El método racional fusiona la ingeniería y las matemáticas basándose en las "leyes del pensamiento", las cuales tienen su origen en la antigua Grecia y han sido influenciadas por filósofos como Aristóteles. Durante el siglo XIX, se diseñaron programas capaces de resolver problemas de lógica. Por consiguiente, el propósito de la Inteligencia Artificial en la vida real es crear sistemas inteligentes que posean estas habilidades. Aun en situaciones de incertidumbre, un "Agente Racional" toma acciones con el fin de obtener el mejor resultado posible. La inteligencia artificial se apoya en diversas disciplinas, tales como la ingeniería computacional, la teoría de control, la cibernética, la lingüística, la filosofía, la economía, la psicología, la neurociencia y las matemáticas, de acuerdo con Russell y Norvig (2004).

Dos investigadores en neurociencia crearon el primer modelo de IA basado en neuronas artificiales en 1943, dando inicio al análisis de la Inteligencia Artificial. McCulloch y Pitts idearon el prototipo que permitía que las neuronas fueran «activadas» o «desactivadas», lo que demostró que una red de neuronas era capaz de realizar cualquier tarea computacional. Posteriormente, Donald Hebb propuso la «Regla de Aprendizaje Hebbiano». John McCarthy, Allen Newell y Herbert Simon desarrollaron un programa que podía tener el pensamiento no numérico en el taller de Dartmouth, aunque no se publicó. El término «Inteligencia Artificial» fue acuñado por McCarthy, (Russell & Norvig, 2004).

La IA comenzó a entrar en la industria en los años 80, especialmente en grandes empresas de países desarrollados, a través de la investigación en sistemas expertos y el desarrollo de computadoras más poderosas.

2.2.2 Aprendizaje Automático

El Machine Learning es un área de la Inteligencia Artificial enfocada en técnicas que permiten a las computadoras aprender a través de algoritmos, convirtiendo muestras de datos en programas sin requerir programación explícita. Según Russell y Norvig (2009), el aprendizaje automático es una división de la inteligencia artificial. Estos algoritmos emplean tecnologías como el procesamiento del lenguaje natural, el aprendizaje profundo y las redes neuronales. Tanto el aprendizaje supervisado como el no supervisado se fundamentan en lecciones extraídas de los datos. La creación de algoritmos capaces de recibir datos de entrada y utilizar análisis estadístico para prever una salida, la cual se ajusta conforme se obtienen nuevos datos, constituye el fundamento del aprendizaje automático. (Alpaydin, 2014)

Se puede clasificar en cuatro tipos principales de la siguiente manera según el objetivo

que se desea alcanzar mediante el uso de ML:

- **Aprendizaje Supervisado:** El Aprendizaje Supervisado se ganó su nombre porque los científicos de datos actúan como una guía para enseñarle al algoritmo las conclusiones a las que debe llegar. Es similar a la forma en que un estudiante aprende aritmética básica de un maestro. Este tipo de aprendizaje requiere datos etiquetados con las respuestas correctas que se esperan del resultado del algoritmo. Para problemas de clasificación y regresión, el aprendizaje supervisado demostró ser preciso y rápido según (Zambrano, 2018).

Los dos tipos de Aprendizaje Supervisado son:

- **La Clasificación:** es la predicción del valor categórico de salida que permite dividir los datos en clases específicas. La clasificación se puede usar para varios propósitos, como determinar el clima, determinar si un correo electrónico es spam o no o identificar tipos de animales después de recibir una educación adecuada, un conjunto de datos con etiquetas de imágenes que incluyen la especie y algunas identificaciones características, según (Zambrano, 2018).
- **La Regresión:** es un tipo de problema en el que la predicción de un valor de respuesta continua es necesaria, como los precios de las acciones y la vivienda, según (Zambrano, 2018).

Por lo tanto, funciona modelando las relaciones y dependencias entre las características de entrada y la salida de predicción objetivo, lo que permite predecir los valores de salida para nuevos datos utilizando las relaciones que aprendió de conjuntos de datos anteriores, según (Alpaydin, 2014).

- **Aprendizaje No Supervisado:** Por otro lado, el Aprendizaje No Supervisado se asemeja más a lo que algunos expertos llaman Inteligencia Artificial real: la idea de que una máquina puede aprender a identificar patrones y procesos complejos sin la supervisión de humanos. Cuando los expertos no saben qué buscar en los datos y los datos en sí no incluyen objetivos, este método es particularmente útil. La agrupación de k-means, el análisis de componentes principales e independientes y las reglas de asociación según (Zambrano, 2018) son algunos de los muchos casos de uso del Aprendizaje Automático No Supervisado.
- **Agrupación K-means:** es un tipo de problema en el que cosas similares están agrupadas, como se muestra en la Figura 16. Comparte el mismo concepto con la clasificación, pero no se proporcionan etiquetas, por lo que el sistema entenderá los

datos y los agrupará. Un uso de esto sería agrupar los artículos y las noticias según su género y contenido, según (Sancho Caparrini, 2018)

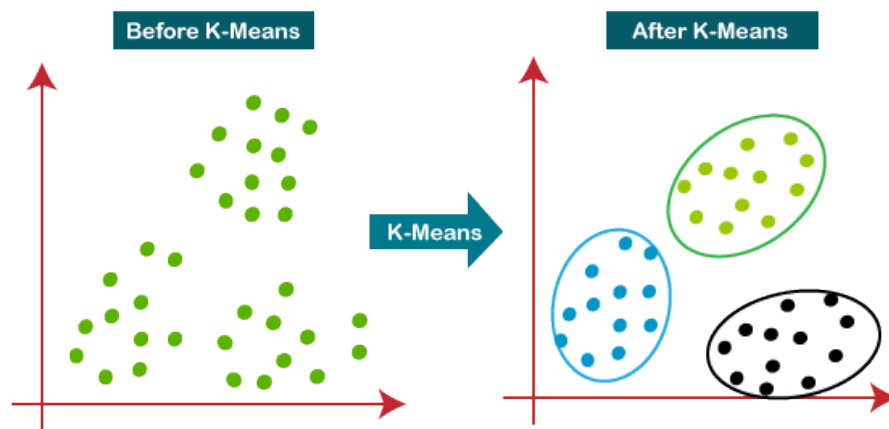


Figura 16. El algoritmo de K medias.

Fuente: Sancho Caparrini (2018). *Clasificación Supervisada y No Supervisada*.

Debido a su complejidad y dificultad de implementación, este tipo de Aprendizaje Automático no se utiliza tan frecuentemente como el Aprendizaje Supervisado, a pesar de que abre las puertas a la resolución de problemas que los humanos normalmente no abordarían, según (Sancho Caparrini, 2018)

- **Aprendizaje Semisupervisado:** Hasta el momento, todos los datos enviados han sido etiquetados con el resultado deseado o no han sido etiquetados en absoluto. El Aprendizaje Automático Semisupervisado utiliza ambos. El costo de etiquetar es bastante alto en muchas situaciones prácticas y, en el caso de grandes conjuntos de datos, se vuelve aburrido y requiere mucho tiempo. Además, proporcionar demasiados datos etiquetados puede hacer que el modelo tenga sesgos humanos. A pesar de que los datos sin etiquetar son desconocidos para la red, ofrecen información útil sobre los parámetros del grupo objetivo. que conduce a la conclusión de que se puede mejorar la precisión del modelo al incluir datos sin etiquetar y, al mismo tiempo, ahorrar tiempo y dinero en su construcción. Por ejemplo, la clasificación de páginas web, el reconocimiento de voz o la secuenciación genética pueden usar Aprendizaje Automático Semisupervisado. En esos casos, los científicos de datos pueden acceder a grandes cantidades de datos sin etiquetarlos, y la tarea de etiquetarlos todos llevaría mucho tiempo, según (Zambrano, 2018).

Se puede comparar estos tres tipos de Aprendizaje Automático para el mismo uso, como clasificación, utilizando los datos recopilados hasta ahora:

- **Clasificación supervisada:** el algoritmo clasificará los tipos de páginas web según las etiquetas proporcionadas desde el principio, según (Zambrano, 2018).
- **Agrupación no supervisada:** el algoritmo buscará patrones y características que ayudan a agrupar páginas web en grupos, según (Zambrano, 2018).
- **Clasificación semi no supervisada:** identificará varios grupos de páginas web utilizando los datos etiquetados, luego utilizará los datos no etiquetados para establecer los límites de esos grupos de páginas web y buscar otros tipos que posiblemente no aparezcan en los datos etiquetados, según (Zambrano, 2018).

2.2.3 Aprendizaje Profundo

Desde que llegó la Inteligencia Artificial hace un tiempo, tiene una amplia gama de aplicaciones y se divide en muchas ramas, como se menciona en (SAS Institute, s.f.). El Aprendizaje Profundo es un subconjunto del Aprendizaje Automático, que es en sí mismo un subcampo de la IA. La Figura 17 es una representación visual de la relación entre AI, ML y DL.

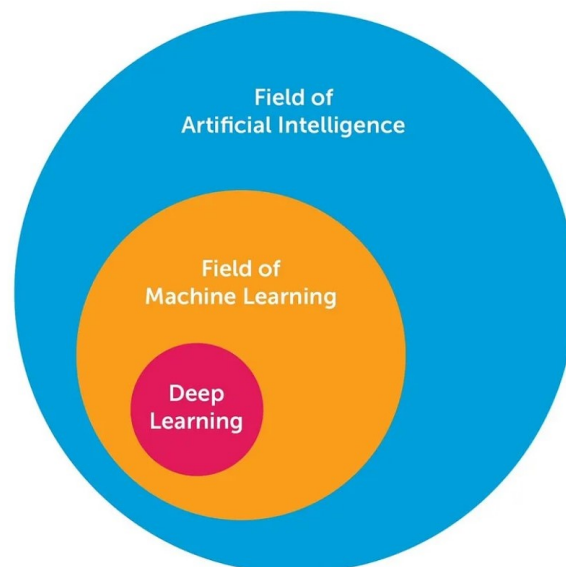


Figura 17. Relación entre IA, ML y DL.

Fuente: House of Bots (2018). *Most Popular 20 Free Online Courses to Learn Deep Learning*.

El Aprendizaje Profundo no solo permite representar datos de la manera correcta, sino que también permite que la computadora aprenda programas informáticos de varios pasos al incluir el concepto de profundidad en sus modelos. Como se muestra en la Figura 18, cada capa de representación puede interpretarse como el estado de la memoria de la computadora. Las

computadoras interpretan las imágenes como una colección de valores de píxeles que representan escenas de nuestra realidad. Según (House of Bots, 2018), identificar un objeto o mapear su identidad a partir de esos valores es una tarea difícil para las máquinas y puede resultar casi imposible cuando se intenta aprender este mapeo directamente.

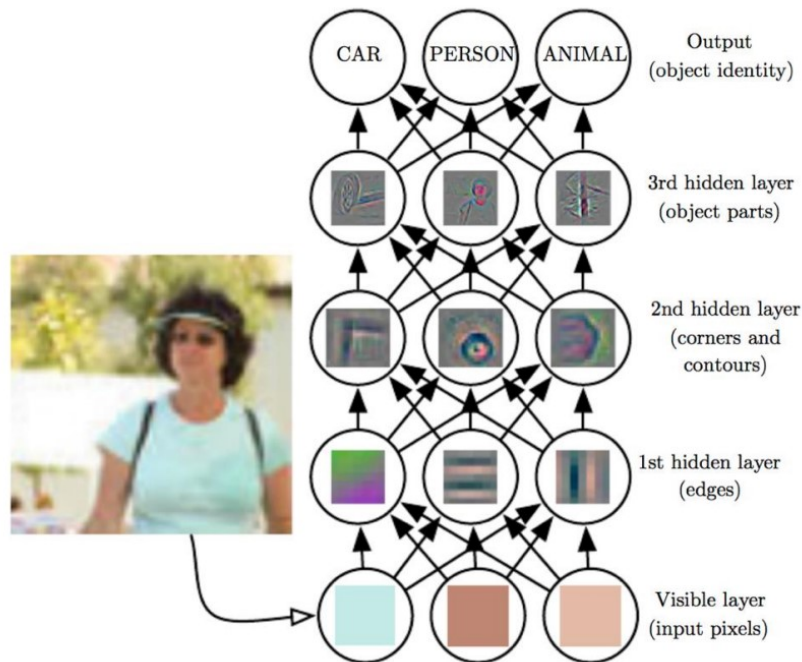


Figura 18. Modelo de aprendizaje profundo.

Fuente: House of Bots (2018). *Most Popular 20 Free Online Courses to Learn Deep Learning*.

2.2.4 Segmentación de Imágenes

2.2.4.1. Definición y objetivos de la segmentación de imágenes

La segmentación de imágenes es un proceso fundamental en el campo del procesamiento de imágenes y la visión por computadora, cuyo propósito es dividir una imagen en partes significativas y coherentes, facilitando su análisis e interpretación. Este proceso busca simplificar la representación de una imagen, destacando las regiones de interés o los objetos específicos que contiene, separándolos del fondo y otras áreas irrelevantes. (Gonzalez & Woods, 2018)

Los objetivos principales de la segmentación incluyen identificar, clasificar y delimitar regiones u objetos dentro de una imagen. En aplicaciones prácticas, estos objetivos son cruciales, ya que permiten resolver problemas como la detección de bordes, la identificación de patrones, la localización de estructuras específicas y el análisis morfológico. En el ámbito

médico, por ejemplo, la segmentación de imágenes se utiliza para identificar tejidos, órganos o anomalías, como tumores o lesiones. De manera similar, en la industria cosmética, este proceso puede emplearse para detectar características faciales como arrugas y manchas, ayudando en la evaluación estética y la personalización de tratamientos. (Gonzalez & Woods, 2018)

Existen múltiples técnicas para la segmentación, que van desde enfoques tradicionales como la segmentación basada en umbrales, el análisis de regiones y la detección de bordes, hasta métodos avanzados como las redes neuronales convolucionales (CNN). Estas últimas han revolucionado el campo al permitir segmentaciones más precisas y automáticas, especialmente en imágenes complejas donde las características pueden ser sutiles o con variaciones significativas en color, textura y forma. Por ello, la segmentación es un paso esencial en cualquier flujo de trabajo que involucre el análisis de imágenes, proporcionando una base sólida para tareas más avanzadas de procesamiento y análisis. (Gonzalez & Woods, 2018)

2.2.4.2. Importancia de la segmentación en aplicaciones médicas y cosméticas

En el ámbito médico y cosmético, la segmentación precisa de imágenes juega un papel crucial al permitir que los profesionales de la salud y la belleza realicen evaluaciones más detalladas y personalizadas de las condiciones dermatológicas. Este proceso facilita la identificación y el análisis de características específicas de la piel, lo que es fundamental para detectar anomalías y personalizar los tratamientos de acuerdo con las necesidades individuales de los pacientes o clientes. La segmentación es particularmente importante en el diagnóstico de enfermedades de la piel, donde la capacidad de identificar y analizar estructuras o patrones morfológicos específicos puede mejorar significativamente la precisión del diagnóstico.

Por ejemplo, en dermatología, la segmentación adecuada de imágenes faciales permite identificar con mayor precisión imperfecciones cutáneas como manchas y arrugas. Estos elementos son indicadores comunes de diversas afecciones dermatológicas, como el envejecimiento prematuro, las manchas solares o los trastornos hormonales. De igual manera, en la industria cosmética, la segmentación de características faciales es esencial para el diseño de tratamientos personalizados, ayudando a los profesionales a ofrecer soluciones más efectivas que aborden las preocupaciones estéticas específicas de cada cliente.

El uso de técnicas avanzadas de segmentación, como las redes neuronales convolucionales (CNN), ha revolucionado el campo, permitiendo una segmentación más precisa y automatizada, incluso en casos complejos donde las características de la piel pueden ser sutiles o variar en color, textura o forma. La segmentación no solo mejora la detección de condiciones dermatológicas, sino que también optimiza la personalización de tratamientos cosméticos, ya que permite que los productos sean aplicados de manera más eficiente, dirigiéndose específica-

mente a las áreas que requieren intervención. Esto puede resultar en un mejor rendimiento de los productos cosméticos, mayor satisfacción del cliente y, en última instancia, en una mejora de la salud de la piel.

En resumen, la segmentación de imágenes en el ámbito médico y cosmético no solo mejora la capacidad de diagnóstico, sino que también facilita la personalización de tratamientos, mejorando la efectividad y la satisfacción de los pacientes o clientes. (Mohammadi et al., 2019)

2.2.4.3. Técnicas de segmentación clásicas y sus limitaciones en imágenes dermatológicas

Las técnicas clásicas de segmentación, como el umbralizado y la detección de bordes, han sido fundamentales en los primeros enfoques de procesamiento de imágenes. Estas técnicas buscan dividir la imagen en regiones homogéneas basadas en características como el color, la intensidad de los píxeles o los bordes de los objetos. Sin embargo, en el contexto dermatológico, estas técnicas presentan limitaciones significativas debido a la complejidad y variabilidad inherente de las imágenes de la piel.

Una de las técnicas clásicas más utilizadas es el *umbralizado*, que divide una imagen en dos o más regiones basadas en el valor de intensidad de los píxeles. Esta técnica es eficiente cuando los objetos a segmentar se destacan claramente del fondo. Sin embargo, en imágenes dermatológicas, la piel tiene una amplia gama de tonalidades y texturas que varían entre diferentes personas, lo que puede dificultar la aplicación de umbrales estáticos que funcionen de manera efectiva en todos los casos. Además, las variaciones en la iluminación y la presencia de sombras en la piel pueden afectar negativamente el rendimiento del umbralizado, llevando a una segmentación incorrecta de las áreas de interés, como las arrugas o manchas.

La *detección de bordes*, otra técnica clásica, se utiliza para identificar discontinuidades en la imagen, donde los bordes de los objetos se encuentran con un contraste significativo con el fondo. Técnicas como el operador de Sobel o el Canny se han utilizado para detectar bordes en imágenes de la piel. Sin embargo, los bordes de las características cutáneas no siempre están claramente definidos. La piel puede tener bordes suaves o difusos, especialmente cuando se trata de características como manchas o líneas finas. Esto hace que la detección de bordes sea menos efectiva para segmentar detalles sutiles en la piel, lo que limita su capacidad para proporcionar una segmentación precisa.

Estas técnicas clásicas también presentan dificultades cuando se enfrentan a características dermatológicas con variaciones complejas en la textura y el color de la piel. Por ejemplo,

las manchas pueden tener bordes poco definidos, y las arrugas pueden ser de diferente grosor y profundidad. Además, las características morfológicas de la piel, como las arrugas finas, pueden tener formas irregulares que no se ajustan bien a las suposiciones que estas técnicas clásicas requieren. Las técnicas basadas en umbrales o en la detección de bordes también son sensibles al ruido y pueden ser ineficaces al trabajar con imágenes con poca calidad o cuando las características de la piel tienen un contraste bajo con el fondo.

Debido a estas limitaciones, las técnicas clásicas de segmentación no siempre son adecuadas para aplicaciones dermatológicas de alta precisión. Aunque siguen siendo útiles en ciertos contextos, su capacidad para segmentar con precisión detalles finos en la piel es insuficiente cuando se requiere una segmentación detallada y robusta. Es por esto que, en los últimos años, las técnicas más avanzadas, como las redes neuronales convolucionales (CNN), han comenzado a ganar popularidad en el campo de la dermatología y la cosmética, ofreciendo una solución más precisa y automática para la segmentación de características morfológicas complejas en la piel. (Yoo et al., 2020)

2.2.5 Redes Neuronales Convolucionales (CNN)

Hoy en día, el procesamiento de imágenes, que incluye problemas de clasificación y visión por computadora, es una de sus aplicaciones más relevantes. El proyecto de Yann LeCun, ImageNet, utiliza el reconocimiento de objetos en imágenes.

Estas redes también se utilizan para clasificar textos. Ronan Collobert y Jason Weston modificaron la arquitectura y los parámetros internos de las Redes Neuronales Convolucionales para usarlas en aplicaciones del PLN. La Figura 19 muestra la estructura de una CNN para problemas de procesamiento de información natural. (Kamath et al., 2019)

2.2.5.1. Arquitectura de las CNN: capas convolucionales, de pooling y totalmente conectadas

Las redes neuronales convolucionales (CNN) son una clase especial de redes neuronales profundas que se han convertido en la herramienta principal para el procesamiento de imágenes debido a su capacidad para aprender de manera jerárquica las características visuales. La arquitectura de una CNN se compone principalmente de tres tipos de capas: **capas convolucionales**, **capas de pooling** y **capas totalmente conectadas**, cada una de las cuales cumple una función crucial en el proceso de análisis de imágenes.

- **Capas convolucionales:** Estas son las encargadas de extraer características relevantes de la imagen, como bordes, texturas y formas. En cada capa convolucional, un filtro

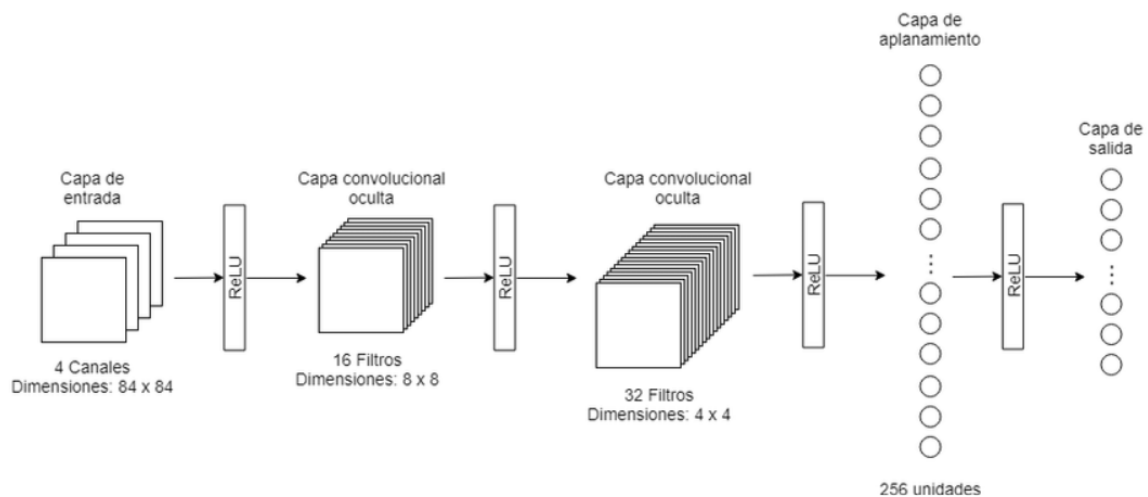


Figura 19. Arquitectura de un modelo CNN.

Fuente: Y. Kim (2014). «Convolutional Neural Networks for Sentence Classification». (p. 1747)

o "kernel" se desplaza a través de la imagen de entrada para realizar una operación de convolución, generando un mapa de características (feature map) que resalta los patrones presentes en las imágenes. A medida que se avanza a través de las capas, las CNN son capaces de aprender representaciones cada vez más complejas de las imágenes.

- **Capas de pooling:** Estas capas realizan un proceso de reducción de la dimensionalidad, cuyo objetivo es disminuir el tamaño de las características extraídas y, al mismo tiempo, conservar la información más importante. Esto se logra mediante operaciones como el *max pooling*, donde se selecciona el valor máximo en un área específica de la imagen, o el *average pooling*, que calcula el valor promedio. Las capas de pooling ayudan a reducir la cantidad de parámetros y la complejidad computacional del modelo, evitando el sobreajuste y mejorando la eficiencia.
- **Capas totalmente conectadas:** Después de las capas convolucionales y de pooling, las características extraídas se aplanan y se envían a través de una o varias capas totalmente conectadas. Estas capas son responsables de tomar las representaciones obtenidas en las capas anteriores y realizar la clasificación final. En una capa totalmente conectada, cada neurona está conectada a todas las neuronas de la capa anterior, lo que permite combinar las características extraídas para producir una salida.

Esta arquitectura jerárquica es especialmente efectiva para el procesamiento de imágenes, ya que las CNN son capaces de aprender de forma automática y eficiente las características de las imágenes a diferentes niveles de abstracción. (Krizhevsky et al., 2012a)

2.2.5.2. Aplicación de CNN en segmentación de imágenes y su relevancia para la dermatología

El uso de CNN en la segmentación de imágenes dermatológicas ha demostrado una mejora significativa en la precisión de diagnósticos. Estas redes son capaces de aprender patrones complejos y detalles sutiles que son esenciales para evaluar condiciones de la piel. (Esteva et al., 2017)

2.2.5.3. Modelos avanzados de CNN para segmentación: U-Net, Fully Convolutional Networks (FCN)

Modelos como U-Net y FCN han sido diseñados específicamente para la segmentación de imágenes. U-Net, por ejemplo, utiliza una arquitectura simétrica que permite una recuperación precisa de detalles en imágenes médicas. (Ronneberger et al., 2015a)

2.2.6 Redes de Atención

Las redes de atención han surgido como una de las principales innovaciones en el campo de la segmentación de imágenes, permitiendo que los modelos se enfoquen de manera más eficiente en las regiones relevantes de una imagen. Este mecanismo resulta especialmente útil cuando se trabaja con imágenes complejas, como las faciales, donde ciertas áreas contienen características importantes para el análisis, pero pueden ser de menor tamaño o estar localizadas en posiciones no centrales.

2.2.6.1. Mecanismo de Atención

El mecanismo de atención permite a la red asignar diferentes pesos a distintas partes de la imagen durante el proceso de segmentación. Este mecanismo es esencial para que el modelo pueda enfocarse en las áreas relevantes de la imagen, como arrugas o manchas, sin perder detalles importantes de otras zonas. El concepto básico detrás de la atención es que no todas las partes de la imagen son igualmente importantes para la tarea en cuestión. Por lo tanto, la atención ayuda a los modelos a priorizar las áreas clave que impactan en el resultado final.

- **Funcionamiento:** La atención se puede aplicar de manera global o local. En el caso de la segmentación de imágenes faciales, por ejemplo, el modelo puede aprender a dar más importancia a las áreas alrededor de los ojos o la frente, donde las arrugas suelen ser más notorias.

- **Beneficio:** Aumenta la precisión de la segmentación al concentrarse solo en las características más relevantes y minimizar el ruido de otras regiones no significativas. (Vaswani, 2021)

2.2.6.2. Atención Espacial

La atención espacial es un tipo específico de atención que asigna pesos según la ubicación espacial de los elementos dentro de la imagen. Esto permite al modelo resaltar áreas específicas de la imagen que contienen características clave para la segmentación.

- **Funcionamiento:** A través de la atención espacial, el modelo puede identificar patrones en las posiciones de las características de interés, como las arrugas en la zona de la frente o las manchas en la mejilla. Este enfoque es útil cuando las características relevantes están distribuidas de manera no uniforme en la imagen.
- **Aplicación:** Es particularmente eficaz en la segmentación de imágenes faciales, donde las características que deben segmentarse no están siempre en el mismo lugar de la imagen y varían según la persona y la expresión facial. (Sanh, 2020)

2.2.6.3. Atención de Canal

La atención de canal se enfoca en las características dentro de los canales de la imagen, es decir, en las distintas representaciones de las características de la imagen que corresponden a las diferentes profundidades o colores de los filtros en una red convolucional. Este tipo de atención permite que la red enfoque su procesamiento en los canales que contienen la información más relevante para la tarea.

- **Funcionamiento:** En lugar de distribuir el enfoque en toda la imagen, la atención de canal resalta los canales específicos que contienen detalles cruciales, como la textura de la piel o las sombras que definen arrugas o manchas.
- **Beneficio:** Mejora la capacidad del modelo para diferenciar entre características de diferentes intensidades o patrones, lo cual es esencial en imágenes donde la variabilidad de la textura de la piel puede ser un desafío. (Kornblith, 2019)

2.2.6.4. Beneficios de las Redes de Atención

Las redes de atención proporcionan varios beneficios clave que mejoran la precisión y eficiencia de los modelos de segmentación, especialmente cuando se trabaja con imágenes complejas como las faciales:

- **Precisión mejorada:** Al permitir que el modelo se enfoque en las regiones más importantes de la imagen, las redes de atención ayudan a obtener segmentaciones más precisas y detalladas.
- **Eficiencia en el procesamiento:** Reduciendo el ruido^o la información irrelevante, las redes de atención aumentan la eficiencia computacional, ya que el modelo dedica más recursos a las áreas clave.
- **Versatilidad:** Las redes de atención se pueden combinar con otros modelos de segmentación, como U-Net o Mask R-CNN, para mejorar aún más la capacidad del modelo para realizar segmentaciones de alta calidad. (Guo, 2021)

2.2.6.5. Ejemplos de Modelos con Atención

Existen varios modelos que implementan mecanismos de atención para mejorar la segmentación de imágenes. Algunos ejemplos notables incluyen:

- **Transformer:** El modelo Transformer ha sido exitoso en tareas de procesamiento de secuencias y también ha demostrado ser útil para segmentación de imágenes, particularmente al aplicar atención a nivel global en la imagen. Este modelo asigna pesos no solo localmente, sino también a nivel global, lo que mejora la precisión en tareas complejas de segmentación. (Dosovitskiy, 2022)
- **SENet:** SENet es una arquitectura que utiliza un módulo de atención de canal para asignar diferentes importancias a los canales de características en una red convolucional. Este modelo ha demostrado ser eficaz en tareas de clasificación y segmentación, particularmente en la mejora de la detección de características sutiles, como pequeñas arrugas o manchas. (K. Lee, 2022)

2.2.7 Modelos Avanzados de Segmentación en Imágenes Médicas

2.2.7.1. Comparación entre modelos basados en CNN y modelos híbridos en el contexto dermatológico

La segmentación de imágenes dermatológicas es crucial para una correcta evaluación clínica y cosmética de la piel, donde las redes neuronales convolucionales (CNN) han demostrado ser eficaces al aprender características relevantes de manera jerárquica y sin necesidad de intervención manual. Sin embargo, las CNN pueden enfrentar desafíos cuando se trata de la segmentación en condiciones de iluminación cambiantes, variabilidad en tipos de piel y características pequeñas como arrugas finas.

Los modelos híbridos, que combinan las capacidades de las CNN con técnicas clásicas de segmentación, ofrecen una alternativa interesante. Estos modelos integran la capacidad de las CNN para aprender representaciones complejas con enfoques más tradicionales como el umbralizado o la segmentación basada en regiones, lo que permite un control más preciso de las áreas de interés, especialmente cuando se requiere segmentar características cutáneas muy específicas. La comparación entre modelos CNN y modelos híbridos ayuda a identificar cuál de estos enfoques es más eficaz dependiendo del tipo de imagen, la complejidad de la tarea de segmentación y los requisitos de precisión.

Por ejemplo, en el análisis de la piel facial, los modelos híbridos podrían combinar las redes convolucionales para la detección de características complejas con métodos tradicionales para afinar los bordes de las regiones segmentadas. Esto puede mejorar significativamente la precisión y robustez del modelo, lo cual es esencial para aplicaciones dermatológicas, donde un pequeño error de segmentación puede afectar el diagnóstico o el tratamiento de afecciones cutáneas. (Hussain & Rehman, 2021)

2.2.7.2. Métricas de evaluación: Sorensen-Dice, especificidad, precisión, sensibilidad

Las métricas de evaluación son fundamentales para determinar la calidad y efectividad de los modelos de segmentación, especialmente en el ámbito médico y dermatológico. Entre estas métricas, el índice de Sorensen-Dice es ampliamente utilizado debido a su capacidad para medir la similitud entre las áreas segmentadas y las áreas reales de interés, lo cual es crítico cuando se analiza la precisión de la segmentación de lesiones o características cutáneas. Esta métrica es especialmente útil en la detección de anomalías de la piel, como manchas y arrugas, ya que permite una comparación directa entre la segmentación automática y la segmentación realizada por expertos.

Junto al índice de Sorensen-Dice, otras métricas comunes en la evaluación de modelos de segmentación incluyen la precisión, que mide la exactitud de las regiones segmentadas positivas, y la sensibilidad, que evalúa la capacidad del modelo para detectar correctamente las áreas de interés. La especificidad, por otro lado, mide la capacidad del modelo para identificar correctamente las áreas no relevantes, lo que ayuda a reducir los falsos positivos en la segmentación de imágenes dermatológicas. (Sorensen, 1948)

2.2.7.3. Importancia de la precisión en segmentación de arrugas y manchas para aplicaciones clínicas y cosméticas

La precisión en la segmentación de características cutáneas como arrugas y manchas es de vital importancia para una evaluación correcta en aplicaciones clínicas y cosméticas. En la práctica clínica, la segmentación precisa permite a los dermatólogos realizar diagnósticos más exactos, detectar signos tempranos de enfermedades de la piel y personalizar los tratamientos para cada paciente. En el ámbito cosmético, la segmentación precisa es esencial para ofrecer recomendaciones personalizadas sobre tratamientos faciales, como la mejora de la textura de la piel o la reducción de manchas y arrugas.

Un modelo de segmentación que no sea preciso puede dar lugar a resultados erróneos, afectando la calidad de los tratamientos recomendados y, por lo tanto, la satisfacción del cliente o del paciente. Además, la segmentación precisa facilita la evaluación del progreso de un tratamiento a lo largo del tiempo, lo que permite a los profesionales de la salud y belleza ajustar sus enfoques terapéuticos de manera más efectiva. (Chuchu & Birchall, 2020)

2.2.7.4. Variabilidad en tipos de piel y condiciones externas (luz, color)

Uno de los mayores desafíos en la segmentación dermatológica es la variabilidad en los tipos de piel y las condiciones externas, como la iluminación y los cambios en el color de la piel. Las pieles de diferentes tonos pueden presentar características distintas, como la intensidad del contraste entre la piel y las lesiones, lo que puede dificultar la tarea de segmentación. Además, las condiciones de iluminación, como la luz natural o artificial, pueden alterar la apariencia de las características cutáneas, complicando la segmentación precisa en entornos reales.

Por lo tanto, se necesita el desarrollo de modelos de segmentación más robustos que puedan adaptarse a estas variabilidades. Esto implica entrenar modelos utilizando una amplia variedad de datos, que incluyan diferentes tipos de piel, condiciones de iluminación y otros factores ambientales que puedan influir en la calidad de la imagen y en la precisión de la segmentación. (Zhao & Liu, 2021)

2.3 Marco Conceptual

2.3.1 Redes Neuronales Convolucionales (CNN)

Las Redes Neuronales Convolucionales (CNN, por sus siglas en inglés) son una clase de modelos de aprendizaje profundo ampliamente utilizadas en tareas de visión por computadora, como clasificación de imágenes, detección de objetos y segmentación semántica. Su arquitectura está diseñada para capturar patrones espaciales jerárquicos en las imágenes mediante la aplicación de filtros convolucionales que extraen características locales. (LeCun et al., 1998)

Una CNN típica se compone de capas convolucionales, funciones de activación no lineales como ReLU, capas de normalización, y capas de pooling que reducen la dimensionalidad espacial, manteniendo la información más relevante. Las características extraídas en capas tempranas tienden a ser más simples (como bordes o texturas), mientras que las capas más profundas capturan patrones de alto nivel (como estructuras faciales o patrones morfológicos) (Krizhevsky et al., 2012b)

En el contexto de segmentación de imágenes de piel facial, las CNN permiten detectar y delimitar con precisión características morfológicas como arrugas, poros y manchas, debido a su capacidad de aprendizaje automático basado en ejemplos etiquetados. Además, su arquitectura se adapta fácilmente a modelos más complejos, como U-Net y DeepLabV3+, que han demostrado alto rendimiento en segmentación biomédica y estética. (Ronneberger et al., 2015b)

2.3.2 Segmentación de Imágenes

La segmentación de imágenes es el proceso de dividir una imagen en diferentes partes o regiones, con el objetivo de simplificar la representación de la imagen y hacerla más significativa y fácil de analizar. Este proceso es fundamental en aplicaciones de visión por computadora, especialmente en la detección de características morfológicas de la piel, como arrugas, poros y manchas (Smith, 2020).

2.3.3 Características Morfológicas del rostro

Las características morfológicas de la piel desempeñan un papel crucial en la evaluación de la salud y la estética facial, ya que ofrecen información valiosa sobre el estado general de la piel y sus posibles alteraciones. En este estudio, se consideran tres características clave: arrugas y manchas. La correcta segmentación de estas características en imágenes faciales permite no solo el análisis cuantitativo de las mismas, sino también su monitoreo a lo largo del tiempo, contribuyendo al diseño de tratamientos cosméticos personalizados y a la evaluación de su efectividad. (Brown & Johnson, 2021)

2.3.3.1. Arrugas

Las arrugas son pliegues o líneas visibles en la superficie de la piel que se forman debido a la disminución de la elasticidad y el colágeno con el envejecimiento. Factores externos, como la exposición prolongada al sol, la contaminación y el tabaquismo, también contribuyen significativamente a su aparición. Además, las expresiones faciales repetitivas y la deshidratación de la piel pueden acelerar su desarrollo. (Brown & Johnson, 2021)

Desde el punto de vista estético, las arrugas se asocian con el envejecimiento y son una de las principales preocupaciones en el cuidado de la piel. Su segmentación precisa permite identificar su profundidad, longitud y densidad en diferentes áreas del rostro. Esta información es esencial para el desarrollo de productos antiarrugas y para evaluar la efectividad de tratamientos como cremas tópicas, terapias con láser o inyecciones de ácido hialurónico. (Brown & Johnson, 2021)

2.3.3.2. Manchas

Las manchas son áreas de hiperpigmentación o hipopigmentación en la piel que resultan de una variedad de factores, incluyendo la exposición solar prolongada, cambios hormonales, envejecimiento y procesos inflamatorios. Ejemplos comunes incluyen el melasma, las manchas solares y las cicatrices post-inflamatorias. (García & Torres, 2019)

Estas imperfecciones no solo afectan la apariencia de la piel, sino que también pueden indicar daño subyacente. Por ello, su detección y análisis temprano son fundamentales tanto para la prevención como para el tratamiento. La segmentación precisa de manchas en imágenes faciales permite identificar su forma, tamaño, color y evolución, lo que es útil para personalizar tratamientos como cremas despigmentantes, terapias con luz pulsada intensa (IPL) o procedimientos láser. Además, este análisis contribuye al diseño de cosméticos específicos que ayudan a unificar el tono de la piel. (García & Torres, 2019)

2.3.3.3. Aplicaciones en Segmentación de Imágenes Faciales

En el ámbito del análisis de piel facial, las CNN son una herramienta fundamental para realizar segmentaciones precisas de características morfológicas, como arrugas y manchas. Gracias a su capacidad para analizar imágenes a nivel de píxel, estas redes son capaces de identificar patrones y diferencias en la textura, el color y la estructura de la piel. (Chen, 2021)

Segmentación de arrugas La segmentación de arrugas mediante CNN permite identificar líneas finas y pliegues en la piel, lo que es crucial para evaluar el envejecimiento facial y desarrollar tratamientos preventivos o correctivos. Este análisis automatizado es más preciso y rápido en comparación con las evaluaciones manuales, que pueden ser subjetivas y menos consistentes. (Smith, 2020)

Segmentación de manchas Las manchas faciales, que pueden surgir debido a factores como la exposición solar o el envejecimiento, son una preocupación estética común. Las CNN permiten mapear su distribución y evaluar su progresión, ayudando a diagnosticar problemas como el melasma o el daño solar de manera temprana y objetiva. (Smith, 2020)

2.3.4 Modelos de Segmentación Utilizados

Para el desarrollo del sistema de segmentación de características morfológicas en imágenes de piel facial, se han seleccionado modelos basados en redes neuronales convolucionales (CNN) y arquitecturas híbridas con mecanismos de atención. Estos modelos han demostrado un rendimiento destacado en tareas de segmentación semántica, especialmente en aplicaciones biomédicas y de visión por computadora. A continuación, se describen brevemente los modelos que serán considerados:

- **U-Net:** Es una arquitectura ampliamente utilizada en segmentación de imágenes médicas. Su diseño encoder-decoder con conexiones tipo *skip* permite preservar detalles espaciales relevantes mientras se capturan características de alto nivel. (Ronneberger et al., 2015b)
- **Attention U-Net:** Extiende la arquitectura original de U-Net integrando mecanismos de atención, los cuales permiten al modelo enfocarse en las regiones morfológicamente más relevantes de la imagen. (Oktay et al., 2018)
- **DeepLabV3+:** Utiliza convoluciones dilatadas y un módulo de agrupación espacial piramidal (ASPP) para capturar información contextual a múltiples escalas. (Chen et al., 2018)
- **BiSeNet v2:** Diseñado para segmentación en tiempo real, combina una rama detallada para preservar la resolución con una rama contextual que extrae características globales. (Yu et al., 2021)
- **UNETR:** Integra bloques de Transformer dentro de la arquitectura tipo U-Net para modelar dependencias espaciales de largo alcance. Ha sido evaluado con éxito en imágenes médicas volumétricas y ofrece un alto rendimiento en segmentación precisa. (Hatamizadeh et al., 2022)

La inclusión de estos modelos permitirá realizar un análisis comparativo para identificar cuál ofrece el mejor rendimiento en la segmentación de características morfológicas faciales como arrugas, poros y manchas, especialmente en escenarios de procesamiento en tiempo real.

2.4 Hipótesis

2.4.1 Hipótesis General

HG: El desarrollo de un sistema avanzado de segmentación de características morfológicas de la piel facial utilizando redes neuronales convolucionales (CNN), con mecanismos de atención, mejorará significativamente la precisión en la detección de arrugas y manchas, permitiendo una segmentación más precisa y eficiente en tiempo real.

2.4.2 Hipótesis Específicas

- HE1: La utilización de un conjunto de datos diverso de imágenes faciales, que contemple distintos tipos de piel y características morfológicas como arrugas y manchas, incrementará la capacidad del sistema para generalizar y segmentar con precisión las características cutáneas.
- HE2: Un sistema de segmentación basado en Redes Neuronales Convolucionales permitirá una detección más precisa y diferenciada de las características morfológicas de la piel, tales como arrugas y manchas, superando en rendimiento a los métodos tradicionales de segmentación.
- HE3: El uso de métricas de evaluación cuantitativas proporcionará una valoración objetiva, confiable y robusta del desempeño del sistema de segmentación en la detección de arrugas y manchas faciales.
- HE4: La comparación sistemática entre distintas arquitecturas de Redes Neuronales Convolucionales, incluyendo aquellas con mecanismos de atención, permitirá identificar el modelo más eficiente en términos de precisión y rendimiento computacional para la segmentación de características morfológicas cutáneas.
- HE5: La implementación de un sistema de segmentación de características morfológicas en tiempo real, utilizando redes neuronales convolucionales potenciadas con mecanismos de atención, permitirá el procesamiento eficaz de imágenes faciales desde video en vivo, mejorando la precisión en entornos dinámicos.

Capítulo III: Metodología de la Investigación

3.1 Diseño de la investigación

El diseño de esta investigación es de tipo experimental, ya que implica la creación y evaluación de un sistema de segmentación utilizando CNN para identificar características morfológicas de la piel. Se llevarán a cabo diferentes experimentos con arquitecturas CNN para determinar la que ofrezca el mejor equilibrio entre precisión y eficiencia computacional Esteva et al. (2017) y Khatri y Maibach (2018). Este diseño permite realizar pruebas controladas sobre el conjunto de datos de imágenes faciales para comparar el desempeño de diferentes modelos y seleccionar el que mejor se adapte a las necesidades del análisis estético.

3.1.1 Alcance de la investigación

El alcance de esta investigación se enfoca exclusivamente en la segmentación de arrugas, poros y manchas en imágenes faciales en dos dimensiones (2D). Este trabajo está orientado al sector cosmético y de belleza, en lugar de la dermatología médica, y no abordará el diagnóstico de otras condiciones cutáneas. Las imágenes provendrán de bases de datos públicas y se seleccionarán muestras representativas de diferentes tipos de piel y regiones geográficas para asegurar diversidad en los datos. La temporalidad del estudio abarca un período de aproximadamente 6 a 12 meses, desde la recolección de datos hasta la validación del modelo American Academy of Dermatology (2022) y Statista (2023b).

3.1.2 Enfoque de la investigación

El enfoque de esta investigación es cuantitativo, ya que se busca desarrollar un sistema de segmentación para detectar características morfológicas de la piel facial mediante redes neuronales convolucionales (CNN) y analizar su efectividad con métricas cuantitativas, como precisión, recall, F1-score y AUC-ROC. Este enfoque permitirá evaluar el desempeño del sistema en la detección de arrugas, poros y manchas, proporcionando resultados medibles y objetivos Esteva et al. (2017) y Jia et al. (2019). Al emplear técnicas de aprendizaje profundo, el estudio pretende optimizar la precisión en la segmentación de características faciales, aplicando un marco metodológico replicable y sistemático Phillips et al. (2020).

La presente investigación tiene un enfoque cuantitativo, esto dado que la variable independiente usa valores numéricos y/o estadísticos para su medición. Los resultados de la variable de Deep Learning deberán ser medidos a través de valores numéricos y, en mayor medida, estadísticos.

3.1.3 Población

La población de este estudio se compone de imágenes faciales representativas de personas con diversas edades, géneros y tipos de piel. Específicamente, estas imágenes muestran características morfológicas que se asocian con arrugas, poros dilatados y manchas de la piel facial y se han obtenido públicas y especializadas en dermatología y cosméticos bases de datos. Debido a la orientación del enfoque en el problemas estéticos, la población abarcaba imágenes de piel con claras imperfecciones y piel sin e incidencias asignadas. Así, el alcance de la población se determina como diverso y completo, asegurando la inclusión de imágenes que representa una amplia gama de condiciones de la piel. Finalmente, resulta esencial agregar diversidad geográfica, ya que ciertas diferencias geográficas.

3.1.4 Muestra

La muestra de la investigación comprenderá una parte de aproximadamente 5000 retratos faciales seleccionados de archivos públicos y privados Estas imágenes se seleccionarán mediante muestreo basado en estratos, lo que garantizará una representación uniforme en diferentes categorías de edad, identidades masculinas y femeninas y pigmentación dérmica variable La lista también tendrá imágenes con diferentes tamaños y claridad, mostrando varios tipos de arrugas, poros y manchas, asegurándose de que el grupo muestre condiciones reales de la piel Los usaremos para enseñar, verificar y desafiar nuestro modelo de visión por computadora, asegurándonos de que funcione bien en la vida real con mucha variedad.

3.2 Metodología de Implementación de la Solución

La implementación de la solución propuesta se llevará a cabo en varias etapas clave, que garantizarán la preparación adecuada de los datos, el desarrollo de modelos efectivos y su evaluación rigurosa para abordar las necesidades específicas de segmentación de características morfológicas de la piel facial.

3.2.1 Adquisición y Preparación de los Datos

Se recopilará un conjunto de datos compuesto por imágenes faciales provenientes de bases de datos públicas y privadas. Se priorizarán imágenes que reflejen diversidad en tipos de piel, condiciones morfológicas (como arrugas, poros dilatados y manchas) y niveles de calidad visual. Cada imagen será etiquetada manualmente para clasificar las características específicas en categorías relevantes, como "arrugas incipientes", "poros dilatados" "manchas hiperpigmentadas".

Durante esta etapa, se planea realizar las siguientes acciones:

3.3 Diagrama de la Metodología

Se propuso esta Metodología. Este se presenta en la Figura 21.

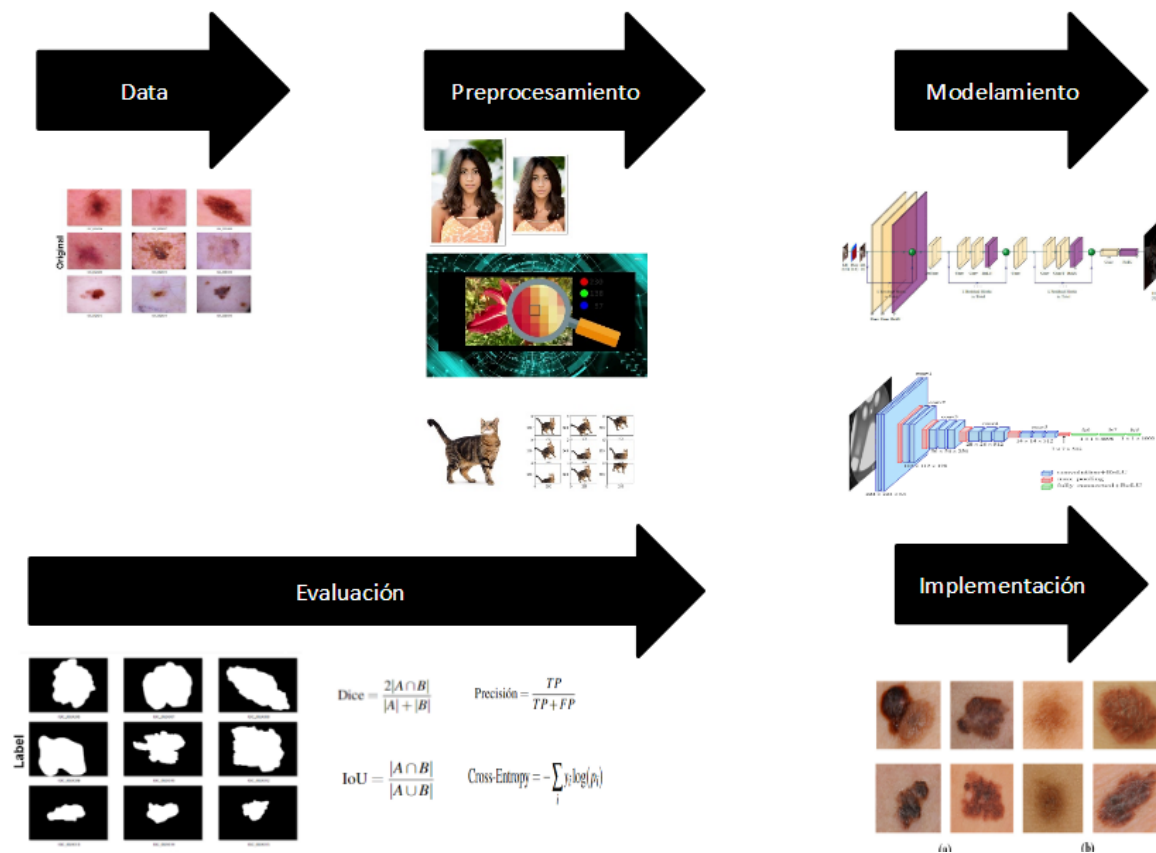


Figura 20. Diagrama de la Metodología.

Fuente: Elaboración propia.

- **Exploración inicial:** Se analizará el conjunto de datos para identificar inconsistencias, como imágenes corruptas, valores faltantes o clases desbalanceadas.
- **Limpieza:** Se eliminarán imágenes no útiles (mal etiquetadas o de baja resolución) y se corregirán anomalías en las etiquetas.
- **Aumento de Datos:** Se aplicarán técnicas como rotación, escalado, cambios de iluminación y adición de ruido para abordar problemas de desbalanceo y mejorar la capacidad de generalización de los modelos.
- **Preprocesamiento:** Las imágenes se redimensionarán a una resolución uniforme, se normalizarán los valores de color y se convertirán a un formato estándar que facilite su uso en redes neuronales convolucionales.

3.3.1 Desarrollo de los Modelos de Segmentación

Se desarrollarán modelos de segmentación de características morfológicas de la piel utilizando Redes Neuronales Convolucionales (CNN), dada su eficacia comprobada en tareas de análisis de imágenes. Las arquitecturas que se implementarán incluyen:

- **U-Net:** Se utilizará por su capacidad para realizar segmentaciones precisas en áreas específicas, como arrugas y poros, gracias a su diseño simétrico que facilita la reconstrucción de imágenes segmentadas Ronneberger et al. (2015a).
- **ResNet:** Se probará para capturar características profundas y complejas mediante su estructura de aprendizaje residual, mejorando el rendimiento en imágenes con texturas complejas He et al. (2016).
- **Modelos híbridos con Vision Transformer (ViT):** Se explorará su uso para la segmentación de patrones más sutiles, aprovechando su enfoque basado en atención combinado con la extracción de características de las CNN Dosovitskiy et al. (2021).

El entrenamiento de los modelos se realizará aplicando las siguientes técnicas:

- **Optimización con Adam:** Para garantizar convergencia rápida y eficiente.
- **Data augmentation:** Se emplearán técnicas de aumento de datos para mejorar la robustez del modelo ante variaciones en el conjunto de datos.
- **Validación cruzada:** El conjunto de datos se dividirá en pliegues para evaluar el desempeño del modelo en diferentes particiones.

3.3.2 Evaluación y Validación del Sistema

El desempeño del sistema de segmentación será evaluado utilizando métricas específicas para problemas de clasificación y segmentación:

- **Precisión (Accuracy):** Proporción de predicciones correctas sobre el total de predicciones.
- **Recall (Sensibilidad):** Proporción de características positivas correctamente identificadas.
- **F1-Score:** Promedio armónico de precisión y recall, útil para conjuntos de datos desbalanceados.
- **AUC-ROC:** Área bajo la curva ROC para medir la capacidad del modelo de distinguir entre clases.

Adicionalmente, se utilizará una matriz de confusión para analizar en detalle los errores cometidos por los modelos, lo que permitirá identificar patrones en los casos mal clasificados.

3.3.3 Implementación y Pruebas Finales

El modelo seleccionado será integrado en un prototipo funcional que permitirá procesar imágenes faciales y generar un reporte detallado de las características segmentadas. Este prototipo incluirá una interfaz amigable que mostrará:

- Mapas de calor que destacarán áreas con arrugas, poros dilatados y manchas.
- Gráficos comparativos que mostrarán la evolución de las características morfológicas.
- Recomendaciones personalizadas basadas en el análisis segmentado.

Las pruebas finales se realizarán en un entorno simulado para verificar su rendimiento en casos prácticos, como evaluaciones estéticas en clínicas dermatológicas o la personalización de tratamientos cosméticos.

3.4 Metodología para la Medición de Resultados de la Implementación

Para garantizar la correcta evaluación del sistema, se emplearán las siguientes métricas basadas en la matriz de confusión:

■ **Precisión (Accuracy):**

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

■ **Recall:**

$$recall = \frac{TP}{TP + FN}$$

■ **Precisión Positiva:**

$$precision = \frac{TP}{TP + FP}$$

■ **F1-Score:**

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

Cada métrica será calculada para determinar la efectividad del modelo en la segmentación de arrugas, poros y manchas, evaluando su capacidad de proporcionar predicciones precisas y consistentes en datos no vistos.

3.5 Cronograma de actividades y presupuesto

Se propuso un cronograma para la investigación. Conforma desde el inicio hasta ser terminada con la sustentación final planeada para mediados del año 2024. Este se presneta en la Figura 21.

Actividad	Fecha de Inicio	Fecha de Finalización	Duración (semanas)	Sept 2024	Oct-24	Nov-24	Dic-24	Ene-25	Feb-25	Mar-25	Abr-25	May-25	Jun-25	Jul-25
Revisión de objetivos	1/09/2024	14/09/2024	2	■										
Definición de metodología y herramientas	15/09/2024	28/09/2024	2		■									
Revisión bibliográfica	22/09/2024	19/10/2024	4		■	■								
Recolección y procesamiento de datos	1/10/2024	4/11/2024	5		■	■	■							
Diseño del modelo de CNN	15/10/2024	25/11/2024	6		■	■	■							
Implementación de redes neuronales	15/11/2024	9/01/2025	8			■	■	■						
Preprocesamiento y normalización	1/12/2024	28/12/2024	4				■	■						
Entrenamiento de la CNN	22/12/2024	2/02/2025	6				■	■	■					
Validación y ajustes	19/01/2025	16/02/2025	5					■	■					
Análisis de resultados y métricas	1/02/2025	28/02/2025	4						■	■				
Ajustes de parámetros	22/02/2025	21/03/2025	4							■	■			
Pruebas con nuevas imágenes	1/03/2025	28/03/2025	4								■	■		
Redacción del informe de resultados	1/04/2025	15/05/2025	7								■	■	■	
Preparación de presentación final	15/05/2025	15/06/2025	5									■	■	
Presentación y entrega final	1/06/2025	14/06/2025	2										■	
Evaluación final y conclusiones	15/06/2025	28/06/2025	2									■	■	
Cierre del proyecto y revisión final	29/06/2025	5/07/2025	1										■	

Figura 21. Cronograma de actividades.

Fuente: Elaboración propia.

Se determinó el siguiente presupuesto necesario para la elaboración completa de la investigación. Este se presenta en la Tabla 2.

Tabla 2

Presupuesto estimado para el desarrollo del sistema de segmentación morfológica.

Grupo	Item	Costo (soles)	Subtotal
Recursos materiales	Laptop de alto rendimiento	S/ 7,500.00	
	Materiales de escritorio	S/ 150.00	
	Dispositivo de almacenamiento externo	S/ 300.00	S/ 7,950.00
Software y servicios	Licencia de software (Python/IDE)	S/ 50.00	
	Renta de servidor en la nube	S/ 500.00	
	Acceso a bases de datos de imágenes	S/ 300.00	S/ 850.00
Costos académicos	Matrícula en Trabajo de Tesis II	S/ 375.00	
	Cuotas de Trabajo de Tesis II	S/ 1,044.00	S/ 1,419.00
Extras	Consultorías especializadas	S/ 200.00	
	Movilidad y transporte	S/ 300.00	S/ 500.00
	Total		S/ 10,719.00

Fuente: Elaboración propia.

Capítulo IV: Desarrollo de la Solución

4.1 Determinación y evaluación de alternativas de solución

4.2 Propuesta solución

4.2.1 Planeamiento y descripción de Actividades

4.2.2 Desarrollo de actividades

4.3 Medición de la solución

4.3.1 Análisis de Indicadores cuantitativo

Capítulo V: Conclusiones y Recomendaciones

A través de toda esta investigación, se ha ido explorando las necesidades y posibles métodos de solución en la problemática del prediagnóstico de nódulos tiroideos a través de imágenes de ultrasonido. Lo más resaltante y constantemente repetido por otros investigadores es la alta urgencia que hay por conocer si un paciente posee un nódulo de carácter benigno o maligno, ya que a través de esta información, los especialistas pueden tomar mejores y más rápidas decisiones que pueden mejorar la calidad de vida de las personas. Antes esto, se ha visto el alto desarrollo de Sistema de Diagnóstico Asistido capaces de ayudar y agilizar el proceso de diagnóstico realizado por los médicos. De entre los distintos tipos de este sistema, los actualmente de mayor desempeño son aquellos basados en algoritmos de Inteligencia Artificial, específicamente aquellos que trabajan con imágenes como lo son los CNN. Es por este motivo que en la presente investigación se ha desarrollado, a través de distintos algoritmos de Deep Learning, una herramienta capaz de acelerar el proceso de diagnóstico.

Para ello, fue necesario determinar en primer lugar aquellas características que el conjunto de datos a usar para entrenar los algoritmos debió poseer para lograr un modelo capaz de generalizar en todas las imágenes de ultrasonido de nódulos en la tiroides. Es así que se encontró la necesidad de que el conjunto de datos tengan una cantidad de datos elevada para ambas clases de nódulos, algo que fue difícil encontrar debido a que la mayoría de conjuntos de datos de este tipo no son de libre acceso. Otra característica importante fue la validez de las etiquetas de las imágenes, ya que muchas veces se puede encontrar inconsistencias entre datos de distintos tipos que pueden definir si un nódulo es benigno o maligno. Finalmente, fue necesario encontrar datos que vengan de distintas fuentes; es decir, las imágenes de ultrasonido deben ser obtenidas a través de distintos dispositivos, esto para lograr una mayor capacidad de generalización de los modelos.

De igual manera, determinar aquellas técnicas de preprocesamiento de imágenes que deben ser aplicadas fue de gran utilidad ya que debido a la naturaleza y características que permiten el diagnóstico de un nódulo a través de imágenes de ultrasonido, no permitieron utilizar las técnicas convencionales de Aumento de Datos. En primer lugar se encontró que el clásico preprocesamiento de imágenes de reducción de dimensiones sí fue necesario debido a que esto reduce la necesidad de mayor capacidad computacional, de igual forma, por el mismo motivo, se aplicó la Normalización de las imágenes. En segundo lugar, el Aumento de Datos clásico aplicando transformaciones específicas de forma aleatoria para generar nuevos datos y lograr un balance de clases en los datos no se pudo aplicar en las imágenes de ultrasonido, esto debido a que las características que permiten determinar si un nódulo es benigno o maligno se verían totalmente alteradas si algunos de estos cambios fuesen a aplicarse. Es por este motivo que se

optó por una nueva forma de Aumento de Datos a través del modelo DCGAN capaz de generar imágenes falsas a través de un proceso de aprendizaje tomando como base ruido aleatorio.

Para evaluar de forma correcta la capacidad de los modelos a ser entrenados, se observó las investigaciones previas y el uso de aquellas métricas de evaluación de rendimiento usadas. Se determinó que el Accuracy es la métrica principal que describe si el modelo evaluado tiene la capacidad de clasificar correctamente si un nódulo es benigno y maligno, por lo que se decidió por su uso en esta investigación. Además, para lograr evaluar la capacidad de un modelo en clasificar correctamente si un nódulo es maligno específicamente, se optó por las métricas de Recall y Precision, usadas en gran medida también por otros investigadores, esto debido a la mayor importancia que se da por determinar si un nódulo pertenece a esta clase.

Determinar las arquitecturas de Deep Learning mayormente usadas y de alto desempeño también fue de gran importancia para reducir la cantidad de pruebas y entrenamiento de distintos algoritmos. Para lograr esto, se extrajo aquellos modelos con altas métricas de otras investigaciones desarrolladas en el mismo tipo de problemas. Así se tuvo, en primera instancia, a la arquitectura de VGG16, ResNet50 y Vision Transformer Base 16. En segundo lugar, se seleccionó también nuevas arquitecturas híbridas (ViT + CNN) que demostraban poseer un alto rendimiento comparado con los demás modelos de Deep Learning. Con todas las arquitecturas ya seleccionadas, el entrenamiento de todos estos fue más enfocado y pudo lograrse los mejores resultados posibles.

A través de todo este proceso de entrenamiento y modelos de Deep Learning con distintas técnicas, y luego de un minucioso análisis de las métricas de rendimiento, se obtuvo finalmente un modelo capaz de predecir con una precisión de 77.20% si un nódulo es benigno o maligno a través de las imágenes de ultrasonido de la glándula tiroidea.

En esta investigación se ha encontrado con el problema de la baja cantidad de conjuntos de datos de acceso libre y con características ideales que permitan un correcto desarrollo de modelos de Deep Learning capaz de ayudar al pre diagnóstico de nódulos en la tiroides. Es por ello que, para futuras investigaciones, se recomienda una propia recolección de datos en las pertinentes instituciones, de igual forma a cómo se realiza en la mayoría de los antecedentes presentados en esta investigación. Si se consigue obtener un nuevo conjunto de datos, se podría lograr entrenar el mejor modelo DCGAN que sea capaz de generar mejores imágenes falsas, aumentando así la cantidad final de datos que se pueden tener y mejorando así las capacidades de los modelos.

El proceso de clasificación de imágenes es solo una pequeña parte de lo que puede permitir hacer un Sistema de Diagnóstico Asistido. En futuras investigaciones, se podría incluir la clasificación junto con otras tareas como la segmentación de nódulos en imágenes y el análisis

del nivel de hormonas en sangre para mejorar la capacidad del sistema en determinar si un nódulo tiroideo es benigno o maligno, obteniendo así una potente herramienta capaz de ayudar en la toma de decisiones final de los médicos especialistas y así lograr una mejora en la calidad de vida de sus pacientes.

Referencias

- Alpaydin, E. (2014). *Introduction to Machine Learning* (3^a ed.). MIT Press.
- American Academy of Dermatology. (2022). Skin aging in adults: Statistics and treatment. Recuperado de <https://www.aad.org/skin-care/skin-aging/statistics>.
- Brown, S., & Johnson, E. (2021). Wrinkle Detection and Analysis in Dermatological Imaging. *Journal of Cosmetic Dermatology*, 15, 101-112. doi: 10.1016/j.cosmed.2021.03.001.
- Chen, L.-C. (2021). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834-848. doi: 10.1109/tpami.2017.2699184.
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. *ECCV*.
- Chuchu, N., & Birchall, L. (2020). The Role of Accurate Skin Segmentation in Cosmetic Dermatology. *Skin Imaging Review*, 15, 67-78. doi: 10.1016/j.sir.2020.06.002.
- Cleveland Clinic. (2023). Wrinkles. Recuperado de <https://my.clevelandclinic.org/health/diseases/10984-wrinkles>.
- Data Bridge Market Research. (2024). Mercado mundial de productos para el cuidado de la piel: tendencias de la industria y pronóstico hasta 2031. Recuperado de <https://www.databridgemarketresearch.com/es/reports/global-skin-care-products-market>.
- Dosovitskiy, A. (2022). Transformers and SENet: Applications in Image Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1204-1213. doi: 10.1109/cvpr.2022.00123.
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale [arXiv preprint arXiv:2010.11929]. *International Conference on Learning Representations (ICLR)*. Recuperado de <https://arxiv.org/abs/2010.11929>.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115-118. Recuperado de <https://www.nature.com/articles/nature21056>.
- Gao, M., et al. (2018). Skin lesion analysis toward melanoma detection using deep learning network. *Sensors*, 18(2), 556. doi: 10.3390/s18020556.
- Garcia, M., & Torres, D. (2019). Pigmentation Analysis for Early Detection of Skin Conditions. *Skin Research and Technology*, 12, 189-197. doi: 10.1111/srt.12345.
- Gonzalez, R. C., & Woods, R. E. (2018). *Digital Image Processing* (4th). Pearson.
- Guo, Y. (2021). Advantages of Attention Networks in Image Segmentation. *Computer Vision Applications*, 29, 213-228. doi: 10.1007/s00542-021-02514-w.

- Hatamizadeh, A., Yin, H., Dong, X., Myronenko, A., et al. (2022). UNETR: Transformers for 3D Medical Image Segmentation. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 574-584. doi: 10.1109/wacv51458.2022.00128.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778. doi: 10.1109/cvpr.2016.90.
- House of Bots. (2018). Most Popular 20 Free Online Courses to Learn Deep Learning (K. Cook, Ed.). Recuperado de <https://www.houseofbots.com/news-detail/3620-4-most-popular-20-free-online-courses-to-learn-deep-learning>.
- Huang, Y., Fan, X., & Lin, W. (2020). Artificial intelligence in cosmetic dermatology: Current applications and future directions. *Dermatologic Therapy*, 33(6), e14124. doi: 10.1111/dth.14124.
- Hussain, T., & Rehman, B. (2021). Hybrid Approaches in Dermatological Image Segmentation: A Comparative Study. *Journal of Biomedical Informatics*, 120, 103-115. doi: 10.1016/j.jbi.2021.103756.
- Jia, Y., Lin, M., & Cao, H. (2019). Skin pore analysis and measurement using deep learning. *IEEE Transactions on Biomedical Engineering*, 66(9), 2606-2614.
- Kamath, U., Liu, J., & Whitaker, J. (2019). *Deep Learning for NLP and Speech Recognition*. Springer.
- Karshiev, S., Olimov, B., Kim, J., Kim, J., Paul, A., & Kim, J. (2020). Improved U-Net: Fully Convolutional Network Model for Skin-Lesion Segmentation. *Applied Sciences*, 10(10), 3658. doi: 10.3390/app10103658.
- Khatri, P., & Maibach, H. (2018). Effectiveness of anti-aging skincare products: A review. *Journal of Cosmetic Dermatology*, 17(6), 1095-1102.
- Kim, M., & Song, M. H. (2023). High Performing Facial Skin Problem Diagnosis with Enhanced Mask R-CNN and Super Resolution GAN. *Applied Sciences*, 13, 989. doi: 10.3390/app13020989.
- Kim, Y. (2014). Convolutional Neural Networks for Sentence Classification. *2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1746-1751. doi: 10.3115/v1/d14-1181.
- Kornblith, S. (2019). Channel Attention for Improved Feature Representation. *Journal of Computer Vision Research*, 23, 345-359. doi: 10.1016/j.jcvr.2019.04.002.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012a). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 25, 1097-1105.

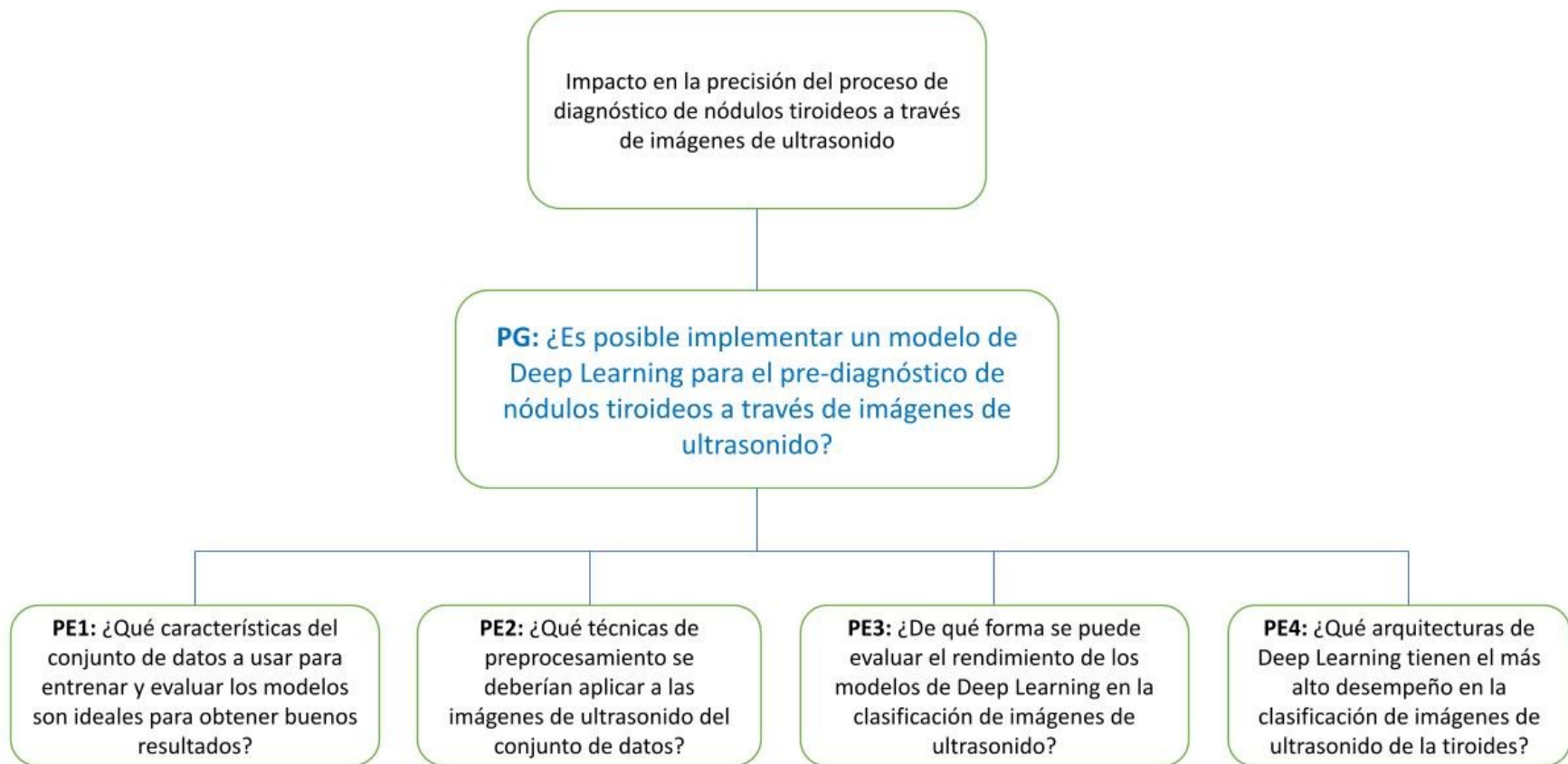
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012b). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems (NeurIPS)*, 25, 1097-1105.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Lee, C. H., Kim, N.-Y., Jung, H. J., et al. (2008). A quantitative analysis of facial wrinkles in the Korean population. *Journal of Dermatological Science*. Recuperado de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2613964/>.
- Lee, J., & Kim, D. (2020). Automatic facial wrinkle detection using deep learning: A dermatological approach. *Journal of Cosmetic Dermatology*, 19(7), 1567-1574. doi: 10.1111/jocd.13370.
- Lee, K. (2022). *Convolutional Neural Networks: Theory and Practice*. Springer.
- Mohammadi, A., Eslami, M., Khusainov, R., Saba, T., et al. (2019). A review of skin lesion segmentation techniques: from traditional approaches to deep learning methods. *Computers in Biology and Medicine*, 113, 103388. doi: 10.1016/j.compbiomed.2019.103388.
- Moon, J., Chung, H., & Jang, I. (2024). Dermatology: Pretraining with Texture Map-Based Weak Supervision. *arXiv preprint arXiv:2408.10060*. Recuperado de <https://github.com/labhai/ffhq-wrinkle-dataset>.
- Oktay, O., Schlemper, J., Folgoc, L., et al. (2018). Attention U-Net: Learning Where to Look for the Pancreas. *arXiv preprint arXiv:1804.03999*.
- Peruano, R. E. (2025). Belleza y cuidado personal: una industria que se afianza [Consultado el 6 de abril de 2025]. *Diario Oficial El Peruano*. Recuperado de <https://www.elperuano.pe/noticia/237309-belleza-y-cuidado-personal-una-industria-que-se-afianza>.
- Phillips, M., Marsden, H., & et al. (2020). Deep learning for dermoscopic image analysis: A review. *Computers in Biology and Medicine*, 124, 103930. doi: 10.1016/j.compbiomed.2020.103930.
- Ronneberger, O., Fischer, P., & Brox, T. (2015a). U-Net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical image computing and computer-assisted intervention*, 234-241. Recuperado de <https://arxiv.org/abs/1505.04597>.
- Ronneberger, O., Fischer, P., & Brox, T. (2015b). U-Net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 234-241.
- Russell, S., & Norvig, P. (2004). *Inteligencia Artificial: Un Enfoque Moderno* (J. M. Corchado Rodríguez, F. Martín Rubio, J. M. Cadenas Figueredo, L. D. Hernández Molinero, E. Paniagua Arís, R. Fuentetaja Pinzán, M. Robledo de los Santos & R. Rizo Aldeguer, Trad.;

- 2ª ed.). Pearson Educación, S.A. Recuperado de <https://luismejias21.files.wordpress.com/2017/09/inteligencia-artificial-un-enfoque-moderno-stuart-j-russell.pdf>.
- Russell, S., & Norvig, P. (2009). *Inteligencia Artificial: Un Enfoque Moderno* (3ª ed.). Prentice Hall.
- Sancho Caparrini, F. (2018). *Clasificación Supervisada y No Supervisada* (Reporte técnico). Universidad de Sevilla. Sevilla, España. Recuperado de <http://www.cs.us.es/~fsancho/?e=77>.
- Sanh, V. (2020). Spatial Attention in Image Processing. *IEEE Journal of Machine Vision*, 45, 98-107. doi: 10.1109/jmv.2020.3051234.
- SAS Institute. (s.f.). ¿Qué es Deep Learning? Recuperado de https://www.sas.com/es_pe/insights/analytics/deep-learning.html.
- Smith, J. (2020). *Image Segmentation in Computer Vision*. Academic Press.
- Sorensen, T. (1948). A Method of Establishing Groups of Equal Amplitude in Plant Sociology Based on Similarity of Species Content. *Biologiske Skrifter*, 5, 1-34.
- Statista. (2023a). Global revenue in the skin care market by segment. Recuperado de <https://www.statista.com/forecasts/1270033/global-revenue-skin-care-market-by-segment>.
- Statista. (2023b). Global skin care market revenue from 2014 to 2023. Recuperado de <https://www.statista.com/statistics/254612/global-skin-care-market-size/>.
- Statista. (2023c). Peru - Beauty & Personal Care. Recuperado de <https://www.statista.com/outlook/cmo/beauty-personal-care/peru>.
- Statista. (2023d). Revenue forecast in the U.S. skin care market (2025-2029). Recuperado de <https://www.statista.com/outlook/cmo/beauty-personal-care/skin-care/united-states>.
- Statista. (2023e). Revenue in the Skin Care market worldwide. Recuperado de <https://www.statista.com/forecasts/1268473/worldwide-revenue-skin-care-market>.
- Statista. (2023f). U.S. skin care market - Statistics & Facts. Recuperado de <https://www.statista.com/topics/4517/us-skin-care-market/>.
- Tamilkodi, R., Chakravarthy, P. K., Kumar, P. H., Narayana, A. V. L., Karthik, A., & Ramabrahmam, I. (2024). A New Approach to Facial Skin Analyzer. *Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024), Advances in Computer Science Research*, 112. doi: 10.2991/978-94-6463-471-6_129.
- Vaswani, A. (2021). Attention Mechanisms in Deep Learning. *IEEE Transactions on Neural Networks*, 32, 405-417. doi: 10.1109/tnn.2021.3051234.
- Yoo, W., Lee, S., & Kim, H. (2020). Limitations of classical segmentation techniques in dermatological images. *Journal of Dermatological Science*, 98(3), 245-256. doi: 10.1016/j.jdermsci.2020.06.012.

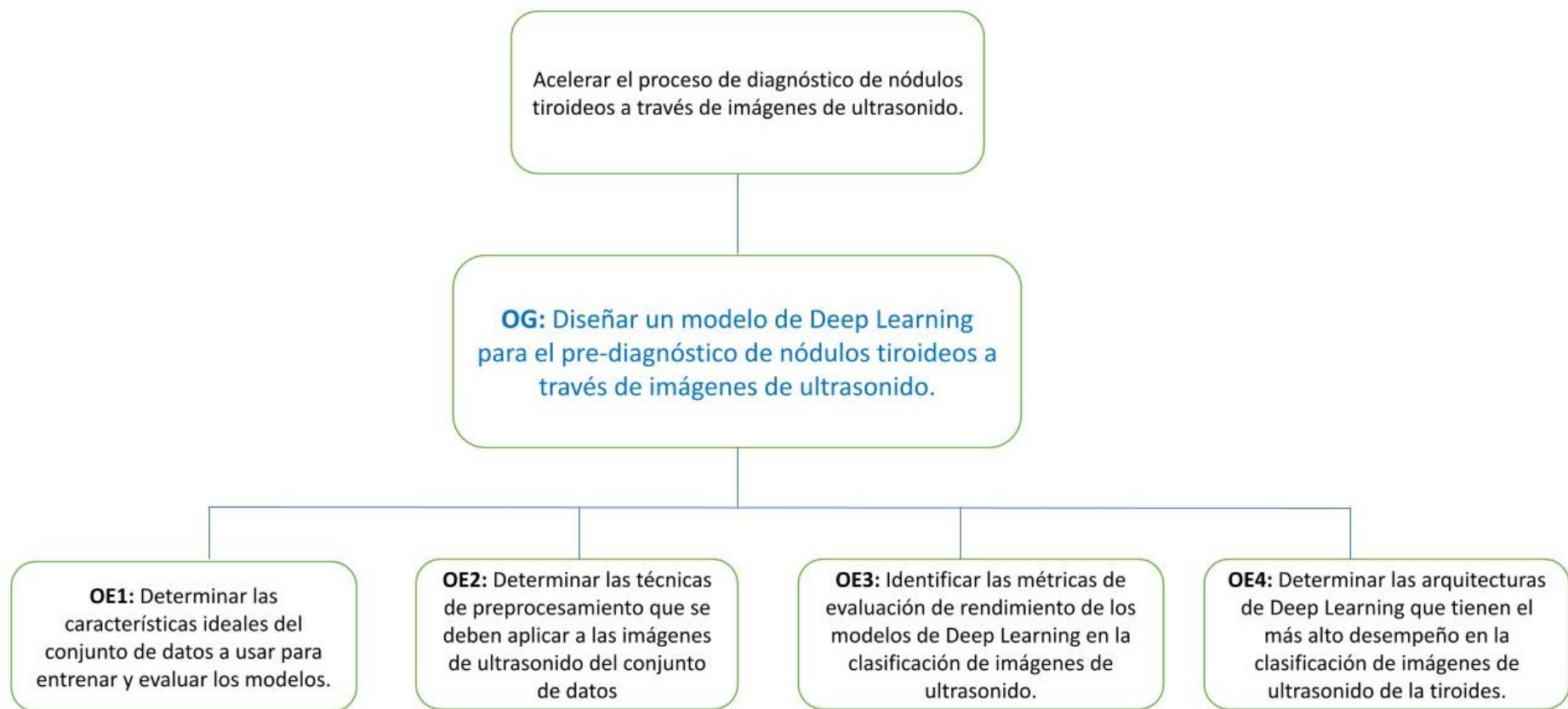
- Yoon, H., Kim, S., Lee, J., & Yoo, S. (2023). Deep-Learning-Based Morphological Feature Segmentation for Facial Skin Image Analysis. *Diagnostics*, 13(1894). doi: 10.3390/diagnostics13111894.
- Yu, C., Gao, J., Wang, J., et al. (2021). BiseNet V2: Bilateral Network with Guided Aggregation for Real-time Semantic Segmentation. *International Journal of Computer Vision (IJCV)*.
- Zambrano, J. (2018, marzo). *¿Aprendizaje supervisado o no supervisado? Conoce sus diferencias dentro del machine learning y la automatización inteligente*. Medium. Recuperado de <https://medium.com/@juanzambrano/aprendizaje-supervisado-o-no-supervisado-39ccf1fd6e7b>.
- Zhao, F., & Liu, X. (2021). Robust Image Segmentation for Dermatological Applications under Varied Skin Tones and Lighting. *Journal of Medical Image Analysis*, 45, 22-34. doi: 10.1016/j.media.2021.01.007.
- Zhong, J., Lang, X., Zhang, X., He, B., Zhang, Z., & Zhang, Y. (2024). Facial Wrinkle Detection Based on DeepLabV3+ and Semi-Automatic Labelling Strategy (G. I. Rashed, Ed.). *Electronic Engineering and Informatics*, 738-743. doi: 10.3233/atde240140.

Anexos

A Árbol de Problemas



B Árbol de Objetivos



C Matriz de Consistencia

Título de la tesis	Diseño de un modelo de Deep Learning para el pre-diagnóstico de nódulos tiroideos a través de imágenes de ultrasonido			
Problema General	Objetivo General	Hipótesis General	Variables	Método

¿De qué manera pueden las Redes Neuronales Convolucionales segmentar de forma eficiente las características morfológicas de la piel del rostro?	Desarrollar un sistema inteligente de segmentación de características morfológicas en imágenes faciales, enfocado en la detección de arrugas y manchas, mediante el uso de Redes Neuronales Convolucionales, con el propósito de optimizar la precisión en la evaluación estética y facilitar la personalización de tratamientos cosméticos.	El desarrollo de un sistema avanzado de segmentación de características morfológicas de la piel facial utilizando redes neuronales convolucionales (CNN), con mecanismos de atención, mejorará significativamente la precisión en la detección de arrugas y manchas, permitiendo una segmentación más precisa y eficiente en tiempo real.	Dependiente: Pre-diagnóstico de nódulos tiroideos. Independiente: Modelo de Deep Learning.
Problemas Específicos	Objetivos Específicos	Hipótesis Específicas	Variables

¿De qué fuentes se obtendrán los datos necesarios para el entrenamiento y la validación del sistema de segmentación?	Recopilar y preprocesar un conjunto representativo de imágenes faciales, que incluya diversidad en tipos de piel y afecciones cutáneas, para su utilización en el entrenamiento y validación del sistema de segmentación.	La utilización de un conjunto de datos diverso de imágenes faciales, que contemple distintos tipos de piel y características morfológicas como arrugas y manchas, incrementará la capacidad del sistema para generalizar y segmentar con precisión las características cutáneas.	Dependiente: Desarrollo del modelo de Deep Learning. Independiente: Las características del conjunto de datos.	Tipo de Investigación: Experimental. Alcance de la investigación: Explicativo.
--	---	--	--	--

¿Cuáles serán las etapas y métodos para el desarrollo del sistema de segmentación basado en Redes Neuronales Convolucionales?	Diseñar, desarrollar e implementar un sistema de segmentación basado en Redes Neuronales Convolucionales, orientado a la detección precisa de características morfológicas faciales, específicamente arrugas y manchas.	Un sistema de segmentación basado en Redes Neuronales Convolucionales permitirá una detección más precisa y diferenciada de las características morfológicas de la piel, tales como arrugas y manchas, superando en rendimiento a los métodos tradicionales de segmentación.	<p>Dependiente: Desempeño del modelo de Deep Learning.</p> <p>Independiente: Técnicas de pre-procesamiento.</p>
---	---	--	---

¿Qué métricas y procedimientos se utilizarán para evaluar la eficiencia del sistema en la segmentación de arrugas y manchas en la piel del rostro?	Definir e implementar métricas de evaluación cuantitativas, con el fin de medir la efectividad del sistema de segmentación en la identificación de patrones morfológicos de la piel facial.	El uso de métricas de evaluación cuantitativas proporcionará una valoración objetiva, confiable y robusta del desempeño del sistema de segmentación en la detección de arrugas y manchas faciales.	Dependiente: Comparación de modelos en la tarea de clasificación de imágenes de ultrasonido. Independiente: Las métricas de evaluación de rendimiento.
--	---	--	---

¿Qué criterios se emplearán para seleccionar el modelo de segmentación más adecuado en la detección de características morfológicas del rostro?	Comparar el desempeño de diversas arquitecturas de Redes Neuronales Convolucionales y técnicas de aprendizaje profundo, con el objetivo de identificar el modelo que ofrezca el mejor compromiso entre exactitud en la segmentación y eficiencia computacional.	La comparación sistemática entre distintas arquitecturas de Redes Neuronales Convolucionales, incluyendo aquellas con mecanismos de atención, permitirá identificar el modelo más eficiente en términos de precisión y rendimiento computacional para la segmentación de características morfológicas cutáneas.	Dependiente: Desempeño en la clasificación de imágenes de ultrasonido. Independiente: Las arquitecturas de Deep Learning.
---	---	---	---
