

**TESIS DOCTORAL**



Doctora en Ingeniería Automática, Electrónica y de  
Telecomunicación

**ANÁLISIS DE IMÁGENES  
DERMATOSCÓPICAS PARA EL DIAGNÓSTICO  
DE LESIONES PIGMENTADAS**

**Autora:**

Paulina Andrea Vélez Núñez

**Directoras:**

Begoña Acha Piñero

Carmen Serrano Gotarredona

Teoría de la Señal y Comunicaciones  
Escuela Técnica Superior de Ingeniería  
Universidad de Sevilla

Sevilla 2022



**Tesis Doctoral**

**Análisis de Imágenes Dermatoscópicas para el Diagnóstico de Lesiones Pigmentadas**

Autora:

**Paulina Andrea Vélez Núñez**

Directoras:

**Begoña Acha Piñero**  
**Carmen Serrano Gotarredona**

Teoría de la Señal y Comunicaciones  
Escuela Técnica Superior de Ingeniería  
Universidad de Sevilla

2022



Tesis Doctoral:           Análisis de Imágenes Dermatoscópicas para el Diagnóstico de Lesiones Pigmentadas

Autora:                    Paulina Andrea Vélez Núñez  
Tutoras:                  Begoña Acha Piñero, Carmen Serrano Gotarredona

El tribunal nombrado para juzgar la Tesis arriba indicada, compuesto por los siguientes doctores:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

El Secretario del Tribunal

Fecha:



*A mi hijo Plácido.*



# Agradecimientos

---

A mi hijo, mi mayor bendición, mi fortaleza, motivación e inspiración, quien me impulsa a ser mejor cada día y quien me ha acompañado durante la realización de esta tesis desde su primer día en este mundo.

A mis padres Lourdes y Pepe por estar siempre a mi lado sin importar la distancia. A mis hermanos Andrea y José Damián, por su apoyo incondicional.

Además, quiero agradecer de forma especial a mis directoras Begoña y Carmen por haberme guiado y tutelado en la realización de esta tesis, por su profesionalidad, entrega y dedicación.

Finalmente quiero agradecer a mis compañeros de despacho, a los demás profesores del Departamento y a mis amigos Ana y José Carlos.

*Paulina Andrea Vélez Núñez*

*Sevilla, 2022*



**E**n este trabajo se realiza un análisis de imágenes dermatoscópicas con el objetivo de ayudar en el proceso de priorización de las consultas en el triaje de los centros de atención primaria.

Específicamente, se plantea la clasificación de lesiones dermatoscópicas centrada en el estudio del carcinoma basocelular (*Basal Cell Carcinoma* - BCC). Para lograr este objetivo se aborda el problema empleando tanto técnicas tradicionales de Aprendizaje Máquina como de Aprendizaje Profundo. Asimismo, en las técnicas de Aprendizaje Profundo propuestas se analiza la conveniencia o no de aplicar un proceso de segmentación previo.

En cuanto al proceso de segmentación, en este trabajo se presentan varios métodos utilizando aprendizaje supervisado y no supervisado, con sus respectivos experimentos y contraste de resultados. Después de aplicar varios algoritmos, el algoritmo supervisado de segmentación semántica basado en la red neuronal profunda denominada SegNet permite obtener los siguientes resultados: un valor del coeficiente DICE igual a 0,8548 y un valor del índice Jaccard de 0,7730 con una exactitud de 0,9357. El resultado obtenido en el proceso de segmentación permite obtener imágenes segmentadas con una alta precisión y la metodología aplicada resulta ser competitiva frente a otras existentes en la literatura.

Para el proceso de extracción de características se consideró el extraer patrones globales presentes en la lesión. De esta forma, la imagen segmentada se divide en superpíxeles, para así obtener regiones con características de textura similares. A cada superpíxel obtenido se le extrajeron sus características de textura utilizando matrices de coocurrencia, y luego se realizó una clasificación basada en árboles de decisión. Como resultado se obtuvo un 99,2% de exactitud en la clasificación de tres patrones globales: globular, empedrado y homogéneo.

Finalmente, se describe el proceso de clasificación de BCC, el cual, después de desarrollar varios algoritmos con sus respectivos experimentos, permitió obtener una exactitud del 98% al clasificar BCC versus nevus y un 95% al clasificar BCC versus todos los demás tipos de lesiones. Estos resultados se obtuvieron al utilizar tres redes neuronales profundas ensambladas y sin segmentar previamente las imágenes. Los resultados obtenidos fueron contrastados, demostrando la fiabilidad del método presentado, ya que éste supera en prácticamente todos los parámetros de evaluación al ganador del reto ISIC-2019 [1]. Además, se concluye que el proceso de segmentación no es necesario al realizar la

clasificación utilizando aprendizaje profundo cuando se cuenta con una base de datos extensa y correctamente balanceada.

Los resultados obtenidos de los modelos desarrollados fueron publicados y se encuentran detallados en los apéndices de esta tesis.

The aim of this paper is to analyse dermatoscopic images in order to assist in the process of prioritization of consultations in the triage of primary care centres.

Specifically, the classification of dermoscopic lesions focused on the study of basal cell carcinoma (BCC) is proposed. To achieve this objective, the problem is approached using both traditional Machine Learning and Deep Learning techniques. Likewise, in the Deep Learning techniques proposed, the convenience or not of applying a prior segmentation process is analysed.

Regarding the segmentation process, several methods using supervised and unsupervised learning are presented in this work, with their respective experiments and contrasting results. After applying several algorithms, the supervised semantic segmentation algorithm based on the deep neural network called SegNet allows obtaining the following results: a value of the DICE coefficient equal to 0.8548 and a value of the Jaccard index of 0.7730 with an accuracy of 0.9357. The result obtained in the segmentation process allows obtaining segmented images with a high precision and the applied methodology is competitive compared to others in the literature.

For the feature extraction process, consideration was given to extracting global patterns present in the lesion. In this way, the segmented image is divided into superpixels in order to obtain regions with similar texture characteristics. Each superpixel obtained had its texture features extracted using co-occurrence matrices, and then a classification based on decision trees was performed. As a result, a 99.2% accuracy was obtained in the classification of three global patterns: globular, cobblestone and homogeneous.

Finally, we describe the BCC classification process, which, after developing several algorithms with their respective experiments, allowed us to obtain an accuracy of 98% when classifying BCC versus nevus and 95% when classifying BCC versus all other types of lesions. These results were obtained by using three deep neural networks assembled without pre-segmenting the images. The results obtained were contrasted, demonstrating the reliability of the method presented, as it outperforms the winner of the ISIC-2019 challenge [1] in practically all evaluation parameters. Furthermore, it is concluded that the segmentation process is not necessary when performing the classification using deep learning, when an extensive and correctly balanced database is available.

The results obtained from the developed models were published and are detailed in the appendices of this thesis.

# Índice

---

Resumen.....	V
Abstract.....	VII
Introducción.....	- 1 -
I.    Motivación.....	- 2 -
II.   Objetivos.....	- 3 -
III.  Descripción del documento.....	- 4 -
CAPÍTULO.....	- 5 -
1.    Bases de Datos Dermatoscópicas.....	- 5 -
1.1.  Bases de datos de imágenes dermatoscópicas.....	- 6 -
1.1.1.  EDRA.....	- 6 -
1.1.2.  ISIC.....	- 7 -
1.1.2.1.  ISIC-2016.....	- 9 -
1.1.2.2.  ISIC-2017.....	- 10 -
1.1.2.3.  ISIC-2018.....	- 11 -
1.1.2.4.  ISIC-2019.....	- 12 -
1.1.2.5.  ISIC-2020.....	- 12 -
CAPÍTULO.....	- 15 -
2.    Aprendizaje profundo.....	- 15 -
2.1.  Aprendizaje Máquina y Aprendizaje Profundo.....	- 15 -
2.1.1.  Aprendizaje supervisado.....	- 16 -
2.2.  Redes Neuronales Convolucionales.....	- 17 -
2.2.1.  AlexNet.....	- 20 -
2.2.2.  VGG-16.....	- 21 -
2.2.3.  GoogLeNet.....	- 22 -
2.3.  Procesos que Benefician el Entrenamiento.....	- 23 -
CAPÍTULO.....	- 25 -
3.    Segmentación de Imágenes Dermatoscópicas.....	- 25 -
3.1.  Metodologías de Segmentación para Imágenes Dermatoscópicas.....	- 25 -
3.2.  Métodos Supervisados.....	- 29 -

3.2.2.1.	Segmentación de imágenes umbralizando capas convolucionales de una CNN .....	- 29 -
3.2.2.2.	Segmentación semántica con la red neuronal profunda SegNet.....	- 31 -
3.3.	Resultados comparativos .....	- 33 -
3.3.1	Métricas para evaluar un algoritmo de segmentación.....	- 33 -
3.3.2.	Base de datos y parámetros de evaluación.....	- 34 -
3.3.3.	Comparación de los métodos propuestos.....	- 35 -
3.3.4.	Comparación del mejor método propuesto con otros de la literatura .....	- 37 -
CAPÍTULO	.....	- 41 -
4.	Aprendizaje Máquina para Detección de Patrones Globales .....	- 41 -
4.1.	Características diseñadas ad-hoc.....	- 43 -
4.1.1.	Características de Textura.....	- 43 -
4.1.2.	División de la imagen en superpíxeles .....	- 46 -
4.2.	Selección de Características.....	- 48 -
4.2.1.	Métodos de Filtro.....	- 49 -
4.2.2.	Métodos envolventes .....	- 50 -
4.3.	Experimentos para la detección de patrones globales .....	- 51 -
4.3.1.	Segmentación de la lesión para identificar características locales.....	- 51 -
4.3.2.	Extracción de características de los recortes de textura.....	- 52 -
4.3.3.	Selección de características y clasificación .....	- 53 -
4.4.	Resultados.....	- 55 -
CAPÍTULO	.....	- 59 -
5.	Clasificación .....	- 59 -
5.1.	Estado del Arte de la Clasificación de lesiones dermatoscópicas .....	- 60 -
5.2.	Métodos .....	- 61 -
5.2.1.	Ensamblado de redes .....	- 64 -
5.2.1.1.	Ensamblador ArgMax .....	- 65 -
Criterio de clasificación	.....	- 65 -
5.2.1.2.	Ensamblador Mean.....	- 65 -
5.3.	Datos y diseño de los experimentos.....	- 66 -
5.3.1.	Datos.....	- 66 -
5.3.2.	Métricas para evaluar la Clasificación.....	- 67 -

5.3.3. Experimentos de clasificación realizados .....	- 68 -
5.3.3.1. Implementación de Ensambladores para mejorar la clasificación....	- 69 -
5.4. Resultados de la clasificación.....	- 70 -
5.4.1. Clasificación de BCC versus todas las lesiones sin utilizar un ensamblador.....	- 70 -
5.4.2. Clasificación en ocho tipos de lesiones sin utilizar un ensamblador .....	- 74 -
5.4.3. Clasificación de BCC versus Nevus sin utilizar un ensamblador .....	- 78 -
5.4.4. Clasificación de BCC versus todas las lesiones con ensamblador.....	- 82 -
5.4.5. Clasificación en ocho tipos de lesiones con ensamblador .....	- 84 -
5.4.6. Clasificación de BCC versus Nevus con ensamblador .....	- 87 -
5.5. Discusión de Resultados .....	- 89 -
5.5.1. Comparación con otros trabajos publicados en la literatura .....	- 92 -
CAPÍTULO .....	- 93 -
6. Conclusiones.....	- 93 -
Apéndices.....	- 95 -
Revistas con índice de impacto (JCR - <i>Journal Citation Report</i> ) .....	- 95 -
Índice de Figuras .....	- 97 -
Índice de Tablas .....	- 101 -
Bibliografía .....	- 103 -
Glosario.....	- 117 -



# Introducción

---

**En** este trabajo se presenta un estudio completo sobre el análisis de imágenes dermatoscópicas enfocado a la clasificación de carcinoma basocelular (BCC). Además, se verifica la importancia del proceso de segmentación de la imagen dermatoscópica al momento de realizar la clasificación de la lesión. Las imágenes que se consideraron para esta investigación fueron tomadas de las bases de datos de ISIC 2017, 2018 y 2019.

La importancia del análisis de las imágenes dermatoscópicas radica en la necesidad de realizar la detección temprana de lesiones malignas, como el melanoma y el carcinoma basocelular. Según la Fundación del Cáncer de Piel [2], el melanoma es uno de los tipos de cáncer de piel más agresivos y el carcinoma basocelular es el tipo de cáncer de piel más frecuente.

El análisis de lesiones de piel ha sido objeto de estudio durante décadas, ya que está ligado a la detección temprana de enfermedades cuya tasa de supervivencia aumenta de forma considerable y proporcional a la rapidez con la que es detectada. Por esta razón, los métodos basados en herramientas de diagnóstico asistido por computador (*Computer Aided Diagnosis* - CAD) han sido propuestos para el apoyo al personal sanitario en el momento de reconocer y clasificar lesiones cutáneas utilizando imágenes dermatoscópicas.

El reto de este trabajo recae en la complejidad del análisis de imágenes dermatoscópicas, por su variedad de colores, tonalidades y texturas. Otro factor importante a considerar son los artefactos presentes en la imagen, siendo estos artefactos pelo o vello, manchas de tinta, cicatrices u otros elementos ajenos a las características propias de la lesión.

## I. Motivación

El cáncer de piel presenta una gran incidencia y mortalidad en la población mundial. Este trabajo está enfocado particularmente al análisis de melanoma y carcinoma basocelular ya que ambos son los tipos de cáncer con la tasa de mortalidad más elevada y con el mayor número de casos, respectivamente. Cuando se trata de melanoma se considera un tipo de cáncer agresivo, capaz de expandirse rápidamente y dañar a otros órganos del cuerpo. Por otra parte, al hablar de BCC se conoce que es una variación de cáncer menos agresivo que el melanoma, pero con una gran incidencia. Según la Sociedad Americana del Cáncer [3] 8 de cada 10 casos de cánceres de piel son casos de BCC.

El aumento anual a nivel mundial de casos de BCC y de melanoma fortalece la premisa sobre la necesidad de desarrollar mejores técnicas de diagnóstico de lesiones de piel con la finalidad de realizar un diagnóstico temprano de la enfermedad y, a su vez, evitar biopsias innecesarias. En la actualidad se manejan dos líneas de trabajo. La primera es la inversión en nuevas técnicas para la adquisición de la imagen dermatoscópica, y la segunda está enfocada al desarrollo de sistemas de diagnóstico asistido (CAD) para el análisis automático de imágenes dermatoscópicas [4].

Las herramientas CAD desarrolladas, en su mayoría, se encuentran enfocadas a la detección de melanomas, dejando de lado al BCC, a pesar de presentar una elevada incidencia. Las aplicaciones CAD enfocadas a la detección de melanoma, generalmente, basan su funcionamiento en la aplicación de aprendizaje profundo, concretamente en redes neuronales profundas, apoyando su entrenamiento en bases de datos dermatoscópicas de carácter público como ISIC y HAM10000 [5, 6].

En la literatura existen varios trabajos que buscan clasificar lesiones de piel utilizando redes neuronales convolucionales (*Convolutional Neural Network* - CNN) y modelos ensamblados [1]. Así también, algunas investigaciones reflejan el interés existente en el reconocimiento del melanoma como, por ejemplo, el trabajo de Winkler et al. [7], que, utilizando modelos de CNN, realizaron la clasificación de melanomas en diversas localizaciones corporales. Otro ejemplo lo describen Vasconcelos y Vasconcelos [8], donde se utiliza la red convolucional profunda GoogLeNet y se modifican sus parámetros para identificar melanomas utilizando una base de datos reducida. Sin embargo, el número de trabajos centrado en el diagnóstico automático del BCC es muy limitado.

La realización de este trabajo está motivada por la necesidad de contar con herramientas CAD que faciliten el diagnóstico del cáncer de piel, específicamente de BCC. Este trabajo no pretende excluir la presencia del médico, al contrario, lo que se desea es proveer al sanitario de atención primaria de una herramienta que permita incrementar la precisión de su diagnóstico y establecer una priorización en la atención de los pacientes.

Un sistema con el que se cuenta actualmente en numerosos hospitales públicos y privados es la teledermatología. En particular, en el Hospital Universitario Virgen Macarena de Sevilla (HUVVM) se dispone de un sistema de teledermatología denominado Telederma. Esta es una aplicación que permite realizar consultas de dermatología a distancia. Con este sistema se pretende que la población en general tenga acceso a un diagnóstico dermatológico temprano, al realizar la consulta en los centros de atención primaria. Actualmente, sólo al Hospital Universitario Virgen Macarena de Sevilla (HUVVM) llega un promedio de 315 consultas de dermatología al mes a través del sistema de teleconsulta. Por otra parte, al Hospital Virgen del Rocío de Sevilla (HUVR), llega un promedio de 938 consultas de dermatología al mes a través del sistema de teleconsulta.

Al considerar el elevado número de teleconsultas que llegan mensualmente a los hospitales públicos, se hace necesario contar con una herramienta de triaje que permita la priorización de atención de las consultas que llegan al servicio de dermatología de los hospitales, y es aquí donde este trabajo de investigación pretende aportar a su solución.

La combinación de sistemas CAD y la intervención humana incrementan la precisión del diagnóstico [9, 10], por lo que se podrían priorizar las consultas por nivel de gravedad de la lesión al mejorar el análisis de la imagen dermatoscópica, ya sea en la segmentación de la lesión con respecto a la piel sana, como en la clasificación de la lesión que lleva a su diagnóstico.

## **II. Objetivos**

El objetivo de este trabajo de investigación es realizar el análisis automático de imágenes dermatoscópicas con especial énfasis en la detección de carcinoma basocelular. Este trabajo se centra en la clasificación de BCC debido al elevado número de casos existentes. Por otro lado, se encontró interesante trabajar con melanomas en el proceso de segmentación debido a la amplia existencia de literatura y de bases de datos que facilitan la comparación de resultados.

Para alcanzar el objetivo general de este trabajo se han desglosado varios objetivos específicos que se describen a continuación:

- Desarrollar un algoritmo de segmentación que permita separar la lesión de la piel sana, cuyo resultado pueda competir con los resultados existentes en la literatura, tanto en exactitud como en coste computacional.
- Verificar la importancia de la segmentación de la imagen dermatoscópica en el proceso de clasificación de la lesión.

- Diseñar un algoritmo que permita la clasificación del carcinoma basocelular, no sólo con respecto al nevus, sino también considerando varios tipos de lesiones de piel.

Estos tres objetivos específicos planteados buscan cubrir el objetivo general de este trabajo y están desarrollados en los cinco capítulos de este trabajo, donde se detallan cada una de las teorías aplicadas y el proceso realizado para alcanzar resultados óptimos.

### **III. Descripción del documento**

Esta tesis está organizada de la siguiente manera:

En el capítulo uno se describen las distintas bases de datos dermatoscópicas existentes, públicas y privadas, así como también los datos que contienen.

En el capítulo dos se describen varias técnicas de aprendizaje profundo, así como el funcionamiento de distintas redes neuronales convolucionales. Además, se realiza una descripción de los procesos de aumento de datos y la transferencia del aprendizaje que benefician el entrenamiento de una red.

El capítulo tres está enfocado a la descripción de las distintas metodologías de segmentación de una imagen dermatoscópica. Esto incluye las metodologías clásicas de segmentación supervisada y no supervisada, así como también las basadas en aprendizaje profundo. En este capítulo se encuentran descritos varios experimentos de segmentación realizados sobre la base de datos ISIC-2017 y sus respectivos resultados.

El capítulo cuatro detalla varios procesos de aprendizaje de máquina para la extracción y selección de características con el objetivo de clasificar patrones globales presentes en una lesión dermatoscópica. Además, se presentan los experimentos realizados y sus respectivos resultados.

En el capítulo cinco se describen los distintos métodos de clasificación de las imágenes dermatoscópicas y se presentan los experimentos y resultados de clasificación de lesiones, enfocándonos en el reconocimiento del carcinoma basocelular, clasificándolo versus nevus y versus distintos tipos de lesiones de piel. La base de datos utilizada para este fin fue ISIC 2019.

Finalmente, se presentan las conclusiones y se añade una sección de Apéndices donde se encuentran adjuntas las publicaciones realizadas que reflejan los resultados obtenidos durante la realización de esta tesis.

# 1. Bases de Datos Dermatoscópicas

---

**E**l análisis de las lesiones de piel ha sido objeto de estudio durante décadas, esto debido a su incidencia en la población, ya que el cáncer de piel es uno de los tipos de cáncer más agresivos y comunes. Éste presenta una tasa de mortalidad muy alta, especialmente el melanoma, cuya tasa de mortalidad es del 1,62% más alta que los demás tipos de cáncer [11]. Según la Asociación Americana del Cáncer, en el 2022 habrá 99.780 nuevos casos de cáncer de piel tan sólo en Estados Unidos de Norte América, de los cuales 7.650 casos terminarán con la vida de los pacientes.

La necesidad de la creación de sistemas CAD para acelerar los tiempos y precisión del diagnóstico de la lesión cutánea, priorizar las consultas y evitar biopsias innecesarias, se convierte en evidente ante la alta demanda en los servicios de dermalotología de los hospitales. La creación de estas herramientas precisa la existencia de bases de datos dermatoscópicas que permitan realizar los experimentos necesarios para su desarrollo.

Una lesión cutánea puede contener distintos tipos de estructuras [12] y una variedad de colores y tonalidades [13, 14], es por esto que existe la necesidad de contar con bases de datos que permitan acceder a la mayor cantidad de casos posibles de imágenes dermatoscópicas. Estas bases de datos, no sólo facilitan la creación y el diseño de herramientas CAD, sino también facilitan el aprendizaje previo a la aplicación de distintas técnicas clínicas de diagnóstico, como la regla ABCDE [15, 16], la escala de Menzies [17] [18] o la regla de los 7 puntos planteada por Argenziano [19, 20].

Estas bases de datos dermatoscópicas, en su mayoría de acceso público, facilitan el trabajo de investigación de las enfermedades de la piel ya que proporcionan un sinnúmero de imágenes capturadas de pacientes reales, sobre las cuales es posible realizar el análisis de características y pruebas de los sistemas CAD desarrollados. Una ventaja de utilizar estas bases de datos es que presentan la información clasificada por tipo de lesión, tipos de estructuras presentes en ellas y en algunos casos incluyen información clínica del paciente.

En este capítulo se realiza la descripción de varias bases de datos dermatoscópicas de acceso libre, que fueron utilizadas para el desarrollo de este trabajo de investigación.

## **1.1. Bases de datos de imágenes dermatoscópicas**

La necesidad de estudiar las lesiones cutáneas y, en particular, las potencialmente mortales para el paciente, ha llevado a especialistas de varios países a interesarse por construir bases de datos dermatoscópicas, ya sea con fines educativos, o con fines de investigación. Es así que, actualmente, existen varias bases de datos dermatoscópicas públicas y privadas.

En este apartado se mencionan dos bases de datos: EDRA, que es una base de datos privada creada con fines formativos/educativos, y la base de datos ISIC, que cada año va incrementado su número de imágenes, es de libre acceso y fue creada con fines de investigación. Además, está la base de datos HAM10000, pero ésta fue incluida dentro de la base de datos ISIC, por lo que no se menciona de forma específica en este apartado.

### **1.1.1. EDRA**

EDRA es una base de datos de imágenes dermatoscópicas creada con fines educativos para formación y entrenamiento del personal médico a cargo del diagnóstico de lesiones de la piel. Su nombre completo es "*The Interactive Atlas of Dermoscopy (EDRA)*" o Atlas interactivo de dermatoscopia [19]. Esta base de datos posee más de 1000 casos clínicos con al menos dos imágenes de cada lesión, una imagen de la fotografía simple (denominada imagen clínica) y la imagen dermatoscópica.

Las imágenes en su mayoría tienen una dimensión de 768×512 píxeles. Las fotografías fueron captadas con una cámara Nikon F3 montada en un microscopio estereoscópico Wild M650. Por otra parte, las dermatoscopias fueron adquiridas con un dermatoscopio Dermaphot / Optotechnik.

Entre los 1000 casos de diagnóstico que incluye esta base de datos están los siguientes: melanoma y sus variaciones, carcinoma basocelular, nevus azul, nevus de Clark, nevus combinados, nevus congénitos, nevus dermales, dermatofibroma, lentigo, melanososis, nevus recurrente, nevus Rees, queratosis seborreica y las lesiones vasculares. Del mismo modo, incluye una sección donde se ubican diversos casos denominada sección de misceláneos [20].

Además de ser una base de datos, EDRA funciona como una enciclopedia interactiva que ofrece varias opciones para la formación de profesionales, como el análisis de patrones, algoritmos de diagnóstico donde se puede aplicar tanto la regla ABCD como la regla de los

7 puntos, algoritmos basado en el grosor de la lesión y finalmente la opción de comparar las lesiones y diferenciar sus diagnósticos.

En las opciones de diagnóstico, esta base de datos permite identificar una lesión y muestra en detalle la ausencia o presencia de característica globales (patrones) y locales (estructuras) de la lesión. En esta opción se permite acceder a la imagen fotografiada y a la dermatoscopia, así como también se indica el diagnóstico de la lesión. La Figura 1.1 muestra una imagen de esta base de datos.

**Database**  
Interactive Atlas of Dermoscopy

**Search for**

Case Number:

**Clinical Data**

Age:  Sex:  male  female

Location:  head-neck  upper limbs  genital area  
 back  lower limbs  
 chest  acral  
 abdomen  buttocks

Diameter:  \* (mm) Elevation:  flat  palpable  nodular

**Dermoscopic Criteria**

**Pigment (pseudo) network**  
 absent  typical  atypical

**Dots and globules**  
 absent  regular  irregular

**Steals**  
 absent  regular  irregular

**Blue-whitish veil**  
 absent  present

**Pigmentation**  
 absent  localized regular  localized irregular  
 diffuse regular  diffuse irregular

**Global pattern**  
 reticular pattern  parallel pattern  
 globular pattern  multicomponent pattern  
 cobblestone pattern  lacunar pattern  
 homogeneous pattern  unspecific pattern  
 starburst pattern

**Hypopigmentation**  
 absent  focal  multifocal  diffuse

**Regression structures**  
 absent  white areas  blue areas  combinations

**Vascular structures**  
 absent  comma  wreath  arborizing  hairpin  dotted  linear, irregular } Atypical vascular pattern  
 within regression

**Other criteria**  
 absent  red lacunas  leaf-like areas  central white patch  
 milia-like cysts  comedo-like openings  exophytic papillary structures

**Diagnosis**

- Basal cell carcinoma
- Blue nevus
- Clark nevus
- Combined nevus
- Congenital nevus
- Dermal nevus
- Dermatofibroma
- Lentigo
- Melanoma
- Melanoma (in situ)
- Melanoma (less than 0.76 mm)
- Melanoma (0.76 - 1.5 mm)
- Melanoma (more than 1.5 mm)
- Melanoma metastasis
- Melanosis
- Miscellaneous
- Recurrent nevus
- Reed/Spitz nevus
- Seborrheic keratosis
- Vascular lesion

**Search**

\* Note that in the age and diameter fields you can also enter the symbol > or < followed by a number

Figura 1.1 EDRA - Base de datos.

### 1.1.2. ISIC

ISIC son las siglas de "International Skin Imaging Collaboration" o Colaboración Internacional en Imágenes de Piel. Éste es un proyecto que incluye los ámbitos académico e industrial y está diseñado para facilitar la aplicación del procesamiento digital de imágenes con el objetivo de reducir la tasa de mortalidad del melanoma. El objetivo del proyecto Melanoma ISIC es colaborar y aportar en la lucha para reducir las muertes relacionadas con el melanoma y las biopsias innecesarias a través de la mejora en la precisión y eficiencia de la detección temprana [21].

Este proyecto está enfocado al entrenamiento de profesionales y demás personas interesadas en trabajar en el reconocimiento y diagnóstico temprano del melanoma. ISIC es una base de datos de acceso libre y no presenta problemas de privacidad de los datos. Además, esta base de datos contiene imágenes estandarizadas y de calidad. ISIC provee de un conjunto de datos de entrenamiento, test y validación para la realización de aplicaciones. En la Figura 1.4 se presenta una muestra de esta base de datos.

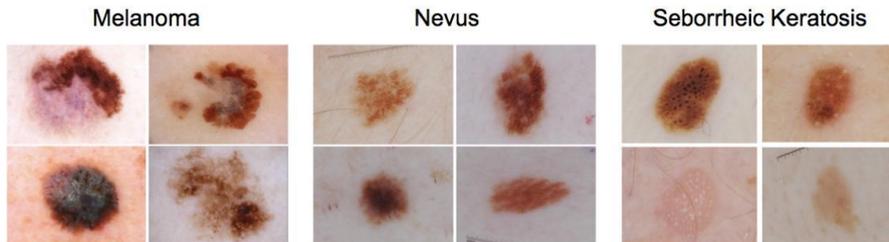


Figura 1.2 ISIC - Base de datos.

Al ser ISIC una base de datos de acceso público es necesario que ésta cumpla con tres estándares básicos que son: calidad, privacidad e interoperabilidad. La calidad hace referencia a que la imagen debe cumplir unos requisitos mínimos de precisión del color, claridad del enfoque y resolución espacial. Por otra parte, la privacidad hace referencia a la protección de los datos incluidos en la imagen, tanto del paciente como del personal que la adquirió. Finalmente, la interoperabilidad se refiere a la facilidad con la que se pueden compartir las imágenes y sus metadatos o datos asociados a ellas.

El cumplimiento de los tres estándares básicos ha llevado a que ISIC desarrolle y cubra los siguientes estándares:

- Estándares de tecnología en cuanto a la calidad de la imagen adquirida.
- Estándares técnicos que hacen referencia a la identificación, documentación de los atributos de la lesión y adquisición de la imagen.
- Estándares de terminología, facilitando así la enseñanza y la comparación de imágenes; metadatos que son los datos de los elementos asociados o embebidos en la imagen como diagnóstico, morfología, y detalles técnicos de la captura de la imagen.
- Privacidad de los metadatos asociados con la imagen, guardando la privacidad del paciente, del personal médico y asegurando que existe el consentimiento para su uso y publicación.

Desde el año 2016, ISIC viene organizando el evento “ISIC: *Skin Lesion Analysis Towards Melanoma Detection*” que incluye algunos retos (*challenges*) para resolver problemas de

segmentación, detección de patrones clínicos y clasificación de la lesión. Para esto incluye una gran cantidad de imágenes validadas por expertos que conforman grupos para entrenamiento, validación y test. A continuación, se mencionan cada una de las bases de datos con sus respectivos retos anuales.

### 1.1.2.1. ISIC-2016

La base de datos ISIC-2016 [22] es la primera base de datos del proyecto ISIC y fue publicada junto con su reto, el ISIC-2016 *Challenge*, el cual está dividido en tres partes: Parte 1, para la segmentación de la lesión; Parte 2, para la detección y localización de características o patrones visuales dermatoscópicos; y Parte 3, destinada a la clasificación de la enfermedad. ISIC-2016 cuenta con 900 imágenes en el conjunto de entrenamiento y 379 imágenes en el conjunto de prueba (*test*). Las imágenes varían su tamaño entre 1022×767 y 2048×1536 píxeles. Los nombres de las imágenes siguen el siguiente formato:

ISIC\_ <image\_id >.jpg

donde <image\_id > es un identificador único de 7 dígitos. A continuación, se describen los datos que esta base de datos entrega según el reto planteado:

- Parte 1: Segmentación de la Lesión. Se entregan 900 imágenes de entrenamiento y 379 de test, acompañadas con su respectiva imagen con la máscara para la comparación (*ground truth*). Los parámetros que se evalúan son: sensibilidad, especificidad, exactitud, índice Jaccard y coeficiente DICE. El ganador se selecciona considerando únicamente el índice Jaccard.
- Parte 2: Extracción de las características dermatoscópicas de la lesión. Se entregan 807 imágenes para el entrenamiento acompañadas con sus respectivas máscaras de superpíxeles (con su nombre en formato: ISIC\_ <image\_id >\_superpixels.png). Además cada imagen de entrenamiento posee asociado un archivo con sus características (llamado ISIC\_<image\_id >.json). En este archivo se han identificado características como glóbulos, líneas o retículos. Los parámetros a medir son sensibilidad, especificidad, exactitud y promedio de exactitud evaluado a una sensibilidad del 100%.
- Parte 3: Clasificación de la lesión. Se entregan 900 imágenes para el entrenamiento junto con su respectivo archivo en formato CSV (*Comma-Separated Value*) el cual contiene dos columnas, una para el nombre de la imagen y otra para identificarla como maligna o benigna. En este reto se evalúan los mismos parámetros que en la Parte 2.

### 1.1.2.2. ISIC-2017

La base de datos ISIC-2017 [23] es la segunda versión de la base de datos ISIC y a su vez lanza también tres retos: segmentación de la lesión, detección y localización de patrones o características visuales dermatoscópicas, y clasificación de lesiones. Esta base de datos presenta 2000 imágenes de entrenamiento, 150 imágenes de validación y 600 imágenes de test. Las imágenes varían de tamaño, pero mantienen su calidad. Los nombres de las imágenes siguen el siguiente formato:

ISIC\_<image\_id> .jpg

donde <image\_id> es un identificador único de 7 dígitos. A continuación, se describen los datos que esta base de datos entrega según el reto planteado [23]:

- Parte 1: Segmentación de la Lesión. Se entregan 2000 imágenes de entrenamiento con su respectiva máscara (*ground truth*) identificadas como:

ISIC\_<image\_id>\_Segmentation.png

donde <image\_id> es el identificador de la imagen y corresponde con la imagen de entrenamiento original. Para evaluar la tarea se consideran los siguientes parámetros: sensibilidad, especificidad, exactitud, índice Jaccard y coeficiente DICE.

- Parte 2: Extracción de las características dermatoscópicas de la lesión. Además de las imágenes de entrenamiento originales, esta base de datos entrega su correspondiente máscara de superpíxeles y las notas del mapa de superpíxeles realizadas por expertos incluyendo las siguientes características: retículo, retículo negativo, líneas y quistes tipo miliar. La imagen de superpíxel tiene el formato ISIC\_<image\_id>\_superpixels.png donde <image\_id> corresponde al nombre de la imagen original. Los parámetros a medir son: sensibilidad, especificidad, exactitud y promedio de exactitud evaluado a una sensibilidad del 100%.
- Parte 3: Clasificación de la lesión. Se entregan 2000 imágenes distribuidas de la siguiente manera: 374 melanomas, 254 queratosis seborreicas y el resto 1372 nevus benignos. A estas imágenes se le adjunta un archivo CSV con tres columnas que incluyen: identificador de la imagen, edad aproximada y sexo. Los parámetros a medir son sensibilidad, especificidad, exactitud, todos estos medidos con un umbral de confianza de 0,5. Por otra parte, el promedio de exactitud es evaluado a una sensibilidad del 100%, la especificidad es evaluada a una sensibilidad del 82%, 89% y 95%; finalmente se evalúa el área bajo la curva característica de funcionamiento (AUC).

### 1.1.2.3. ISIC-2018

La base de datos ISIC-2018 [6] [24] agrega la base de datos pública HAM-10000 ("*Human Against Machine with 10000 training images*"). HAM-10000 provee de 10015 imágenes en su conjunto de entrenamiento. Estas imágenes son el producto de la recoplicación de imágenes dermatoscópicas durante 20 años en el Departamento de Dermatología de la Universidad Médica de Viena-Austria y del Cliff Rosendahl en Queensland-Australia. Como las imágenes se empezaron a almacenar antes de la era digital, muchas de ellas fueron almacenadas como diapositivas y fueron digitalizadas utilizando un escáner Nikon Coolscan 5000ED, junto con un escaneo doble con DigitalICE, y almacenadas en formato JPEG (8 bits de profundidad del color) con una calidad de 300 DPI (15×17 cm), para finalmente ser convertidas a 800×600 píxeles a 72 DPI. Por otra parte, las imágenes que fueron capturadas de forma digital se obtuvieron con una cámara DermLite™ FOTO (3Gen™). Esta base de datos presenta las siguientes imágenes: 327 imágenes de queratosis actínicas, 514 imágenes de carcinomas basocelular, 1099 imágenes de queratosis benigna, 115 imágenes de dermatofibroma, 1113 imágenes de melanomas, 6705 imágenes de nevus melanocíticos, y 142 imágenes de lesiones vasculares de la piel [6]

Al igual que las versiones anteriores de ISIC, todas las imágenes fueron estandarizadas y esto incluyó el formato de sus nombres, el cual fue el mismo que en ISIC-2017 (ISIC\_<image\_id>.jpg). En el ISIC-2018 se realizó un reto, el cual estaba formado por tres partes que se describen a continuación[25]:

- Parte 1: Segmentación de la Lesión. Para esta tarea se entregó el siguiente conjunto de imágenes: 2594 imágenes con sus respectivas máscaras, 100 imágenes de validación y 1000 imágenes de test. Para evaluar la tarea se consideran los siguientes parámetros sensibilidad, especificidad, precisión, coeficiente DICE y se evaluará específicamente el índice Jaccard.
- Parte 2: Detección de los atributos de la lesión. La idea es realizar una segmentación semántica de características en la imagen y para esto se entregan 2594 imágenes de entrenamiento acompañadas de cinco imágenes que corresponden a los datos verdaderos (*ground truth*) de las máscaras de las cinco clases de patrones a analizar, estos son retículo pigmentado, retículo negativo, líneas (*streaks*), quistes miliares y glóbulos incluidos los puntos; en total 12970 imágenes con su verdad de referencia. Para realizar la validación de los modelos realizados se entregan 100 imágenes y para las pruebas o test se entrega un conjunto con 1000 imágenes.
- Parte 3: Clasificación de la enfermedad. Se entregan 10015 imágenes de entrenamiento con su correspondiente archivo CSV con información relacionada a cada imagen, además se entregan 193 imágenes de validación y 1512 imágenes de test. En esta tarea se pretende clasificar entre siete enfermedades: melanoma,

nevus melanocítico, carcinoma basocelular, queratosis actínica, queratosis benigna, dermatofibroma y lesiones vasculares.

#### **1.1.2.4. ISIC-2019**

ISIC-2019 [6] [23] [25] incrementa la extensión de la base de datos ISIC-2018 agregando al conjunto de datos de ISIC-2017 [23] y HAM10000 [6], el conjunto de datos recopilado por el Departamento de Dermatología del Hospital Clínic de Barcelona [25], incrementando la base de datos a un total de 33126 imágenes dermatoscópicas con su respectivo análisis clínico realizado por expertos.

Al igual que en las bases de datos de ISIC de años anteriores, con la ISIC-2019 se lanzó también un reto, esta vez con una única tarea: clasificar las imágenes en nueve categorías diferentes. Las nueve clases en las que se deberían identificar las imágenes son: melanoma, nevus melanocítico, carcinoma basocelular, queratosis actínica, queratosis benigna, dermatofibroma, lesión vascular, carcinoma de células escamosas, y finalmente ninguna de las clases anteriores. Junto con las imágenes se entregó también la verdad de referencia y su correspondiente archivo CSV con información relacionada a cada imagen detallando edad, sexo y localización de la lesión.

El conjunto de imágenes de entrenamiento está formado por 25331 imágenes divididas en las ocho clases conocidas. Con estos datos la tarea principal se dividió en dos, clasificar las imágenes utilizando la información de los metadatos adicionales y clasificar las imágenes sin utilizar los metadatos. El conjunto de test presenta 8238 imágenes y viene acompañado de su respectivo archivo CSV con información sobre edad, sexo y localización de la lesión.

Las métricas de evaluación de la tarea fueron las siguientes: sensibilidad, especificidad, exactitud, área bajo la curva característica de funcionamiento (AUC), media del promedio de la exactitud, puntuación F1, AUC integrado entre 80% al 100% de sensibilidad (AUC80), valor predictivo positivo y el valor predictivo negativo.

#### **1.1.2.5. ISIC-2020**

La base de datos ISIC-2020 [26] contiene 33126 imágenes de entrenamiento y 10982 imágenes de test, con lesiones benignas y malignas, cuya malignidad ha sido confirmada por medio de pruebas histopatológicas y su benignidad ha sido confirmada por expertos. Para la obtención de esta base de datos, que es la extensión de ISIC-2019, se recogieron imágenes de las siguientes fuentes: Hospital Clínic de Barcelona, Universidad Médica de Viena, Centro del Cáncer Memorial Sloan Kettering, Instituto de Melanoma de Australia, Universidad de Queensland, y la Escuela de Medicina de la Universidad de Atenas.

Esta base de datos fue utilizada por la *Society for Imaging Informatics in Medicine* (SIIM) *ISIC Melanoma Classification*, cuya tarea fue clasificar melanomas dentro del conjunto de imágenes.



## 2. Aprendizaje profundo

---

**En** esta sección se describen varios conceptos relacionados con el Aprendizaje Profundo (*Deep-Learning* - DL). El Aprendizaje Profundo es un tipo de aprendizaje automático que se basa en el uso de redes neuronales artificiales. Se denomina profundo porque emplear un número elevado de capas de neuronas. En particular, cuando se trata del procesamiento de imágenes, las redes neuronales convolucionales (*Convolutional Neural Network* - CNN) son las más utilizadas.

Las redes neuronales convolucionales (CNN) están formadas por varias capas que transforman sus entradas utilizando filtros convolucionales de pequeña extensión. De esta forma se han creado modelos exitosos que se presentaron en el “*ImageNet Challenge*” [45], como el modelo de AlexNet y los modelos VGG y GloogleNet, también llamado Inception, de los cuales se hablará más adelante

El éxito de las redes neuronales en el análisis de imágenes se debe a factores como: la capacidad de procesamiento de la unidad central de procesamiento (CPU) y de la unidad gráfica (GPU), la disponibilidad de bases de datos con gran cantidad de imágenes y el desarrollo de los algoritmos de aprendizaje.

### 2.1. Aprendizaje Máquina y Aprendizaje Profundo

El Aprendizaje Máquina (*Machine-Learning*) es un conjunto de sistemas utilizados para identificar objetos, texto, palabras o seleccionar opciones de interés de un conjunto. Las técnicas convencionales de aprendizaje de máquina, en sus inicios, se limitaban a procesar datos naturales en su forma cruda (como los valores de los píxeles de una imagen) con una representación interna adecuada o vector de características, cuyo subsistema de aprendizaje podía detectar o clasificar patrones en la entrada [27]. El aprendizaje consiste en un conjunto de métodos que permiten a la máquina ser alimentada con datos crudos y descubrir de forma automática las representaciones necesarias para la detección o clasificación.

El Aprendizaje Profundo puede ser definido como un tipo de aprendizaje máquina basado en redes neuronales artificiales (*Artificial Neural Networks – ANNs*) con un incremento en el número de sus capas ocultas [28]. Las arquitecturas basadas en redes neuronales convolucionales (*Convolutional Neural Network – CNN*) son las preferidas para las aplicaciones de visión por computador, ya que incluye extracción de características y poseen capacidad de clasificación.

Una red neuronal con múltiples capas no lineales (con una profundidad de entre 5 y 20) puede implementar funciones extremadamente complejas, permitiendo que sea la máquina quien realice las tareas de extracción de características y clasificación. La primera capa típicamente representa la presencia o ausencia de bordes en una orientación y localización particular de la imagen. La segunda capa típicamente detecta los detalles fijando arreglos particulares en los bordes, sin importar las pequeñas variaciones en la posición de los mismos. La tercera capa puede ensamblar los detalles en combinaciones que corresponden a una parte de objetos familiares. Las capas siguientes pueden detectar objetos como una combinación de estas partes [29]. La clave del aprendizaje profundo radica en que estas capas de características no están diseñadas por el humano, sino que han sido obtenidas de los datos utilizando procedimientos de aprendizaje de propósito general.

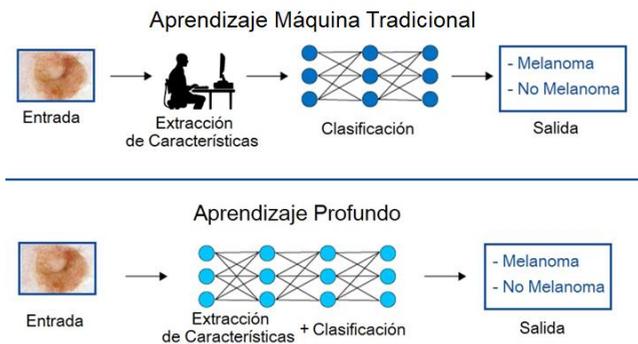


Figura 2.1 Aprendizaje Máquina Tradicional vs Aprendizaje Profundo.

Los métodos de aprendizaje máquina están divididos en algoritmos de aprendizaje supervisado y no supervisado, y una buena descripción de los mismos se encuentra en el trabajo de Suganyadevi et al. [30]. Por otro lado, es importante considerar que la forma común del aprendizaje máquina es el aprendizaje supervisado, que se explica a continuación.

### 2.1.1. Aprendizaje supervisado

En los algoritmos de aprendizaje supervisado se presenta un modelo con un conjunto de datos  $D = \{x, y\}_1^N$  donde  $x$  es el conjunto de características e  $y$  es el conjunto de etiquetas

correspondientes, y  $N$  es el número total de características. El entrenamiento supervisado equivale a encontrar los parámetros  $\theta$  que mejor predicen los datos basados en una función de pérdida  $L(y, \hat{y})$ .  $\hat{y}$  representa la salida del modelo obtenido al alimentar con  $x$  a la función  $f(x; \theta)$  que representa el modelo [31].

Para construir un sistema que sea capaz de clasificar imágenes es necesario entrenar dicho sistema con una cantidad de imágenes adecuada, etiquetadas según su clase. Durante el entrenamiento, la máquina toma una imagen de entrada y genera como salida un vector con puntuaciones, y así se obtiene un vector por cada categoría. Se busca que la categoría deseada tenga la mayor puntuación entre todas las categorías, pero es poco probable que ocurra antes del entrenamiento.

Para cumplir el objetivo de la clasificación es necesario calcular una función objetivo que mida el error (o distancia) entre las puntuaciones de salida y el patrón de puntuaciones deseado. La máquina modifica sus parámetros ajustables internos (llamados pesos) para reducir el error. En el caso de un sistema típico de aprendizaje profundo hay cientos de millones de estos pesos ajustables y será necesario un número elevado de ejemplos etiquetados para entrenar la máquina [29].

La optimización de los pesos se realiza mediante algoritmos basados en el Descenso del Gradiente y empleando la técnica de retropropagación.

Para el correcto funcionamiento de los distintos algoritmos basados en el descenso del gradiente, es necesario fijar adecuadamente la **tasa de aprendizaje  $h$** . Este parámetro es un factor decisivo para el tamaño del ajuste de los pesos realizado en cada iteración y por lo tanto afecta a la tasa de convergencia. En muchos casos puede necesitar varios procesos de prueba y error antes de encontrar un valor adecuado. Si el valor elegido para  $h$  es muy elevado, puede descender rápidamente hacia el mínimo, pero quedar oscilando alrededor de éste. Por otra parte, si el valor elegido de  $h$  es muy pequeño, el descenso va a progresar en pasos muy pequeños, incrementando significativamente el tiempo total de convergencia, aunque se acercará más al valor mínimo.

## 2.2. Redes Neuronales Convolucionales

Las Redes Neuronales Convolucionales (*Convolutional Neural Network* - CNN) aprenden a realizar tareas de clasificación directamente de las imágenes, vídeos, texto o sonidos. Su utilidad se ha incrementado debido a los tres factores siguientes [32]:

- Eliminan la necesidad de realizar una extracción manual de características, ya que éstas se aprenden directamente de la CNN.
- Producen buenos resultados de clasificación.

- Pueden ser reentrenadas para nuevas tareas de reconocimiento, permitiendo la utilización de redes preexistentes.

Las Redes Neuronales Convolucionales están diseñadas para procesar datos que vienen en forma de matrices múltiples, como es el caso de las imágenes en color, las cuales están compuestas por tres matrices de dos dimensiones (2D) que contienen información sobre la intensidad de los píxeles en los distintos espacios de color (RGB, LAB, HSV).

Al igual que cualquier red neuronal, las redes neuronales convolucionales están compuestas por una capa de entrada, varias capas ocultas y una capa de salida. Una arquitectura típica está estructurada como una serie de etapas. Las primeras etapas están compuestas por capas convolucionales y capas de agrupamiento (*pooling*), con las que se realiza la extracción de características. Las últimas capas son capas totalmente conectadas y, por último, existe una capa *Softmax* se genera la salida como se muestra en la Figura 2.3.

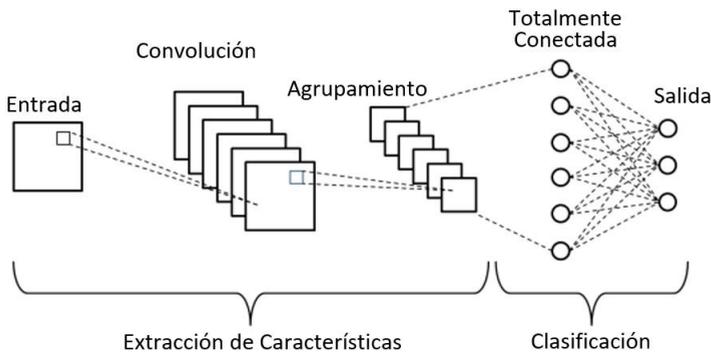


Figura 2.2 Diagrama esquemático de una Red Neuronal Convolutional.

Las unidades en las capas convolucionales están organizadas en mapas de características. Internamente cada unidad está conectada a un área local en el mapa de características de las capas previas a través de un conjunto de pesos llamado banco de filtros [33]. A continuación, se describen las tareas de las capas más comunes [34, 35]:

- **Capa Convulsional:** en esta capa las imágenes de entrada pasan a través de un conjunto de filtros convolucionales, cada uno de los cuales activan ciertas características de la imagen.
- **Unidad Lineal Rectificada (Rectified Linear Unit - ReLu):** se trata de una función de activación que permite un entrenamiento más rápido y efectivo, ya que asigna un cero a los valores negativos y mantiene los valores positivos. Este proceso se conoce como proceso de activación, ya que sólo las características

activadas (no nulas) pasan a la siguiente capa. El comportamiento de esta función se muestra en la Figura 2.4.

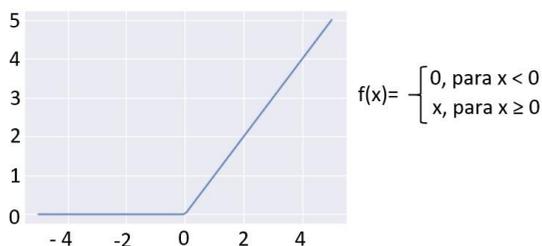


Figura 2.3 Función de activación ReLu.

- **Agrupación (*pooling*):** esta capa realiza una disminución no lineal de la tasa de muestreo (*downsampling*), reduciendo el número de parámetros necesarios para que la red aprenda, simplificando la salida [36].
- **Capas Totalmente Conectadas (*fully connected*):** generalmente se divide en dos, capa de entrada a la capa totalmente conectada, y la capa totalmente conectada. La primera toma las salidas de las capas anteriores y las "aplana" generando un vector de  $n$  dimensiones, donde  $n$  es el número de características de las capas de la etapa anterior; la segunda toma las entradas del análisis de características y aplica los pesos para realizar la predicción correctamente antes de pasar los datos a la capa *Softmax* [37].
- **Softmax:** esta capa recibe los datos de la última capa y calcula la probabilidad para la ocurrencia de cada clase. Esta capa posee la misma cantidad de nodos que la capa de salida de resultados.

En la Figura 2.5 se muestra un esquema del funcionamiento de las capas de una red neuronal convolucional.

Si se quiere resumir el funcionamiento de una red neuronal convolucional en el procesamiento de imágenes, se puede decir que es necesario contar con una base de datos con miles de imágenes para realizar un entrenamiento óptimo y que la red sea capaz de capturar las características de cada objeto y generalizarlo. La entrada de la red son todos y cada uno de los píxeles que posee una imagen, es decir, si una imagen en escala de grises tiene como dimensiones  $40 \times 40$  píxeles, se tendrán 1600 neuronas, pero si se trata de una imagen en color se tendrán 4800 neuronas en la entrada (al tener tres canales, ya sea el espacio de color RGB, HSV, LAB, o cualquier otro).

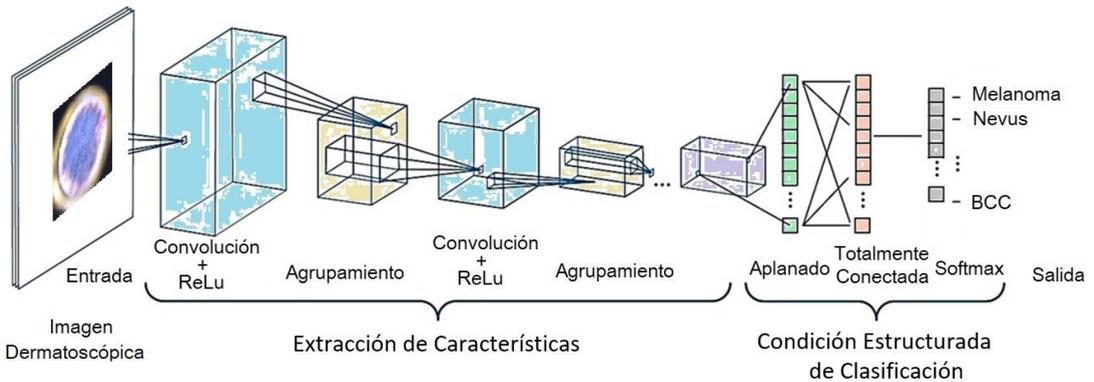


Figura 2.4 Diagrama del funcionamiento de las capas de una CNN. Editado de: [38]

En la realización de esta tesis se utilizaron principalmente tres redes neuronales convolucionales: AlexNet, VGG-16 y GoogLeNet. Estas tres redes fueron utilizadas para procesos de segmentación, y clasificación de imágenes dermatoscópicas. A continuación, se detalla el funcionamiento y arquitectura de cada una de ellas.

### 2.2.1. AlexNet

AlexNet es una red neuronal convolucional con 8 capas de profundidad que ha sido preentrenada con la base de datos ImageNet para el “*ImageNet Large-Scale Visual Recognition Challenge*” (ILSVRC) [39], pudiendo clasificar más de 1000 categorías de objetos y aprender las características que los representan. El tamaño de la imagen de entrada que acepta esta red es de 227×227 píxeles.

Esta red está formada por 8 capas, de las cuales las primeras 5 son capas convolucionales, algunas de ellas seguidas por capas de máximo agrupamiento, siendo las últimas capas totalmente conectadas. Esta red utiliza la función de activación ReLu. La salida de la última capa totalmente conectada alimenta a una capa Softmax, la cual se encarga de la distribución entre las 1000 clases etiquetadas [40]. En la Figura 2.6 se muestra el esquema de la arquitectura de red AlexNet [40].

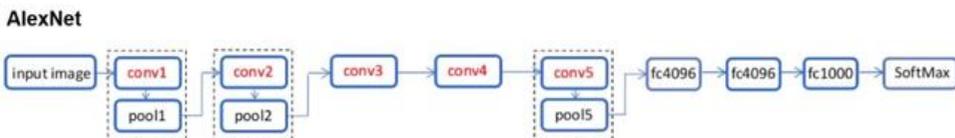


Figura 2.5 Esquema representativo de la arquitectura de red AlexNet. Fuente: [41]

### 2.2.2. VGG-16

VGG es uno de los varios modelos de redes neuronales convolucionales diseñados para el “*ImageNet Large-Scale Visual Recognition Challenge*” (ILSVRC) [39]. La arquitectura del modelo VGG está descrita en [42].

La imagen pasa a través de una pila de capas. El paso de la convolución se ha fijado a un píxel, lo que garantiza que la resolución espacial sea preservada después de la convolución. Se realiza submuestreo aplicando 5 capas de máximo agrupamiento (*max-pooling*) intercaladas entre las capas convolucionales, con una ventana de  $2 \times 2$  píxeles y un paso de 2.

La pila de las capas convolucionales está seguida por 3 capas totalmente conectadas (*fully connected* - FC), cuya configuración es la misma para toda la red; de estas capas, las primeras 2 tienen 4096 canales cada una y la tercera realiza la clasificación de las 1000 imágenes del ILSVRC, por lo que contiene 1000 canales diferentes, uno para cada clase. La capa final es una capa Softmax, además es importante mencionar que todas las capas ocultas poseen una capa ReLu.

En la Figura 2.7 se muestra un resumen de la arquitectura.

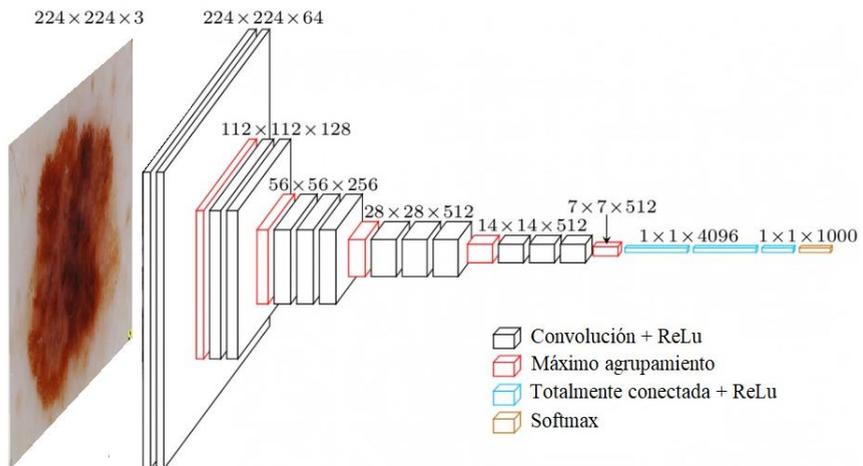


Figura 2.6 Esquema representativo de la arquitectura de Red VGG-16. Editado de: [42]

### 2.2.3. GoogLeNet

GoogLeNet [43] es una red neuronal convolucional, también conocida como **Inception V1**, existiendo las versiones 2, 3 y 4. Es una red con 22 capas de profundidad que ha sido preentrenada con la base de datos ImageNet y Places365. La red preentrenada con ImageNet es capaz de clasificar más de 1000 categorías de objetos y aprender las características que los representan. Por otra parte, la red preentrenada con Places365 es capaz de clasificar entre 365 categorías de lugares diferentes (parques, corredor, salón, etc.).

La red GoogLeNet se encuentra descrita en [42]. Su característica principal es el empleo de módulos **Inception**, como los mostrados en la Figura 2.8. Las características principales de estos módulos son, por un lado, el empleo en paralelo de convoluciones de distinto tamaño y concatenación de las salidas de dichas convoluciones y, por otro lado, el empleo de convoluciones  $1 \times 1$  para reducir la dimensionalidad.

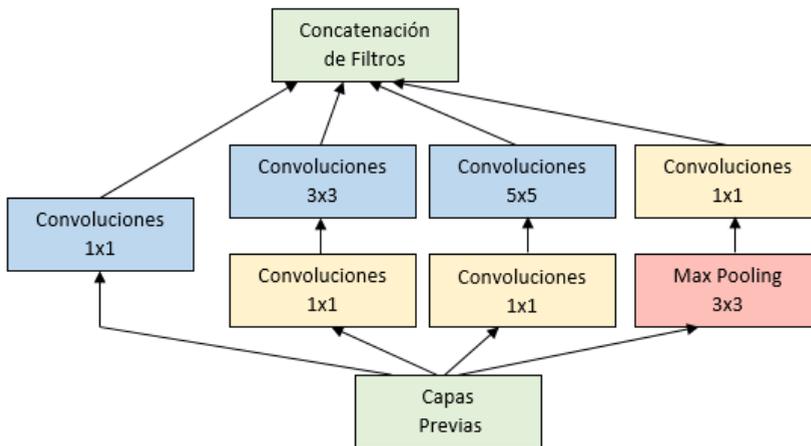


Figura 2.7 Esquema representativo del módulo Inception. Editado de: [44]

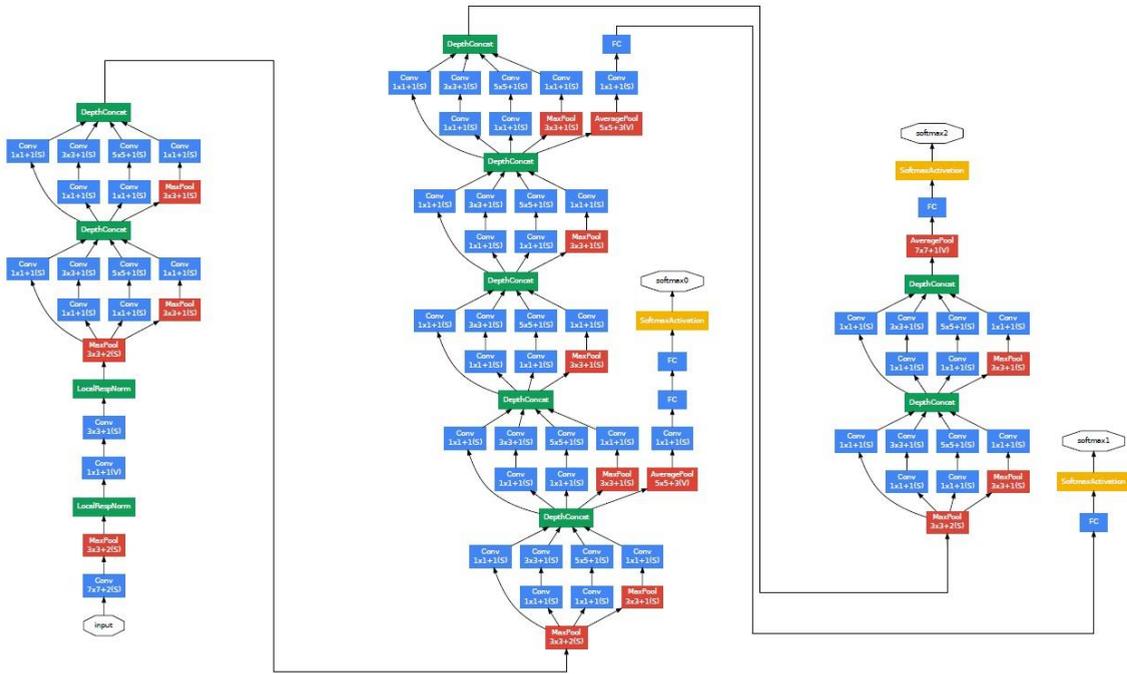


Figura 2.8 Esquema de la red GoogLeNet. Editado de: [44]

GoogLeNet tiene 22 capas de profundidad, si se cuentan tan sólo sus capas con parámetros, y 27 capas si se cuentan las capas de agrupamiento. Si se cuentan todas las capas utilizadas para la construcción de la red, independientemente del número de bloques, serán alrededor de 100.

### 2.3. Procesos que Benefician el Entrenamiento

Los procesos que permiten optimizar el entrenamiento de una red neuronal y previenen que el modelo caiga en sobreajuste (*overfitting*) y el subajuste (*underfitting*) de datos son [45]: Aumento de Datos (*Data Augmentation*), Transferencia del Aprendizaje (*Transfer Learning*) y Ajuste de Precisión (*Fine Tuning*), que se describen a continuación:

- **Aumento de datos:** permite evitar el sobreajuste. Este proceso se realiza al incrementar de forma artificial el conjunto de datos de entrenamiento. El aumento de datos se puede alcanzar al aplicar transformaciones geométricas y de espacio de color o fotométricas a las imágenes [46]. A la base de datos ISIC-2017 se le

aplicaron procesos de volteado (*flipping*), recorte (*cropping*), rotación, traslación, y adición de ruido. En la Figura 2.10 se muestra un ejemplo del aumento de datos realizado.

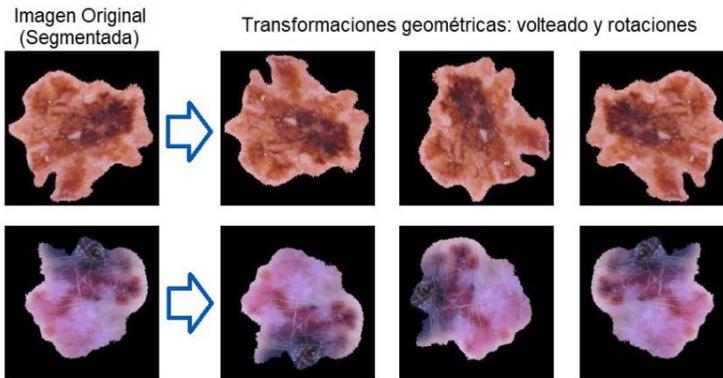


Figura 2.9 Aumento de datos realizado sobre la base de datos ISIC-2017 utilizando transformaciones geométricas.

- **Transferencia de aprendizaje** (*Transfer Learning*): es otra forma de prevenir el sobreajuste y se realiza entrenando la red con un conjunto grande de datos para después utilizar los pesos iniciales de esa red en una nueva tarea de clasificación [46, 47]. Esto permite aplicar de forma inteligente un conocimiento aprendido previamente y resolver nuevos problemas [48]. Los diferentes escenarios que se pueden dar en la transferencia de aprendizaje se encuentran descritos en [49].
- **Ajuste de precisión** (*Fine-Tuning*): consiste en entrenar una CNN desde un conjunto de pesos preentrenados. Al tener la CNN un gran número de pesos y una cantidad limitada de datos etiquetados disponibles, la actualización iterativa de pesos puede alcanzar un mínimo local para la función de coste, lo que sería indeseable. Es por esto, que los pesos de las capas convolucionales pueden ser inicializados con los pesos de una CNN preentrenada que tenga la misma arquitectura y luego ajustados a los datos del problema a resolver [50]. Un ajuste de precisión efectivo comienza en la última capa y va incluyendo capas más internas según vaya siendo el comportamiento de la red para alcanzar el objetivo deseado.

### 3. Segmentación de Imágenes Dermatoscópicas

---

**E**l proceso de segmentación consiste en dividir un conjunto de píxeles en subconjuntos, donde los píxeles de cada subconjunto comparten alguna característica común [51]. Cuando se trata de segmentación de imágenes dermatoscópicas el primer objetivo es aislar la lesión de la piel sana. En el desarrollo de este capítulo se describen varias técnicas de segmentación aplicadas a imágenes dermatoscópicas de la base de datos ISIC-2017 [23]. Además, se describen los resultados obtenidos como parte de los experimentos realizados en la búsqueda de una técnica adecuada de segmentación de lesiones pigmentadas, concretamente melanomas. Para ello se encuentran descritas las métricas consideradas en el proceso de validación de la segmentación de la imagen.

Finalmente, para verificar la eficacia del método de segmentación propuesto en este trabajo, se realiza la comparación de resultados considerando varios trabajos existentes en la literatura que aplicaron sus metodologías de segmentación a la base de datos ISIC-2017, incluyendo al ganador del reto ISIC-2017.

#### 3.1. Metodologías de Segmentación para Imágenes Dermatoscópicas

La segmentación de la lesión se convierte en una tarea desafiante debido a la presencia de artefactos como pueden ser vello, burbujas, distintas condiciones de iluminación, bordes borrosos, vasos sanguíneos, cicatrices o diferentes colores de piel. Algunos ejemplos se muestran en la Figura 3.1. Por este motivo, la segmentación se convierte en un proceso complejo y delicado.

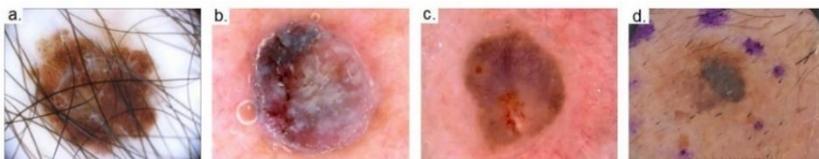


Figura 3.1 Artefactos presentes en imágenes dermatoscópicas. a. Vello. b. Burbujas. c. Vasos Sanguíneos. d. Marcas de tinta.

A través de los años, se han desarrollado muchas técnicas que superan los desafíos de la segmentación. Métodos de segmentación basados en regiones, basados en minimización de energía, máquinas de soporte de vectores (SVM), segmentación basada en Bayes o basada en métodos de aprendizaje profundo (DLBSM) han sido probados con éxito sobre todo tipo de imágenes [52].

Existen un sinnúmero de algoritmos para realizar segmentación de imágenes dermatoscópicas. A continuación, se describen algunos de éstos, los cuales han sido seleccionados ya que su eficacia fue probada sobre la base de datos ISIC 2017.

Entre los algoritmos no supervisados que presentan un buen rendimiento en la segmentación de imágenes dermatoscópicas está GrabCut y K-medias [53], el cual utiliza aprendizaje semisupervisado e incluye el preprocesamiento y la segmentación de la imagen dermatoscópica utilizando información de los bordes y de la región de la lesión. García-Arroyo y García-Zapirain desarrollaron un algoritmo de clasificación borrosa y umbralización de histograma [54], el cual basa su funcionamiento en tres pasos: extracción de características de textura y color, obtención de un modelo borroso para la clasificación de píxeles entre lesión y piel sana y segmentación basada en la umbralización de histograma. El algoritmo SDI para la segmentación automática de una lesión dermatoscópica oscura [55] realiza la segmentación siguiendo tres pasos: preprocesamiento seleccionando una región de interés (*Region of Interest* – ROI), segmentación con el algoritmo de umbralización de Otsu considerando los espacios de color HSV y RGB y, finalmente, realiza el postprocesamiento de la imagen aplicando la operación morfológica de dilatación.

En cuanto a los algoritmos supervisados, existe una gran variedad de ellos que presentan un buen comportamiento. Entre éstos se encuentra el algoritmo de redes completamente convolucionales – deconvolucionales (*Fully Convolutional Deconvolutional Network* - FCDN) [56], el cual aprovecha la arquitectura de bloques convolucionales para extraer características semánticas y mejora la representación a nivel de píxeles. Tan et al. presentan un algoritmo de segmentación que genera un modelo híbrido entre una red profunda y un modelo de agrupamiento (*clustering*) basado en CNN y finalmente un agrupamiento híbrido de segmentos. El algoritmo descrito en [57], basa su funcionamiento en el preprocesamiento de la imagen y realiza la extracción de características al combinar la información contextual, local y global de cada píxel para finalmente realizar postprocesamiento de la imagen previo a su segmentación final. El algoritmo de Red de Rotación Equivalente Profundamente Supervisada descrito en [58] utiliza operaciones de sobremuestreo y proyecciones para adaptar una red de rotación equivariante que resuelve el problema de segmentación de la lesión. En [59] se propone la utilización de una red de indexación de la lesión (*Lesion Indexing Network* - LIN) para solventar el problema de la segmentación de la imagen, la cual está formada por dos redes residuales totalmente convolucionales FCRN-88 (*Fully Convolutional Residual Network* - FCRN) a cuyas salidas se les aplica interpolación bilineal antes de ser sumadas y obtener el resultado final. Otro

algoritmo supervisado es el presentado en [60], el cual consiste en entrenar una FCN basada en la red ResNet para diferenciar las clases (*Deep class-specific learning* - DCL) y después utilizar una aproximación probabilística basada en integración paso a paso (*Probabilistic based Step-wise Integration* - PSI) como paso complementario a la segmentación de la imagen dermatoscópica. El algoritmo descrito en [61] consiste en la aplicación de una red convolucional de resolución completa (FrCN), de tal forma que la red aprende de la información de cada píxel y la segmentación no requiere un preprocesamiento de la imagen. Finalmente, el algoritmo de segmentación descrito en [62] basa su funcionamiento en la aplicación de una red ResNet34 cuyas capas son reutilizadas para el codificador de una arquitectura U-Net, tras el entrenamiento con la base de datos HAM10000.

### **3.2. Algoritmos propuestos**

Entre los algoritmos de segmentación de imágenes se pueden encontrar dos grandes clases: algoritmos no supervisados y algoritmos supervisados. A continuación, se describen los algoritmos que fueron considerados para los experimentos realizados en este trabajo.

#### **3.2.1. Método no supervisado: Algoritmo Basado en Minimización de Energía**

Los métodos no supervisados no requieren un conjunto de datos etiquetados para el entrenamiento. Una de sus principales ventajas es su bajo coste computacional [63]. Otra ventaja de estos métodos es que no requieren una base de datos extensa debido a que no se realiza ningún entrenamiento. En contraste, su rendimiento puede no ser robusto para imágenes de baja calidad, o será necesario alguna interacción con el usuario para lograr un buen rendimiento.

En la literatura existen numerosos algoritmos no supervisados y algunos de ellos se revisan en este trabajo. No obstante, para el análisis final se seleccionó un método de minimización de energía. En este tipo de algoritmos, una medida de energía, la cual incluye información de la región y de sus bordes, es minimizada para resolver el problema de segmentación. Esta tesis presenta la adaptación de un algoritmo de minimización de energías utilizando relajación convexa basada en un trabajo previo de los autores [64, 65]. La idea original fue propuesta por Papadakis y Rabin [66]. Ésta consiste en plantear el problema de segmentación como un problema de minimización de la función de energía convexa. Esta función de energía a minimizar depende de la distancia entre histogramas de cada región dentro de la imagen y el histograma del modelo.

Para extraer los histogramas de la piel sana y de la lesión fue necesario seleccionar un área en ambas regiones como se muestra en la Figura 3.2.

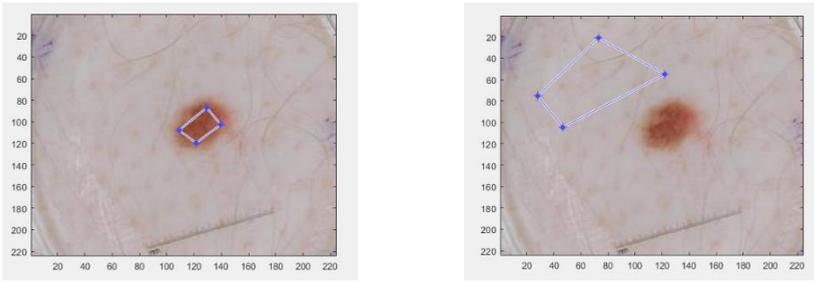


Figura 3.2 Selección de regiones, tanto de piel sana como de la lesión previa a la definición de sus respectivos histogramas.

La segmentación se resuelve como un problema de minimización convexa. Sea  $\lambda$  una función que idealmente tendrá valor 1 dentro de la lesión y 0 fuera de ella. El proceso de segmentación se puede identificar con el siguiente problema de minimización:

$$\min_{\lambda(x) \in [0,1]} \int_{\Omega} (1 - \lambda(x)) C_s(x) dx + \int_{\Omega} \lambda(x) C_t(x) dx + \alpha \int_{\Omega} |\nabla \lambda(x)| dx \quad (3.1)$$

La función de coste  $C_s$  debe tomar valores altos dentro de la lesión y bajos fuera de ella. Por otra parte,  $C_t$  debe tomar valores bajos dentro de la lesión y altos fuera de ella. El término más a la derecha de la expresión (3.1) corresponde al término de regularización y  $\alpha(x)$  es la función de penalización. Idealmente ésta debe ser baja en el borde de la lesión.  $C_s$  y  $C_t$  se denominan términos regionales.

La función de penalización  $\alpha(x)$  se describe en la ecuación (3.2). Los valores  $a$  y  $b$ , se fijaron de forma empírica:  $a=100$ ;  $b=0,003$ .

$$\alpha(x) = \frac{b}{1 + a \cdot |\nabla(\lambda(x))|} \quad (3.2)$$

En la Figura 3.3 se muestran las diferentes máscaras que fueron obtenidas durante el proceso de optimización hasta llegar a la obtención de la máscara final utilizada para la segmentación de la imagen.

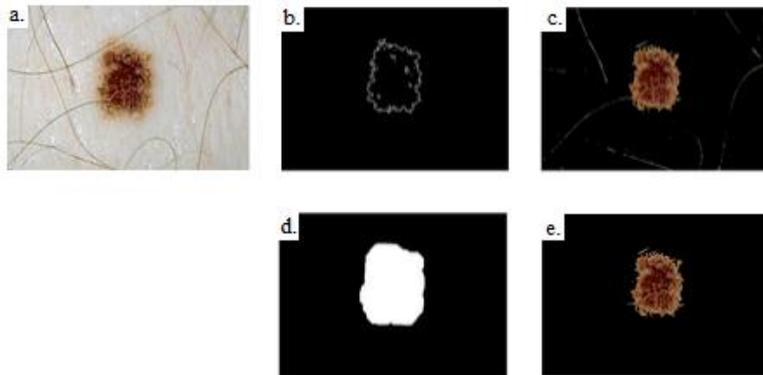


Figura 3.3 Ejemplo de segmentación incluyendo postprocesado de la imagen segmentada.  
 a) Imagen original b) Borde de la lesión. c) Lesión extraída por el algoritmo.  
 d) Nueva máscara invertida. e) Resultado final de la segmentación.

### 3.2.2. Métodos Supervisados

Los métodos supervisados necesitan un conjunto de datos de entrenamiento para fijar los parámetros del clasificador. Tal como describen Al-masni et al. [61], algunos de estos métodos de segmentación están basados en máquinas de vectores de soporte (*Support Vector Machines - SVMs*), clasificador de Bayes, árboles de decisión (*Decision Trees - DTs*) o redes neuronales artificiales (*Artificial Neural Network - ANN*).

Como algoritmos supervisados se seleccionaron dos métodos diferentes. El primer método emplea características profundas de una CNN. Esto presenta la ventaja de requerir una base de datos pequeña. El segundo método de segmentación supervisada es una red neuronal convolucional, cuya eficiencia en la segmentación de imágenes médicas ha sido demostrada. Más específicamente, se ha elegido la red neuronal SegNet [67], dada su probada eficiencia en problemas de segmentación.

A continuación se describen los dos métodos aplicados.

#### 3.2.2.1. Segmentación de imágenes umbralizando capas convolucionales de una CNN

Una CNN posee capas convolucionales que proveen una amplia información sobre las características locales y globales de una imagen. Las capas convolucionales más profundas contienen información de las características globales, abstractas y conceptuales. Por otra parte, las capas convolucionales más bajas dan información sobre las estructuras locales, las cuales son relevantes para el proceso de segmentación [68] [69].

Para este trabajo se utilizó una red VGG-16 preentrenada con la base de datos ImageNet [45]. Se extrajo un conjunto de imágenes de la cuarta capa convolucional de la red VGG-16. Estas imágenes fueron normalizadas y filtradas aplicando un filtro Gaussiano con una desviación estándar igual a 2, antes de ser sumadas. Fue necesario redimensionarlas, ya que después de cada salida del filtro convolucional, éstas poseen distintas dimensiones ( $224 \times 224 \times 64$  en las dos primeras capas y  $112 \times 112 \times 128$  en las dos capas siguientes). Un ejemplo de estas imágenes, después de haber sido redimensionadas, se puede observar en la Figura 3.4.

Finalmente, se aplicó una umbralización utilizando el método de Otsu y operaciones morfológicas (dilatación y relleno de huecos) a la máscara cruda (*Raw Mask* - RM), para obtener el resultado de segmentación final. Un esquema de este algoritmo se muestra en la Figura 3.5.

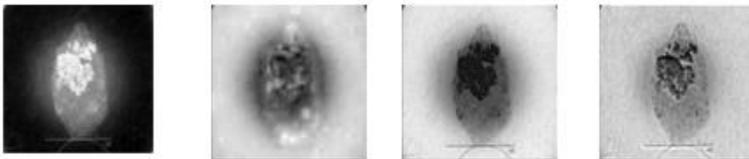


Figura 3.4 Imágenes de las 4 primeras capas convolucionales de la red VGG-16 una vez normalizadas.

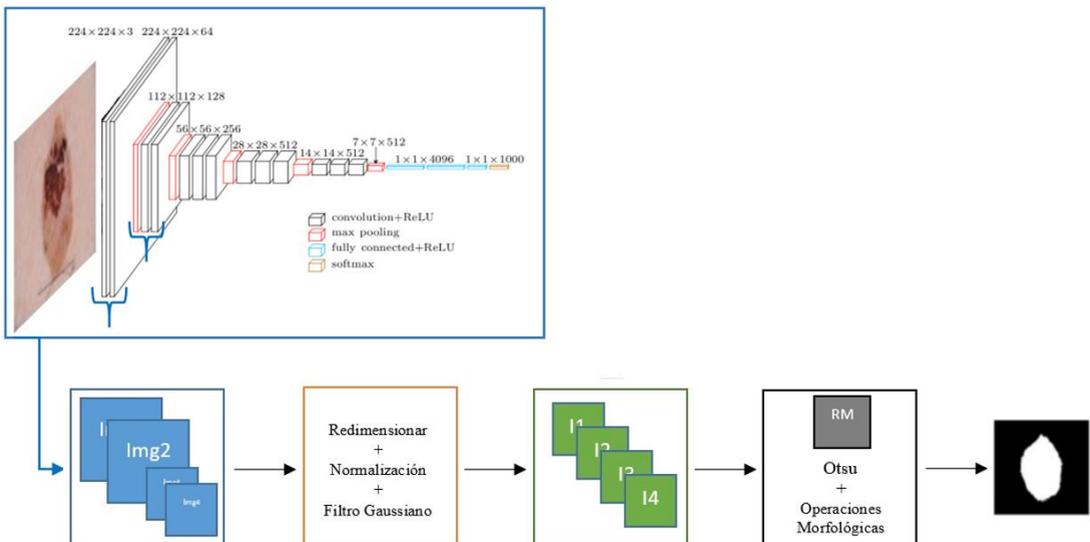


Figura 3.5 Esquema del proceso de obtención de una máscara utilizando imágenes convolucionales (RM: *Raw Mask*).

### 3.2.2.2. Segmentación semántica con la red neuronal profunda SegNet

La segmentación semántica permite identificar un objeto en la imagen al clasificar cada píxel en clases etiquetadas. Como fue descrito por Badrinarayanan et al. [67], SegNet es una arquitectura de Red Totalmente Convolutiva (*Fully Convolutional Network* - FCN) cuyo codificador es topológicamente similar a las capas convolutivas de una VGG-16, pero sin sus capas totalmente conectadas. Las convoluciones son realizadas con un banco de filtros. La última capa del decodificador trabaja como un clasificador *Softmax*, el cual permite obtener etiquetas segmentadas predichas para cada píxel como salida, donde cada etiqueta está asociada a una clase existente. SegNet admite como entrada un mapa de características o una imagen.

Para este trabajo se definieron dos clases, piel y lesión. Las imágenes fueron redimensionadas de su tamaño original  $4288 \times 2848 \times 3$  píxeles a  $224 \times 224 \times 3$  en el espacio RGB, siendo éste el tamaño de la imagen de entrada requerido por la DNN.

El primer paso fue crear una red DAG [70], que utiliza un grafo acíclico dirigido (*Directed Acyclic Graph* – DAG), el cual es un tipo especial de grafo que permite representar datos relacionados de forma única y sin ciclos.

Las capas de la red DAG se convierten en las capas de la arquitectura SegNet. A la entrada de la red DAG llega cada una de las imágenes a segmentar, con una dimensión de  $224 \times 224 \times 3$ . La red DAG fue preinicializada con capas y pesos obtenidos de una CNN VGG-16 preentrenada.

La arquitectura de la red DAG creada se muestra en la Figura 3.6, donde se observa la imagen de entrada, la estructura del DAG configurado con los pesos de la CNN VGG-16 y sus respectivas conexiones.

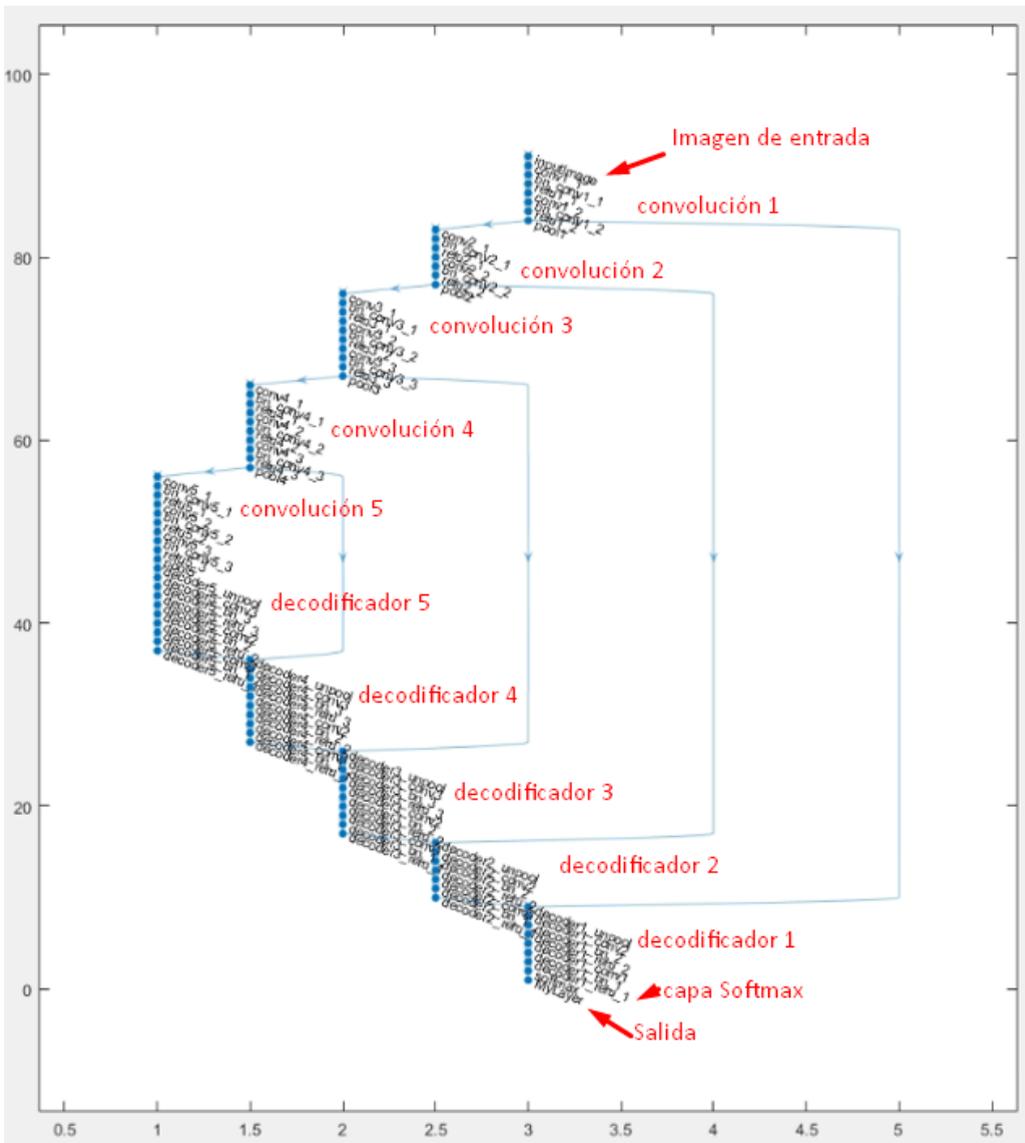


Figura 3.6 DAG configurado para el proceso de segmentación.

### 3.3. Resultados comparativos

Para comparar las metodologías de segmentación descritas en este capítulo, se utilizó la base de datos ISIC-2017, ya que ésta posee la imagen dermatoscópica junto con su máscara de verdad de referencia (*ground truth mask*). Asimismo, se establecieron las métricas de comparación, las cuales se describen a continuación.

#### 3.3.1 Métricas para evaluar un algoritmo de segmentación

Una vez que una imagen ha sido segmentada es necesario evaluar la calidad de dicho proceso, y para esto se han desarrollado distintas formas de medir la similitud entre la imagen original y la imagen segmentada. A estas métricas se las conoce como coeficientes de similitud y a continuación, en la Tabla 3.1 se describe cada una de ellas utilizando los valores de verdadero positivo (*True Positive* - TP), falso positivo (*False Positive* - FP), verdadero negativo (*True Negative* - TN) y falso negativo (*False Negative* - FN).

Tabla 3.1 Coeficientes de similitud [71].

Coeficientes	Expresión	Intervalo de Ocurrencia
Jaccard	$\frac{TP}{TP + FP + TN}$	[0,1]
Sorensen-Dice	$\frac{2TP}{2TP + FP + TN}$	[0,1]

Los coeficientes descritos en la Tabla 3.1 son los que se utilizan con mayor frecuencia para medir la exactitud del proceso de segmentación. Estos coeficientes se describen a continuación:

- Sorensen-DICE [72] conocido como DICE: sirve para medir el número de muestras que compartan dos conjuntos de muestras. Los valores de este coeficiente varían dentro del rango de [0, 1].
- Jaccard o índice Jaccard, descrito en [73], permite medir la similitud de dos conjuntos de muestras binarias. Los valores de este índice pueden variar entre 0 y 1.

Otros coeficientes que se utilizan para verificar el correcto funcionamiento de un algoritmo, no solo de segmentación sino también de clasificación, son los coeficientes de Sensibilidad (*Sensitivity* - Se), Especificidad (*Specificity* - Sp) y Exactitud (*Accuracy* - Acc) [74].

$$\text{Sensibilidad} = \frac{TP}{TP + FN} \quad (3.1)$$

$$\text{Especificidad} = \frac{TN}{TN + FP} \quad (3.2)$$

$$\text{Exactitud} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.3)$$

La Sensibilidad también es conocida como *Recall* y sirve para detectar el número de casos reales positivos, mientras que la Especificidad detecta el número de casos reales negativos.

### 3.3.2. Base de datos y parámetros de evaluación

La base de datos ISIC-2017 provee 2.000 imágenes de entrenamiento, 600 imágenes de prueba y 200 imágenes de validación, aunque estas últimas no fueron utilizadas para la medición de las metodologías presentadas.

Esta base de datos fue considerada debido a que no sólo posee las imágenes dermatoscópicas de las lesiones, sino también ofrece las imágenes de segmentación de la verdad de referencia o “*ground-truth*” que permiten evaluar la metodología de una forma eficaz.

Los parámetros considerados para la evaluación de la metodología de segmentación fueron las siguientes: DICE, Jaccard, exactitud, sensibilidad y especificidad. Las métricas DICE y Jaccard se consideraron ya que, como se explicó anteriormente, son las métricas que mejor miden la exactitud de un proceso de segmentación. Por otro lado, los valores de exactitud, sensibilidad y especificidad fueron considerados por su capacidad de evaluar el funcionamiento de un algoritmo.

Considerando el parámetro de exactitud, DICE y Jaccard los resultados de segmentación resaltan el comportamiento de la SegNet preinicializada con los pesos de las capas de la CNN VGG-16, mejorando a los demás métodos utilizados para realizar el proceso de segmentación de imágenes dermatoscópicas.

En cuanto al comportamiento de cada metodología resulta importante mencionar que la segmentación semántica presenta un alto coste computacional en contraste con la segmentación basada en umbralizar las capas convolucionales, que consume un 72,6% de este coste, mientras que la segmentación basada en minimización de energías consume un 12,33% del total del coste computacional de una segmentación semántica.

Esta metodología utiliza una VGG-16, y extrae información de las primeras 4 capas convolucionales.

### 3.3.3. Comparación de los métodos propuestos

A continuación, se realiza la comparación de resultados obtenidos después de aplicar el método no supervisado basado en minimización de energía y los dos métodos de segmentación basados en aprendizaje profundo.

En el caso del método de segmentación basado en la red neuronal Segnet se utilizó un optimizador del gradiente estocástico descendiente con momentum – SGDM (*Stochastic gradient descent with momentum*). Los demás parámetros de entrenamiento se fijaron con los siguientes valores:

- Momentum: 0,9
- Tamaño del mini-batch: 5
- Número de épocas: 200
- Tasa de aprendizaje inicial: 0,001
- Disminución de pesos (*L2Regularization –weight decay*): 0,005

La red fue entrenada con 2000 imágenes, que después del proceso de aumento de datos se incrementaron a 18000 imágenes obtenidas del conjunto de entrenamiento de la base de datos ISIC-2017. La red DAG entrenada fue utilizada como arquitectura de red de una CNN SegNet con un codificador de profundidad 5, obteniendo una SegNet con 91 capas y 100 conexiones.

El aumento de datos que se aplicó consistió en rotaciones aleatorias en el rango angular de -30 a 30 grados, traslaciones sobre los ejes *X* e *Y* en intervalos de -10 y 10 píxeles, reflexiones horizontales y verticales, y escalado con factor aleatorio entre 0,9 y 1,1. Todas estas operaciones seleccionaron sus valores aleatorios de una distribución continua uniforme.

Más allá de los resultados de la segmentación en cuanto a los valores de las métricas, resulta interesante discutir las ventajas y desventajas de cada metodología. El coste computacional es uno de estos parámetros, y como se mencionó anteriormente existe una gran diferencia en cada caso. Otro aspecto interesante, es el análisis del nivel de exactitud requerido según cada aplicación, ya que la segmentación basada en histogramas permite obtener resultados aceptables considerando el mínimo coste computacional y es útil si se trabaja con una cantidad de imágenes reducida.

En la Tabla 3.2 se muestran los parámetros de evaluación de la segmentación para cada una de las imágenes de la Figura 3.7, donde se representa gráficamente la comparación de las metodologías propuestas, aplicadas a cuatro imágenes dermatoscópicas. Además, en la Tabla 3.3 se muestran los parámetros de evaluación de cada método de segmentación obtenidos de forma global después de evaluar la segmentación del conjunto de prueba de la base de datos ISIC 2017.

Tabla 3.2 Resultados de la segmentación de cuatro imágenes de la base de datos ISIC 2017, según las metodologías aplicadas.

<b>Imagen</b>	<b>Metodología</b>	<b>DICE</b>	<b>Jac</b>	<b>Acc</b>	<b>Se</b>	<b>Sp</b>
a.ISIC 0014392	Segmentación semántica	<b>0,9326</b>	<b>0,8738</b>	<b>1,0000</b>	<b>0,9685</b>	0,8738
	Segmentación por minimización de energías	0,9116	0,8376	0,9594	0,8382	<b>0,9997</b>
	Segmentación por umbralización de imágenes convolucionales	0,8938	0,8081	0,9503	0,8383	0,9875
b.ISIC 0012134	Segmentación semántica	<b>0,9460</b>	<b>0,8976</b>	<b>0,9829</b>	<b>0,9654</b>	0,9291
	Segmentación por minimización de energías	0,7466	0,5956	0,8678	0,5956	<b>1,0000</b>
	Segmentación por umbralización de imágenes convolucionales	0,8351	0,7168	0,9075	0,7168	<b>1,0000</b>
c.ISIC 0012272	Segmentación semántica	<b>0,8438</b>	<b>0,7299</b>	<b>0,9891</b>	<b>0,9046</b>	0,7449
	Segmentación por minimización de energías	0,7487	0,5984	0,6703	0,6100	0,9197
	Segmentación por umbralización de imágenes convolucionales	0,5720	0,4005	0,5172	0,4005	<b>1,0000</b>
d.ISIC 0015436	Segmentación semántica	<b>0,9551</b>	<b>0,9141</b>	<b>0,9672</b>	0,9757	<b>0,9999</b>
	Segmentación por minimización de energías	0,9181	0,8486	0,9539	<b>0,9979</b>	0,9386
	Segmentación por umbralización de imágenes convolucionales	0,8534	0,7442	0,9191	0,9094	0,9223

Tabla 3. 3 Comparación de las metodologías de segmentación para el conjunto de test de la base de datos de ISIC-2017.

<b>Metodología</b>	<b>DICE</b>	<b>Jac</b>	<b>Acc</b>	<b>Se</b>	<b>Sp</b>
Segmentación semántica	<b>0,8548</b>	<b>0,7730</b>	<b>0,9557</b>	<b>0,8533</b>	<b>0,9632</b>
Segmentación por minimización de energías	0,5937	0,4927	0,8647	0,6020	0,9131
Segmentación por umbralización de imágenes convolucionales	0,5853	0,4700	0,8170	0,8627	0,8690

### 3.3.4. Comparación del mejor método propuesto con otros de la literatura

En la literatura existen un sinnúmero de metodologías de segmentación supervisados y no supervisados. En este apartado se han considerado trabajos que han utilizado la base de datos ISIC-2017 y que evaluaron su proceso de segmentación con los mismos parámetros considerados en este trabajo.

En la Tabla 3.4 se muestra la comparación de 14 metodologías de segmentación, tanto supervisada como no supervisada, donde se encuentran incluidos los resultados del proceso de segmentación desarrollado en este trabajo. Esta tabla muestra el resultado de segmentar las 600 imágenes del conjunto de prueba (*test*) de la base de datos ISIC-2017.

En términos del coeficiente DICE, el método que mejor valor obtuvo fue el presentado por Al-masni et al. [61] con un valor de 0,8708 frente a un 0,8504 obtenido por el método propuesto en este trabajo. Por otra parte, en términos del coeficiente Jaccard el método con un mejor valor fue el propuesto por Tang et al. [57] con un valor de 0,7926 frente a un 0,7730 obtenido por el método propuesto en este trabajo. Como se puede observar, los valores obtenidos por la metodología planteada en este trabajo, si bien no superan a otros métodos, sí permite obtener parámetros de segmentación muy cercanos a los óptimos en todas las métricas.

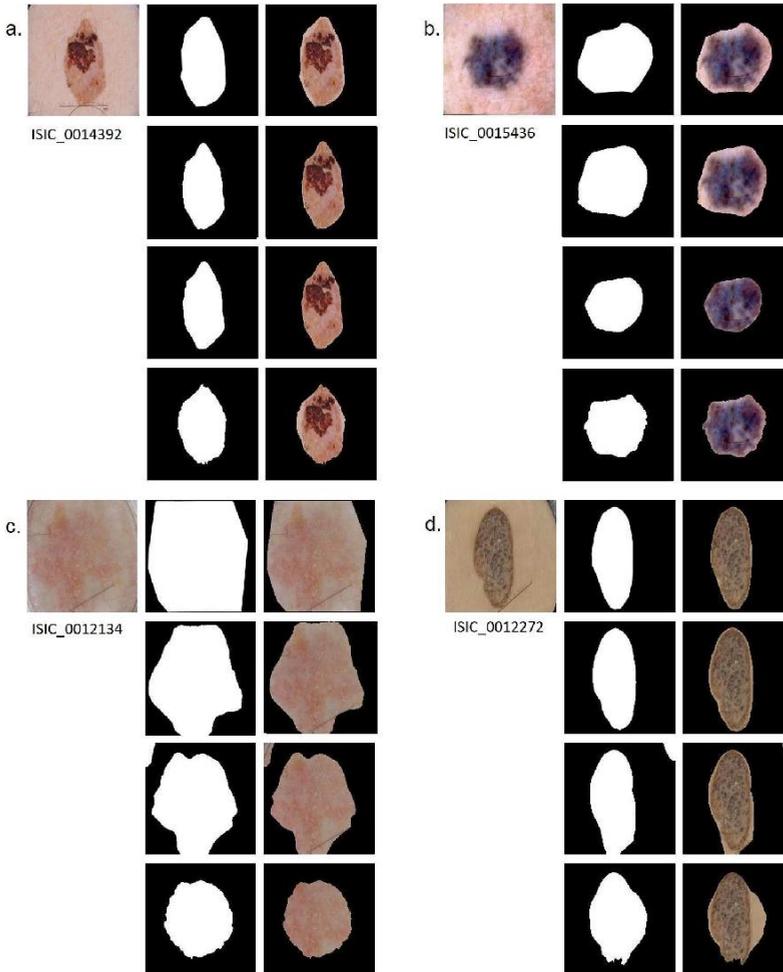


Figura 3.7 Imágenes segmentadas. Para las imágenes a, b, c, d, la primera fila muestra la imagen original, la máscara y el resultado segmentado. Para las 4 imágenes, las filas 2, 3 y 4 representan la máscara y la segmentación obtenidas por el método de segmentación semántica, segmentación por minimización de energías y segmentación utilizando imágenes convolucionales.

Tabla 3.4 Resultados de segmentación tras aplicar distintos procesos.

Ref.	Metodología	DICE	Jac	Acc	Se	Sp
[56]	FCDN- Primer lugar en el reto ISBI-2017	0,8490	0,7650	0,9340	0,8250	0,9750
[53]	GrabCut y k-medias	0,7141	0,6048	0,8779	0,7568	0,9613
[75]	Optimización por enjambre	0,7677	-	-	-	-
[54]	Clasificación Fuzzy e histogramas	0,7600	0,6650	0,8840	0,8690	0,9230
[57]	Separable-UNet modelo con esquema de pesos estocástico	0,8693	<b>0,7926</b>	0,9431	<b>0,8953</b>	0,9632
[76]	Pre procesamiento + CNN + Postprocesamiento	0,8550	-	0,9360	0,8290	0,9510
[59]	LIN (Red de lesión indexada en una FCRN-88)	0,8390	0,7530	<b>0,9500</b>	0,8550	<b>0,9740</b>
[58]	Red de rotación equivalente profundamente supervisada	0,8560	0,7723	0,9355	0,8540	0,9715
[55]	Modelo SDI	0,7820	0,6920	0,8880	0,8130	0,9270
[77]	DCL-PSI	0,8566	0,7773	0,9408	0,8620	0,9671
[61]	Arquitectura FrCN	<b>0,8708</b>	0,7711	0,9403	0,8540	0,9669
[62]	Modelo LinkNet152	0,8530	0,7700	-	-	-
**	Esta contribución: SegNet	0,8548	0,7730	0,9357	0,8533	0,9632



## 4. Aprendizaje Máquina para Detección de Patrones Globales

---

La extracción de características es una etapa fundamental en la clasificación de imágenes dermatoscópicas mediante aprendizaje máquina, ya que la clasificación se basa en el análisis de las características presentes en la imagen [78]. Ésta puede definirse como la colección de técnicas que procesan y extraen datos específicos como textura, forma o posición [79]. Una imagen dermatoscópica contiene gran cantidad de información debido a la existencia de varias estructuras y colores. Además, se debe considerar la existencia de posibles artefactos que deben ser filtrados de la imagen para no generar características erróneas de la lesión.

Existen varias técnicas para el análisis de patrones de una imagen dermatoscópica. Entre las más conocidas en el ámbito clínico están la Regla ABCDE, Regla de Menzies y la Regla de los 7 puntos. Desde el punto de vista técnico, para implementar técnicas automáticas que emulen el análisis de lesiones de la piel, existen las técnicas basadas en aprendizaje profundo, así como también las características diseñadas ad-hoc (*handcrafted features*) que permiten cuantificar las características clínicas de la lesión.

Una imagen dermatoscópica, según se indica en el Atlas Interactivo de la Dermatoscopia EDRA [19], puede presentar 9 tipos distintos de **patrones globales**, siendo éstos: reticular, globular, empedrado, homogéneo, estrellado (*starburst*), paralelo, lagunar e inespecífico. Por otra parte, estos patrones globales pueden tener **características locales** las cuales pueden ser: red pigmentada, ramales, estructuras de regresión, puntos o glóbulos, velo azul blanquecino, hipopigmentación y estructuras vasculares.

Las características o patrones globales describen a la imagen como un objeto completo generalizado, mientras que las características locales describen sectores en un objeto de la imagen. En el caso de las imágenes dermatoscópicas, se dice que las características locales describen sectores en la lesión.

La correcta identificación de los patrones globales y de sus características locales resulta fundamental en el diagnóstico de una lesión, donde lo primero es definir si ésta es melanocítica o no melanocítica, para luego verificar su malignidad o benignidad. En la

Figura 4.1 se muestran los recortes de varias imágenes dermatoscópicas, donde se puede observar las características locales de una lesión de piel. Por otra parte, en la Figura 4.2 se muestra un ejemplo de imágenes dermatoscópicas clasificadas según un patrón global.

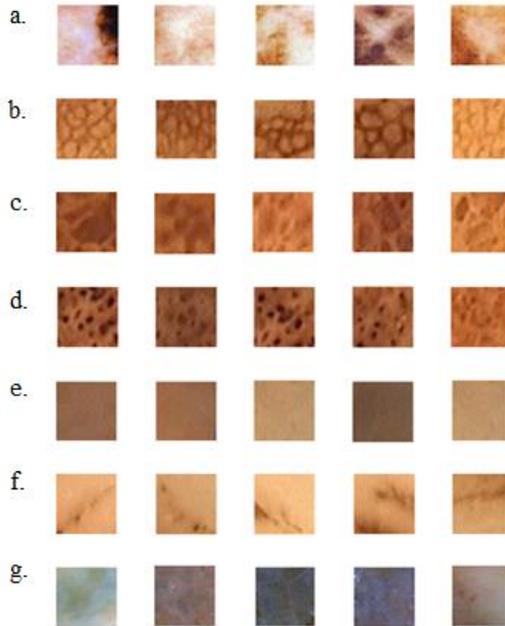


Figura 4.1 Características Locales. Descripción por fila: a) Regresión b) Reticular c) Empedrado d) Globular e) Homogéneo f) Paralelo g) Velo Blanco azulado.

El objetivo de este capítulo es identificar distintos patrones globales en la imagen dermatoscópica de una lesión pigmentada. La detección de patrones globales sirve para el diagnóstico de una lesión realizado mediante “análisis de patrones” [80, 81]. En este capítulo, se plantea un método para realizar la detección de tres patrones globales, empedrado, homogéneo y reticular. Este proceso se realizó sobre las imágenes de la base de datos ISIC 2017 e ISIC 2018.

La organización de este capítulo se describe a continuación. En primer lugar, se describe el proceso de división de la lesión utilizando superpíxeles, en el cual se utilizó la imagen de la lesión previamente segmentada de la piel sana. Los superpíxeles agrupan píxeles con características similares. En segundo lugar, se describe el método aplicado para realizar la extracción de las características de textura de la imagen, ya que la textura es una propiedad nativa de la piel [82]. En tercer lugar, se describe el proceso de selección de características significativas dentro del total de características extraídas, y que fueron utilizadas en el proceso de clasificación y etiquetado de patrones. Finalmente se presentan los resultados

obtenidos después del proceso de clasificación de los tres patrones mencionados anteriormente.

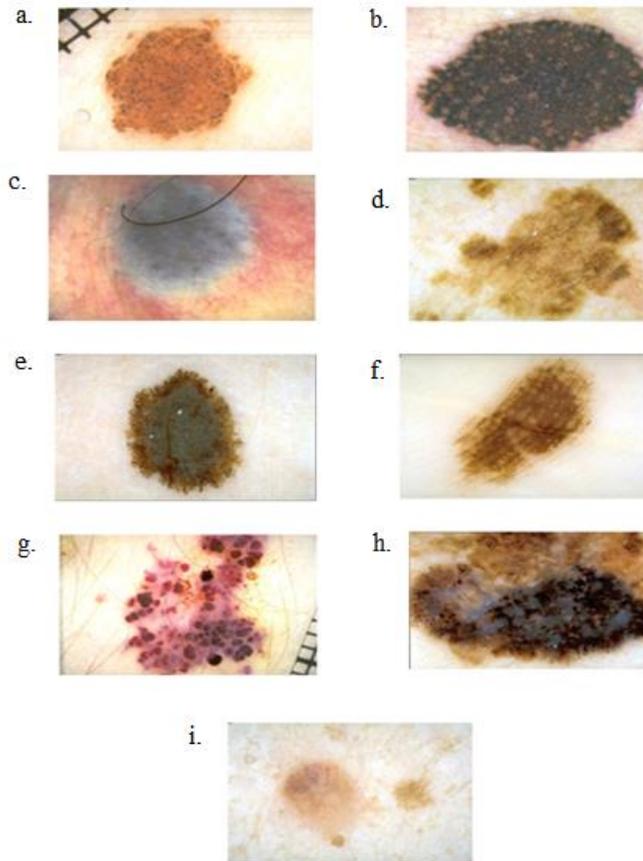


Figura 4.2 Patrones Globales. a) Empedrado b) Globular c) Homogéneo d) Reticular e) Estrellado f) Paralelo g) Lagunar h) Patrón múltiple i) Patrón no especificado. Fuente: [20]

#### 4.1. Características diseñadas ad-hoc

Las características diseñadas ad-hoc son aquellas que tratan de cuantificar propiedades visuales de la imagen. Es necesario obtener estas características cuando se va a clasificar la información de la imagen dermatoscópica utilizando métodos tradicionales de aprendizaje máquina como los árboles de decisión o las SVMs.

##### 4.1.1. Características de Textura

Las características de textura de una imagen tratan de extraer de ésta propiedades referentes a la suavidad y rugosidad de los objetos, así como detectar la existencia de patrones

repetitivos, tales como líneas o puntos. Para definir una textura es necesario cuantificar propiedades de ésta tales como tamaño, forma, color y orientación de los elementos del patrón que se repite en la imagen [83, 84].

En esta sección se describen los métodos para realizar la extracción de características de textura de una imagen dermatoscópica. Asumiendo que existen  $n$  muestras, denotadas como  $\{x^i\}_{i=1}^n$ , de las cuales se desea extraer características, se puede decir que el proceso de extracción de características pretende extraer un conjunto de características (dimensiones) de un patrón de datos denotado como  $\{y^j\}_{j=1}^d$ , siendo éste un nuevo espacio de dimensiones y no un subconjunto de  $\{x^i\}_{i=1}^n$  [85].

Los métodos para el análisis de textura pueden ser categorizados de la siguiente manera: estadísticos, basados en modelos, transformadas o métodos de filtrado en el dominio espacial como se detalla en [86]. A continuación, se detallan brevemente algunos de estos métodos descritos en la literatura.

**Métodos estadísticos**, representan la textura al estimar momentos estadísticos que gobiernan la distribución y relaciones entre los valores de los píxeles y que están variando lenta o constantemente sobre la región a la que se calculan las características de textura. Dentro de estos métodos se encuentran los siguientes:

- Matriz de Co-ocurrencia a nivel de escala de grises (*Gray Level Co-occurrence Matrix – GLCM*), descrita en [87]. Es una de las más conocidas y aplicadas en el análisis de textura. Esta matriz indica la cantidad de pares de píxeles que, teniendo una posición relativa determinada, poseen ciertos niveles de gris dentro de un área específica.

Los parámetros que se pueden extraer de la matriz GLCM son: contraste, correlación, energía y homogeneidad, como se detalla en [88]. A continuación, se detallan cada uno de estos valores.

- Homogeneidad: permite medir la proximidad de la distribución de los elementos en la matriz GLCM a su diagonal.

$$\text{Homogeneidad} = \sum_{i,j}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (4.1)$$

siendo  $P_{i,j}$  la probabilidad de coocurrencia de los valores de gris  $i$  y  $j$  para una distancia dada.

- Contraste: también conocida como varianza, representa la medida del contraste de intensidad entre un píxel y su respectivo píxel vecino en toda la imagen.

$$\text{Contraste} = \sum_{i,j=0}^{N-1} P_{i,j} (i - j)^2 \quad (4.2)$$

- Entropía: obtiene información sobre la aleatoriedad de los elementos de la matriz. Utiliza la distribución de los valores de gris en la imagen.

$$\text{Entropía} = - \sum_{i,j}^{N-1} P_{i,j} \ln(P_{i,j}) \quad (4.3)$$

- Correlación: indica cuantitativamente la medida de correlación de un píxel con su vecino en la imagen.

$$\text{Correlación} = \sum_{i,j=0}^{N-1} P_{i,j} \left[ \frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right] \quad (4.4)$$

- Media: no es el valor promedio de los valores originales de los niveles de gris dentro de la ventana, ya que el valor del píxel no es ponderado por su propia frecuencia sino por la frecuencia de su coocurrencia combinado con un valor del píxel vecino.

$$\text{Media} = \sum_{i,j}^{N-1} iP_{i,j} \quad (4.5)$$

- Desviación estándar: su medida se basa en la media y dispersión alrededor de la media de los valores de la celda de la GLCM.

$$\sigma_i = \sqrt{\sigma_i^2} \quad (4.6)$$

$$\sigma_j = \sqrt{\sigma_j^2} \quad (4.7)$$

- Energía: también conocida como uniformidad corresponde a la raíz cuadrada del Segundo Momento Angular (*Angular Second Moment – ASM*)

$$\text{ASM} = \sum_{i,j=0}^{N-1} P_{i,j}^2 \quad (4.8)$$

$$\text{Energía} = \sqrt{\text{ASM}} \quad (4.9)$$

- Matriz de dependencia del nivel de gris adyacente (*Neighboring Gray Level Dependence Matrix - NGDLM*), descrito en [89], es un método de rotación invariante para determinar la textura. Éste considera la relación entre un píxel y todos sus vecinos, dentro de una determinada distancia del píxel indexado, en lugar de sólo una dirección y dentro de un rango predeterminado de nivel de gris del píxel estudiado.

**Métodos de filtrado del dominio espacial o métodos espaciales**, se encuentran principalmente enfocados en medir la densidad de los bordes por unidad de área. Este proceso se realiza al aplicar filtros en una imagen en el dominio espacial y, a través de un análisis inverso, la información de la imagen original es recuperada. Esta categoría une a los métodos que están basados en filtros lineales y *textons*, así como también otros no lineales, como los patrones binarios locales [86]. En este grupo de métodos se incluyen:

- Leyes de energía de las máscaras (*Laws Energy Masks* – LEM) [90].
- *Textons*, [91].
- Patrones binarios locales (*Local Binary Pattern* – LBP) [92] [93].
- Histograma de gradientes orientados (*Histogram of Oriented Gradients* – HOG) [94].

### **Métodos basados en modelos**

Describen o sintetizan la textura de una imagen con un modelo matemático. Los parámetros estimados del modelo capturan la naturaleza de la textura. Su desventaja es la complejidad computacional usualmente envuelta en la estimación de sus parámetros. Algunos de estos métodos son:

- Campos aleatorios de Markov [95].
- Modelos fractales [96].
- Modelos autorregresivos [97].

### **Modelos de Transformadas**

Generan un grupo de características desde imágenes transformadas a un espacio cuyo sistema de coordenadas tiene una interpretación específica, como frecuencia o espacio de escala. Entre estos modelos se encuentran los siguientes:

- La Transformada Discreta de Fourier (Discrete Fourier Transform – DFT) [98].
- Wavelets [115].
- Filtros Gabor y la transformada Gabor Wavelet – GWT [99].

#### **4.1.2. División de la imagen en superpíxeles**

La extracción de características está ligada al proceso de segmentación de la imagen. Como paso inicial se segmentó la lesión de la piel sana. Para la identificación de los patrones locales de la lesión (características) se realizó la segmentación de la misma utilizando superpíxeles. Un superpíxel agrupa conjuntos de píxeles con características similares dentro de una imagen.

Para dividir la lesión en superpíxeles se aplicó el algoritmo de agrupación iterativa lineal simple – SLIC (*Simple Linear Iterative Clustering*).

El algoritmo SLIC crea superpíxeles basados en el espacio 5D (5 dimensiones), el cual está definido por el espacio de color LAB y las coordenadas  $(x,y)$  para cada píxel.

SLIC requiere un parámetro externo  $K$ , el cual indica el número de superpíxeles a crear, de tal manera que los grupos pueden ser creados de un tamaño lo suficientemente grande para considerar toda la información de la región. A partir de este parámetro, se calcula el intervalo de malla,  $S$ , como:

$$S = \sqrt{\frac{N}{K}} \quad (4.10)$$

$S^2$  puede considerarse como una aproximación de área de los superpíxeles.

También se puede especificar el valor de compacidad de los superpíxeles,  $m$ . El algoritmo sugiere un intervalo entre 1 y 20.

El algoritmo requiere el cálculo del centroide de cada superpíxel, que se denota por  $C_k = [l_k, a_k, b_k, x_k, y_k]^T$ .

A partir de estos parámetros, el algoritmo SLIC calcula la distancia normalizada  $D_s$  según la ecuación 4.11:

$$D_s = d_{lab} + \frac{m}{S} d_{xy} \quad (4.12)$$

en donde  $d_{lab}$  representa la distancia euclídea en el espacio CIELAB entre el píxel y el centroide del superpíxel, como se expresa en la ecuación (4.12.) y  $d_{xy}$  representa la distancia espacial entre el píxel y el superpíxel, tal y como se expresa en la ecuación 4.13.

$$d_{lab} = ((l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2)^{1/2} \quad (4.12)$$

$$d_{xy} = ((x_k - x_i)^2 + (y_k - y_i)^2)^{1/2} \quad (4.13)$$

SLIC aplica un algoritmo de agrupamiento que, iterativamente, asigna cada píxel al superpíxel para el que su distancia normalizada es menor y recalcula los centroides de cada superpíxel.

Una vez obtenidos los superpíxeles dentro de la lesión, se analiza cada uno, para así identificar por medio de un proceso de clasificación el tipo de patrón o característica que posee.

La Figura 4.3 ilustra el funcionamiento del algoritmo SLIC y la agrupación de superpíxeles en una imagen dermatoscópica. Además, esta figura contiene una muestra de los superpíxeles extraídos.

Tras la división de la lesión en superpíxeles, se realizó el análisis de cada superpíxel utilizando la matriz GLCM, de la cual se obtuvieron los valores de contraste, correlación, energía, entropía, homogeneidad y máxima probabilidad.

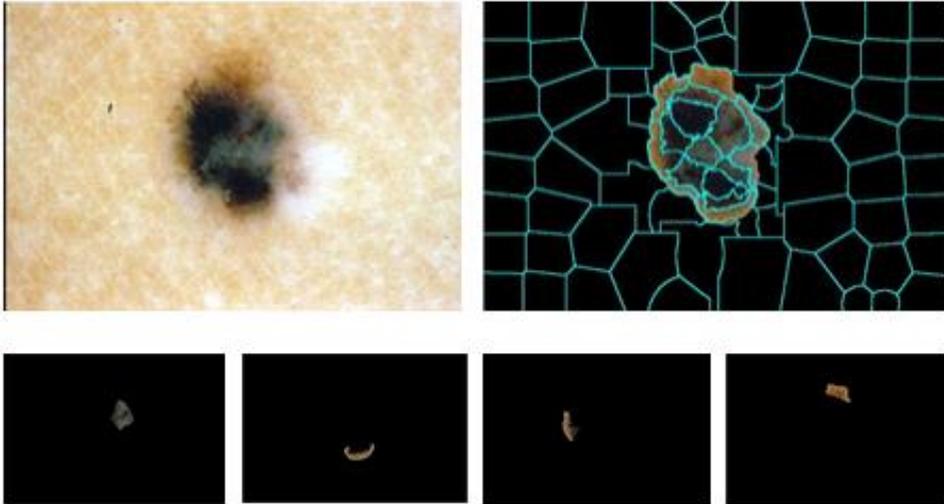


Figura 4.3 Imagen dermatoscópica con la lesión segmentada en superpíxeles. En la fila inferior se expone una muestra de 4 superpíxeles pertenecientes a la lesión.

## 4.2. Selección de Características

La selección de características tiene como objetivo reducir el mapa de dimensiones  $\{y^j\}_{j=1}^d$  obtenido en el proceso de extracción de características [85]. Este proceso implica obtener un subconjunto optimizado de características con un alto poder de discriminación, sin redundancias y sin datos irrelevantes, simplificando los procesos computacionales al momento de aplicar un clasificador [100].

Para obtener un subconjunto de características que cumpla con el objetivo de la selección de características, es necesario descartar todos los elementos de  $\{y^j\}_{j=1}^d$ , a los que llamaremos variables, que no sean relevantes. Para lograr esto, tal como se detalla en [101], es posible determinar la influencia de cada variable sobre el resultado de la inferencia del bloque de aprendizaje, o intentar asignar una puntuación a cada característica de entrada según una métrica de selección determinada y elegir las variables con las puntuaciones más altas. Una vez que se realiza este proceso, es posible que existan variables redundantes al tener una alta correlación entre sí, por lo que varios algoritmos de selección de características categorizan las variables de  $\{y^j\}_{j=1}^d$  en dos categorías: características con una alta relevancia, ya que si se descarta una de estas variables el resultado de la

clasificación disminuye su precisión; categorías con relevancia baja, en cuyo caso, si se elimina esta variable la clasificación no empeora pero si son descartadas del subconjunto de características seleccionado, se generará una disminución en el rendimiento del clasificador. Además, es necesario que el algoritmo de selección de características sea capaz de identificar y eliminar todas aquellas características cuya relevancia no sea alta ni baja y que no afectarán a la clasificación final.

Si se considera un número  $d$  de características, el número total posible de subconjuntos que se podrán generar a partir de  $\{y^j\}_{j=1}^d$  será de  $2^d$ , y sería inviable enumerar el número exponencial de subconjuntos de características si  $d$  es demasiado grande. Por lo tanto, es necesario encontrar un método que evalúe los subconjuntos en un tiempo razonable [100]. La evaluación de los subconjuntos está basada en un criterio, el cual puede ser categorizado como uno de los siguientes métodos: filtro y envolvente.

#### 4.2.1. Métodos de Filtro

Se los conoce como métodos de filtro porque filtran las características antes de alimentar un modelo de aprendizaje. Estos métodos están basados en dos conceptos, relevancia y redundancia. La relevancia es la dependencia (correlación) de la característica con el objetivo y la redundancia se encarga de verificar si las características comparten información redundante [85]. Dentro de estos métodos, los más comunes son:

- Criterio de Correlación también conocido como Dependencia Medida y está basado en la relevancia (poder de predicción) de cada característica [102].
- Información Mutua, también conocida como Ganancia de Información, es la medida de la dependencia o información compartida entre dos variables aleatorias [103].
- Estadísticos  $\chi^2$ , mide la dependencia (relevancia) de la aparición de características en el valor objetivo, y se basa en la distribución de probabilidad  $\chi^2$  [104].
- Markov Blanket, es un método basado en la relevancia y que considera a cada característica y a cada objetivo como un nodo en una red Bayesiana [105].
- Filtros basados en consistencia, que usan una medida de consistencia basada en la relevancia y la redundancia. s un criterio de selección que pretende retener el poder de discriminación de los datos definidos por las características originales [106].
- Filtro basado en Correlación rápida [107], utiliza una medida basada en la entropía llamada incertidumbre simétrica para encontrar correlación tanto para la relevancia como para la redundancia.
- Interacción [108]. Puesto que algunas características podría no tener ningún efecto de forma individual pero junto a otras características dan alta correlación con el objetivo y aumentan el rendimiento de la clasificación, este método encuentra de

forma eficiente las interacciones entre las características para llevar a cabo la selección de características.

- Redundancia Mínima – Máxima Relevancia [109], que está basado en maximizar la relevancia y minimizar la redundancia de las características.

#### 4.2.2. Métodos envolventes

Estos métodos integran el modelo dentro de la búsqueda de subconjuntos de características, generando diferentes subconjuntos de características y los evalúa entrenando y probándolo en el modelo. De esta forma, el algoritmo para la búsqueda del mejor subconjunto subóptimo del conjunto de características, está esencialmente envuelto alrededor del modelo. Para generar los distintos subconjuntos de características a explorar, estos métodos emplean la búsqueda heurística. Estos métodos de búsqueda se pueden dividir en dos categorías: Secuenciales y Metaheurísticos.

- Métodos de selección secuencial, también llamados secuenciales debido a la naturaleza iterativa del algoritmo. Su funcionamiento se basa en añadir o eliminar características del subconjunto de éstas, basándose en un criterio de selección. Este método es una forma común de selección de características y posee dos componentes que son: función objetivo y un algoritmo de búsqueda secuencial [100].
  - Función Objetivo: también conocida como **criterio**, es el método que busca minimizar el subconjunto de características sobre todas las opciones posibles. Los criterios comunes son el error mínimo cuadrado y la tasa de error de clasificación, utilizados para modelos de regresión y modelos de clasificación respectivamente.
  - Algoritmo de búsqueda secuencial: este algoritmo se encargará de añadir o eliminar características del subconjunto candidato mientras evalúa el criterio. Debido a que realizar una comparación del valor del criterio a los  $2^d$  subconjuntos donde  $d$  es el número de características factibles muy costoso computacionalmente, este método realiza búsquedas secuenciales en una sola dirección (siempre crece o siempre decrece el número de características).

Considerando que este método realiza la búsqueda tan solo en dos direcciones posibles y una a la vez, se tienen dos variantes: selección secuencial hacia adelante, SFS (*Sequential Forward Selection*), y selección secuencial hacia atrás, SBS (*Sequential Backward Selection*) [110, 111].

- Métodos Metaheurísticos, estos algoritmos tienen una baja complejidad de implementación y se pueden adaptar a una gran cantidad de problemas. Además, estos algoritmos son menos propensos a atascarse en un óptimo local en comparación con los otros métodos secuenciales, donde la función objetivo es el desempeño del modelo.

### 4.3. Experimentos para la detección de patrones globales

En este apartado se describe la metodología basada en Aprendizaje Máquina que se ha seguido para determinación de los patrones globales presentes en una lesión. .

En primer lugar, se segmentó la lesión, tal y como se describe en el capítulo 3. A continuación, la lesión se divide en regiones de textura homogénea mediante el algoritmo SLIC, tal y como se detalla en el apartado 4.3.1. A continuación se extrajeron características de textura, como se describe en el apartado 4.3.2, y, tras una selección de característica, se emplean para clasificar los patrones presentes en la lesión (apartado 4.3.3).

#### 4.3.1. Segmentación de la lesión para identificar características locales

Habiendo aislado a la lesión de la piel sana se procedió a segmentar la lesión, utilizando los superpíxeles obtenidos al aplicar el algoritmo SLIC. Para esto se utilizaron como valores de los parámetros:  $K=100$  para el número máximo de superpíxeles en cada lesión, y  $m=1000$  como valor de la compacidad con el objetivo de obtener superpíxeles con formas regulares.

Una vez obtenidos los superpíxeles, cada uno de ellos fue analizado como una imagen individual En la Figura 4.4 se muestra la división de una lesión en superpíxeles.

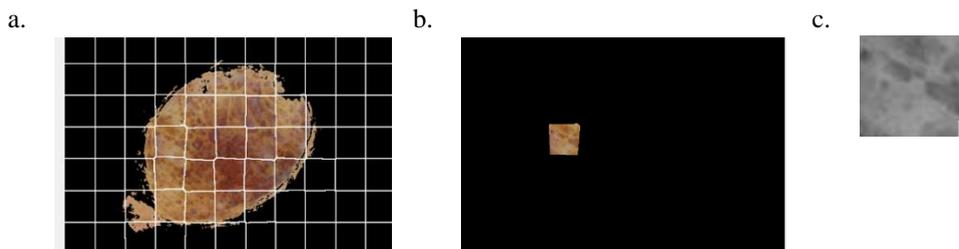


Figura 4.4 Obtención del recorte. a) SLIC aplicado a la imagen dermatoscópica. b) Superpíxel aislado. c) Recorte en escala de grises.

Cada recorte fue postprocesado realizando un reescalado de los niveles de gris para realzar los bordes dentro de la lesión. El postprocesamiento del recorte  $I$ , se realizó siguiendo los siguientes pasos:

- Primero, se aplicó el filtro de mediana de la imagen  $I$  y la salida se asignó a  $J$ .
- Segundo, a la nueva imagen  $J$  se le aplicó un filtro de realce (*unsharp masking*). Los parámetros aplicados en el enmascaramiento que permiten obtener una imagen con los bordes de la lesión realzados fueron: desviación estándar del filtro paso baja que se sustrae de la imagen original igual a 7, y el factor por el que se multiplica el término de realce, que se suma a la imagen original, que fue 1,5.

De esta manera se pudo disponer de una región con aumento del contraste, como se observa en la Figura 4.5.

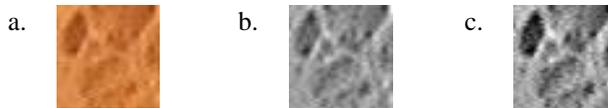


Figura 4.5 a) Recorte original. b) Imagen filtrada. c) Realce.

#### 4.3.2. Extracción de características de los recortes de textura

Para obtener las características de textura se aplicaron métodos estadísticos. En primer lugar, se extrajeron los primeros cuatro parámetros estadísticos del histograma, que son: media, varianza, torsión y curtosis; un ejemplo de estos histogramas se muestra en la Figura 4.6.

Por otra parte, el segundo conjunto de parámetros fue obtenido utilizando información de la matriz GLCM. Para esto se establecieron los siguientes valores de distancia  $d$ , y orientación,  $\theta$  :

$$d = [1, 3, 6, ]$$
$$\theta = [0^\circ, 45^\circ, 90^\circ, 135^\circ]$$

De cada una de las matrices GLCM se extrajeron los siguientes parámetros: contraste, correlación, energía, homogeneidad, entropía y máxima probabilidad (ecuaciones 4.1 a 4.9),

Por último, se creó una matriz de características por con los parámetros obtenidos de los estadísticos de primer y segundo orden. Cada matriz está compuesta por  $n$  filas por 72 columnas, siendo 72 el número de parámetros extraídos a cada recorte y,  $n$  el número de recortes (superpíxeles) de la lesión.

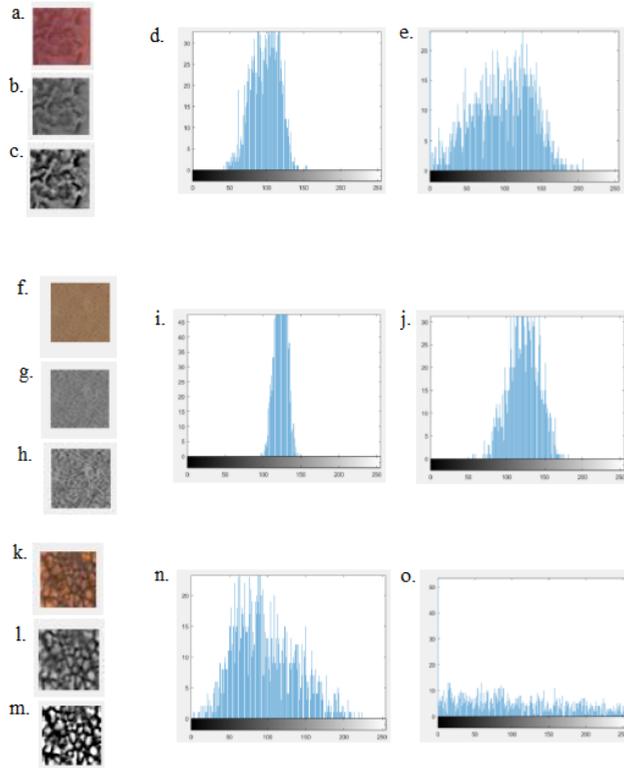


Figura 4.6 Recortes. a,b,c,d,e) Recorte Empedrado: original, escala de grises, recorte postprocesado, histograma del recorte en escala de grises e histograma del recorte postprocesado. f,g,h,i,j) Recorte Homogéneo: original, escala de grises, recorte postprocesado, histograma del recorte en escala de grises e histograma del recorte postprocesado. k,l,m,n,o) Recorte Reticular: original, escala de grises, recorte postprocesado, histograma del recorte en escala de grises e histograma del recorte postprocesado.

### 4.3.3. Selección de características y clasificación

Se realizó la extracción de 126 recortes y de todos éstos se procedió a extraer sus características en una matriz descrita en la ecuación 4.14, formada por  $126 \times 76$  elementos. Como se observa, es una matriz con un volumen extenso de datos y es por esto que se procedió a la selección de características.

$$V = \begin{bmatrix} v_{1,1} & \cdots & v_{1,76} \\ \vdots & \ddots & \vdots \\ v_{126,1} & \cdots & v_{126,76} \end{bmatrix} \quad (4.14)$$

Para la selección de características se utilizaron los algoritmos de Selección Secuencial hacia Adelante (SFS) y Selección Secuencial hacia Atrás (SBS).

Cada variable de la matriz  $V$  se normalizó antes de aplicar los algoritmos SFS y SBS. Con las características seleccionadas por ambos algoritmos, se creó una nueva matriz de características  $S$ , cuyas dimensiones son  $126 \times n$ , siendo  $n$  el número de características seleccionadas.

$$\begin{aligned} \text{SBS} &= [\text{sbs}_1, \dots, \text{sbs}_a] \\ \text{SFS} &= [\text{sfs}_1, \dots, \text{sfs}_b] \\ N &= a+b \end{aligned} \tag{4.15}$$

$$S = \begin{bmatrix} s_{1,1} & \dots & s_{1,n} \\ \vdots & \ddots & \vdots \\ s_{126,1} & \dots & s_{126,n} \end{bmatrix} \tag{4.16}$$

Las características finalmente seleccionadas se describen en la Tabla 4.1, donde se indica las características seleccionadas a una distancia e inclinación determinadas. Se observa la gran reducción en el número de características que han pasado de un conjunto de 76 a 10.

Tabla 4. 1 Características seleccionadas.

Característica	Distancia ( $d$ )	Inclinación ( $\theta$ )
Contraste	1	0°
Contraste	1	90°
Contraste	3	90°
Correlación	1	0°
Correlación	6	0°
Correlación	1	45°
Energía	1	90°
Energía	1	135°
Máxima Probabilidad	3	0°
Máxima Probabilidad	6	0°

El clasificador utilizado consiste en un árbol de decisión con análisis de componente principales (*Principal Components Analysis* - PCA) [112]. Para el diseño del árbol de decisión, se le introduce como entrada  $i_1=S$  e  $i_2=F$ , siendo  $F$  un vector que contiene la clase asociada a cada recorte relacionado con la variable categórica.

$$F = \begin{bmatrix} \text{Empedrado} \\ \dots \\ \text{Empedrado} \\ \text{Homogéneo} \\ \dots \\ \text{Homogéneo} \\ \text{Reticular} \\ \dots \\ \text{Reticular} \end{bmatrix} \quad (4.17)$$

Como resultado de aplicar el árbol de decisión con PCA, se obtuvo un clasificador denotado como  $E$ . Este clasificador  $E$  está compuesto por dos elementos principales: el árbol de decisión y los coeficientes PCA. A su vez, el árbol de decisión tiene tres componentes principales: la matriz  $Z$  que contiene las características seleccionadas de  $S$ ; el vector  $F$  con el nombre de la clase a la que pertenece cada muestra de entrenamiento; y finalmente el vector  $c$  que contiene los nombres de las clases etiquetadas después de la clasificación. Por otro lado, los coeficientes PCA muestran la correlación de las variables originales con las componentes principales, que recogen la mayor parte de la variabilidad. Esto se puede observar en la Figura 4.7.

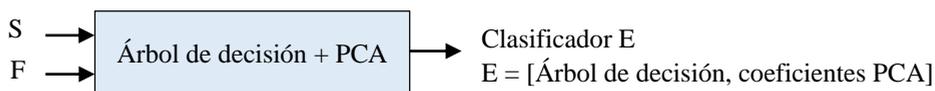


Figura 4.7 Esquema del clasificador

#### 4.4. Resultados

Al aplicar el clasificador  $E$  se obtuvo una exactitud del 99,2% en la clasificación entre tres patrones: homogéneo, empedrado y reticular. El número de muestras que se utilizaron se describen en la Tabla 4.1 y a partir de éstas, el algoritmo encontró las regiones de decisión sobre el conjunto de muestras suministrado. En la Figura 4.8 se muestran las predicciones realizadas por el modelo.

Tabla 4. 2 Muestras de entrenamiento.

Valor	Cantidad	Porcentaje
Empedrado	28	22,22%
Homogeneo	53	42,06%
Reticular	45	35,71%

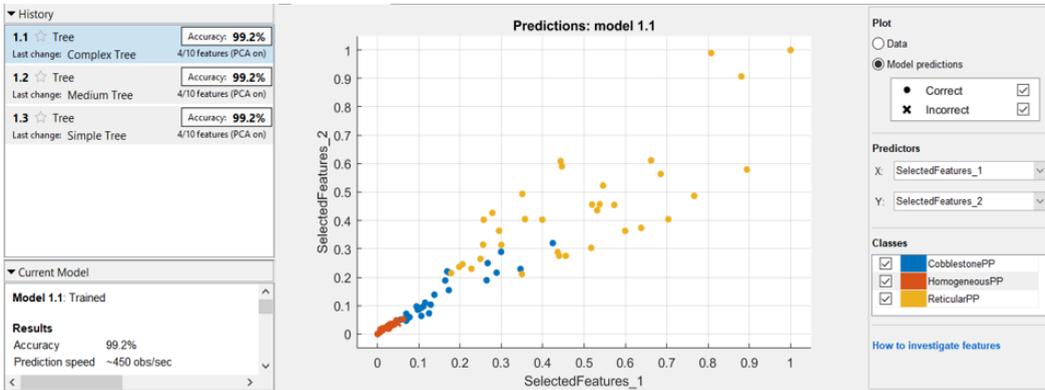


Figura 4.8 Modelo predictivo utilizando el subconjunto de características seleccionadas.

La componente PCA mantuvo una cantidad suficiente de componentes para explicar el 95% de varianza total. De esta forma, la varianza por componente obtenida fue la siguiente: 46,4%, 27,0%, 14,9%, 6,8%, 2,2%, 1,3%, 0,8%, 0,4%, 0,2%, 0,1%. Después de realizar el entrenamiento la PCA mantuvo 4 componentes y se generó el árbol de decisión que se muestra en la Figura 4.9.

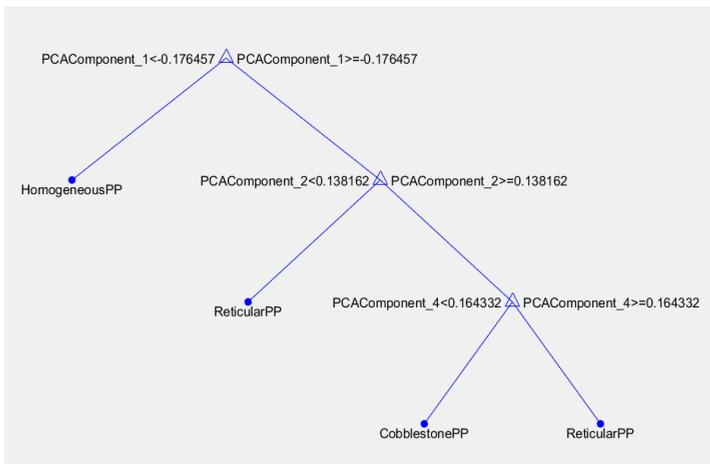


Figura 4.9 Árbol de decisión

En la Figura 4.10 se muestra la matriz de confusión que representa la eficiencia del modelo y finalmente en la Tabla 4.3 se muestran los resultados de la clasificación de tres imágenes correspondientes a los tres patrones dermatoscópicos analizados.

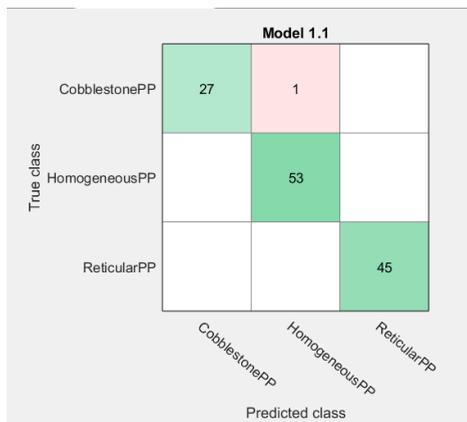
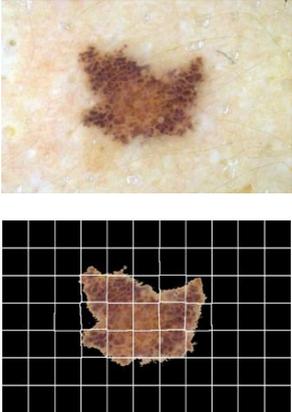
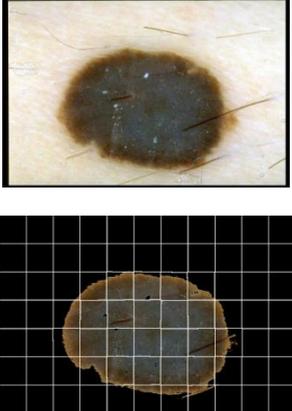
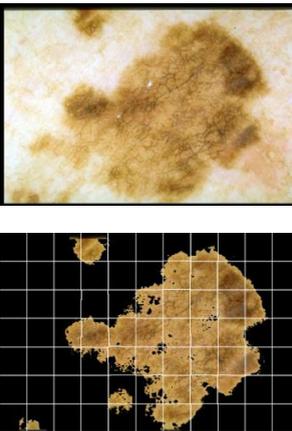


Figura 4.10 Matriz de confusión

Tabla 4. 3 Clasificación.

	Lesión original y segmentada	Análisis por superpíxeles	Clasificación por superpíxel	Lesión clasificada
Empedrado		Superpíxel 1:  Superpíxel 2:  Superpíxel 3:  Superpíxel 4: 	Superpíxel 1: Empedrado  Superpíxel 2: Empedrado  Superpíxel 3: Reticular  Superpíxel 4: Empedrado	Empedrado
Homogéneo		Superpíxel 1:  Superpíxel 2:  Superpíxel 3:  Superpíxel 4:  Superpíxel 5: 	Superpíxel 1: Homogéneo  Superpíxel 2: Homogéneo  Superpíxel 3: Homogéneo  Superpíxel 4: Homogéneo  Superpíxel 5: Homogéneo	Homogéneo
Reticular		Superpíxel 1:  Superpíxel 2:  Superpíxel 3: 	Superpíxel 1: Empedrado  Superpíxel 2: Reticular  Superpíxel 3: Reticular	Reticular

## 5. Clasificación

---

El proceso de clasificación es el último paso para lograr el objetivo obtener un diagnóstico preciso de la misma. En este capítulo, más allá de describir las técnicas de clasificación supervisadas y no supervisadas, se encuentran descritos los procesos que se siguieron para alcanzar la identificación precisa de un tipo de lesión cutánea particular, el carcinoma basocelular (*Basal Cell Carcinoma – BCC*), dada su alta incidencia. Según la Sociedad Americana del Cáncer 8 de cada 10 casos de cáncer de piel son BCC [3].

En la actualidad y desde hace varios años se vienen utilizando técnicas de aprendizaje máquina supervisado como parte del procesamiento de imágenes dermatoscópicas. Las técnicas de aprendizaje supervisado, las cuales aprenden de un banco de imágenes de entrenamiento y permiten obtener los datos etiquetados en la salida, poseen un potencial de exactitud prometedor en cuanto al análisis de imágenes médicas, concretamente de imágenes dermatoscópicas. Esto es, facilitan la caracterización y detección de posible lesiones malignas [113].

En la literatura se mencionan tres técnicas de clasificación aplicadas a imágenes médicas que utilizan los beneficios de las CNNs [114]:

- Entrenamiento desde cero, es decir se entrena la red sin contar con información de una red preentrenada con datos.
- Utilización de una CNN “fuera de plataforma”, es decir valerse de la CNN para extracción de características y manejar estos datos como información complementaria a las características obtenidas ad-hoc (*handcrafted features*).
- Realizar un preentrenamiento con imágenes generales y realizar una transferencia de aprendizaje y ajuste fino en imágenes médicas objetivo, utilizando una CNN o cualquier otro tipo de modelo de aprendizaje profundo.

Este capítulo se encuentra organizado de la siguiente manera: como primer punto se expone el estado de arte sobre los métodos de clasificación y sus aplicaciones a imágenes dermatoscópicas; en segundo lugar, se presentan los datos y el diseño de los experimentos realizados para alcanzar la clasificación del BCC; en tercer lugar, se muestran los clasificadores diseñados y los distintos resultados que se obtuvieron al clasificar BCC versus distintos tipos de lesiones y aplicar distintos métodos de segmentación de la imagen; como cuarto punto, este capítulo incluye la implementación de dos ensambladores y finalmente se muestra la discusión de los resultados obtenidos.

### 5.1. Estado del Arte de la Clasificación de lesiones dermatoscópicas

Para realizar la clasificación de imágenes dermatoscópicas se han desarrollado diversas metodologías enfocadas mayoritariamente en la clasificación binaria de melanoma versus otro tipo de lesiones como las descritas por Kenneth et al. [115], las cuales en su mayoría utilizan herramientas de reconocimiento de imágenes para realizar la detección del melanoma. Albahar logra una clasificación binaria de imágenes dermatoscópicas utilizando una arquitectura con dos capas convolucionales seguida por capas de agrupamiento y “*drop-out*”, y donde se aplican métodos de regularización de forma embebida a cada capa convolucional, obteniendo un 97,49% de exactitud y un 0,98 de área bajo la curva (AUC) [116]. Maron et al. presentan la comparación del funcionamiento de clasificadores binarios de imágenes dermatoscópicas basados en el modelo ResNet-50 [117].

Achim et al. presentan la combinación de CNN y la intervención humana para realizar una clasificación multiclase y una clasificación binaria de lesiones dermatoscópicas, alcanzado una precisión de 82,95% , frente a 81,59% con respecto a un clasificador exclusivamente formado por CNN [118]. Por otro lado, Gessert et al. presentan un método de clasificación de imágenes dermatoscópicas basado en la red *EfficientNet* [1]. Este trabajo fue el ganador del ISIC-2019 *Challenge* y, por lo tanto, en este capítulo se considera este trabajo al momento de validar los resultados obtenidos con el método aquí propuesto.

Existen numerosos trabajos enfocados a la detección de melanoma y sus características, mientras que muy pocos autores se han centrado en la identificación y clasificación del carcinoma basocelular. Huang et al. analizan la traslucidez de lesiones BCC mediante el entrenamiento de un clasificador *Softmax* con características de alto nivel obtenidas al aplicar un *autoencoder* [128]. Serrano et al. realizan la detección de las características dermatoscópicas del BCC y además detectan BCC frente otras lesiones, obteniendo una sensibilidad de 0,99, una especificidad de 0,94 y una precisión de 0,97 [129]. Marka et al. presentan una revisión de artículos dedicados a detectar carcinomas [119].

## 5.2. Métodos

Para la realización de este trabajo, se realizaron varios experimentos buscando la detección de carcinoma basocelular utilizando la base de datos ISIC-2019:

- 1) BCC versus Nevus.
- 2) BCC versus cualquier otro tipo de lesión.
- 3) Clasificación en ocho lesiones diferentes, siendo una de ellas el BCC.

Los experimentos que se realizaron basaron su funcionamiento en la precisión que brindan las redes profundas en las tareas de clasificación. Se utilizaron tres tipos de DNN:

- 1) Red VGG-16, sobre la cual se realizaron modificaciones de su arquitectura;
- 2) Red Inception-V3;
- 3) Red ResNet50.

La arquitectura de la CNN VGG-16 fue modificada para optimizar su funcionamiento en la tarea de clasificación de una imagen dermatoscópica, por lo que de aquí en adelante se denominará VGG-16 Ajustada. A continuación, se presenta un análisis breve del funcionamiento de cada una de estas redes:

- ResNet-50 es una red residual (*Residual Network - ResNet*), la cual está basada en una red plana a la que se le han insertado conexiones de acceso directo (*shortcuts*) para convertirla en su versión residual equivalente [120]. Estos accesos directos permiten realizar saltos sobre algunas capas, para así evitar la desaparición de gradientes y reutilizar las activaciones de las capas anteriores, hasta que las capas adyacentes aprendan sus pesos. ResNet-50 posee 50 capas de profundidad y admite como entrada a imágenes de 224×224 píxeles. Los creadores de esta red demostraron que la exactitud en los procesos de clasificación de imágenes es mucho mayor que si se utilizan redes planas como VGG-19. A continuación en la Figura 5.1 se muestra el modelo de esta arquitectura de red.

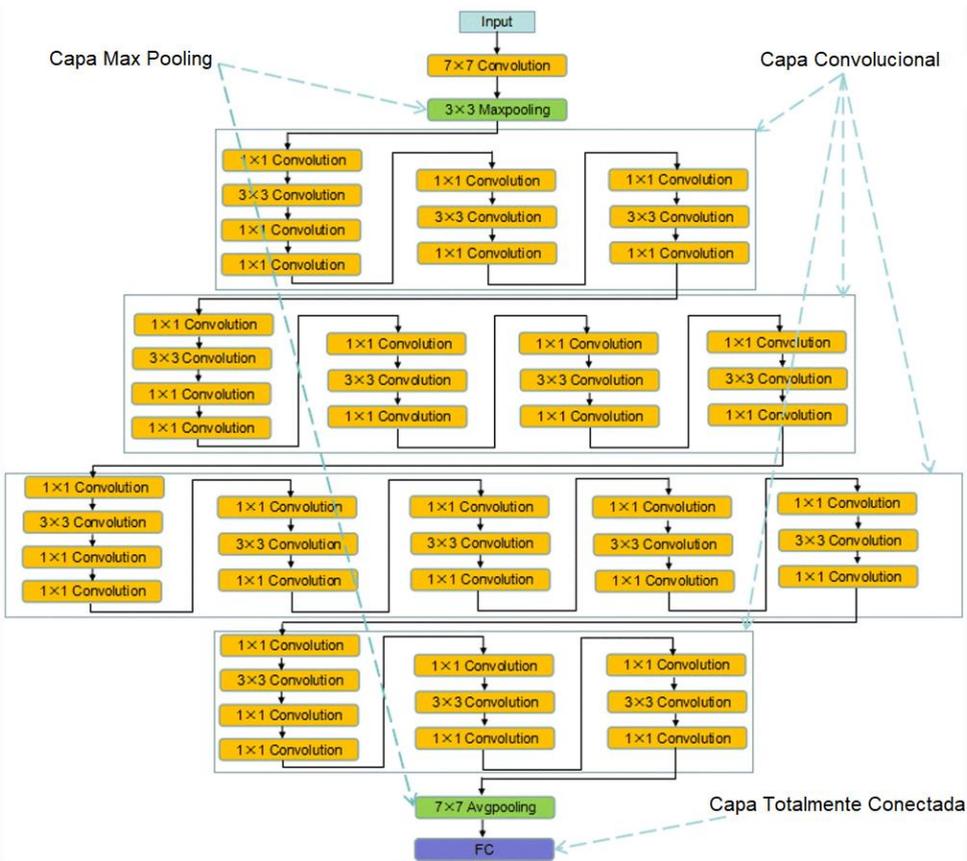


Figura 5.1 Arquitectura ResNet50. Fuente: [114].

- Inception-V3. Como se menciona en el Capítulo 2, en sus inicios eran módulos que forman la red GoogLeNet. A partir de estos módulos se desarrolla la DNN Inception-V3, descrita en [121, 122]. Esta red profunda consiste en la repetición de bloques llamados Inception, los cuales son altamente ajustables. Cada módulo consiste en un conjunto de filtros convolucionales en un rango de  $5 \times 5$  a  $1 \times 1$ , que entregan un resultado en cascada en cada paso y aseguran que no se pierda la localización de la información, garantizando la perseverancia de la información de las características generales de cada imagen. Los autores de este modelo, que posee 48 capas de profundidad, aseguran que esta red reduce el número de parámetros a ser entrenados. En la Figura 5.2 se muestra el esquema de la arquitectura de esta DNN.

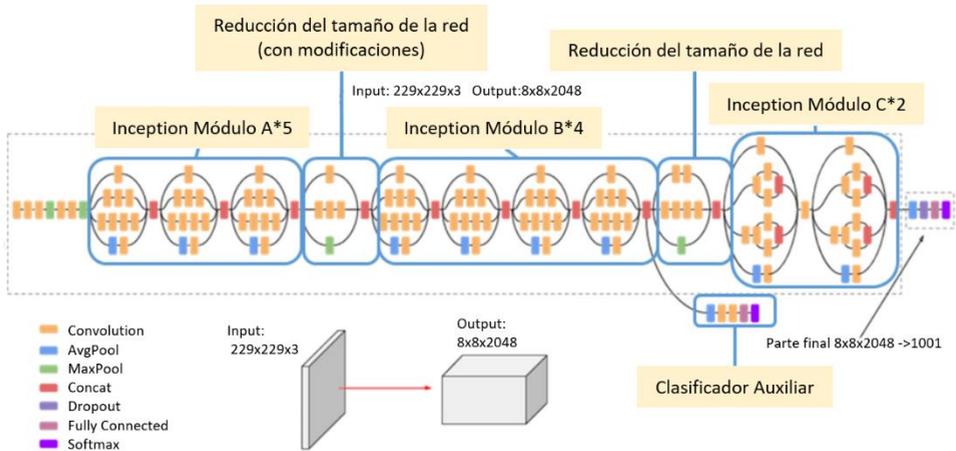


Figura 5.2 Arquitectura de la Red Inception-V3 . Fuente: [44].

- VGG-16 Ajustada, es un modelo basado en la red VGG-16 pero con una modificación en su arquitectura. Esta modificación se basó en añadir una capa de *Zero Padding* a cada bloque de capas convolucionales, como se muestra en la Figura 5.3. Al realizar esta modificación se aplican diferentes ajustes a los límites de la imagen dermatoscópica permitiendo conservar el tamaño original de la misma.

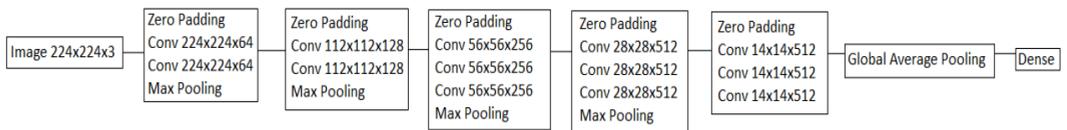


Figura 5.3 VGG-16 Ajustada (A-VGG16).

Para el entrenamiento de cada una de las redes mencionadas, se utilizó el conjunto de entrenamiento de la base de datos ISIC 2019. Para los experimentos que se detallan en esta sección, se realizó la clasificación de las lesiones aplicando, por una parte, el proceso de segmentación explicado en el Capítulo 3, y en contraste se realizó la clasificación con las imágenes sin segmentar. Adicionalmente, se realizaron procesos de balanceo de datos como paso previo al entrenamiento.

Las salidas predichas por estos tres modelos de redes neuronales fueron enlazadas utilizando dos modelos de ensambladores denominados ArgMax y Mean, respectivamente, que permitieron obtener un modelo optimizado de clasificación y aplicar lo que se conoce como aprendizaje ensamblado.

### 5.2.1. Ensamblado de redes

El ensamblado de redes sirve para reducir la varianza y mejorar el comportamiento de la predicción, combinando las predicciones de múltiples modelos de redes neuronales. Las técnicas para aprendizaje ensamblado pueden ser agrupadas por elementos que varían, como por ejemplo: datos de entrenamiento, el modelo y la combinación de predicciones [123].

Los modelos de redes neuronales son métodos no lineales, esto implica que pueden aprender relaciones no lineales complejas, siendo esto a su vez una desventaja, ya que las hace dependientes a sus condiciones iniciales, tanto en sus pesos aleatorios como en términos del ruido estadístico en el conjunto de entrenamiento. De esta forma, se puede ver a una red neuronal como un método con un bajo margen de error, pero con una varianza elevada.

Un método para reducir la varianza es entrenar múltiples modelos y combinar sus predicciones. Combinando las predicciones de varios modelos con un bajo margen de error se logra obtener predicciones menos sensibles a datos de entrenamiento específicos. Este proceso reduce la varianza en las predicciones, por lo que el modelo ensamblado puede generar mejores resultados en términos de predicciones y se le denomina **aprendizaje ensamblado**.

Generalmente el aprendizaje ensamblado involucra el entrenamiento de más de una red neuronal con el mismo conjunto de datos. Utiliza cada modelo entrenado para obtener las predicciones antes de combinarlas, obteniendo un solo resultado de predicciones finales y generando así un conjunto de predicciones más estable y mejorado [123].

El proceso de ensamblado de redes neuronales cuenta con tres elementos principales que pueden variar y que derivan en las tres técnicas descritas a continuación [123]:

- Variación de los datos de entrenamiento: cambiar la selección de datos para entrenar cada modelo, utilizando, por ejemplo, validación cruzada k-fold (*k-fold cross-validation*).
- Variación de los modelos ensamblados: cambiar los modelos que van a ser ensamblados. Para esto se puede variar la configuración de cada modelo ensamblado utilizando, por ejemplo, redes con distintas capacidades o modelos entrenados bajo diferentes condiciones.
- Variación de las combinaciones: variar la forma en la que se combinarán los resultados de los miembros de cada conjunto.

### 5.2.1.1. Ensamblador ArgMax

Este ensamblador toma para cada imagen un vector de probabilidades de cada modelo. En este caso: VGG-16 Ajustada, ResNet-50 e Inception-V3. De las tres matrices de probabilidades resultantes de cada modelo, se elige la coordenada con la probabilidad más alta y, dependiendo de la posición de esta coordenada, se comprueba un criterio y se asigna un diagnóstico. A continuación, la ecuación 5.1 muestra este ensamblador.

$$f(V) = \text{yArgMax}(V) \quad (5.1)$$

donde  $V$  es la salida de la última capa (*Softmax*) del modelo que concatena los tres vectores: el vector  $v_1$  es generado por VGG-16 Ajustada, el vector  $v_2$  es generado por ResNet50, y el vector  $v_3$  es generado por Inception V3.

$$V = [v_1; v_2; v_3] \quad (5.2)$$

#### Criterio de clasificación

Para definir el criterio de diagnóstico que permite clasificar imágenes dermatoscópicas en un tipo específico de lesión, se toma la probabilidad más alta del vector  $V$  cuya dimensión es  $3 \times m$ , donde  $m$  es el número de clases.

Cada imagen genera un vector de  $m$  probabilidades para cada modelo, por lo tanto, para una imagen se crea un nuevo vector de  $3m$  probabilidades. Se selecciona el valor con la probabilidad más alta de  $V$ , cuya coordenada se pasa a denominar  $n$ .

$$n = \text{ArgMax}_m V \quad (5.3)$$

Realizando estas consideraciones, la clase elegida dentro del vector  $V$  quedará definida por la ecuación 5.4 y donde  $V_n$  puede ser cualquier elemento de  $V$ :

$$\text{clase} = \text{mod}(n, m) \quad (5.4)$$

donde *mod* se refiere a la operación módulo, es decir, el resto de dividir  $n$  y  $m$ .

### 5.2.1.2. Ensamblador Mean

El ensamblador Mean toma cada vector de probabilidades  $v_1$ ,  $v_2$ , y  $v_3$ , de tal forma que el vector  $V$  es calculado como la media de los tres vectores:

$$V = \frac{1}{3} (v_1 + v_2 + v_3) \quad (5.5)$$

donde  $v_i$  es un vector de dimensión  $1 \times m$ , para  $i=1,2,3$ , indexando las tres redes usadas (VGG16 ajustada, ResNet50 e Inception V3).

### 5.3. Datos y diseño de los experimentos

En este capítulo se encuentra detallado el proceso de clasificación para la detección del carcinoma basocelular, así como también de distintos tipos de lesiones. En este proceso se considera la utilización de las tres redes neuronales profundas VGG-16 Ajustada, ResNet-50 e Inception-V3, aplicando los ensambladores ArgMax y Mean. Además, se muestra la comparación entre los clasificadores aplicados al conjunto de imágenes segmentadas y no segmentadas.

#### 5.3.1. Datos

Cada una de las redes probadas fue entrenada utilizando la base de datos ISIC-2019 [6] [23] [25], la cual contiene un total de 25.331 imágenes dermatoscópicas, de las cuales 17.731 pertenecen al conjunto de entrenamiento, 2.534 al conjunto de datos de validación y 5.066 pertenecen al conjunto de datos de prueba. El contenido de esta base de datos viene detallado como sigue, y en la Figura 5.4 se muestra un ejemplo:

- Queratosis actínica: 867 imágenes
- Carcinoma basocelular: 3.323 imágenes
- Queratosis benigna: 2.624 imágenes
- Dermatofibroma: 239 imágenes
- Melanoma: 4.522 imágenes
- Nevus: 12.875 imágenes
- Carcinoma celular escamoso: 628 imágenes
- Lesión vascular: 253 imágenes

Para el proceso de clasificación, que se encuentra descrito en las Figuras 5.5 y 5.6, se realizaron varias pruebas divididas de la siguiente manera:

- Pruebas con las imágenes sin una segmentación previa.
- Pruebas realizadas con las lesiones previamente segmentadas de la piel sana.

Se realizaron distintas pruebas utilizando los dos tipos de ensambladores diferentes, Mean y ArgMax.

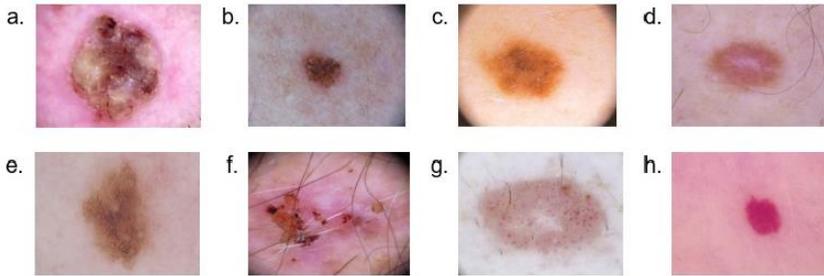


Figura 5.4 Imágenes dermatoscópicas de la base de datos ISIC-2019. a. Carcinoma celular escamoso. b. Nevus. c. Melanoma. d. Dematorfibroma. e. Queratosis benigna. f. Carcinoma basocelular. g. Queratosis actínica. h. Lesión vascular.

Finalmente es importante mencionar que las pruebas se implementaron en lenguaje de programación Python, utilizando un procesador Intel Core I9 – 3.6 GHz, 32 GB de memoria RAM y una tarjeta gráfica NVIDIA TITAN RTX.

### 5.3.2. Métricas para evaluar la Clasificación

Se realizaron tres tipos de clasificaciones: BCC versus cualquier otro tipo de lesión, clasificación en ocho tipos de lesiones distintas, y finalmente, BCC versus nevus. Las métricas que se utilizaron para medir la eficiencia de los experimentos realizados fueron las siguientes: sensibilidad ( $Se$ ), especificidad ( $Sp$ ), exactitud ( $Acc$ ), precisión ( $Pre$ ) y tasa de falsos positivos ( $FPR$ ), siendo las dos últimas obtenidas directamente de la matriz de confusión extraída de cada método.

La tasa de falsos positivos está definida por la siguiente ecuación:

$$\frac{FP}{N} = \frac{FP}{FP+TN} \quad (5.6)$$

Donde  $FP$  es el número de falsos positivos,  $TN$  es el número de verdaderos negativos y  $N$  el número total de negativos comprobados:

$$N = FP + TN \quad (5.7)$$

La métrica de precisión hace referencia a la consistencia del resultado obtenido una vez que se han realizado varias mediciones y es una forma de observar el error en cada medición.

$$Pre = \frac{TP}{TP+FP} \quad (5.8)$$

### 5.3.3. Experimentos de clasificación realizados

La clasificación de imágenes dermatoscópicas se realizó utilizando la base de datos ISIC-2019. La lesión objetivo de la clasificación fue el carcinoma basocelular (BCC), el cual fue identificado entre tres conjuntos de datos, que se describen a continuación:

- Clasificación de carcinoma basocelular versus cualquier otra lesión.
- Clasificación en 8 tipos de lesiones, siendo una de ellas el carcinoma basocelular.
- Clasificación de carcinoma basocelular versus nevus.

Al no contar con una base de datos balanceada, fue necesario realizar el balanceo de datos y para esto se aplicó un proceso de aumento de datos. El proceso de aumento de datos se basó en la aplicación de operaciones de reflexión en espejo y rotaciones, considerando un ángulo de  $0,2\pi$ . De esta forma se garantizó un correcto funcionamiento del clasificador. Después del proceso de balanceo de datos, aplicado sólo a las imágenes de entrenamiento, se obtuvo una base de datos con el siguiente número de imágenes de entrenamiento:

- Clasificación de carcinoma basocelular versus cualquier otro tipo de lesión:
  - 2.199 imágenes de validación.
  - 15.419 imágenes de entrenamiento.
- Clasificación en los ocho tipos de lesiones presentes en ISIC-2019:
  - Queratosis actínica (AK): 8.914 imágenes de entrenamiento.
  - Carcinoma basocelular (BCC): 8.958 imágenes de entrenamiento.
  - Queratosis benigna (BKL), 9.006 imágenes de entrenamiento.
  - Dermatofibroma (DF), 8.873 imágenes de entrenamiento.
  - Melanoma (MEL), 8.997 imágenes de entrenamiento.
  - Nevus (NEVUS), 9.012 imágenes de entrenamiento.
  - Carcinoma celular escamoso (SCC), 8.978 imágenes de entrenamiento.
  - Lesiones vasculares (VASC), 8.757 imágenes de entrenamiento.
- Clasificación de carcinoma basocelular versus nevus:
  - Carcinoma basocelular: 8.982 imágenes de entrenamiento.
  - Nevus: 8.982 imágenes de entrenamiento.

En las Figuras 5.5 y 5.6 se muestra un esquema de la clasificación seguida.

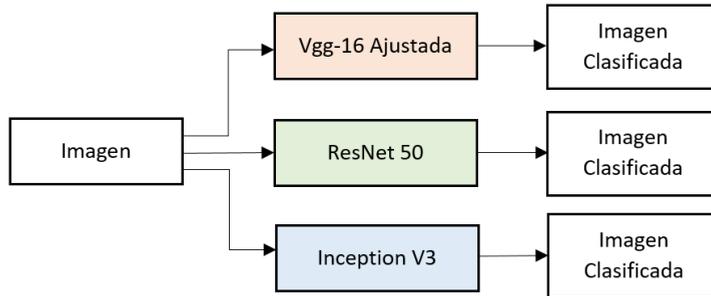


Figura 5.5 Proceso de Clasificación de imágenes dermatoscópicas sin segmentar previamente la imagen y sin utilizar un ensamblador.

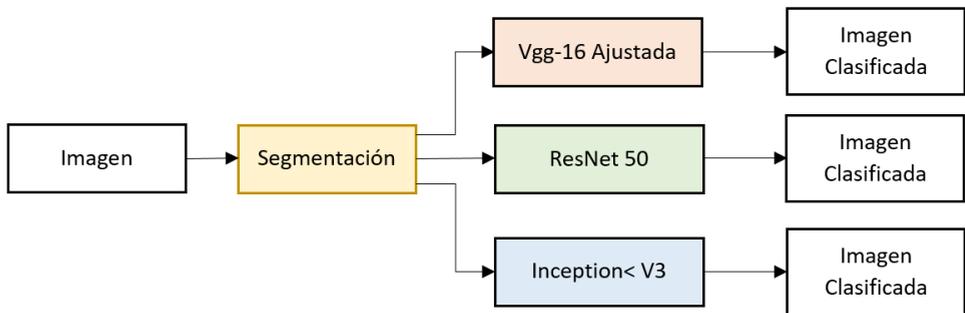


Figura 5.6 Proceso de Clasificación de imágenes dermatoscópicas segmentadas sin utilizar un ensamblador.

### 5.3.3.1. Implementación de Ensambladores para mejorar la clasificación

Con el objetivo de mejorar los resultados de clasificación se utilizaron dos ensambladores que permitieron enlazar las predicciones realizadas por cada una de las tres redes y así obtener una predicción con un mayor nivel de precisión. De esta manera, se consideraron los ensambladores ArgMax y Mean para realizar los experimentos de clasificación. Todos estos procesos se detallan en las Figuras 5.7 y 5.8.

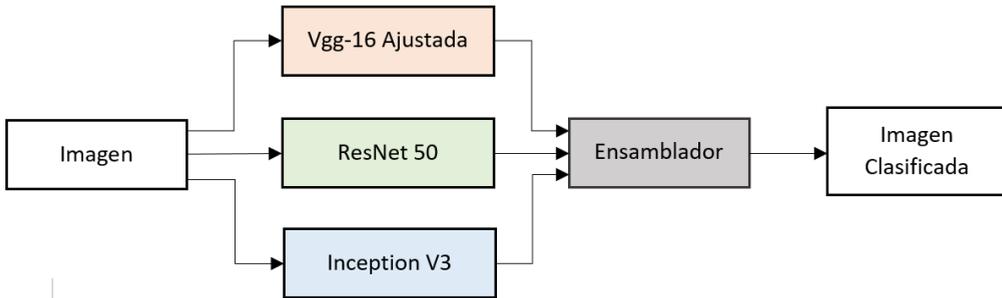


Figura 5.7 Proceso de clasificación de la imagen dermatoscópica utilizando un ensamblador pero sin realizar una segmentación previa.

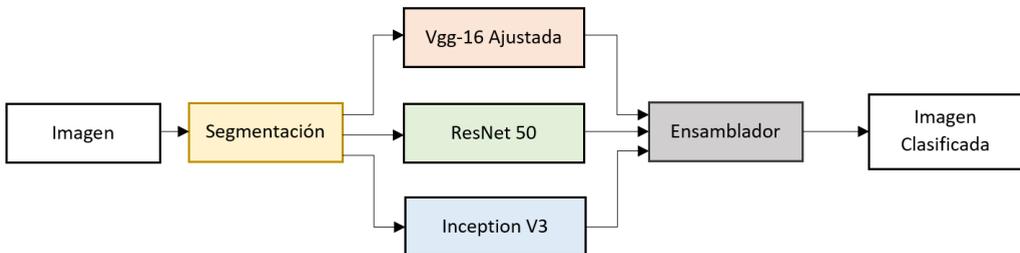


Figura 5.8 Proceso de clasificación de la imagen dermatoscópica utilizando un ensamblador y aplicando una segmentación previa a la imagen.

## 5.4. Resultados de la clasificación

### 5.4.1. Clasificación de BCC versus todas las lesiones sin utilizar un ensamblador

Las tres redes fueron entrenadas con un total de 40.302 imágenes obtenidas después de balacear el conjunto de entrenamiento. El proceso de entrenamiento se completó después de 100 épocas. Los resultados de los experimentos se muestran en la Tabla 5.1.

Se realizaron 3 pruebas sobre el conjunto de imágenes balanceadas:

- Clasificación de imágenes dermatoscópicas sin segmentar. Las matrices de confusión de cada experimento realizado con cada una de las redes se muestran en la Figura 5.9.

- Clasificación de imágenes dermatoscópicas segmentadas utilizando segmentación semántica. Las matrices de confusión de cada experimento realizado con cada una de las redes se muestran en la Figura 5.10.
- Clasificación de imágenes dermatoscópicas segmentadas utilizando segmentación con umbralización de imágenes de las capas convolucionales de una CNN. Las matrices de confusión de cada experimento realizado con cada una de las redes se muestran en la Figura 5.11.

Tabla 5.1 Resultados de clasificación de BCC versus todas las lesiones sin utilizar un ensamblador.

Imágenes	Metodología	Métricas				
		Se	Sp	Pre	FPR	Acc
Sin Segmentar	VGG-16 Ajustada	<b>0,84</b>	0,96	0,75	0,75	<b>0,94</b>
	ResNet 50	0,81	0,95	0,75	0,75	0,93
	Inception V3	0,78	0,95	0,70	0,70	0,93
Segmentación Semántica	VGG-16 Ajustada	0,68	0,96	0,74	0,95	0,93
	ResNet 50	0,65	0,96	0,70	0,95	0,92
	Inception V3	0,60	0,96	0,67	0,94	0,91
Segmentación con imágenes convolucionales	VGG-16 Ajustada	0,72	<b>0,97</b>	<b>0,77</b>	<b>0,96</b>	<b>0,94</b>
	ResNet 50	0,67	<b>0,97</b>	0,74	0,95	0,93
	Inception V3	0,66	0,96	0,72	0,95	0,92

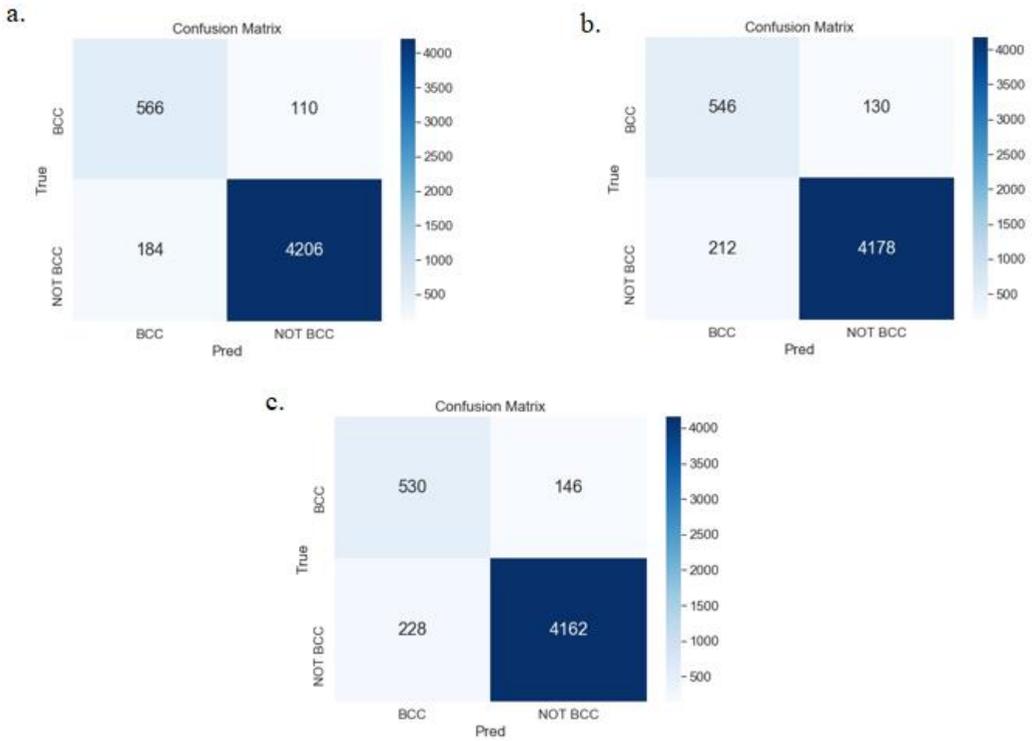


Figura 5.9 Matriz de confusión para la clasificación de BCC versus todas las lesiones, sin segmentación. a. VGG-16 Ajustada. b. ResNet-50. c. Inception-V3.

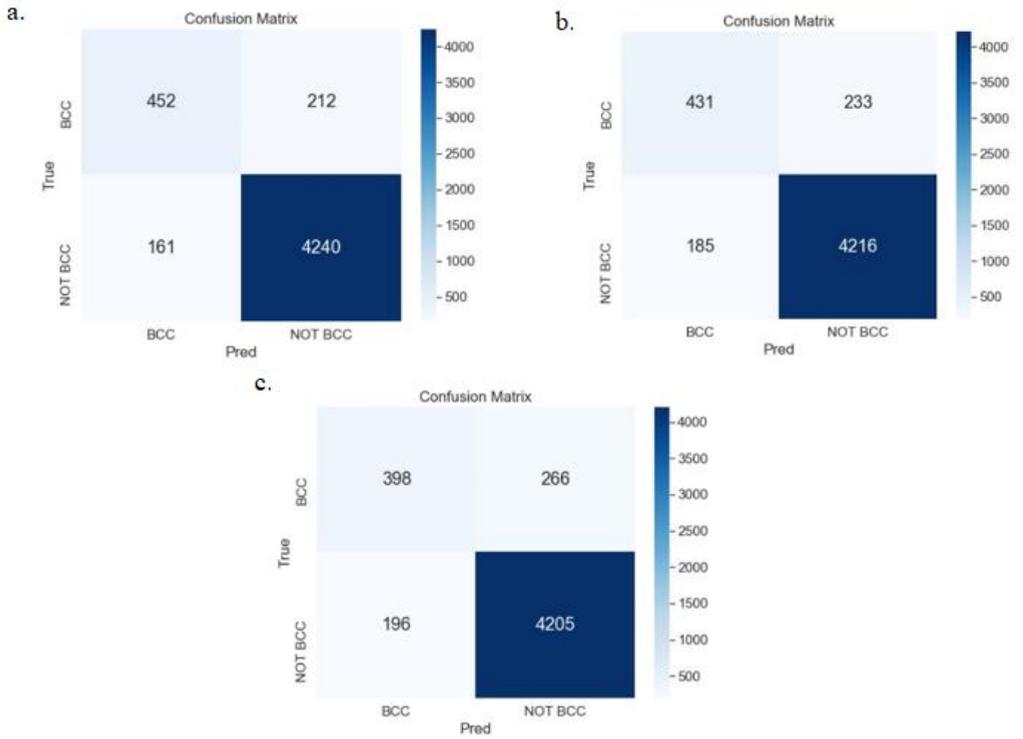


Figura 5.10 Matriz de confusión de la clasificación de BCC versus todas las lesiones habiendo aplicado segmentación semántica. a. VGG-16 Ajustada. b. ResNet-50. c. Inception-V3.

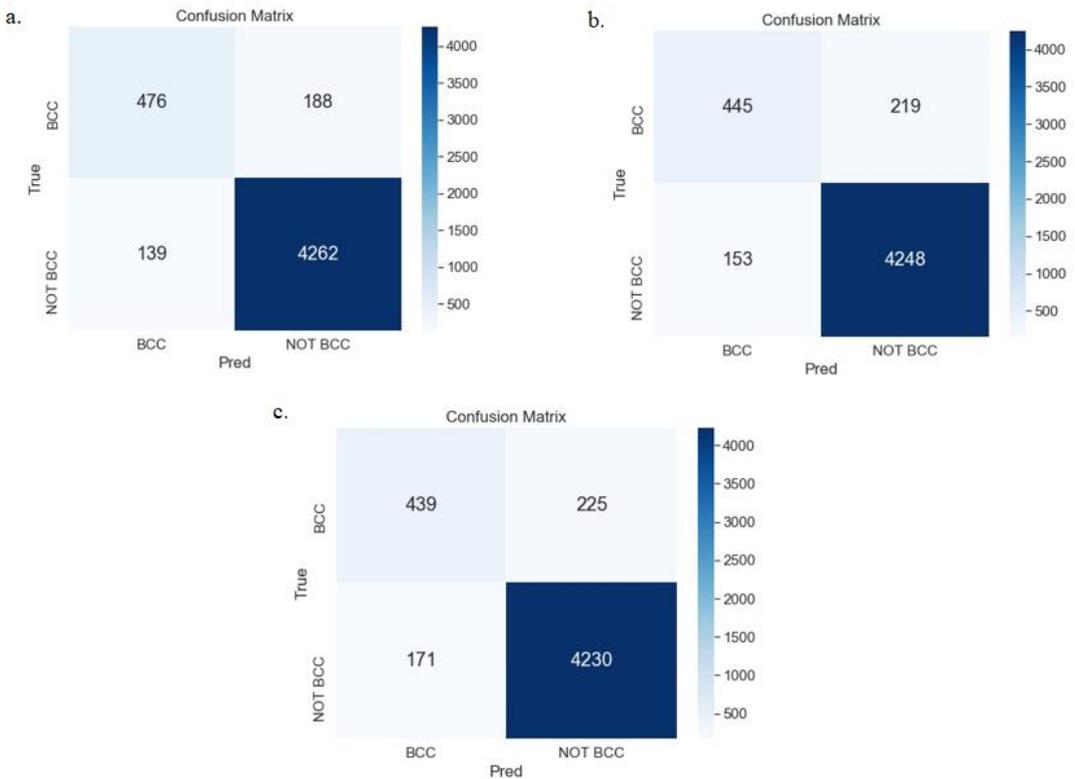


Figura 5.11 Matriz de confusión de la clasificación de BCC versus todas las lesiones habiendo aplicado segmentación con umbralización de imágenes convolucionales. a. VGG-16 Ajustada. b. ResNet-50. c. Inception-V3.

#### 5.4.2. Clasificación en ocho tipos de lesiones sin utilizar un ensamblador

Las tres redes fueron entrenadas con un total de 86.857 imágenes obtenidas después de balancear el conjunto de entrenamiento. El proceso de entrenamiento se completó después de 10 épocas.

Se realizaron tres clasificaciones, como se describe a continuación:

- Clasificación de imágenes dermatoscópicas sin aplicar un proceso de segmentación previo, cuyos resultados de las métricas obtenidas se muestran en la Tabla 5.2, mientras que las matrices de confusión de cada experimento realizado con cada una de las redes se muestran en la Figura 5.12.
- Clasificación de los ocho tipos de lesiones previamente segmentadas aplicando segmentación semántica, cuyos valores de las métricas obtenidas se muestran en

la Tabla 5.3, mientras que las matrices de confusión de cada experimento realizado con cada una de las redes se muestran en la Figura 5.13.

- Clasificación de imágenes dermatoscópicas previamente segmentadas aplicando segmentación con umbralización de imágenes convolucionales, cuyas métricas obtenidas se muestran en la Tabla 5.4, mientras que las matrices de confusión de cada experimento realizado con cada una de las redes se muestran en la Figura 5.14.

Tabla 5.2 Resultados de clasificación de los 8 tipos de lesiones sin segmentación.

Lesión	VGG-16 Ajustada			ResNet-50			Inception-V3		
	Se	Pre	Acc	Se	Pre	Acc	Se	Pre	Acc
AK	<b>0,66</b>	<b>0,82</b>	<b>0,81</b>	0,37	0,73	0,77	0,27	0,64	0,73
BCC	<b>0,83</b>	<b>0,84</b>		0,77	0,82		0,60	0,81	
BKL	<b>0,62</b>	<b>0,75</b>		0,48	<b>0,75</b>		0,38	0,73	
DF	<b>0,52</b>	0,82		0,37	<b>1,00</b>		0,40	0,68	
MEL	<b>0,65</b>	<b>0,77</b>		0,60	0,75		0,55	0,72	
NEVUS	0,93	<b>0,82</b>		<b>0,96</b>	0,77		<b>0,96</b>	0,72	
SCC	<b>0,52</b>	<b>0,81</b>		0,33	0,66		0,30	0,71	
VASC	<b>0,84</b>	0,85		0,71	0,95		0,50	<b>1,00</b>	

Tabla 5.3 Resultados de clasificación de los 8 tipos de lesiones segmentadas aplicando segmentación semántica.

Lesión	VGG-16 Ajustada			ResNet 50			Inception V3		
	Se	Pre	Acc	Se	Pre	Acc	Se	Pre	Acc
AK	<b>0,46</b>	0,61	<b>0,81</b>	0,37	<b>0,73</b>	0,77	0,27	0,64	0,73
BCC	<b>0,78</b>	0,74		0,77	<b>0,82</b>		0,60	0,81	
BKL	<b>0,62</b>	0,64		0,48	<b>0,75</b>		0,38	0,73	
DF	<b>0,47</b>	0,76		0,37	<b>1,00</b>		0,40	0,68	
MEL	<b>0,65</b>	0,70		0,60	<b>0,75</b>		0,55	0,72	
NEVUS	0,89	<b>0,85</b>		<b>0,96</b>	0,77		<b>0,96</b>	0,72	
SCC	<b>0,41</b>	0,49		0,33	0,66		0,30	<b>0,71</b>	
VASC	<b>0,74</b>	0,77		0,71	0,95		0,50	<b>1,00</b>	

Tabla 5. 4 Resultados de clasificación de los 8 tipos de lesiones segmentadas aplicando segmentación con umbralización de imágenes convolucionales.

Lesión	VGG-16 Ajustada			ResNet 50			Inception V3		
	Se	Pre	Acc	Se	Pre	Acc	Se	Pre	Acc
AK	0,49	<b>0,54</b>	0,75	<b>0,50</b>	0,52	0,74	0,39	0,46	0,72
BCC	<b>0,76</b>	<b>0,73</b>		<b>0,76</b>	0,69		0,75	0,69	
BKL	<b>0,55</b>	0,60		0,50	<b>0,62</b>		0,49	0,56	
DF	<b>0,40</b>	0,54		<b>0,40</b>	<b>0,58</b>		0,30	0,42	
MEL	<b>0,58</b>	<b>0,68</b>		0,56	0,66		0,57	0,64	
NEVUS	<b>0,89</b>	<b>0,82</b>		<b>0,89</b>	<b>0,82</b>		<b>0,87</b>	0,80	
SCC	<b>0,48</b>	<b>0,63</b>		0,43	0,53		0,38	0,62	
VASC	<b>0,90</b>	<b>0,85</b>		0,82	0,77		0,68	<b>0,85</b>	

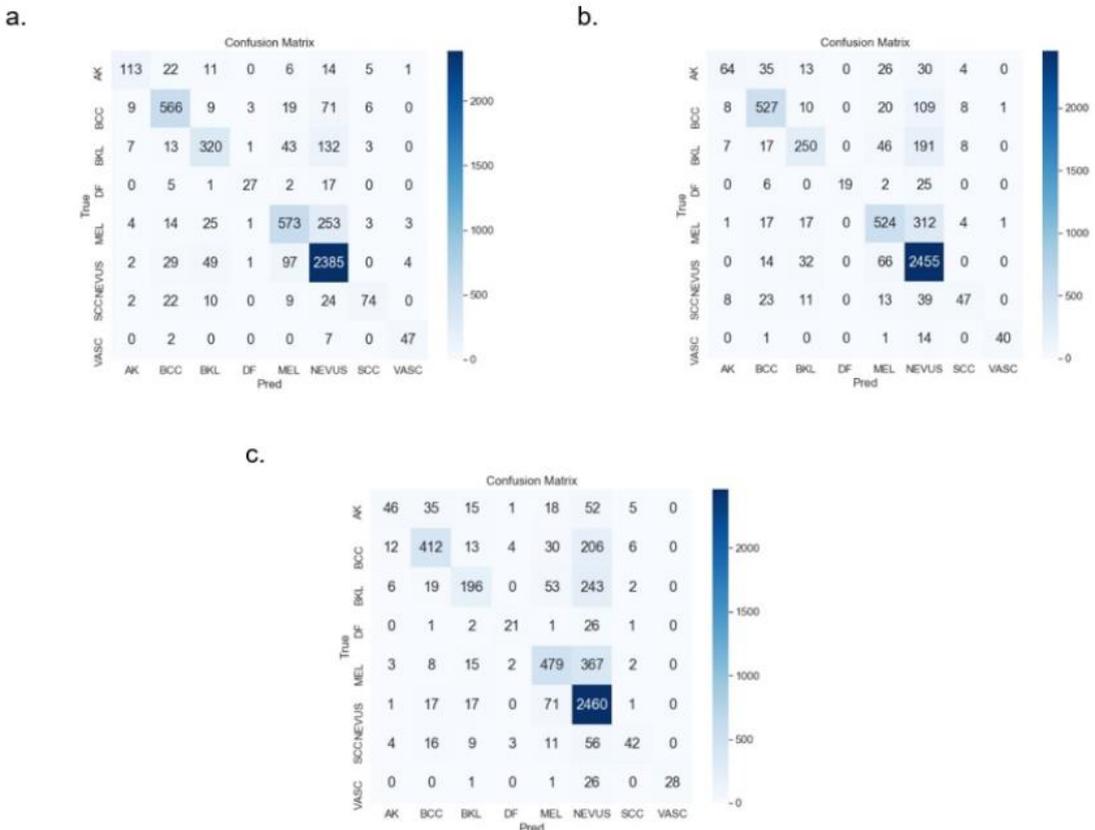


Figura 5.12 Matrices de confusión de la clasificación en ocho tipos de lesiones sin segmentación. a. VGG-16 Ajustada. b. ResNet-50. c. Inception-V3.

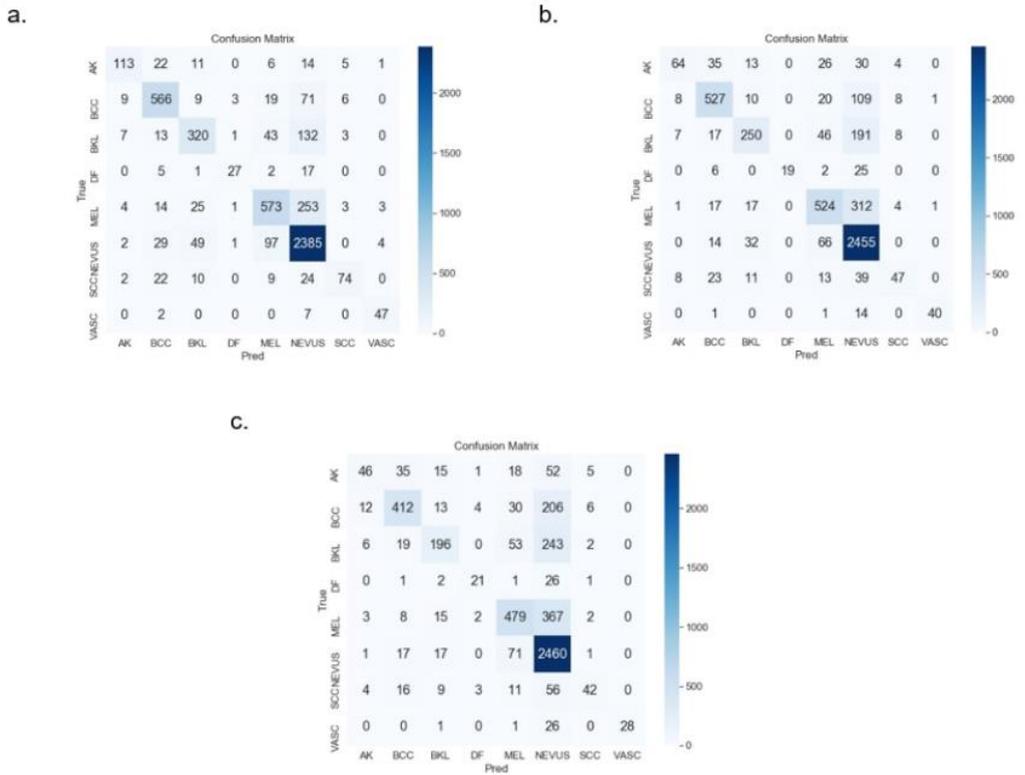


Figura 5.13 Matrices de confusión de la clasificación en ocho tipos de lesiones habiendo aplicado Segmentación Semántica. a. VGG-16 Ajustada. b. ResNet-50. c. Inception-V3.

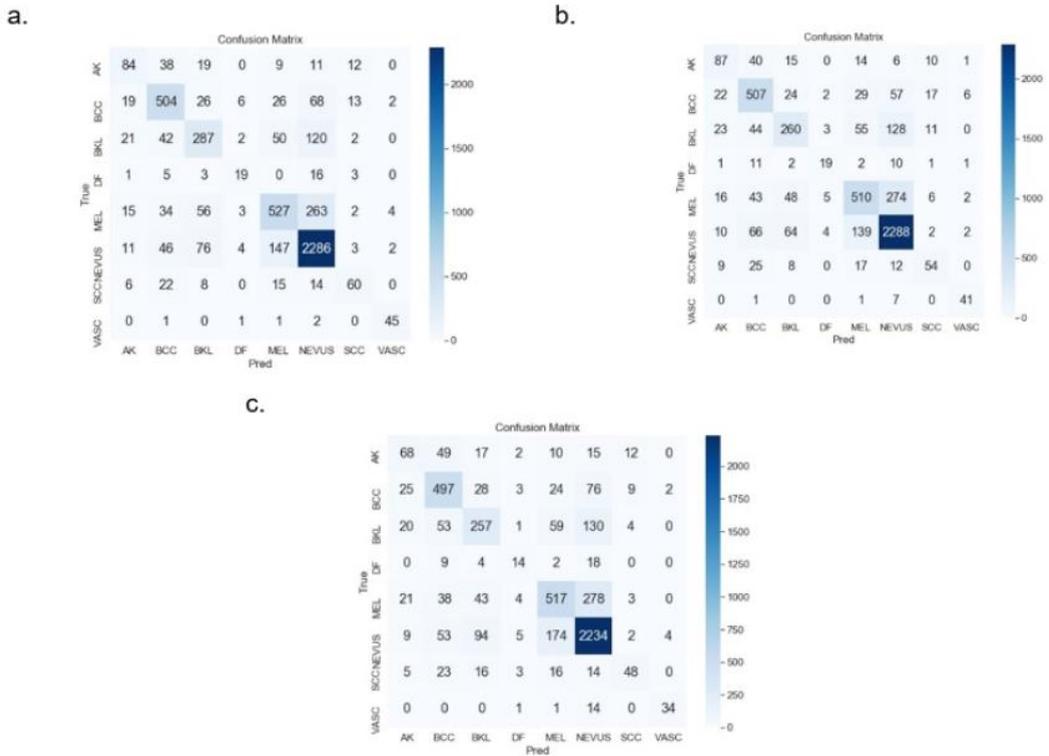


Figura 5.14 Resultados de clasificación en ocho tipos de lesiones, con segmentación utilizando umbralización de imágenes convolucionales. a. VGG-16 Ajustada. b. ResNet-50. c. Inception-V3.

### 5.4.3. Clasificación de BCC versus Nevus sin utilizar un ensamblador

Las tres redes fueron entrenadas con un total de 17.964 imágenes obtenidas después de balancear el conjunto de entrenamiento. El proceso de entrenamiento se completó después de 25 épocas.

Se realizaron tres experimentos de clasificación, cuyos resultados se muestran en la Tabla 5.5, y se describen a continuación:

- Clasificación BCC versus nevus sin una segmentación previa. Las matrices de confusión de cada experimento realizado con cada una de las redes se muestran en la Figura 5.15.

- Clasificación BCC versus nevus aplicando segmentación semántica. Las matrices de confusión de cada experimento realizado con cada una de las redes se muestran en la Figura 5.16.
- Clasificación BCC versus nevus segmentadas utilizando segmentación por umbralización de imágenes convolucionales. Las matrices de confusión de cada experimento realizado con cada una de las redes se muestran en la Figura 5.17.

Tabla 5.5 Resultados de clasificación de BCC vs Nevus sin utilizar un ensamblador.

Imágenes	Metodología	Métricas				
		Se	Sp	Pre	FPR	Acc
Sin Segmentar	VGG-16 Ajustada	<b>0,95</b>	<b>0,97</b>	<b>0,88</b>	<b>0,99</b>	<b>0,97</b>
	ResNet 50	<b>0,95</b>	<b>0,97</b>	<b>0,88</b>	<b>0,99</b>	0,96
	Inception V3	<b>0,95</b>	<b>0,97</b>	0,87	0,96	0,94
Segmentación Semántica	VGG-16 Ajustada	0,87	<b>0,97</b>	0,87	0,97	0,95
	ResNet 50	0,88	0,96	0,86	0,97	0,95
	Inception V3	0,85	<b>0,97</b>	0,87	0,96	0,94
Segmentación con umbralización de imágenes convolucionales	VGG-16 Ajustada	0,88	<b>0,97</b>	<b>0,88</b>	0,97	0,95
	ResNet 50	0,89	<b>0,97</b>	0,87	0,97	0,95
	Inception V3	0,85	<b>0,97</b>	0,87	0,96	0,94

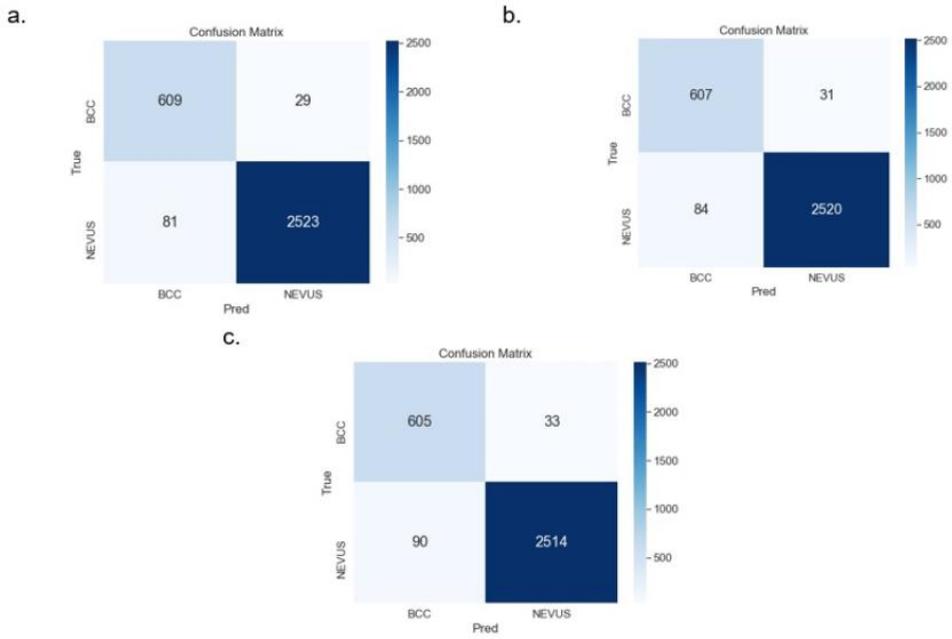


Figura 5.15 Resultados de clasificación de BCC versus Nevus sin realizar una segmentación previa. a. VGG-16 Ajustada. b. ResNet-50. c. Inception-V3.

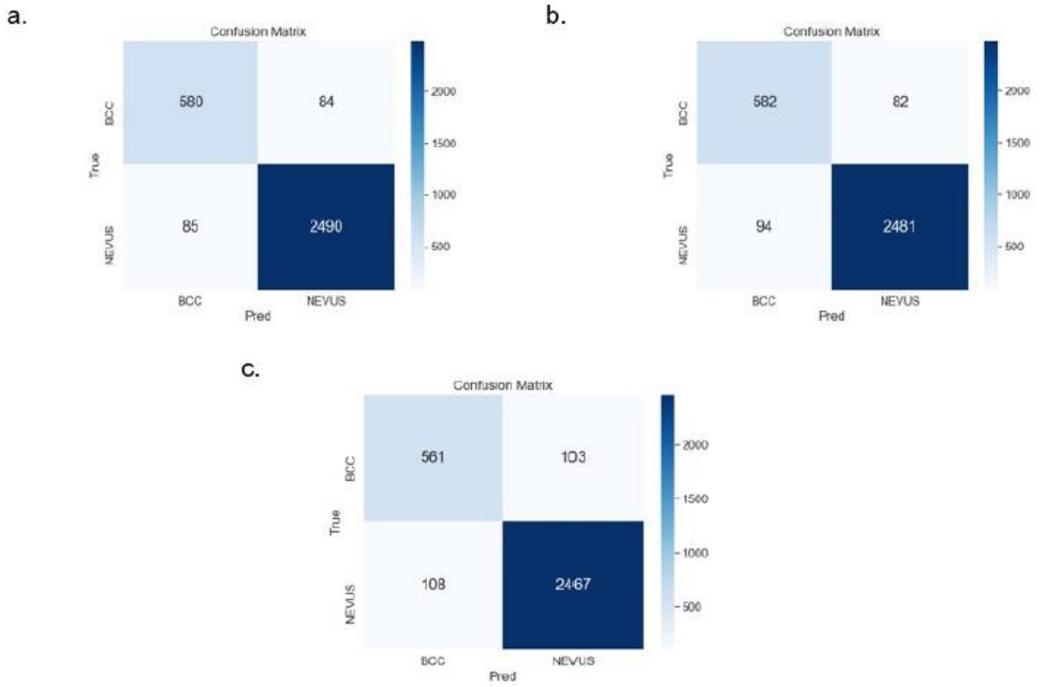


Figura 5.16 Matrices de confusión de la clasificación de BCC versus Nevus habiendo aplicado Segmentación Semántica. a. VGG-16 Ajustada. b. ResNet50. c. Inception V3.

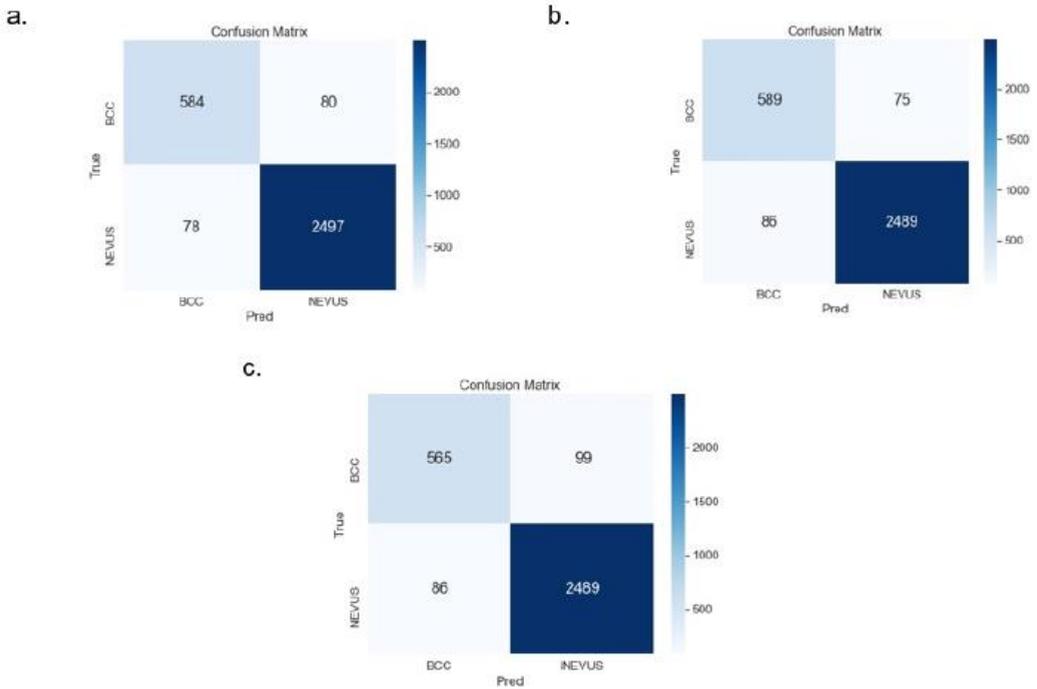


Figura 5. 17 Matrices de confusión de la clasificación de BCC versus Nevus habiendo aplicado segmentación basada en umbralización de imágenes convolucionales. a. VGG-16 Ajustada. b. ResNet50. c. Inception V3.

#### 5.4.4. Clasificación de BCC versus todas las lesiones con ensamblador

En la Tabla 5.6 se muestran los resultados obtenidos de la clasificación de BCC versus todas las clases aplicando dos ensambladores, el ensamblador ArgMax y el ensamblador Mean.

En la Figura 5.18 se presenta la matriz de confusión resultante de la clasificación realizada utilizando el ensamblador ArgMax. Por otra parte, en la Figura 5.19 se presenta la matriz de confusión resultante de la clasificación realizada utilizando el ensamblador Mean.

Tabla 5.6 Resultados de Clasificación de BCC vs todas las clases utilizando distintos ensambladores.

Ensamblador	Método de Segmentación	Se	Sp	Pre	APR	Acc
ArgMax	Sin Segmentación	<b>0,84</b>	0,96	0,78	<b>0,98</b>	<b>0,95</b>
	Segmentación Semántica	0,68	0,97	0,77	0,95	0,93
	Segmentación con imágenes convolucionales	0,73	0,97	<b>0,81</b>	0,96	0,94
Mean	Sin Segmentación	0,83	0,96	0,78	0,97	<b>0,95</b>
	Segmentación Semántica	0,68	0,97	0,78	0,95	0,93
	Segmentación con umbralización de imágenes convolucionales	0,70	<b>0,98</b>	<b>0,81</b>	0,96	0,94

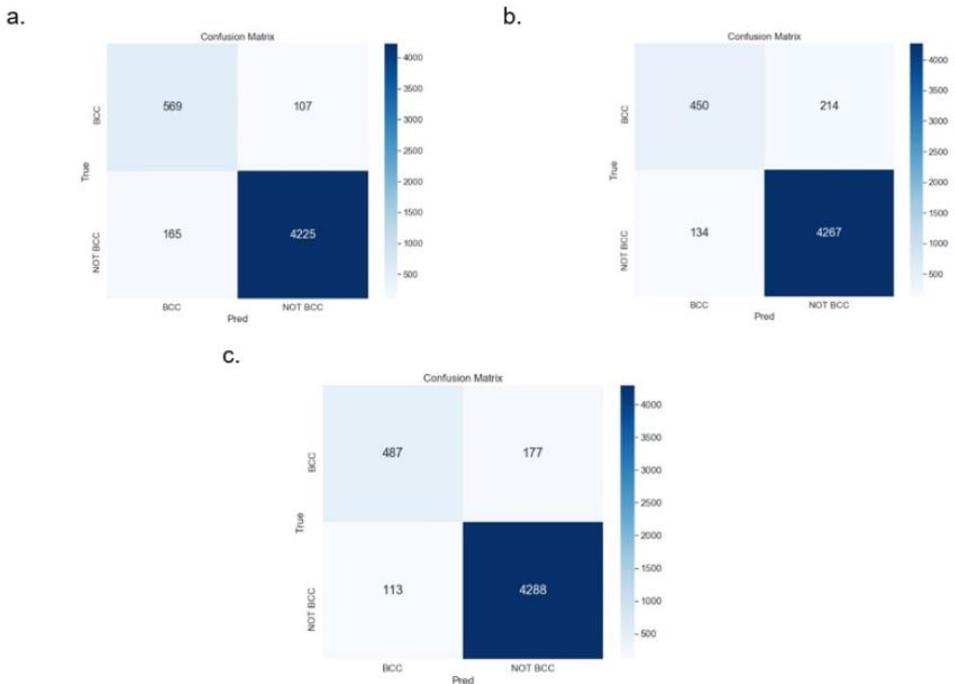


Figura 5.18 Matrices de confusión después de aplicar el ensamblador ArgMax a la clasificación de BCC versus todas las lesiones. a. Sin segmentación. b. Aplicando segmentación Semántica. c. Segmentación con umbralización de imágenes convolucionales.

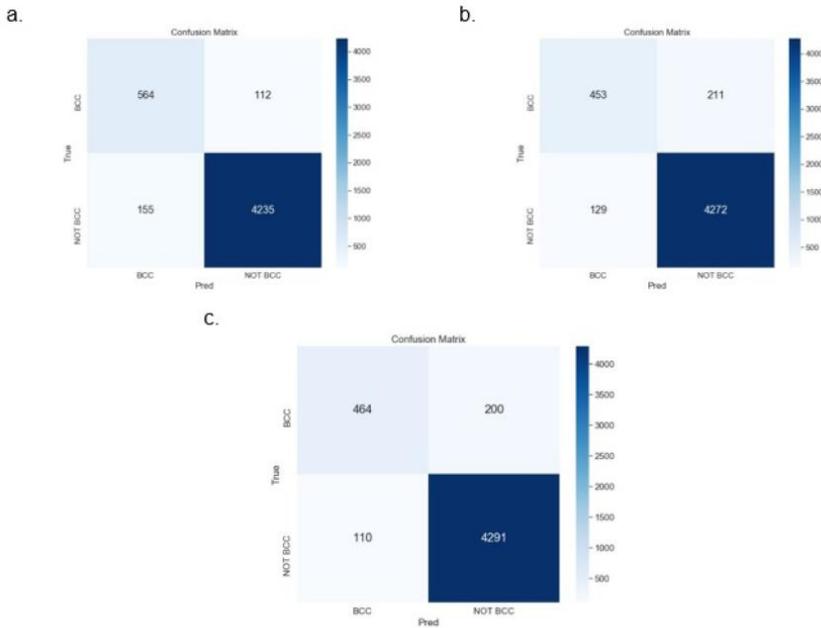


Figura 5.19 Matrices de confusión después de aplicar el ensamblador Mean a la clasificación de BCC versus todas las lesiones. a. Sin segmentación. b. Aplicando segmentación Semántica. c. Segmentación con umbralización de imágenes convolucionales.

#### 5.4.5. Clasificación en ocho tipos de lesiones con ensamblador

Los resultados de las métricas de clasificación obtenidas al aplicar el ensamblador ArgMax se muestran en la Tabla 5.7, mientras que las matrices de confusión de cada experimento de clasificación con el ensamblador ArgMax, realizado con cada una de las redes aplicadas para segmentar, se muestran en la Figura 5.20.

Por otra parte, la Tabla 5.8 muestra los resultados de la clasificación utilizando el ensamblador Mean y cuyos valores fueron obtenidos de las matrices de confusión que se muestran en la Figura 5.21.

Tabla 5.7 Métricas de clasificación en ocho tipos de lesiones aplicando el Ensamblador ArgMax.

Lesión	Sin segmentación			Con segmentación Semántica			Segmentación con umbralización de Imágenes Convolucionales		
	Se	Pre	Acc	Se	Pre	Acc	Se	Pre	Acc
AK	0,48	<b>0,84</b>	<b>0,80</b>	0,49	0,62	0,79	<b>0,53</b>	0,58	0,78
BCC	<b>0,83</b>	<b>0,86</b>		0,80	0,74		0,82	0,74	
BKL	0,50	<b>0,84</b>		<b>0,61</b>	0,67		0,53	0,66	
DF	0,48	<b>0,86</b>		<b>0,55</b>	0,79		0,36	0,55	
MEL	0,61	<b>0,82</b>		<b>0,64</b>	0,75		0,60	0,74	
NEVUS	0,97	0,78		0,92	<b>0,85</b>		0,91	0,83	
SCC	0,43	<b>0,87</b>		0,46	0,56		<b>0,52</b>	0,68	
VASC	0,79	<b>0,90</b>		0,74	0,86		<b>0,84</b>	0,84	

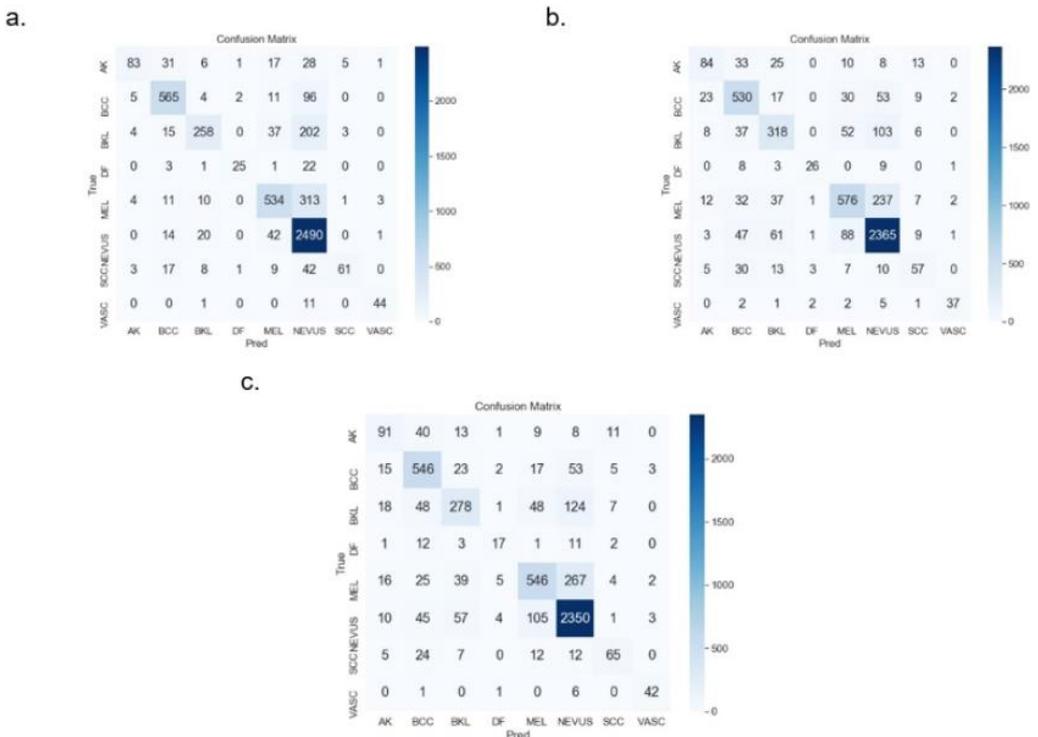


Figura 5.20 Matrices de confusión de aplicar el ensamblador Argmax a la clasificación en ocho tipos de lesiones. a. Sin segmentación. b. Con Segmentación Semántica. c. Segmentando con umbralización de Imágenes Convolucionales.

Tabla 5.8 Métricas de clasificación en ocho tipos de lesiones aplicando el Ensamblador Mean.

Lesión	Sin segmentación			Segmentación Semántica			Segmentación con umbralización de Imágenes Convolucionales		
	Se	Pre	Acc	Se	Pre	Acc	Se	Pre	Acc
AK	0,48	<b>0,84</b>	0,80	0,49	0,62	0,79	<b>0,53</b>	0,58	0,78
BCC	<b>0,83</b>	<b>0,86</b>		0,80	0,74		0,82	0,74	
BKL	0,50	<b>0,84</b>		<b>0,61</b>	0,67		0,53	0,66	
DF	0,48	<b>0,86</b>		<b>0,55</b>	0,79		0,36	0,55	
MEL	0,61	<b>0,82</b>		<b>0,64</b>	0,75		0,60	0,74	
NEVUS	<b>0,97</b>	0,78		<b>0,92</b>	0,85		<b>0,91</b>	0,83	
SCC	0,43	<b>0,87</b>		0,46	0,56		<b>0,52</b>	0,68	
VASC	0,79	<b>0,90</b>		0,74	0,86		<b>0,84</b>	0,84	

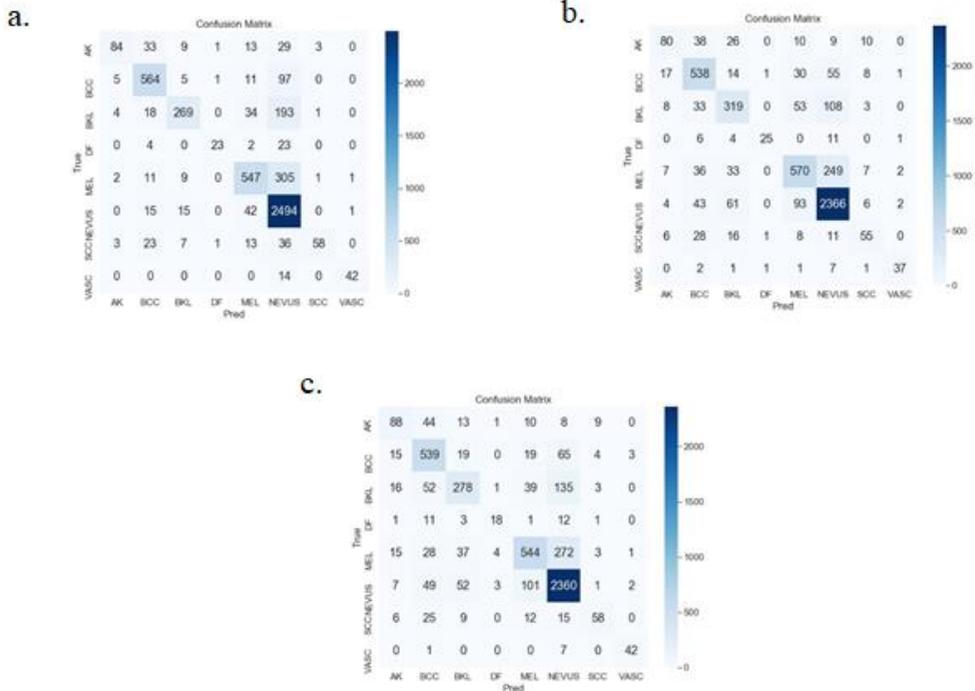


Figura 5.21 Matrices de confusión de aplicar el ensamblador Mean a la clasificación en ocho tipos de lesiones. a. Sin segmentación. b. Con Segmentación Semántica. c. Segmentando con umbralización de Imágenes Convolucionales.

**5.4.6. Clasificación de BCC versus Nevus con ensamblador**

Los resultados de las métricas obtenidas de la clasificación de BCC versus Nevus después de aplicar los ensambladores ArgMax y Mean se muestran en la Tabla 5.9, mientras que las matrices de confusión de esta clasificación se muestran en la Figura 5.22 y en la Figura 5.23.

Tabla 5.9 Resultados de Clasificación de BCC Vs Nevus utilizando distintos ensambladores.

<b>Ensamblador</b>	<b>Método de Segmentación</b>	<b>Se</b>	<b>Sp</b>	<b>Pre</b>	<b>APR</b>	<b>Acc</b>
ArgMax	Sin Segmentación	<b>0,97</b>	<b>0,98</b>	<b>0,92</b>	<b>0,99</b>	<b>0,98</b>
	Segmentación Semántica	0,90	0,97	0,88	0,97	0,95
	Segmentación con imágenes convolucionales	0,90	0,97	0,90	0,97	0,96
Mean	Sin Segmentación	<b>0,97</b>	0,97	0,90	0,99	0,97
	Segmentación Semántica	0,89	0,97	0,88	0,97	0,95
	Segmentación con umbralización de imágenes convolucionales	0,89	<b>0,98</b>	0,90	0,97	0,96

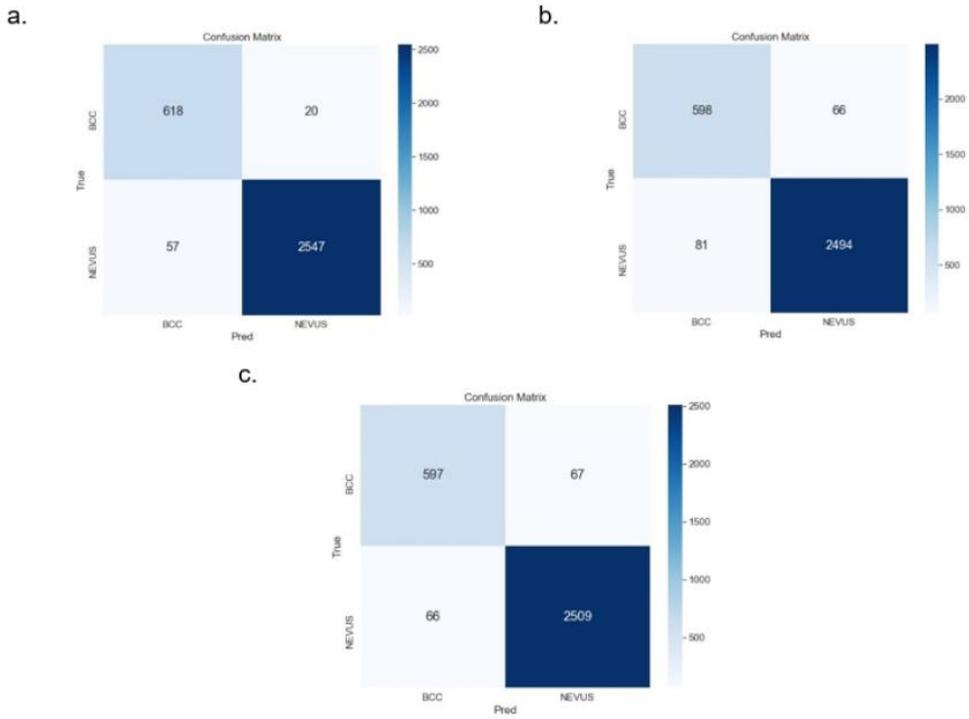


Figura 5.22 Matrices de confusión de aplicar el ensamblador Argmax a la clasificación de BCC versus Nevus. a. Sin segmentación. b. Con Segmentación Semántica. c. Segmentando con umbralización de Imágenes Convolucionales.

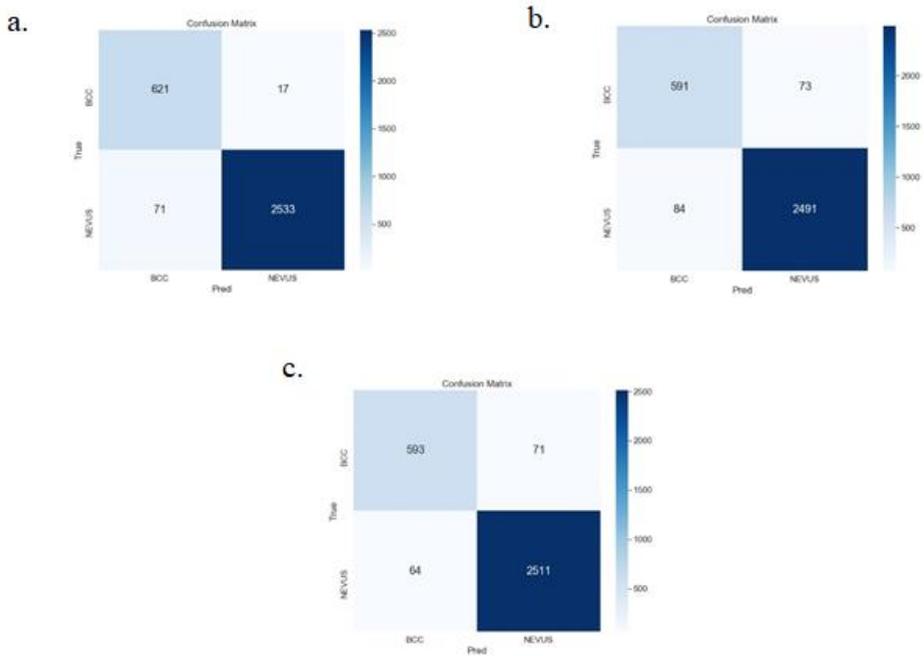


Figura 5.23 Matrices de confusión de aplicar el ensamblador Mean a la clasificación de BCC versus Nevus. a. Sin segmentación. b. Con Segmentación Semántica. c. Segmentando con umbralización de Imágenes Convolucionales.

### 5.5. Discusión de Resultados

Analizando los resultados obtenidos tras los experimentos detallados anteriormente, se deducen una serie de conclusiones, que se describen a continuación:

- Tal como se detalla en [124], los modelos de aprendizaje profundo ensamblados combinan las ventajas de los modelos de aprendizaje profundo con las del aprendizaje ensamblado, de forma que el modelo final posee un mejor rendimiento. Por lo tanto, en los modelos ensamblados el entrenamiento es más rápido que si se entrenaran múltiples modelos de aprendizaje profundo por separado, y además existe una disminución de la varianza. El éxito del aprendizaje ensamblado se ve justificado por los siguientes aspectos: descomposición del sesgo-varianza, aspectos estadísticos, computacionales y representacionales y finalmente el incremento en la diversidad entre los clasificadores. En el caso de estudio de esta tesis, donde se busca etiquetar la lesión BCC, se ha demostrado que la aplicación de un ensamblador obtuvo los mejores resultados.

En este caso se observa que el ensamblador ArgMax, permitió obtener los mejores resultados de clasificación.

- En cuanto a la aplicación del proceso de segmentación de imágenes dermatoscópicas, se encontró un resultado interesante, ya que los mejores resultados se obtuvieron al ingresar al clasificador las imágenes sin segmentar. Esto puede explicarse porque se contó con una base de datos extensa y correctamente balanceada para realizar el entrenamiento a partir de las imágenes completas, permitiendo a las redes aprender correctamente lo que pertenece a lesión o a piel sana, o extraer información relevante de la piel sana para la clasificación.

En este trabajo se presenta el análisis de tres clasificaciones distintas: BCC versus todas las lesiones presentes en la base de datos, BCC versus nevus y finalmente clasificación en ocho tipos de lesiones presentes en la base de datos. La motivación de esta elección recae principalmente en tres aspectos:

- En el primer caso, el etiquetado de las lesiones BCC versus el resto de lesiones se debió a la alta incidencia de BCC en los centros de atención primaria.
- En el segundo caso, el etiquetado de BCC versus nevus, se realizó debido a que el nevus es el tipo de lesión benigna más común entre la población mundial.
- En el tercer caso, el etiquetado en ocho tipos de lesiones distintas se realizó para poder comparar los resultados de la tesis con los resultados publicados con otros autores.

En cuanto a los resultados obtenidos, los mejores resultados de clasificación se obtuvieron al utilizar el ensamblador ArgMax sin segmentar previamente las imágenes, resultando en una exactitud de 0,98 en la clasificación de BCC versus Nevus frente a un 0,95 de exactitud utilizando imágenes segmentadas previamente. Por otro lado, se obtuvo una exactitud de 0,95 en la clasificación de BCC frente a todas las lesiones y este valor se mantuvo con las imágenes segmentadas, sin embargo los demás parámetros evaluados fueron superiores al utilizar imágenes sin segmentar. Finalmente, un resultado similar se obtuvo en la clasificación en ocho tipos de lesiones, ya que el mejor resultado se obtuvo al aplicar el clasificador con el ensamblador ArgMax y sin realizar una segmentación previa de las imágenes. En las Tablas 5.10 – Tabla 5.13 se muestran de forma detallada los resultados obtenidos en cada uno de los parámetros de evaluación presentes en cada caso de clasificación.

Tabla 5.10 Métricas de clasificación de BCC versus Nevus sin realizar una segmentación previa a las imágenes.

Método de clasificación	Se	Sp	Pre	FPR	Acc
VGG-16	0,95	0,97	0,88	<b>0,99</b>	0,97
ResNet 50	0,95	0,97	0,88	<b>0,99</b>	0,96
InceptionV3	0,95	0,97	0,87	0,96	0,94
Ensamblador ArgMax	<b>0,97</b>	<b>0,98</b>	<b>0,92</b>	<b>0,99</b>	<b>0,98</b>
Ensamblador Mean	<b>0,97</b>	0,97	0,90	<b>0,99</b>	0,97

Tabla 5.11 Métricas de clasificación de BCC versus Nevus habiendo segmentado las imágenes previamente

Método de clasificación	Se	Sp	Pre	FPR	Acc
VGG-16	0,87	<b>0,97</b>	0,87	<b>0,97</b>	<b>0,95</b>
ResNet50	0,88	0,96	0,86	<b>0,97</b>	<b>0,95</b>
InceptionV3	0,85	<b>0,97</b>	0,87	0,96	0,94
Ensamblador ArgMax	<b>0,90</b>	<b>0,97</b>	<b>0,88</b>	<b>0,97</b>	<b>0,95</b>
Ensamblador Mean	0,89	<b>0,97</b>	<b>0,88</b>	<b>0,97</b>	<b>0,95</b>

Tabla 5.12 Métricas de clasificación de BCC versus todas las lesiones sin realizar una segmentación previa a las imágenes.

Método de clasificación	Se	Sp	Pre	FPR	Acc
VGG-16	<b>0,84</b>	<b>0,96</b>	0,75	0,97	0,94
ResNet50	0,81	0,95	0,72	0,97	0,93
InceptionV3	0,78	0,95	0,70	0,97	0,93
Ensamblador ArgMax	<b>0,84</b>	<b>0,96</b>	<b>0,78</b>	<b>0,98</b>	<b>0,95</b>
Ensamblador Mean	0,83	<b>0,96</b>	<b>0,78</b>	0,97	<b>0,95</b>

Tabla 5.13 Métricas de clasificación de BCC versus todas las lesiones habiendo segmentado las imágenes previamente.

Método de clasificación	Se	Sp	Pre	FPR	Acc
VGG-16	<b>0,68</b>	0,96	0,74	<b>0,95</b>	0,94
ResNet50	0,65	0,96	0,70	<b>0,95</b>	0,93
InceptionV3	0,60	0,96	0,67	0,94	0,93
Ensamblador ArgMax	<b>0,68</b>	<b>0,97</b>	0,77	<b>0,95</b>	<b>0,95</b>
Ensamblador Mean	<b>0,68</b>	<b>0,97</b>	<b>0,78</b>	<b>0,95</b>	<b>0,95</b>

Como se mencionó en el Capítulo 3, la mayoría de trabajos dedicados a la segmentación de imágenes dermatoscópicas resaltan la importancia de la precisión de este proceso para obtener una adecuada extracción de características y por lo tanto una correcta caracterización de la lesión [125]. Por otra parte, en este trabajo se demuestra que para clasificar una imagen dermatoscópica utilizando algoritmos de aprendizaje profundo junto

a una base de datos extensa, no se requiere de un proceso de segmentación. Como se muestra en las tablas 5.10 y 5.11, las métricas de clasificación son mejores cuando no se han segmentado previamente las imágenes.

### 5.5.1. Comparación con otros trabajos publicados en la literatura

En la literatura existen pocos trabajos dedicados a la detección de BCC [126, 127] [119] y la principal razón para esto es la falta de bases de datos públicas con lesiones BCC que contengan delineación de contornos y que se encuentren correctamente etiquetadas con un criterio dermatoscópico.

Según la búsqueda en la literatura realizada en este trabajo de investigación, existen pocos trabajos dedicados a clasificar BCC de otras lesiones, ni a clasificar BCC versus nevus. Sin embargo, para poder comparar la eficacia del método propuesto en este trabajo, se realizó la clasificación propuesta por el reto de ISIC-2019.

El reto ISIC-2019 consiste en clasificar ocho tipos distintos de lesiones, entre las que se encuentra el BCC: nevus (NEVUS), queratosis benigna (BKL), dermatofibroma (DF), melanoma (MEL), carcinoma celular escamoso (SCC), queratosis actínica (AK) y lesiones vasculares (VASC), carcinoma basocelular (BCC).

La Tabla 5.14 muestra los valores de las métricas obtenidas al aplicar el algoritmo de clasificación que se propone en este trabajo de investigación, comparadas con el algoritmo ganador del reto ISIC-2019 [5, 1]. Como se puede ver, el algoritmo de clasificación propuesto en este trabajo supera la marca obtenida por el ganador del reto en prácticamente todas las métricas.

Tabla 5.14 Comparación del funcionamiento de los algoritmos de clasificación, para clasificar en ocho tipos de lesiones. Algoritmo ganador del reto ISIC 2019 [1] y algoritmo propuesto (ensamblador ArgMax sin segmentación previa de las imágenes).

Lesión	Ganador del reto ISIC 2019 [1]		Método Propuesto	
	Se	Sp	Se	Sp
AK	<b>0,48</b>	0,97	<b>0,48</b>	<b>1,00</b>
BCC	0,72	0,94	<b>0,83</b>	<b>0,98</b>
BKL	0,39	<b>0,99</b>	<b>0,50</b>	<b>0,99</b>
DF	0,58	0,98	<b>0,48</b>	<b>1,00</b>
MEL	0,59	0,96	<b>0,61</b>	<b>0,97</b>
NEVUS	0,71	<b>0,98</b>	<b>0,97</b>	0,71
SCC	<b>0,44</b>	0,99	0,43	<b>1,00</b>
VASC	0,64	0,99	<b>0,79</b>	<b>1,00</b>

## 6. Conclusiones

---

**En** este trabajo se presenta el desarrollo de varios algoritmos para analizar imágenes dermatoscópicas. En este sentido, se detallan las metodologías utilizadas para desarrollar algoritmos de segmentación, extracción de características y clasificación de imágenes dermatoscópicas, habiendo enfocado el análisis principalmente en el carcinoma basocelular.

Como primer paso para el análisis de lesiones de piel se desarrollaron varios algoritmos con el fin de encontrar la mejor técnica para la segmentación de imágenes dermatoscópicas. Para esto, se realizaron tres experimentos, de los cuales uno fue basado en aprendizaje no supervisado y dos de ellos basados en CNNs. Como resultado de estos experimentos se desarrolló una metodología de segmentación de imágenes dermatoscópicas basada en segmentación semántica utilizando la CNN SegNet, la que presentó la exactitud (Acc) más alta, un 94%, habiendo sido probada con la base de datos de ISIC-2017. En cuanto a los índices de similitud Jaccard y DICE, también se obtuvieron los índices más altos, 77% y 85% respectivamente. La utilización de la CNN SegNet para segmentar imágenes dermatoscópicas, además de ser la que presenta el mejor comportamiento aún con imágenes de baja calidad, es también la de mayor sencillez y su coste computacional es aceptable frente a otras metodologías existentes en la literatura actual.

En segundo lugar, se consideró una herramienta de aprendizaje máquina para realizar el análisis de patrones globales presentes en la imagen dermatoscópica. Es así que se desarrolló un algoritmo que permitió extraer superpíxeles, los cuales aislaron secciones con características similares dentro de la lesión cutánea. Una vez obtenidos los superpíxeles de cada lesión se obtuvieron estadísticos de primer y segundo orden y se seleccionaron las características relevantes, dando como resultado el contraste, la correlación, la energía y la máxima probabilidad obtenidos utilizando un vector de distancias de  $d=[1,3,6]$  y un vector de ángulos  $\theta=[0^\circ, 45^\circ, 90^\circ, 135^\circ]$ .

Las pruebas de clasificación realizadas aplicando un árbol de decisión configurado con análisis de componentes principales (PCA), cuyas predicciones se realizaron con los recortes obtenidos de los superpíxeles de cada imagen, permitieron obtener una exactitud del 99%, habiendo clasificado entre tres tipos de patrones: homogéneo, empedrado y

reticular. Esta clasificación se realizó considerando que la presencia de ciertos patrones en una lesión cutánea son determinantes al momento de diagnosticar una lesión como maligna.

Finalmente se realizó el proceso de clasificación de lesiones de la piel, enfocando el análisis a la detección de BCC frente a otros tipos de lesiones y en particular frente al nevus. Para realizar esta tarea se consideró la base de datos ISIC-2019. En cuanto a la metodología utilizada, se desarrolló la clasificación aplicando el ensamblado de la CNN VGG-16 con algunas modificaciones en su arquitectura y de las DNNs ResNet-50 e Inception-V3. Como resultado se obtuvo un 95% de exactitud al clasificar entre BCC y el resto de lesiones presentes en la base de datos y un 98% en la clasificación de BCC y nevus. Un hallazgo importante en el proceso de clasificación de BCC es que, cuando se trabaja con una base de datos extensa y se aplican algoritmos de aprendizaje profundo, la segmentación deja de ser un paso importante y no beneficia al proceso de clasificación. Esto a su vez sugiere que la piel sana que rodea la lesión puede contener información importante para la clasificación.

Como líneas futuras se plantea el diagnóstico de las lesiones de la piel con una explicación del mismo, es decir, en lugar de dar una decisión final, detectar los patrones dermatoscópicos indicativos de cada lesión, dando un diagnóstico de la misma forma que lo realiza el médico. También es interesante realizar un estudio prospectivo de estas herramientas CAD, implantadas en un entorno clínico y realizando evaluaciones comparativas con los juicios clínicos.

## **Revistas con índice de impacto (JCR - *Journal Citation Report*)**

**1) Título:** Does a Previous Segmentation Improve the Automatic Detection of Basal Cell Carcinoma Using Deep Neural Networks?

**Autores:** Paulina Vélez Núñez, Manuel Mirando, Carmen Serrano, Begoña Acha

**Revista:** Applied Sciences

**Doi:** <https://doi.org/10.3390/app12042092>

**Cuartil JCR:** Q2

**Año:** 2022

## **Publicaciones en congresos internacionales**

**1) Título:** An Evaluation of Segmentation and Classification Strategies on Melanomas

**Autores:** Paulina Vélez Núñez, Carmen Serrano, Begoña Acha

**Conferencia:** 2019 IEEE 16<sup>th</sup> International Symposium on Biomedical Imaging (ISBI)

**Lugar:** Venecia – Italia

**Mes:** Abril

**Año:** 2019

**2) Título:** Dermoscopic Image Segmentation: A comparison of methodologies

**Autores:** Paulina Vélez Núñez, Carmen Serrano, Begoña Acha, J.A. Pérez-Carrasco

**Conferencia:** 15<sup>th</sup> Mediterranean Conference on Medical and Biological Engineering and Computing. – MEDICOM 2019

**Lugar:** Coimbra – Portugal

**Mes:** Septiembre

**Año:** 2019

**Publicaciones en congresos nacionales**

**1) Título:** An evaluation of segmentation and classification strategies on Melanomas

**Autores:** Paulina Vélez Núñez, Begoña Acha, Carmen Serrano

**Conferencia:** XXXIV Simposium Nacional de la Unión Científica Internacional de Radio  
- URSI 2019

**Lugar:** Sevilla – España

**Mes:** Septiembre

**Año:** 2019

# Índice de Figuras

## Capítulo 1.

Figura 1.1 EDRA – Base de datos.....	- 7 -
Figura 1.2 ISIC – Base de datos. ....	- 8 -

## Capítulo 2.

Figura 2.1 Aprendizaje Máquina Tradicional vs Aprendizaje Profundo.....	- 16 -
Figura 2.2 Diagrama esquemático de una Red Neuronal Convolutacional.....	- 18 -
Figura 2.3 Función de activación ReLu. ....	- 19 -
Figura 2.4 Diagrama del funcionamiento de las capas de una CNN. Editado de: [44]-	20 -
Figura 2.5 Esquema representativo de la arquitectura de red AlexNet. Fuente: [47]..-	20 -
Figura 2.6 Esquema representativo de la arquitectura de Red VGG-16. Editado de: [48].-	21 -
Figura 2.7 Esquema representativo del módulo Inception. Editado de: [49].....	- 22 -
Figura 2.8 Esquema de la red GoogLeNet. Editado de: [49].....	- 23 -
Figura 2.9 Aumento de datos realizado sobre la base de datos ISIC-2017 utilizando transformaciones geométricas. ....	- 24 -

## Capítulo 3.

Figura 3.1 Artefactos presentes en imágenes dermatoscópicas. a. Vello. b. Burbujas. c. Vasos Sanguíneos. d. Marcas de tinta. ....	- 25 -
Figura 3.2 Selección de regiones, tanto de piel sana como de la lesión previa a la definición de sus respectivos histogramas.....	- 28 -
Figura 3.3 Ejemplo de segmentación incluyendo postprocesado de la imagen segmentada. a) Imagen original b) Borde de la lesión. c) Lesión extraída por el algoritmo. d) Nueva máscara invertida. e) Resultado final de la segmentación.....	- 29 -
Figura 3.4 Imágenes de las 4 primeras capas convolucionales de la red VGG-16 una vez normalizadas. ....	- 30 -
Figura 3.5 Esquema del proceso de obtención de una máscara utilizando imágenes convolucionales (RM: <i>Raw Mask</i> ). ....	- 30 -
Figura 3.6 DAG configurado para el proceso de segmentación. ....	- 32 -
Figura 3.7 Imágenes segmentadas. Para las imágenes a, b, c, d, la primera fila muestra la imagen original, la máscara y el resultado segmentado. Para las 4 imágenes, las filas 2, 3 y	

4 representan la máscara y la segmentación obtenidas por el método de segmentación semántica, segmentación por minimización de energías y segmentación utilizando imágenes convolucionales. .... - 38 -

#### Capítulo 4.

Figura 4.1 Características Locales. Descripción por fila: a) Regresión b) Reticular c) Empedrado d) Globular e) Homogéneo f) Paralelo g) Velo Blanco azulado. .... - 42 -

Figura 4.2 Patrones Globales. A) Empedrado b) Globular c) Homogéneo d) Reticular e) Estrellado f) Paralelo g) Lagunar h) Patrón múltiple i) Patrón no especificado. Fuente: EDRA. .... - 43 -

Figura 4.3 Imagen dermatoscópica con la lesión segmentada en superpíxeles. En la fila inferior se expone una muestra de 4 superpíxeles pertenecientes a la lesión. .... - 48 -

Figura 4.4 Obtención del recorte. A) SLIC aplicado a la imagen dermatoscópica. B) Superpíxel aislado. C) Recorte en escala de grises. .... - 51 -

Figura 4.5 a) Recorte original. B) Imagen filtrada. C) Realce. .... - 52 -

Figura 4.6 Recortes. A,b,c,d,e) Recorte Empedrado: original, escala de grises, recorte postprocesado, histograma del recorte en escala de grises e histograma del recorte postprocesado. F,g,h,i,j) Recorte Homogéneo: original, escala de grises, recorte postprocesado, histograma del recorte en escala de grises e histograma del recorte postprocesado. K,l,m,n,o) Recorte Reticular: original, escala de grises, recorte postprocesado, histograma del recorte en escala de grises e histograma del recorte postprocesado. .... - 53 -

Figura 4.7 Esquema del clasificador ..... - 55 -

Figura 4.8 Modelo predictivo utilizando el subconjunto de características seleccionadas. - 56 -

Figura 4.9 Árbol de decisión ..... - 56 -

Figura 4.10 Matriz de confusión ..... - 57 -

#### Capítulo 5.

Figura 5.1 Arquitectura ResNet50. Fuente: [114]. .... - 62 -

Figura 5.2 Arquitectura de la Red Inception-V3 . Fuente: [133]. .... - 63 -

Figura 5.3 VGG-16 Ajustada (A-VGG16). .... - 63 -

Figura 5.4 Imágenes dermatoscópicas de la base de datos ISIC-2019. A. Carcinoma celular escamoso. B. Nevus. C. Melanoma. D. Dematorfibroma. E. Queratosis benigna. F. Carcinoma basocelular. G. Queratosis actínica. H. Lesión vascular. .... - 67 -

Figura 5.5 Proceso de Clasificación de imágenes dermatoscópicas sin segmentar previamente la imagen y sin utilizar un ensamblador. .... - 69 -

Figura 5.6 Proceso de Clasificación de imágenes dermatoscópicas segmentadas sin utilizar un ensamblador. .... - 69 -

Figura 5.7 Proceso de clasificación de la imagen dermatoscópica utilizando un ensamblador pero sin realizar una segmentación previa. ....	- 70 -
Figura 5.8 Proceso de clasificación de la imagen dermatoscópica utilizando un ensamblador y aplicando una segmentación previa a la imagen. ....	- 70 -
Figura 5.9 Matriz de confusión para la clasificación de BCC versus todas las lesiones, sin segmentación. A. VGG-16 Ajustada. B. ResNet-50. C. Inception-V3.....	- 72 -
Figura 5.10 Matriz de confusión de la clasificación de BCC versus todas las lesiones habiendo aplicado segmentación semántica. A. VGG-16 Ajustada. B. ResNet-50. C. Inception-V3. ....	- 73 -
Figura 5.11 Matriz de confusión de la clasificación de BCC versus todas las lesiones habiendo aplicado segmentación con umbralización de imágenes convolucionales. A. VGG-16 Ajustada. B. ResNet-50. C. Inception-V3. ....	- 74 -
Figura 5.12 Matrices de confusión de la clasificación en ocho tipos de lesiones sin segmentación. A. VGG-16 Ajustada. B. ResNet-50. C. Inception-V3.....	- 76 -
Figura 5.13 Matrices de confusión de la clasificación en ocho tipos de lesiones habiendo aplicado Segmentación Semántica. A. VGG-16 Ajustada. B. ResNet-50. C. Inception-V3. ....	- 77 -
Figura 5.14 Resultados de clasificación en ocho tipos de lesiones, con segmentación utilizando umbralización de imágenes convolucionales. A. VGG-16 Ajustada. B. ResNet-50. C. Inception-V3.....	- 78 -
Figura 5.15 Resultados de clasificación de BCC versus Nevus sin realizar una segmentación previa. A. VGG-16 Ajustada. B. ResNet-50. C. Inception-V3.....	- 80 -
Figura 5.16 Matrices de confusión de la clasificación de BCC versus Nevus habiendo aplicado Segmentación Semántica. A. VGG-16 Ajustada. B. ResNet50. C. Inception V3. -	81 -
Figura 5.17 Matrices de confusión de la clasificación de BCC versus Nevus habiendo aplicado segmentación basada en umbralización de imágenes convolucionales. A. VGG-16 Ajustada. B. ResNet50. C. Inception V3.....	- 82 -
Figura 5.18 Matrices de confusión después de aplicar el ensamblador ArgMax a la clasificación de BCC versus todas las lesiones. A. Sin segmentación. B. Aplicando segmentación Semántica. C. Segmentación con umbralización de imágenes convolucionales.....	- 83 -
Figura 5.19 Matrices de confusión después de aplicar el ensamblador Mean a la clasificación de BCC versus todas las lesiones. A. Sin segmentación. B. Aplicando segmentación Semántica. C. Segmentación con umbralización de imágenes convolucionales.....	- 84 -
Figura 5.20 Matrices de confusión de aplicar el ensamblador Argmax a la clasificación en ocho tipos de lesiones. A. Sin segmentación. B. Con Segmentación Semántica. C. Segmentando con umbralización de Imágenes Convolucionales.....	- 85 -

Figura 5.21 Matrices de confusión de aplicar el ensamblador Mean a la clasificación en ocho tipos de lesiones. A. Sin segmentación. B. Con Segmentación Semántica. C. Segmentando con umbralización de Imágenes Convolucionales. .... - 86 -

Figura 5.22 Matrices de confusión de aplicar el ensamblador Argmax a la clasificación de BCC versus Nevus. A. Sin segmentación. B. Con Segmentación Semántica. C. Segmentando con umbralización de Imágenes Convolucionales. .... - 88 -

Figura 5.23 Matrices de confusión de aplicar el ensamblador Mean a la clasificación de BCC versus Nevus. A. Sin segmentación. B. Con Segmentación Semántica. C. Segmentando con umbralización de Imágenes Convolucionales. .... - 89 -

# Índice de Tablas

## Capítulo 3.

Tabla 3.1 Coeficientes de similitud [58].	-35-
Tabla 3.2 Resultados de la segmentación de cuatro imágenes de la base de datos ISIC 2017, según las metodologías aplicadas.	- 36 -
Tabla 3.3 Comparación de las metodologías de segmentación para el conjunto de test de la base de datos de ISIC-2017.	- 37 -
Tabla 3.4 Resultados de segmentación tras aplicar distintos procesos.	- 39 -

## Capítulo 4.

Tabla 4.1 Características seleccionadas.	- 54 -
Tabla 4.2 Muestras de entrenamiento.	- 55 -
Tabla 4.3 Clasificación.	- 58 -

## Capítulo 5.

Tabla 5.1 Resultados de clasificación de BCC versus todas las lesiones sin utilizar un ensamblador.	- 71 -
Tabla 5.2 Resultados de clasificación de los 8 tipos de lesiones sin segmentación.	- 75 -
Tabla 5.3 Resultados de clasificación de los 8 tipos de lesiones segmentadas aplicando segmentación semántica.	- 75 -
Tabla 5.4 Resultados de clasificación de los 8 tipos de lesiones segmentadas aplicando segmentación con umbralización de imágenes convolucionales.	- 76 -
Tabla 5.5 Resultados de clasificación de BCC vs Nevus sin utilizar un ensamblador.	- 79 -
Tabla 5.6 Resultados de Clasificación de BCC vs todas las clases utilizando distintos ensambladores.	- 83 -
Tabla 5.7 Métricas de clasificación en ocho tipos de lesiones aplicando el Ensamblador ArgMax.	- 85 -

Tabla 5.8 Métricas de clasificación en ocho tipos de lesiones aplicando el Ensamblador Mean.....	- 86 -
Tabla 5.9 Resultados de Clasificación de BCC Vs Nevus utilizando distintos ensambladores. ....	- 87 -
Tabla 5.10 Métricas de clasificación de BCC versus Nevus sin realizar una segmentación previa a las imágenes.....	- 91 -
Tabla 5.11 Métricas de clasificación de BCC versus Nevus habiendo segmentado las imágenes previamente .....	- 91 -
Tabla 5.12 Métricas de clasificación de BCC versus todas las lesiones sin realizar una segmentación previa a las imágenes.....	- 91 -
Tabla 5.13 Métricas de clasificación de BCC versus todas las lesiones habiendo segmentado las imágenes previamente.....	- 91 -
Tabla 5.14 Comparación del funcionamiento de los algoritmos de clasificación, para clasificar en ocho tipos de lesiones. Algoritmo ganador del reto ISIC 2019 y Algoritmo propuesto (ensamblador ArgMax sin segmentación previa de las imágenes) .....	- 92 -

# Bibliografía

---

- [1] N. Gessert, M. Nielsen, M. Shaikh, R. Werner y A. Schlaefer, «Skin lesion classification using ensembles of multi-resolution EfficientNets with meta data,» *MethodsX*, vol. 7, n° 100864, 2020.
- [2] Skin Cancer Foundation, «Skin Cancer Facts & Statistics,» 13 Enero 2021. [En línea]. Available: <https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/>.
- [3] American Cancer Society, «Estadísticas importantes sobre los cánceres de piel de células basales y de células escamosas,» cancer.org, 8 julio 2020. [En línea]. Available: <https://www.cancer.org/es/cancer/cancer-de-piel-de-celulas-basales-y-escamosas/acerca/estadisticas-clave.html>.
- [4] C. Barata, M. E. Celebi y J. S. Marques, «Explainable skin lesion diagnosis using taxonomies,» *Pattern Recognition*, vol. 110, p. 107413, 2021.
- [5] «International Skin Imaging Collaboration,» [En línea]. Available: <https://www.isic-archive.com>.
- [6] P. Tschandl y C. R. a. H. Kittler, «The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions,» *Scientific Data*, vol. 5, n° 1, 2018.
- [7] J. K. Winkler, K. Sies, C. Fink, F. Toberer, A. Enk, T. Deinlein, R. H.-W. a. L. Thomas, A. Lallas, A. Blum, W. Stolz, M. S. Abassi, T. Fuchs, A. Rosenberger y H. A., «Melanoma recognition by a deep learning convolutional neural network—Performance in different melanoma subtypes and localisations,» vol. 127, pp. 21--29, 2020.
- [8] C. N. Vasconcelos y B. N. Vasconcelos, «Experiments using deep learning for dermoscopy image analysis,» *Pattern Recognition Letters*, vol. 139, pp. 95--103, 2020.
- [9] H. Haenssle, C. F. a. F. Toberer, J. Winkler, W. Stolz, T. Deinlein, R. Hofmann-Wellenhof, A. Lallas, S. Emmert, T. Buhl, M. Zutt, A. Blum, M. Abassi, L. Thomas, I. Tromme, P. Tschandl, A. Enk, A. Rosenberger, C. Alt y Mari, «Man

- against machine reloaded: performance of a market-approved convolutional neural network in classifying a broad spectrum of skin lesions in comparison with 96 dermatologists working under less artificial conditions,» *Annals of Oncology*, vol. 31, n° 1, p. Annals of Oncology, 2020.
- [10] A. Hekler, J. S. Utikal, A. H. Enk, A. Hauschild, M. Weichenthal, R. C. M. a. C. B. a. S. Haferkamp, J. Klode, D. Schadendorf, B. Schilling, T. Holland-Letz y B. Izar, «Superior skin cancer classification by the combination of human and artificial intelligence,» *European Journal of Cancer*, vol. 120, pp. 114--121, 2019.
- [11] R. L. Siegel, K. D. Miller y A. Jemal, «Cancer statistics, 2020,» *CA: Cancer Journal for Clinicians*, vol. 70, n° 1, pp. 7--30, 2020.
- [12] R. P. Braun, A. A. Marghoob y A. Lallas, «Level 1: Nevi,» *dermoscopia*, 2020. [En línea]. Available: [https://dermoscopia.org/w/index.php?title=Level\\_1:\\_Nevi&oldid=13536](https://dermoscopia.org/w/index.php?title=Level_1:_Nevi&oldid=13536).
- [13] A. Oakley, «Dermoscopy. Dermatoscopic features,» *DermNet NZ*, 2008. [En línea]. Available: <https://dermnetnz.org/cme/dermoscopy-course/dermoscopic-features/>.
- [14] A. Marghoob y R. Braun, «Two-step algorithm: Differentiating melanocytic from nonmelanocytic lesions,» de *Atlas of Dermoscopy*, CRC Press, 2012, pp. 33--39.
- [15] F. Nachbar, W. Stolz, T. Merkle, A. B. Cagnetta, T. Vogt, M. Landthaler, P. Bilek, O. Braun-Falco y G. Plewig, «The ABCD rule of dermatoscopy: High prospective value in the diagnosis of doubtful melanocytic skin lesions,» *Journal of the American Academy of Dermatology*, vol. 30, n° 4, pp. 551--559, 1994.
- [16] J. D. Jensen y B. E. Elewski, «The ABCDEF Rule: Combining the “ABCDE Rule” and the “Ugly Duckling Sign” in an Effort to Improve Patient Self-Screening Examinations,» *The Journal of clinical and aesthetic dermatology*, vol. 8, n° 2, p. 15, 2015.
- [17] S. W. Menzies, «Menzies method,» de *Atlas of Dermoscopy*, 2 ed., CRC Press, 2012, pp. 118--123.
- [18] R. H. Johr, «Dermoscopy: alternative melanocytic algorithms-the ABCD rule of dermatoscopy, Menzies scoring method, and 7-point checklist,» *Clinics in dermatology*, vol. 20, n° 3, pp. 240--247, 2002.

- [19] G. Argenziano, P. Soyer, V. D. Giorgi, D. Piccolo, P. Carli, M. Delfino, A. Ferrari, R. Hofmann-Wellenhof, D. Massi, G. Mazzocchetti, M. Scalvenzi y I. Wolf, «Interactive atlas of dermoscopy,» 2020.
- [20] Robust Melanoma Screening project, «Datasets,» 2020. [En línea]. Available: <https://sites.google.com/site/robustmelanomascreeing/dataset>.
- [21] ISIC Project, «ISIC Challenge,» 2020. [En línea]. Available: <https://challenge.isic-archive.com/>.
- [22] D. Gutman, N. C. F. Codella, E. Celebi, B. Helba, M. Marchetti, N. Mishra y A. Halpern, «Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016 hosted by the International Skin Imaging Collaboration (ISIC),» *arXiv:1605.01397 [cs.CV]*, 2016.
- [23] N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, H. K. Nabin Mishra y A. Halpern, «Skin Lesion Analysis Toward Melanoma Detection: A Challenge at the 2017 International Symposium on Biomedical Imaging (ISBI), Hosted ty the International Skin Imaging Collaboration (ISIC),» *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 168-172, 2018.
- [24] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, H. Kittler y A. Halpern, «Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC),» *arxiv.org/abs/1902.03368*, 2018.
- [25] M. Combalia, N. C. F. Codella, V. Rotemberg, B. Helba, V. Vilaplana, O. Reiter, A. C. Halpern, S. Puig y J. Malvehy, «BCN20000: Dermoscopic Lesions in the Wild,» *arXiv:1908.02288*, 2019.
- [26] V. Rotemberg, N. Kurtansky, B. Betz-Stablein, L. Caffery, E. Chousakos, N. Codella, M. Combalia, S. Dusza, P. Guitera, D. Gutman, A. Halpern, B. Helba, H. Kittler, K. Kose, S. Langer, K. Lioprys, J. Malvehy, S. Musthaq, J. Nanda y Reiter, «A patient-centric dataset of images and metadata for identifying melanomas using clinical context.,» *Sci Data* 8, vol. 34, 2021.

- [27] D. Shen, G. Wu y H.-I. Suk, «Deep Learning in Medical Image Analysis,» *Medical Image Analysis. Annual review of biomedical engineering*, vol. 19, pp. 221--248, 2017.
- [28] R. Kaymak, C. Kaymak y A. Ucar, «Skin lesion segmentation using fully convolutional networks: A comparative experimental study,» *Expert Systems with Applications*, vol. 161, p. 113742, 2020.
- [29] Y. LeCun, Y. Bengio y G. Hinton, «Deep learning,» *Nature*, vol. 521, n° 7553, pp. 436--444, 2015.
- [30] S. Suganyadevi, V. Seethalakshmi y K. Balasamy, «A review on deep learning in medical image analysis,» *International Journal of Multimedia Information Retrieval*, vol. 11, p. 19--38, 2022.
- [31] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. v. d. L. Laak, B. v. Ginneken y C. I. Sánchez, «A survey on deep learning in medical image analysis,» *Medical Image Analysis*, vol. 42, pp. 60--88, 2017.
- [32] MathWorks, «Redes Neuronales Convolucionales,» 2020. [En línea]. Available: <https://es.mathworks.com/solutions/deep-learning/convolutional-neural-network.html>.
- [33] Q. Zhang, L. T. Yang, Z. Chen y P. Li, «A survey on deep learning for big data,» *Information Fusion*, vol. 42, pp. 146--157, 2018.
- [34] O. Wanli, W. Xiaogang, F. Paul, C. Jie, L. Xinwang y P. Matti, «Deep Learning for Generic Object Detection: A Survey,» *International Journal of Computer Vision*, vol. 128, p. 261--318, 2020.
- [35] A. Voulodimos, N. Doulamis, A. Doulamis y E. Protopapadakis, «Deep Learning for Computer Vision: A Brief Review,» *Computational Intelligence and Neuroscience*, vol. 2018.
- [36] Y. L. Boureau, J. Ponce y Y. LeCun, «A theoretical analysis of feature pooling in visual recognition,» *Proceedings of the 27th International Conference on International Conference on Machine Learning*, p. 111--118, 2010.
- [37] R. Girshick, J. Donahue, T. Darrell y J. Malik, «Rich feature hierarchies for accurate object detection and semantic segmentation,» *Proceedings of the 27th IEEE Conference on Computer Vision and Pattern Recognition (CVPR'14)*, pp. 580-587, 2014.

- [38] S. Saha, «A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way,» 15 Diciembre 2015. [En línea]. Available: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>.
- [39] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg y L. Fei-Fei, «ImageNet Large Scale Visual Recognition Challenge,» *International Journal of Computer Vision*, vol. 115, n° 3, pp. 211--252, 2015.
- [40] A. Krizhevsky, I. Sutskever y G. E. Hinton, «ImageNet Classification with Deep Convolutional Neural Networks,» *Communications of the ACM*, vol. 60, n° 6, pp. 84-90, 2017.
- [41] W. Yu, K. Yang, Y. Bai, H. Yao y Y. Rui, «Visualizing and Comparing AlexNet and VGG using Deconvolutional Layers,» *arXiv:1412.6631*, 2016.
- [42] K. Simonyan y A. Zisserman, «Very Deep Convolutional Networks for Large-Scale Image Recognition,» *CoRR*, vol. abs/1409.1556, 2015.
- [43] R. Olga, D. Jia, S. Hao, K. Jonathan, S. Sanjeev, M. Sean, H. Zhiheng, K. Andrej, K. Aditya, B. Michael, B. Alexander C. y F.-F. Li, «ImageNet Large Scale Visual Recognition Challenge,» *International Journal of Computer Vision (IJCV)*, vol. 115, n° 3, pp. 211-252, 2015.
- [44] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke y A. Rabinovich, «Going deeper with convolutions,» *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1-9, 2015.
- [45] J. I. Bagnato, «¿Qué es overfitting y underfitting y cómo solucionarlo?,» 02 2020. [En línea]. Available: <https://www.aprendemachinelearning.com/que-es-overfitting-y-underfitting-y-como-solucionarlo/>.
- [46] C. Shorten y T. M. Khoshgoftaar, «A survey on Image Data Augmentation for Deep Learning,» *Journal of Big Data*, vol. 6, n° 60, 2019.
- [47] S. Hoo-Chang, «Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning,» *IEEE Transactions on Medical Imaging*, vol. 35, n° 5, pp. 1285-1298, 2016.
- [48] S. J. Pan y Q. Yang, «A Survey on Transfer Learning,» *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, n° 10, pp. 1345--1359, 2010.

- [49] S. Ruder, «Transfer Learning - Machine Learning's Next Frontier,» 02 2020. [En línea]. Available: <https://ruder.io/transfer-learning/>.
- [50] N. Tajbakhsh, J. Y. Shin, S. Gurudu, R. T. Hurst, C. B. Kendall, M. Gotway y J. Liang, «Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?,» *IEEE Transactions on Medical Imaging*, vol. 35, n° 5, pp. 1299--1312, 2016.
- [51] I. Aganj, M. G. Harisinghani, R. Weissleder y B. Fischl, «Unsupervised Medical Image Segmentation Based on the Local Center of Mass,» *Scientific Reports*, vol. 8, n° 1, 2018.
- [52] D. Kaur y Y. Kaur, «Various Image Segmentation Techniques:A Review,» *International Journal of Computer Science and Mobile Computing*, vol. 3, n° 5, pp. 809--814, 2014.
- [53] S. M. Jaisakthi, P. Mirunalini y C. Aravindan, «Automated skin lesion segmentation of dermoscopic images using GrabCut and k-means algorithms,» *IET Computer Vision*, vol. 12, n° 8, pp. 1088-1095, 2018.
- [54] J. L. Garcia-Arroyo y B. Garcia-Zapirain, «Segmentation of skin lesions in dermoscopy images using fuzzy classification of pixels and histogram thresholding,» *Computer Methods and Programs in Biomedicine*, vol. 168, pp. 11-19, 2019.
- [55] M. R. Guarracino y L. Maddalena, «SDI+: A Novel Algorithm for SegmentingDermoscopic Images,» *IEEE Journal of Biomedical and Health Informatics*, vol. 23, n° 2, pp. 481--488, 2019.
- [56] Y. Yuan, M. Chao y Y.-C. Lo, «Automatic skin lesion segmentation with fully convolutional-deconvolutional networks,» *arXiv:1703.05165 [cs.CV]*, 2018.
- [57] P. Tang, Q. Liang, X. Yan, S. Xiang, W. Sun, D. Zhang y G. Coppola, «Efficient skin lesion segmentation using separable-Unet with stochastic weight averaging,» *Computer Methods and Programs in Biomedicine*, vol. 178, pp. 289 -- 301, 2019.
- [58] X. Li, L. Y. C.-W. Fu y Pheng-AnnHeng, «Deeply Supervised Rotation Equivariant Network for Lesion Segmentation in Dermoscopy Images,» *arXiv:1807.02804 [cs.CV]*, vol. abs/1807.02804, pp. 235--243, 2018.
- [59] Y. Li y L. Shen, «Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network,» *Sensors (Basel)*, vol. 18, n° 2, 2018.

- [60] L. Bi, J. Kim, E. Ahn, A. Kumar, D. Feng y M. Fulham, «Step-wise integration of deep class-specific learning for dermoscopic image segmentation,» *Pattern Recognition*, vol. 85, pp. 78--89, 2019.
- [61] M. A. Al-masni, M. A. Al-antari, M.-T. Choi, S.-M. Han y T.-S. Kim, «Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks,» *Computer Methods and Programs in Biomedicine*, vol. 162, pp. 221--231, 2018.
- [62] P. Tschandl, C. Sinz y H. Kittler, «Domain-specific classification-pretrained fully convolutional network encoders for skin lesion segmentation,» *Computers in Biology and Medicine*, vol. 104, pp. 111--116, 2019.
- [63] P. Carcagnì, M. Leo, A. Cuna, P. L. Mazzeo, P. Spagnolo, G. Celeste y C. Distance, «Classification of Skin Lesions by Combining Multilevel Learnings in a DenseNet Architecture,» *Image Analysis and Processing – ICIAP 2019. ICIAP 2019. Lecture Notes in Computer Science*, vol. 11751, p. 335–344, 2019.
- [64] J.-A. Pérez-Carrasco, B. Acha, C. Suárez-Mejías, J.-L. López-Guerra y C. Serrano, «Joint segmentation of bones and muscles using an intensity and histogram-based energy minimization approach,» *Computer Methods and Programs in Biomedicine*, vol. 156, pp. 85-95, 2018.
- [65] C. Suárez-Mejías, J. A. Pérez-Carrasco, C. Serrano, J. L. ópez-Guerra, C. Parra-Calderón, T. Gómez-Cía y B. Acha, «Three-dimensional segmentation of retroperitoneal masses using continuous convex relaxation and accumulated gradient distance for radiotherapy planning,» *Med Biol Eng Comput*, vol. 55, n° 1, pp. 1-15, 2017.
- [66] N. Papadakis y J. Rabin, «Convex Histogram-Based Joint Image Segmentation with Regularized Optimal Transport Cost,» *Journal of Mathematical Imaging and Vision*, vol. 59, n° 2, pp. 161--186, 2017.
- [67] V. Badrinarayanan, A. Kendall y R. Cipolla, «SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, n° 12, pp. 2481--2495, 2017.
- [68] L. Yu, H. Chen, Q. Dou, J. Qin y P. Heng, «Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks,» *EEE Transactions on Medical Imaging*, vol. 36, n° 4, pp. 994-1004, 2017.

- [69] A. Kwasigroch, A. Mikołajczyk y M. Grochowski, «Deep convolutional neural networks as a decision support tool in medical problems – malignant melanoma case study,» *Trends in Advanced Intelligent Control, Optimization and Automation*, vol. 577, pp. 848--856, 2017.
- [70] V. Thost y J. Chen, «Directed acyclic graph neural networks,» *arXiv:2101.07965*, 2021.
- [71] A. d. S. Meyer, A. Garcia, A. P. d. Souza y C. L. d. S. Jr., «Comparison of similarity coefficients used for cluster analysis with dominant markers in maize (*Zea mays* L),» *Genetics and Molecular Biology*, vol. 27, n° 1, pp. 83--91, 2004.
- [72] L. R. Dice, «Measures of the Amount of Ecologic Association Between Species,» *Ecology*, vol. 26, n° 3, pp. 297--302, 1945.
- [73] R. Real y J. M. Vargas, «The Probabilistic Basis of Jaccard's Index of Similarity,» *Systematic Biology*, vol. 45, n° 3, pp. 380--385, 1996.
- [74] H. Fan, F. Xie, Y. Li, Z. Jiang y J. Liu, «Automatic segmentation of dermoscopy images using saliency combined with Otsu threshold.,» *Computers in Biology and Medicine*, vol. 85, p. 75–85, 2017.
- [75] T. Y. Tan, L. Zhang, C. P. Lim, B. Fielding, Y. Yu y E. Anderson, «Evolving Ensemble Models for Image Segmentation Using Enhanced Particle Swarm Optimization,» *IEEE Access*, vol. 7, pp. 34004--34019, 2019.
- [76] M. Ammar, S. G. Khawaja, A. Atif, M. U. Akram y M. Sakeena, «Learning Based Segmentation of Skin Lesion from Dermoscopic Images,» *2018 IEEE*, pp. 1--6, 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom).
- [77] L. Bi, J. Kim, E. Ahn, A. Kumar, D. Feng y M. Fulham, «M. Step-wise integration of deep class-specific learning for dermoscopic image segmentation,» *Pattern Recognition*, vol. 85, pp. 78--89, 2019.
- [78] T. Akram, H. M. J. Lodhi, S. R. Naqvi, S. Naeem, M. Alhaisoni, M. Ali, S. A. Haider y N. N. Qadri, «A multilevel features selection framework for skin lesion classification,» *Human-centric Computing and Information Sciences*, vol. 10, n° 12, 2020.

- [79] R. B. Fisher, T. P. Breckon, K. Dawson-Howe, A. Fitzgibbon, C. Robertson, E. Trucco y C. K. I. Williams, *Dictionary of Computer Vision and Image Processing*, New Jersey: Wiley, 2014.
- [80] A. Giuseppe, S. Peter, S. Chimenti, N. Takeji, W. Klaus y K. Alfred, «Dermoscopy of pigmented skin lesions: Results of a consensus meeting via the Internet,» *Journal of the American Academy of Dermatology*, vol. 48, n° 5, pp. 679-693, 2003.
- [81] K. Korotkov y R. García, «Computarized analysis of pigmented skin lesion: A review,» *Artificial Intelligence in Medicine*, vol. 56, pp. 69--90, 2012.
- [82] R. Javed, M. S. M. Rahim, T. Saba y A. Rehman, «A comparative study of features selection for skin lesion detection from dermoscopic images,» *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 9, n° 4, 2020.
- [83] J. Parker, «Texture and Color,» de *Algorithms for Image Processing and Computer Vision*, Segunda ed., John Wiley & Sons, Incorporated, 2011, pp. 177-208.
- [84] A. Sáez, C. Serrano y B. Acha, «Model-Based Classification Methods of Global Patterns in Dermoscopic Images,» *IEEE Transactions on Medical Imaging*, vol. 33, n° 5, pp. 1137-1147, 2014.
- [85] B. Ghogh, M. N. Samad, S. Asif Mashhadi, T. Kapoor, W. Ali, F. Karray y M. Crowley, «Feature Selection and Feature Extraction in Pattern Analysis: A Literature Review,» *arXiv:1905.02845 [cs.LG]*, 2019.
- [86] L. Talavera-Martínez, P. Bibilón y M. González-Hidalgo, «Computational texture features of dermoscopic images and their link to the descriptive terminology: A survey,» *Computer Methods and Programs in Biomedicine*, vol. 182, n° 105049, 2019.
- [87] R. M. Haralick, K. Shanmugam y I. Dinstein, «Textural Features for Image Classification,» *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, n° 6, pp. 610-621, 1973.
- [88] M. Presutti, «La Matriz de Co-Ocurrencia en la Clasificación Multiespectral: Tutorial para la Enseñanza de Medidas Texturales en Cursos de Grado Universitario,» *4ª Jornada de Educação em Sensoriamento Remoto no Âmbito do Mercosul*, 2004.

- [89] C. Sun y W. GWee, «Neighboring gray level dependence matrix for texture classification,» *Computer Vision, Graphics, and Image Processing*, vol. 23, n° 3, pp. 341-352, 1983.
- [90] K. Laws, «Textured image segmentation,» *Technical Report, University of Southern California Los Angeles Image Processing INST*, 1980.
- [91] B. Julesz, «Textons, the elements of texture perception, and their interactions,» *Nature*, vol. 290, pp. 91-97, 1981.
- [92] T. Ojala y M. Pietikainen, «Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,» *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, n° 7, pp. 971-987, 2002.
- [93] Y. He, N. Sang y C. Gao, «Pyramid-Based Multi-structure Local Binary Pattern for Texture Classification,» *Computer Vision – ACCV 2010 .ACCV 2010. Lecture Notes in Computer Science*, vol. 6494, pp. 133-144, 2011.
- [94] D. G. Lowe, «Distinctive image features from scale-invariant keypoints,» *International journal of computer vision*, vol. 60, n° 2, pp. 91-110, 2004.
- [95] C. Serrano y B. Acha, «Pattern analysis of dermoscopic images based on Markov random fields,» *Pattern Recognition*, vol. 42, n° 6, pp. 1052 - 1057, 2009.
- [96] A. P. Pentland, «Fractal-based description of natural scenes,» *IEEE transactions on pattern analysis and machine intelligence*, n° 6, pp. 661-674, 1984.
- [97] B. H. McCormick y S. N. Jayaramamurthy, «Time series model for texture synthesis,» *International Journal of Computer & Information Sciences*, vol. 3, n° 4, pp. 329 - 343, 1974.
- [98] R. C. Gonzalez, «Digital image processing. Pearson education india,» 2009.
- [99] J. G. Daugman, «Two-dimensional spectral analysis of cortical receptive field profiles,» *Vision Research*, vol. 20, n° 10, pp. 847-856, 1980.
- [100] S. Pathan, K. G. Prabhub y P.C.Siddalingaswamy, «Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—A review,» *Biomedical Signal Processing and Contro*, vol. 39, pp. 237-262, 2018.
- [101] D. Castillo-Barnes, *Tesis Doctoral: Caracterización de imágenes cerebrales mediante distribuciones alfa-estables e isosuperficiales. Combinación de*

- marcadores heterogéneos*, Granada: Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación. Universidad de Granada, 2021.
- [102] M. Mall, *Correlation-based feature selection for machine learning*, Hamilton: PhD thesis, University of Waikato, 1999.
- [103] T. M. Cover y J. A. Thomas, *Elements of Information* 2nd edn, New Jersey: Wiley, 2006.
- [104] N. Nicolosiz, *Feature selection methods for text classification*, Department of Computer Science, Rochester Institute of Technology, 2008.
- [105] F. Shunkai y M. Desmarais, «Markov Blanket based Feature Selection: A Review of Past Decade,» *Proceedings of the World Congress on Engineering 2010*, vol. 1, 2010.
- [106] M. Dash y H. Liu, «Consistency-based search in feature selection,» *Artificial Intelligence*, vol. 151, n° 1-2, pp. 155-176, 2003.
- [107] L. Yu y H. Liu, «Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution,» *Machine Learning, Proceedings of the Twentieth International Conference (ICML 2003)*, pp. 856-863, 2003.
- [108] Z. Zhao y H. Liu, «Searching for Interacting Features,» *Proceedings of the 20th international joint conference on Artificial intelligence*, vol. 7, p. 1156–1161, 2007.
- [109] C. Ding y H. Peng, «Minimum redundancy feature selection from microarray gene expression data,» *J Bioinform Comput Biol*, vol. 3, n° 2, pp. 185-205, 2005.
- [110] D. W. Aha y R. L. Bankert, «A Comparative Evaluation of Sequential Feature Selection Algorithms,» *Learning from Data. Lecture Notes in Statistics*, vol. 112, p. 199–206, 1996.
- [111] A. Idris, A. Khan y Y. S. Lee, «Intelligent churn prediction in telecom: employing mRMR feature selection and RotBoost based ensemble classification,» *Applied Intelligence*, vol. 36, n° 3, p. 659–672, 2013.
- [112] G. Dunteman, *Principal components analysis*, Sage, 1989.
- [113] M. Bruijne, «Machine learning approaches in medical image analysis: From detection to diagnosis,» *Medical Image Analysis*, vol. 33, pp. 94--97, 2016.

- [114] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura y R. M. Summers, «Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning,» *IEEE Trans Med Imaging*, vol. 35, n° 5, pp. 1285--1298, 2016.
- [115] T. Kenneth, I. Lars, T. Therese Louise y W. Ole, «Systematic review of machine learning for diagnosis and prognosis in dermatology,» *Journal of Dermatological Treatment*, vol. 31, n° 5, pp. 496-510, 2020.
- [116] M. A. Albahar, «Skin Lesion Classification Using Convolutional Neural Network With Novel Regularizer,» *EEE Access*, vol. 7, pp. 38306--38313, 2019.
- [117] R. C. Maron, M. Weichenthal, J. S. Utikal, A. Hekler, C. Berking, A. Hauschild, A. H. Enk, S. Haferkamp, J. Klode, D. Schadendorf y otros, «Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks,» *Eur J Cancer*, vol. 119, pp. 57--65, 2019.
- [118] H. Achim, U. Jochen S., E. Alexander H., F. Stefan y T. J. Brinker, «Superior skin cancer classification by the combination of human and artificial intelligence,» *European Journal of Cancer*, vol. 120, pp. 114-121, 2019.
- [119] A. Marka., J. Carter, E. Toto. y S. Hassanpour, «Automated detection of nonmelanoma skin cancer using digital images: A systematic review,» *BMC Med. Imaging*, vol. 19, pp. 1-12, 2019.
- [120] K. He, X. Zhang, S. Ren y J. Sun, «Deep Residual Learning for Image Recognition,» *arXiv:1512.03385 [cs.CV]*, 2015.
- [121] C. Szegedy, V. Vanhoucke, S. Ioffe y J. Shlens, «Rethinking the Inception Architecture for Computer Vision,» *arXiv:1512.00567v3 [cs.CV]*, pp. 2818--2826, 2016.
- [122] K. V. Sai Sundar, L. Rao Bonta, A. .B, P. Baruah y S. S. Sankara, «Evaluating Training Time of Inception-v3 and Resnet-50,101 Models using TensorFlow across CPU and GPU,» *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1964--1968, 2018.
- [123] J. Brownlee, «Machine Learning Mastery. Ensemble Learning Methods for Deep Learning Neural Networks,» 6 Agosto 2019. [En línea]. Available:

<https://machinelearningmastery.com/ensemble-methods-for-deep-learning-neural-networks/>.

- [124] G. M.A., H. Minghui, M. A.K., T. M. y S. P.N., «Ensemble deep learning: A review,» *arXiv:2104.02395 [cs.LG]*, 2021.
- [125] C. Barata, M. E. Celebi y J. S. Marques, «A survey of feature extraction in dermoscopy image analysis of skin cancer,» *IEEE Journal of Biomedical and Health Informatics*, vol. 23, n° 3, pp. 1096-1109, 2019.
- [126] M. A. Wahba, A. S. Ashour, S. A. Napoleon, M. M. Abd Elnaby y Y. Guo, «Combined empirical mode decomposition and texture features for skin lesion classification using quadratic support vector machine,» *Health Inf Sci Syst*, vol. 5, n° 10, 2017.
- [127] R. Sarkar, C. C. Chatterjee y A. Hazra, «A novel approach for automatic diagnosis of skin carcinoma from dermoscopic images using parallel deep residual networks,» *Advances in Computing and Data Sciences. ICACDS 2019. Communications in Computer and Information Science*, vol. 1045, 2019.
- [128] G. Argenziano, «Seven-point checklist and the seven rules not to miss melanoma incognito,» de *Atlas of Dermoscopy*, CRC Press, 2012, pp. 124--130.
- [129] K. Rungta, «Back Propagation Neural Network: Explained With Simple Example,» 2020. [En línea]. Available: <https://www.guru99.com/backpropogation-neural-network.html>.
- [130] D. Shen, G. Wu y H.-I. Suk, «Deep Learning in Medical Image Analysis,» *Medical Image Analysis. Annual review of biomedical engineering*, vol. 19, pp. 221--248, 2017.
- [131] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua y S. Süsstrunk, «SLIC Superpixels Compared to State-of-the-Art Superpixel Methods,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, n° 11, pp. 2274--2282, 2012.
- [132] H. Huang, P. Kharazmi, D. I. McLean, H. Lui, Z. J. Wang y T. K. Lee, «Automatic detection of translucency using a deep learning method from patches of clinical basal cell carcinoma images,» *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, pp. 685--688, 2018.

- [133] G. Castellano, L. Bonilha, L. Li y F. Cendes, «Texture analysis of medical images,» *Clinical Radiology*, vol. 59, n° 12, pp. 1061-1069, 2004.
- [134] L. Bottou, «Stochastic Gradient Descent Tricks,» *Neural Networks: Tricks of the Trade. Lecture Notes in Computer Science*, vol. 7700, 2012.
- [135] D. Rumelhart, G. Hinton y R. Williams, «Learning internal representations by error propagation,» de *Parallel distributed processing: explorations in the microstructure of cognition*, vol. 1: foundations, MIT Press, 1986, pp. 318-362.
- [136] M. Irfan, S. Malaji, C. Patsa, S. Rangarajan, R. Collins y T. Senjyu, «Online Learning-Based ANN Controller for a Grid-Interactive Solar PV System,» *Applied Sciences*, vol. 11, n° 18, p. 8712, 2021.
- [137] N. Codella, Q.-B. Nguyen, S. Pankanti, D. Gutman, B. Helba, A. Halpern y J. R. Smith, «Images, Deep Learning Ensembles for Melanoma Recognition in Dermoscopy,» *Computer Vision and Pattern Recognition (cs.CV)*, vol. 61, n° 4/5, 2017.
- [138] H. Sofian, J. C. Ming Than, M. Suraya y N. M. Noor, «Calcification Detection of Coronary Artery Disease in Intravascular Ultrasound Image: Deep Feature Learning Approach,» *International Journal of Integrated Engineering*, 2018.
- [139] C. Serrano, M. Lazo, A. Serrano, T. Toledo-Patrana, R. Barros-Tornay y B. Acha, «Clinically Inspired Skin Lesion Classification through the Detection of Dermoscopic Criteria for Basal Cell Carcinoma,» *Journal of Imaging*, vol. 8, n° 197, 2022.

# Glosario

---

<b>2D</b>	<i>2 Dimensional space</i>	Espacio 2 dimensional
<b>ABCDE</b>	<i>Asymetry, borer, color, diameter and evolving</i>	Asimetría, borde, color, diámetro y evolución
<b>Acc</b>	<i>Accuracy</i>	Exactitud
<b>AK</b>	<i>Actinic Keratosis</i>	Queratosis Actínica
<b>ANN</b>	<i>Artificial Neural Networks</i>	Redes neuronales artificiales
<b>AUC</b>	<i>Area Under Curve</i>	Área bajo la curva
<b>BCC</b>	<i>Basal Cell Carcinoma</i>	Carcinoma baso celular
<b>BKL</b>	<i>Bening Keratosis</i>	Queratosis Benigna
<b>CAD</b>	<i>Computer Aided Diagnosis</i>	Diagnóstico asistido por computadora
<b>CNN</b>	<i>Convolutional Neural Network</i>	Red neuronal convolucional
<b>CPU</b>	<i>Central Process Unit</i>	Unidad central de procesamiento
<b>CSV</b>	<i>Comma Separated Value</i>	Valor separado por coma
<b>DAG</b>	<i>Directed Acyclic Graph</i>	Grafo acíclico dirigido
<b>DCN</b>	<i>Deep Convolutional Network</i>	Red convolucional profunda
<b>DF</b>	<i>Dermatofibroma</i>	Dermatofibroma
<b>DFT</b>	<i>Discrete Fourier transform</i>	Transformada discreta de Fourier
<b>DIC</b>	<i>DICE</i>	DICE
<b>DNN</b>	<i>Deep Neural Network</i>	Red neuronal profunda
<b>DP</b>	<i>Deep Learning</i>	Aprendizaje profundo
<b>DPI</b>	<i>Dot per inch</i>	Puntos por pulgada

<b>EDRA</b>	<i>The Interactive Atlas of Dermoscopy</i>	Atlas interactivo de dermatoscopia
<b>FC</b>	<i>Fully Connected</i>	Totalmente conectada
<b>FCM</b>	<i>Fuzzy C-Means</i>	Fuzzy C-Medias
<b>FCN</b>	<i>Fully Convolutional Networks</i>	Redes totalmente convolucionales
<b>FCRN</b>	<i>Fully Convolutional Recurrent Network</i>	Red recurrente totalmente convolucional
<b>FFT</b>	<i>Fast Fourier Transform</i>	Transformada Rápida de Fourier
<b>FPR</b>	<i>False Positive Rate</i>	Tasa de falsos positivos
<b>GANs</b>	<i>Generatives Adversarial Networks</i>	Redes adversarias generativas
<b>GMM</b>	<i>Gaussian Mixture Model</i>	Modelo Gaussiano mixto
<b>GPU</b>	<i>Graphic Process Unit</i>	Unidad de procesamiento gráfico
<b>HAM</b>	<i>Human Against Machine with 10000 training images</i>	Humanos vs máquina con 10000 imágenes de entrenamiento
<b>HSV</b>	<i>Hue, Saturation, Brightness</i>	Matiz, saturación, brillo
<b>IHP</b>	<i>Independent Histogram Pursuit</i>	Búsqueda del histograma independiente
<b>IoU</b>	<i>Intersection Over Union</i>	Intersección sobre la unión
<b>ISIC</b>	<i>ISIC: Skin Lesion Analysis Towards Melanoma Detection</i>	Análisis de la lesión en búsqueda de la detección del melanoma
<b>Jac</b>	<i>Jaccard</i>	Jaccard
<b>LAB</b>	<i>CIELAB</i>	CIELAB
<b>LEM</b>	<i>Laws Energy Mask</i>	Leyes de energía de las máscaras
<b>MEL</b>	<i>Melanoma</i>	Melanoma
<b>MRF</b>	<i>Markov Random Field</i>	Canales aleatorios de Markov
<b>PCA</b>	<i>Principal Component Analysis</i>	Análisis de componentes principales
<b>Pre</b>	<i>Precision</i>	Precisión
<b>ReLU</b>	<i>Rectified Linear Unit</i>	Unidad lineal rectificada

---

<b>RGB</b>	<i>Red, Green, Blue</i>	Rojo, verde, azul
<b>RM</b>	<i>Raw Mask</i>	Máscara cruda
<b>RNN</b>	<i>Recurrent Neural Network</i>	Red neuronal recurrente
<b>ROI</b>	<i>Region of Interest</i>	Región de Interés
<b>SCC</b>	<i>Squamous Cell Carcinoma</i>	Carcinoma celular escamoso
<b>Se</b>	<i>Sensitibity</i>	Sensibilidad
<b>SGD</b>	<i>Stochastic Gradient Descent</i>	Gradiente estocástico descendiente
<b>SGDM</b>	<i>Stochastic gradient descent with momentum</i>	Gradiente estocástico descendiente con momentum
<b>Sp</b>	<i>Especificity</i>	Especificidad
<b>SVM</b>	<i>Suppor Vector Machine</i>	Máquina de soporte de vectores
<b>VASC</b>	<i>Vascular Lesion</i>	Lesiones Vasculares