

DETECCIÓN DE MELANOMAS DE PIEL MALIGNOS MEDIANTE PROCESAMIENTO DIGITAL DE IMÁGENES USANDO REDES NEURONALES CONVOLUCIONALES

Proyecto de Grado de Maestría

Sebastian Riaño Borda

Universidad Santo Tomás

Facultad de Ingeniería Electrónica

Maestría en Ingeniería Electrónica

Bogotá, 2022

DETECCIÓN DE MELANOMAS DE PIEL MALIGNOS MEDIANTE PROCESAMIENTO DIGITAL DE IMÁGENES USANDO REDES NEURONALES CONVOLUCIONALES

Proyecto de Grado de Maestría

Sebastian Riaño Borda

Ingeniero José Guillermo Guarnizo (Director)

Ingeniero Armando Mateus (Co-Director)

Ingeniero Edgar Camilo Camacho (Co-Director Externo)

Universidad Santo Tomás

Facultad de Ingeniería Electrónica

Maestría en Ingeniería Electrónica

Bogotá, 2022

Contenido

Resumen.....	1
1 Introducción	2
2 Objetivos	4
2.1 General	4
2.2 Específicos	4
3 Marco Conceptual	5
4 Estado del Arte.....	11
5 Metodología de Investigación Desarrollada	18
6 Arquitectura Para Sistema Detector De Melanomas	21
6.1 Arquitectura Propuesta	21
6.2 Configuración Experimental	22
6.2.1 Fase 1: Experimentos previos	23
6.2.2 Fase 2	24
7 Proceso de Evaluación (Resultados).....	28
7.1 Fase 1: Experimentos previos	28
7.2 Fase 2.....	31
8 Discusión.....	35
9 Conclusiones	37
10 Trabajos Futuros	39
11 Referencias.....	40
12 Anexos	46

Índice de Figuras

Figura 1. Lunares [7]	5
Figura 2. Melanomas [7].....	6
Figura 3. Porcentaje de búsqueda a nivel mundial [17].....	11
Figura 4. Desglose de la metodología propuesta.	18
Figura 5. Fototipos de piel [55].	19
Figura 6. Arquitectura propuesta para la CNN	21
Figura 7. Arquitectura propuesta para el experimento 5.....	24
Figura 8. Entrenamiento del modelo 104. La gráfica de la izquierda muestra la curva de aprendizaje del modelo, mientras que la segunda cómo la pérdida decrece con cada época.	29
Figura 9. Entrenamiento del modelo 706. La gráfica de la izquierda muestra la curva de aprendizaje del modelo, mientras que la segunda cómo la pérdida decrece con cada época.	30

Índice de Tablas

Tabla 1. Weighted Glasgow 7-point checklist.....	6
Tabla 2. Rangos de los valores utilizados en las configuraciones aplicadas para los diferentes entrenamientos de los experimentos de la fase 1	23
Tabla 3. Resumen de las configuraciones aplicadas para los 3 entrenamientos del experimento 1	24
Tabla 4. Resumen de las configuraciones aplicadas para los 3 entrenamientos del experimento 2	25
Tabla 5. Resumen de las configuraciones aplicadas para los 3 entrenamientos del experimento 3	26
Tabla 6. Resumen de las configuraciones aplicadas para los 3 entrenamientos del experimento 4	26
Tabla 7. Resumen de las configuraciones aplicadas para los 3 entrenamientos del experimento 5	27
Tabla 8. Matriz de confusión del modelo entrenado con mejor respuesta (Modelo 10).....	28
Tabla 9. Matriz de confusión del modelo entrenado con mejor respuesta (Modelo 401).....	29
Tabla 10. Matriz de confusión del modelo entrenado con mejor respuesta (Modelo 706).....	30
Tabla 11. Matrices de confusión de los 3 modelos del experimento 1. El modelo 1 que fue entrenado con un dropout de 0.2, el modelo 2 con 0.3 y el modelo 3 con 0.4	31
Tabla 12. Matrices de confusión de los 3 modelos del experimento 2. El modelo 4 que fue entrenado con un dropout de 0.4, el modelo 5 con 0.3 y el modelo 6 con 0.2	31
Tabla 13. Matrices de confusión de los 3 modelos del experimento 3. El modelo 7 que fue entrenado con un dropout de 0.2, el modelo 8 con 0.3 y el modelo 9 con 0.4	32
Tabla 14. Matrices de confusión de los 3 modelos del experimento 4. El modelo 10 que fue entrenado con un dropout de 0.4, el modelo 11 con 0.3 y el modelo 12 con 0.2.....	33
Tabla 15. Matrices de confusión de los 3 modelos del experimento 5. En los modelos 13, 14 y 15 la diferencia respecto de los casos anteriores se encuentra en el pooling utilizado, este fue inicialmente seleccionado de 3x3 en todas las capas, sin embargo, debido a las limitaciones del dispositivo utilizado se combinó con pooling de 2x2	34
Tabla 16. Resumen de los mejores modelos entrenados en cada experimento.	35
Tabla 17. Resumen de los mejores modelos entrenados en cada experimento de la fase 2	36

Resumen

Este proyecto se propone diseñar una red neuronal para el reconocimiento de melanomas (un tipo de cáncer de piel), mediante el uso de una técnica conocida como redes neuronales convolucionales, mayormente utilizada en visión artificial (una rama de la inteligencia artificial), aplicada en el reconocimiento de patrones sobre lunares en la piel y determinar la existencia de un melanoma maligno, o no, a partir de un dataset limitado. Para esto, la red convolucional diseñada y entrenada para clasificar los melanomas está formada por unas capas de convolución y pooling apiladas entre sí para formar la red propuesta, una “fully connected layer” y un clasificador con 1 o 2 salidas, y es parametrizada con diferentes valores en características como el dropout, el tamaño de los filtros, entre otros, realizando los entrenamientos en 5 diferentes etapas o experimentos. El dataset propuesto para el entrenamiento de la CNN (Convolutional Neural Networks) es la colección pública más grande de imágenes demoscópicas de lesiones en la piel, proveída de manera gratuita por “International Skin Imaging Collaboration (ISIC)”, un esfuerzo por mejorar el diagnóstico de melanomas, patrocinado por la “International Society for Digital Imaging of the Skin (ISDIS)”. El propósito de este proyecto es diseñar una red neuronal convolucional con alto nivel de precisión que ayude a los profesionales en medicina con el diagnóstico de melanomas, en este caso fue posible conseguir una precisión de hasta 87.82% con la red diseñada con mejor rendimiento.

1 Introducción

La necesidad por mejorar las condiciones y calidad de vida ha traído a la humanidad a una era donde la tecnología es capaz de interactuar con otras áreas de estudio, permitiendo que algunas tareas sean desarrolladas por dispositivos inanimados, gastando menos tiempo, manteniendo una mayor precisión, etc. Cómo en el campo de la salud algunas aplicaciones van desde examinación por medio de la recopilación de información, la investigación, los tratamientos, hasta la creación y producción de fármacos para diferentes enfermedades, o equipos de soporte en intervenciones quirúrgicas simplificando procesos [1] [2].

Una de las ciencias, y que mantiene en constante desarrollo es la medicina, esto debido a la gran variedad de diferentes enfermedades que afectan a la salud humana. Estas enfermedades pueden ser tan evidentes como un resfriado común que puede ser tratado en casa sin ninguna clase de complicación; o, por el contrario, como otro tipo de afecciones que pueden pasar desapercibidas hasta que los síntomas empiezan a ser obvios y es entonces cuando la falta de un tratamiento oportuno puede desencadenar en una incapacidad, discapacidad o incluso en la muerte del paciente. Esta es la razón más significativa por la cual la importancia de un diagnóstico temprano, debido a que hay enfermedades silenciosas cuyo avance puede detenerse o ser controlado, si no hay posibilidad de curarlas, todo esto con ayuda de rigurosos controles para identificar heridas, signos, síntomas, cambios en el organismo, etc [3]. Aunque es cierto que estas enfermedades pueden ser encontradas en un primer chequeo realizado por un profesional de la salud, también es cierto que pueden pasar desapercibidas debido a que sus características en etapas tempranas suelen ser sutiles y casi invisibles, hasta que su progreso empieza a ser evidente como sucede con enfermedades como el Alzheimer, diabetes, cáncer, entre otros. Una de las enfermedades más estudiadas, debido a su alto índice de mortalidad e incidencia es el cáncer, inicialmente puede ubicarse en cualquier parte del cuerpo humano y en algunos casos su existencia es desconocida hasta el momento cuando el tumor empieza a crecer de manera descontrolada o dispersarse a otras partes del cuerpo. De acuerdo con [4], uno de los tipos de cáncer de piel más agresivo es conocido como melanoma maligno causante de hasta el 90% de las muertes por cáncer de piel (2285 muertes en 2016 en el Reino Unido), esto debido a que es propenso a la metástasis y posee gran similitud con lesiones de piel benignas, principalmente este luce como una lesión similar a un lunar, que dependiendo de ciertas características puede ser detectado y estudiado en tiempo para evitar una dispersión (también conocida como metástasis) y tratarlo para disminuir su riesgo de mortalidad.

Usualmente el primer diagnóstico dado por un profesional, para determinar si un lunar es o no un melanoma maligno, es por medio de una técnica visual conocida como la regla del ABCDE, esta es usada para revisar diferentes características físicas del lunar. Sin embargo, como esta es una técnica visual, esta puede fácilmente generar falsos positivos (casos que no son melanomas, pero tras ser analizados son diagnosticados como si lo fueran) y se cree que la sensibilidad de un diagnóstico clínico hecho por un médico general varía entre un 70% a 88%, mientras que en el caso de un dermatólogo con experiencia se estima alrededor del 82% al 100% [5] [6], específicamente en el artículo [5] mencionan que los expertos al usar dermatoscopia como medio para valorar las lesiones alcanzaron una sensibilidad del 89 % y una especificidad del 79 %. Debido a este porcentaje y la existencia de los falsos positivos, este proyecto propone desarrollar un sistema basado en inteligencia artificial y visión por computadora, construido a base de una red neuronal convolucional capaz de aprender por sí misma a diferenciar entre melanomas y lesiones benignas a partir de un conjunto de imágenes tomadas de la página oficial de ISIC [7], como

determinar si la lesión es maligna o no, esperando obtener mejores resultados a la de los profesionales en medicina, mediante un tamizaje que ayude a evitar la necesidad de una posible biopsia o hacer más sencilla la obtención de un diagnóstico temprano y el tratamiento respectivo, de esta manera proveer una herramienta hecha para profesionales en medicina les ayude a determinar si es necesario remitir al paciente a ser valorado por un especialista en dermatología o no.

A lo largo de los últimos años, la inteligencia artificial (AI por sus siglas en inglés) se ha convertido en una de las tecnologías más usadas gracias a su gran variedad de aplicaciones, versatilidad y adaptabilidad a cualquier tipo de proceso que pueda ser automatizado, controlado y aprendido, todo esto esperando minimizar el tiempo de respuesta y obtener resultados más precisos que aquellos obtenidos por los sentidos del hombre. AI puede también ser aplicada a la clasificación de imágenes por reconocimiento de patrones como es la aplicación que se presenta en este proyecto, mediante el desarrollo de una red neuronal convolucional, diseñada para identificar melanomas mediante un 1 o 0 a su salida, lo cual significa que 1 es melanoma y 0 no lo es, de esta manera se obtuvo un porcentaje de precisión de este método de clasificación un poco mayor al aproximado dado por un dermatólogo con experiencia [5], con hasta 90 % de las lesiones benignas exitosamente diagnosticadas luego de entrenar la arquitectura propuesta conformada por 3 etapas. Primero se encontró la entrada, la cual pasa por un preprocesamiento para asegurar el mismo tamaño a todas las imágenes y debido a la limitada cantidad de fotos que se tienen para entrenar la red, y las dificultades causadas por el bajo contraste entre la lesión y la piel normal, también se aplicaron giros y rotaciones para de esta manera tener un set de imágenes más amplio. La segunda etapa consiste en las capas convolucionales cada una seguida de una capa de pooling para extraer y filtrar únicamente las características más relevantes en cada imagen. Finalmente, la tercera etapa se conoce como “*fully connected layer*”, en esta capa es donde la clasificación toma lugar, donde el resultado nos dice si la lesión de la imagen es o no maligna.

2 Objetivos

2.1 General

Desarrollar un sistema para el análisis de melanomas de piel, mediante el uso de redes neuronales convolucionales entrenadas utilizando fotografías, con el fin de obtener un modelo que permita la detección de tumores cutáneos malignos en seres humanos.

2.2 Específicos

- Diseñar la arquitectura de la red neuronal convolucional que permita el proceso de clasificación de melanomas.
- Entrenar la red neuronal convolucional propuesta a partir de una base de datos de fotografías de melanomas.
- Evaluar el correcto funcionamiento de la red neuronal convolucional mediante bases de datos y la creación de una matriz de confusión.

3 Marco Conceptual

¿Qué es un lunar?

Los lunares son manchas comunes que crecen sobre la piel (Figura 1) y suelen ser de color rosado, oscuros o cafés, pueden ser lisos o con relieve, por regla general son circulares o en forma de óvalo, además su diámetro no suele ser mayor al de la goma de un lápiz. Un lunar aparece cuando las células productoras del pigmento en la piel conocidas como melanocitos, se acopian y crecen en grupos. Una persona puede tener entre 10 y 40 lunares en todo el cuerpo y además seguir desarrollando nuevos lunares hasta los 40 años aproximadamente, sin embargo, estos también suelen ir desapareciendo a medida que pasan los años. Se cree que más o menos una por cada diez personas presenta un lunar atípico, es decir, de aspecto inusual o distinto a los lunares frecuentes, a este caso atípico también se le denomina nevo displásico y es este el que tiene más probabilidades de transformarse en un melanoma, con respecto a los demás lunares [8].

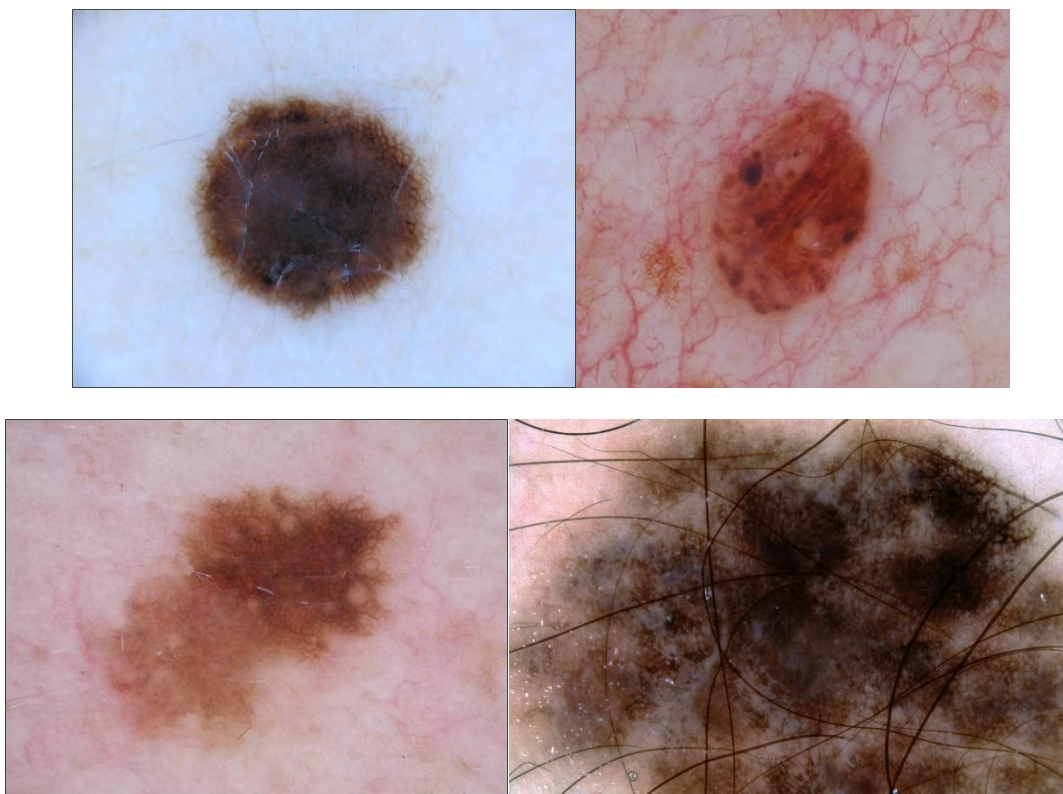


Figura 1. Lunares [7]

¿Qué es un Melanoma?

Es el tipo de cáncer de piel más serio (Figura 2), puesto que, aunque representa el 1% de los casos registrados de cáncer en la piel, es el que mayor número de muertes causa [9]. Por lo general, las primeras señales de alerta que se presentan respecto de la existencia de un melanoma son el cambio en su tamaño, forma, color o textura de la lesión. La gran mayoría de los melanomas poseen un área de color negra o negra azulada, pueden surgir como un nuevo lunar inofensivo. El melanoma puede ser negro, anormal o "de aspecto irritante a la vista". Una vez se ha identificado este tipo de

cáncer, los especialistas del área de la salud tratarán de identificar la etapa en que se encuentra. Aunque la cirugía es el tratamiento más efectivo en los distintos estadios del melanoma, existen otros tratamientos que incluyen quimioterapia, radiación, terapia biológica y terapia dirigida. La terapia biológica consiste en aumentar la capacidad propia del cuerpo humano para combatir el cáncer, mientras que la terapia dirigida utiliza medicinas u otras sustancias como tratamiento para combatir directamente células cancerosas específicas y que van a causar menos daños a las células normales [10] [11].



Figura 2. Melanomas [7]

Métodos de detección de Melanomas

Uno de los métodos utilizados para realizar el diagnóstico y que además sirve como ayuda para realizar un seguimiento o autodiagnóstico de melanomas es el "ABCDE" [10]:

- Asimetría: el contorno de una parte de la lesión no es igual al resto.
- Bordes: tiene bordes desiguales, de bajo contraste o irregulares.
- Color: no es uniforme y puede incluir tonalidades negras, cafés y canela.
- Diámetro: cambian de tamaño y generalmente se van haciendo más grandes.
- Evolución: se presentan cambios en el lunar en las últimas semanas o meses.

Otro método utilizado para la evaluación visual de los melanomas es una lista de verificación ponderada de 7 puntos, también conocida como "*Weighted Glasgow 7-point checklist*" [4], que se muestra en la **Tabla 1**:

Tabla 1. Weighted Glasgow 7-point checklist.

(a) Lista de verificación ponderada de 7 puntos de Glasgow	
	La puntuación ≥ 3 sugiere derivación
Principales características (2 puntos)	
	Cambio en el tamaño de la lesión.
	Pigmentación irregular
	Borde irregular
Características menores (1 punto)	
	Inflamación
	Picazón o sensación alterada
	Más grande que otras lesiones (diámetro > 7 mm)
	Exudación / costra de la lesión
(b) Características clínicas tranquilizadoras que sugieren una lesión cutánea benigna	
	Color, superficie y borde regulares
	Crecimiento durante días, no semanas (esto sugiere un granuloma piógeno)
	Apariencia "atascada" con tapones queratónicos en la superficie (esto sugiere una queratosis seborreica)
	Hoyuelos si se aplica presión bilateral a la lesión, esta característica sugiere un dermatofibroma)
	Una lesión pigmentada en un niño.
Debido a la apariencia variable y la dificultad para diagnosticar los melanomas, las características de "descarte" como estas pueden ser extremadamente útiles en la clasificación de las lesiones cutáneas.	

Adicionalmente, se han propuesto otros métodos como por ejemplo el de Jose Antonio Avilés-Izquierdo [12], en el cual para detectar metástasis en melanomas de piel se realiza una clasificación con base en cuatro patrones de colores: azul, rosado, café y patrones mezclados.

¿Qué es inteligencia Artificial?

Según [13]- [14]- [15], alrededor de los años 50's científicos empezaron a explorar las capacidades de los computadores para resolver los problemas por su cuenta, sin embargo, el término IA fue acuñado como tal en 1956. El concepto de IA ha cambiado con el paso del tiempo, sin embargo, en su núcleo siempre se ha tratado de construir máquinas capaces de pensar como humanos y tener la habilidad de ejecutar funciones cognitivas como observar, aprender, razonar y resolver

problemas. De esta manera la inteligencia artificial se puede categorizar en 3 niveles de abordaje según su funcionalidad [14], denominados:

IA General: que puede ejecutar tareas cognitivas con el mismo nivel de precisión del cerebro humano.

IA Estrecha: se refiere a aplicaciones donde la máquina puede ejecutar una tarea específica mejor que un humano. La mayoría de las investigaciones actuales en IA suceden en esta área.

IA Fuerte: este nivel de IA permite a las máquinas superar al cerebro humano en varias tareas.

Por otro lado, existen barreras que separan las máquinas de los humanos, estos límites definen 4 tipos de IA, [16]:

Puramente reactiva: no pueden formar memorias, por eso no poseen experiencias pasadas o datos con los cuales trabajar. Ellos se especializan en ejecutar una tarea específica.

Memoria limitada: mantienen recolectando datos que encuentran y los agregan a su memoria. Confían en experiencias anteriores para llevar a cabo una ejecución, sin embargo, solo pueden usar esta experiencia para tomar una decisión, pero no pueden formar ideas basadas en ella.

Teoría de la mente: estas máquinas están más cerca del actuar del cerebro humano en el sentido que pueden interpretar emociones y tienen entendimiento cognitivo basado en el ambiente, características faciales, etc.

Conciencia: esta es la clase de máquinas super inteligentes y conscientes que impulsan el futuro de la inteligencia artificial y son vistos como la siguiente generación de la tecnología. Son capaces de imitar muchas reacciones humanas y probablemente usarán modelos internos para simular sus acciones.

¿Qué es visión por computadora?

En *A Guide to Convolutional Neural Networks for Computer Vision* [15], Salman Khan afirma que la visión por computadora se define como la ciencia que tiene como objetivo dar a los computadores una capacidad similar al ojo humano, si es que no superior, a la hora de ver e interpretar el mundo 3D, es decir, la capacidad de categorizar, localizar, etiquetar y segmentar cada uno de los objetos que están presentes en una imagen. Exactamente, con la visión por computadora lo que se busca es desarrollar sistemas capaces de imitar una de las habilidades más asombrosas que tiene el sistema visual humano como, por ejemplo, inferir características de los objetos del mundo real en 3D utilizando solamente la luz reflejada en los ojos por diferentes objetos.

Basados en el nivel de abstracción de la información de salida, las tareas de visión por computadora se dividen en 3 diferentes categorías, conocidas en inglés como low-level, mid-level, and high level [15].

Visión de bajo nivel (low-level): La coincidencia de imágenes es un ejemplo de tareas de visión de bajo nivel. Se define como la identificación automática de puntos de imagen correspondientes en un par dado de la misma escena desde diferentes puntos de vista, o una escena en movimiento capturada por una cámara fija.

Visión de nivel medio (mid-level): Inferir la geometría de los objetos es uno de los principales aspectos de la visión de nivel medio. La visión geométrica incluye geometría de múltiples vistas, estéreo y estructura a partir del movimiento, que infieren la información de la escena en 3D a partir de imágenes en 2D, de modo que la reconstrucción en 3D podría ser posible.

Visión de alto nivel (high-level): Basados en una adecuada representación segmentada de la estructura 2D y / o 3D de la imagen, extraída utilizando una técnica de visión por computador de nivel inferior, la visión de alto nivel completa la tarea de entregar una interpretación coherente de la imagen. La visión de alto nivel determina qué objetos están presentes en la escena e interpreta sus interrelaciones. Por ejemplo, el reconocimiento de objetos y entendimiento de escenas son dos tareas de visión de alto nivel que infieren la semántica de objetos y escenas, respectivamente.

¿Qué son redes neuronales artificiales?

Los modelos de redes neuronales artificiales, [15], pueden entenderse como un conjunto de unidades básicas de procesamiento, que están estrechamente interconectadas y operan en las entradas dadas para procesar la información y generar los resultados deseados. Las redes neuronales se pueden agrupar en dos categorías genéricas según la forma en que se propaga la información en la red:

Redes Feed-Forward: El flujo de la información va en una sola dirección. Si la red es considerada como una gráfica con neuronas como sus nodos, las conexiones entre los nodos son tales que no hay bucles o ciclos en el gráfico, por ejemplo, MPL (Multilayer Perceptron) y CNNs (Convolutional Neural Networks).

Redes Feed-Back: Como su nombre implica, las redes feed-back tienen conexiones que forman ciclos dirigidos (o bucles). Esta arquitectura permite operar y generar secuencias de tamaños arbitrarios. Las redes feed-back exhiben capacidad de memorización y pueden almacenar información, y relaciones de secuencia en su memoria interna. Ejemplos de tales arquitecturas incluyen RNN (Recurrent Neural Network) y LSTM (Long-Short Term Memory)

¿Qué son redes neuronales artificiales convolucionales?

Según Salman Khan, en [15], son una de las categorías de redes neuronales más conocidas y usadas, especialmente para datos con alta dimensionalidad, como es el caso de las imágenes y los videos. Ellas operan de manera similar a las redes neuronales estándar. Una de las diferencias claves, es que cada unidad en una capa es un filtro de 2 dimensiones (o superior) que es convolucionado con la entrada de dicha capa. La que es esencial para casos donde se quieren aprender patrones de entradas de altas dimensiones, por ejemplo, las imágenes o videos.

Las CNN son un tipo útil de modelo para paradigmas de aprendizajes supervisados y no supervisados. Un mecanismo de aprendizaje supervisado es aquel en el que se conocen la entrada de la red y las salidas requeridas (etiquetas verdaderas), aprendiendo de esta manera el mapeo entre los dos. Mientras que, en el mecanismo de aprendizaje no supervisado, no se tiene conocimiento sobre las etiquetas reales para un conjunto dado de entrada y, es entonces que el modelo entrenado tiene como objetivo estimar la distribución subyacente de las muestras de datos de entrada. Una CNN aprende a fijar una imagen dada a su categoría correspondiente mediante la detección de una serie de representaciones de características abstractas, que van desde las más básicas hasta las más compuestas. Estas características discriminatorias se usan luego dentro de la red entrenada para poder realizar la predicción la categoría correcta a la que una imagen de entrada pertenece.

4 Estado del Arte

Con los recientes avances tecnológicos alrededor del mundo y las exigencias del mercado durante los últimos años, se ha empezado a popularizar cada vez más, según el generador de tendencias de *Google Trends* en la Figura 3 [17], el uso de las redes neuronales y otras técnicas de este estilo con el fin de automatizar algunos procesos como la detección de características o patrones en imágenes o señales de audio, partiendo de que el sistema aprenderá a identificar y clasificar la información de importancia para luego dar resultados, a este proceso se le llama *machine learning* y con el tiempo a partir de este se han ido trabajando otras técnicas que han dejado mejores resultados como son el caso del aprendizaje profundo, que es un grupo de algoritmos aplicados al aprendizaje de máquina que pueden o no ser supervisados y las CNN, que han mostrado un asombroso desempeño y resultados en el tratamiento y procesamiento de imágenes de gran tamaño.

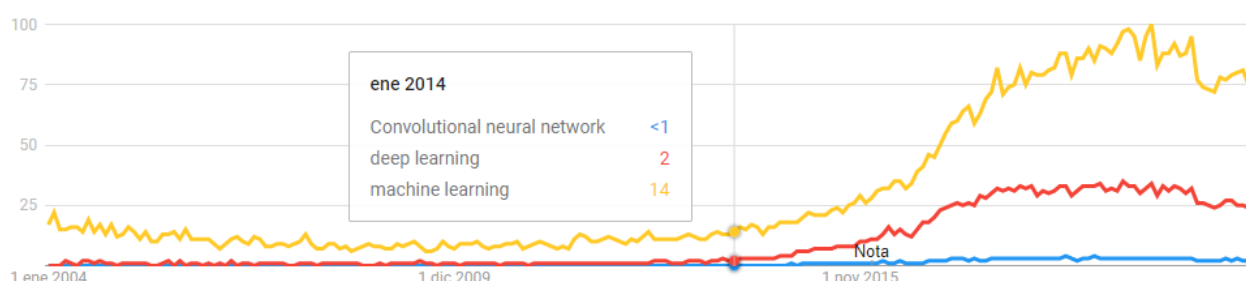


Figura 3. Porcentaje de búsqueda a nivel mundial [17].

A finales de la década de 1950, Rosenblatt tuvo como objetivo crear un cerebro analógico útil para tareas de análisis. Esto marcó el inicio de la investigación en el campo de *machine learning*, es decir, permitir que las máquinas aprendan y clasifiquen de la manera que lo hace un ser humano. Las técnicas de *machine learning* han sido ampliamente aplicadas en áreas como el reconocimiento de patrones, procesamiento del lenguaje natural y aprendizaje computacional, de esta manera las computadoras son equipadas con la capacidad de actuar sin ser programadas explícitamente, construyendo algoritmos que puedan aprender de los datos y tomando decisiones o predicciones basadas en los mismos. Para este tipo de sistemas, se deben cumplir 2 condiciones: En primer lugar, que tenga la capacidad de reconocer y procesar patrones complejos de información relacionada entre sí desde todas las dimensiones como lo hace el cerebro humano y, en segundo lugar, para lograr una inteligencia similar a la humana, una máquina debe requerir acceso a un gran volumen de información. Un sistema que podría operar con algunas de las nociones anteriores, llamado perceptrón se sugirió en Rosenblatt (1957). Este perceptrón se convirtió en la base para la creación de redes de aprendizaje de múltiples capas que han formado la base de lo que popularmente se llama *Deep Learning* [18], [19].

Deep learning es un subcampo del *machine learning* el cual se caracteriza por tener múltiples capas de procesamiento no lineal mediante las que trata aprender de manera no supervisada las representaciones de alto nivel mediante el uso de arquitecturas jerárquicas de datos. Hay principalmente tres razones importantes por las cuales el actual auge del *deep learning*: el

increíble aumento de la capacidad de procesamiento de los chips, el significativamente bajo costo del hardware de computación y los considerables avances en algoritmos de *machine learning* [20], [21].

La detección temprana de melanomas de piel malignos mediante imágenes demoscópicas es un reto debido al bajo contraste entre el lunar y la piel que lo rodea, las similitudes entre un melanoma y un no-melanoma, la presencia de artefactos en la imagen (vello corporal), la forma irregular y poco definida de los bordes. Con base en estos detalles, se han realizado una serie de propuestas para el procesamiento de las imágenes con el fin de obtener un reconocimiento preciso de estas lesiones de manera digital, algunas de ellas implementan métodos de bajo nivel, aplicando algunas técnicas de reconocimiento de patrones en imágenes basadas en *machine learning*, y específicamente con *deep learnig* como sucede en *Neural Network Diagnosis of Malignant Melanoma From Color Images* [22], donde desarrolla una técnica para el reconocimiento y diagnóstico de melanomas mediante el uso de redes neuronales, teniendo inconvenientes en la estimación del tamaño de la lesión haciendo que esta sea omitida debido a su imprecisión, o como Patrick Vannoorenberghe en *Dempster-Shafer's theory as an aid to color information processing. Application to melanoma detection in dermatology* [23], los autores proponen un método de segmentación de una imagen a color basado en la teoría de *Dempster-Shafer* aplicado a la detección de cáncer de piel tipo melanoma, con el cual se obtuvieron un 23% de no detecciones y 15% de falsas alarmas, razón por la cual los autores aseguran que es necesario reforzar los resultados utilizando una base de datos más grande; y específicamente redes convolucionales neuronales se han desarrollado sistemas de hasta 50 *layers* (capas) que permiten dar un diagnóstico, aunque sigue presentando algunas dificultades en el reconocimiento de las lesiones, como propone Lequan Yu en [24].

Esta propuesta sobre las redes neuronales artificiales convolucionales ha tomado fuerza recientemente y debido a esto se encuentran en la literatura algunos estudios aplicados al reconocimiento de melanomas como en *Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks with Jaccard Distance* [25], donde se aplicó una CNN con 19 capas (*layers*) de profundidad que es entrenada de extremo a extremo y no se basa en el conocimiento previo de los datos, además de diseñar una función de pérdida novedosa basada en la distancia Jaccard para eliminar la necesidad de volver a ponderar la muestra y finalmente, con ayuda de dos bases de datos disponibles públicamente; se evaluó la efectividad, la eficiencia y la capacidad de generalización del marco propuesto; adicionalmente Lequan y demás autores de *Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks* [24], proponen una CNN con más de 50 capas de profundidad para adquirir características más ricas y discriminativas para un reconocimiento más preciso, proponiendo una serie de esquemas para asegurar la efectividad del entrenamiento y aprendizaje bajo una limitada cantidad de información, luego se aplicó un *Full Convolutional Residual Network* (FCRN) para mejorar la segmentación de las lesiones y mejorar aún más su capacidad mediante la incorporación de un esquema de integración de información contextual de escala múltiple; o en [26], *Deep learning ensembles for melanoma recognition in dermoscopy images*, en el que proponen un sistema que combina los desarrollos recientes en *Deep learning* con enfoques de *machine learning* establecidos, creando conjuntos de métodos capaces segmentar las lesiones de la piel, así como analizar el área detectada y el tejido circundante para la detección del melanoma. En un artículo más reciente *DermaKNet: Incorporating the knowledge of dermatologists to Convolutional Neural Networks for skin lesion diagnosis* [27], Iván González-Díaz presenta un sistema basado

en CNNs que incorpora subsistemas para el modelado de las propiedades de las lesiones en la piel. Por otro lado, en *Melanoma Detection by Analysis of Clinical Images Using Convolutional Neural Network* [28] se propone una CNN que usa imágenes no dermatoscópicas, las cuales son preprocesadas con el fin de reducir interferencias como la iluminación u otros artefactos con el fin de desarrollar una herramienta que asista al dermatólogo para dar un diagnóstico temprano de este tipo de cáncer. Mobeen ur Rehman en *Classification of Skin Lesion by Interference of Segmentation and Convolution Neural Network* [29], presenta las bases para la detección de lesiones de piel con inclinación cancerosa mediante segmentación y posterior aplicación de redes de convolución en imágenes dermatoscópicas, extrayendo características que luego son clasificadas usando una red neuronal artificial clasificadora. En *Hybrid dermoscopy image classification framework based on Deep convolutional neural network and Fisher vector* [30], propone un método híbrido para la clasificación de imágenes combinando las CNN, *Fisher Vector* (FV) y *support vector machine* (SVM). Estas y otras propuestas han sido desarrolladas y evaluadas, usando técnicas y herramientas diferentes, como “Raman Spectra” en *Detection of skin cancer by classification of Raman spectra* [31], técnicas basadas en análisis de histograma en *Automatic lesion segmentation for melanoma diagnostics in macroscopic images* [32], entre otros tipos de herramientas de procesamiento para finalmente obtener la clasificación, como en [33], [34], [35], [36]

En publicaciones más recientes, se encontraron trabajos como *Progressive Transfer Learning and Adversarial Domain Adaptation for Cross-Domain Skin Disease Classification* el autor Yanyang Gu propone dos métodos para una nueva tarea de reconocimiento de enfermedades de la piel de dominios cruzados, partiendo por un clasificador pre-entrenado en ImageNet una red neuronal convolucional profunda completamente supervisada, se explora una técnica de aprendizaje de transferencia progresiva de dos pasos ajustando la red en dos conjuntos de datos de enfermedades de la piel; además proponen adoptar el *Adversarial Learning* como una técnica de adaptación de dominio para realizar la traducción de atributos invariantes desde el dominio fuente al del objetivo con el fin de mejorar el rendimiento de reconocimiento [37]. En *Classifying Melanoma Skin Lesions Using Convolutional Spiking Neural Networks With Unsupervised STDP Learning Rule* [38], se hace una clasificación de lesiones de piel de melanomas malignos y nevus melanocitos benignos mediante el uso de SNNs convolucionales con la regla de aprendizaje no supervisado *spike-timing-dependent plasticity (STDP)*, además de utilizar selección de características para seleccionar más características diagnósticas para mejorar el rendimiento de clasificación de las redes. En *Automatic Skin Cancer Detection in Dermoscopy Images based on Ensemble Lightweight Deep Learning Network* [39], proponen un modelo ligero de reconocimiento de cáncer de piel con discriminación de características basada en el principio de clasificación detallada, este modelo incluye dos módulos de extracción de características comunes de la red de clasificación de lesiones y una red de discriminación de características, de esta manera se construyó un modelo de segmentación semántica ligero del área de la lesión de una imagen dermoscópica, el cual puede alcanzar una alta precisión en la segmentación del área de la lesión de extremo a extremo sin complicados preprocesamientos de imágenes.

Qian Zhou y demás autores de *An Imbalanced R-STDP Learning Rule in Spiking Neural Networks for Medical Image Classification* [40] proponen una regla de aprendizaje para SNNs denominada “*imbalanced reward-modulation spike-timing-dependent plasticity (R-STDP)*” para resolver el problema del desbalance entre las clases de imágenes médicas, de esta manera se introduce un coeficiente de recompensa de desbalance para la R-STDP para compensar la clase más pequeña y que sea mayor a su contraparte, este coeficiente puede ayudar a establecer la compensación de las clases de acuerdo a la estadística de datos del set de entrenamiento. Adekanmi A. Adegun y Serestina Viriri en su artículo titulado *Deep Learning-Based System for Automatic Melanoma Detection* [41], proponen un método basado en el aprendizaje profundo que supera limitaciones debidas a las complejas características visuales para la detección y segmentación automática de lesiones de melanoma, basado en una red mejorada de codificador-decodificador con subredes de codificador y decodificador conectadas a través de una serie de rutas de salto que acerca el nivel semántico de los mapas de características del codificador al de los mapas de características del decodificador para un aprendizaje y extracción de características eficientes, este sistema emplea un enfoque de múltiples etapas y múltiples escalas y utiliza el clasificador Softmax para la clasificación de las lesiones de melanoma por píxeles. En *Melanoma Lesion Detection and Segmentation Using YOLOv4-DarkNet and Active Contour* [42] Saleh Albahli presenta un enfoque de detección y segmentación del melanoma que aporta una mejora significativa en términos de precisión frente a los enfoques más avanzados, inicialmente los artefactos como pelos, burbujas de gel y marcas clínicas se eliminan de las imágenes dermatoscópicas aplicando las operaciones morfológicas y las regiones de la imagen se definen; posteriormente, para la detección de la región infectada, utilizan el detector de objetos YOLOv4 ajustándolo para la detección de melanoma para discriminar las regiones infectadas y no infectadas altamente correlacionadas, una vez que obtienen los cuadros delimitadores contra las regiones de melanoma, las regiones de melanoma infectadas se extraen aplicando el enfoque de segmentación de contorno activa. En *Identification of Melanoma From Hyperspectral Pathology Image Using 3D Convolutional Networks* [43], Qian Wang y demás autores aplicaron las redes profundas de convolución en imágenes de patología hiperespectral para ejecutar la segmentación del melanoma; para hacer el mejor uso de las propiedades espectrales de los datos hiperespectrales tridimensionales, propusieron una red 3D convolucional llamada “Hyper-net” para segmentar melanomas desde imágenes de patología hiperespectral y, en orden de mejorar la sensibilidad del modelo, hicieron una modificación especial a la función de pérdida teniendo precaución de obtener falsos diagnósticos en el diagnóstico. Un conjunto de redes neuronales convolucionales mejoradas combinadas con una técnica de desplazamiento espaciada regularmente en el tiempo de prueba para la clasificación de lesiones cutáneas es propuesta por Karl Thurnhofer-Hemsi en *Skin Lesion Classification by Ensembles of Deep Convolutional Networks and Regularly Spaced Shifting* [44], donde la técnica de desplazamiento crea varias versiones de la imagen de entrada de prueba, que se desplazan mediante vectores de desplazamiento que se encuentran en una red regular en el plano de posibles desplazamientos, estas versiones desplazadas de la imagen de prueba se pasan

posteriormente a cada uno de los clasificadores de un conjunto y, finalmente todas las salidas de los clasificadores se combinan para producir el resultado final.

Como la segmentación de bordes de la lesión es un inconveniente por solucionar, Manu Goyal en *Skin Lesion Segmentation in Dermoscopic Images With Ensemble Deep Learning Methods* [45], proponen métodos de conjunto de aprendizaje profundo totalmente automatizado para lograr una alta sensibilidad y especificidad, con entrenamientos basados en Mask R-CNN y DeeplabV3+, usando como set de entrenamiento ISIC-2017 y evaluaron el funcionamiento del conjunto de redes con un dataset de prueba formado con imágenes de ISIC-2017 y PH2. En *Melanoma Detection Using an Objective System Based on Multiple Connected Neural Networks* [46], Loretta Ichim y Dan Popescu aseguran que la detección de melanomas realizada a simple vista por dermatólogos puede estar sujeta a errores, por lo tanto la implementación de dispositivos de procesamiento de imágenes con inteligencia artificial pueden actuar como un soporte para el profesional durante la examinación y la toma de decisiones, para esto proponen un nuevo tipo de sistema inteligente el cual está basado en varias redes neuronales conectadas en 2 niveles de clasificación, donde el primer nivel contiene 5 clasificadores (nivel subjetivo), mientras que en el segundo nivel (nivel objetivo) se usó un clasificador tipo perceptrón que decide si la lesión es in melanoma, basado en aprendizaje de pesos ajustados y las decisiones del primer nivel. Dado que al analizar imágenes dermoscópicas, los bellos corporales y sus sombras en la piel pueden obstruir información importante de la lesión al momento de un diagnóstico, en *Hair Segmentation and Removal in Dermoscopic Images Using Deep Learning* [47] Lidia Talavera-Martínez, Pedro Bibiloni y Manuel González-Hidalgo presentan un nuevo enfoque para la tarea de remoción de estos distractores de la imagen dermoscópica, basados en técnicas de aprendizaje profundo, donde el modelo propuesto consiste en una arquitectura tipo codificador-decodificador, con redes neuronales convolucionales, para la detección y posterior restauración de los píxeles del cabello a partir de las imágenes, además introducen una nueva función de pérdida combinada en la fase de entrenamiento de la red que combina la distancia L1, la pérdida de variación total y una función de pérdida basada en la métrica del índice de similitud estructural.

An Gong, Xinjie Yao y Wei Lin proponen en *Dermoscopy Image Classification Based on StyleGANs and Decision Fusion* [48] un método de fusión de decisiones, donde a través de transferencia de aprendizaje, basado en múltiples redes neuronales convolucionales pre-entrenadas, usaron el bloque para combinar múltiples CNNs y finalmente tomaron decisiones a través de múltiples bloques; el método de fusión de decisión puede resolver la capacidad de generalización de un modelo individual de una red neuronal convolucional, y es más estable y robusto que la estrategia tradicional de fusión, finalmente, basados en el dataset de ISIC 2019 usaron StyleGANs para generar imágenes de alta calidad, para aliviar el problema de una distribución menor y desigual del conjunto de datos de imágenes de dermatoscopia y mejorar el efecto de clasificación de las CNN. El estudio realizado por Tri-Cong Pham en *Improving Skin-Disease Classification Based on Customized Loss Function Combined With Balanced Mini-Batch Logic and Real-Time Image Augmentation* [49], propone un método híbrido para manejar el

existente desequilibrio entre las clases para clasificar las enfermedades de piel, este método combina el método de nivel de datos de lógica de mini lotes balanceados seguido de aumento de imagen en tiempo real con el método de nivel de algoritmo para diseñar una nueva función de pérdida. En *Short-Term Lesion Change Detection for Melanoma Screening With Novel Siamese Neural Network* [50] los autores afirman que la detección del cambio en una lesión se formula como una tarea que mide la similitud entre dos imágenes dermatoscópicas tomadas de una lesión en un período de tiempo corto, para esto se propone una red profunda basada en una novedosa estructura siamesa que producirá la decisión (cambió o sin cambios); además, bajo el marco siamés, se propone un proceso de regresión tensorial para extraer las características globales de las imágenes de las lesiones, conjuntamente con las características convolucionales profundas; por otro lado, para imitar el proceso de toma de decisiones de los médicos, se diseña una pérdida de segmentación (SegLoss) y se incorpora a la red propuesta como término de regularización, finalmente para evaluar el método propuesto, se estableció un conjunto de datos interno con 1000 pares de imágenes de lesiones tomadas en un período corto de tiempo en un centro clínico de melanoma. Homa Arab en *Early-Stage Detection of Melanoma Skin Cancer Using Contactless Millimeter-Wave Sensors* [51] propone un diseño de un radar de onda continua multitono de baja potencia a 77 GHz para aplicaciones de imágenes biomédicas, que se utilizará para la detección de cáncer de piel; la implementación de este sensor está basada en un diseño de circuito integrado de microondas híbrido en miniatura (MHMIC) de bajo costo, con esto el diseño de radar propuesto tiene la capacidad de detectar las propiedades dielécticas de los tejidos, lo que lo hace adecuado para la detección de melanomas con una precisión del orden de decenas de micrones. El framework propuesto por Lei Song en *An End-to-End Multi-Task Deep Learning Framework for Skin Lesion Analysis* [52] puede ejecutar tareas de detección, clasificación y segmentación de lesiones cutáneas de manera simultánea, para abordar el problema del desequilibrio que hay entre las clases en el dataset de imágenes médicas utilizado y, mientras tanto, mejorar el rendimiento de la segmentación, se propone una función de pérdida basada en la pérdida focal y la distancia de Jaccard; en ese orden de ideas, durante el entrenamiento, se empleó una estrategia de entrenamiento conjunta de tres fases para garantizar la eficiencia del aprendizaje de características.

En *Skin Melanoma Detection in Microscopic Images Using HMM-Based Asymmetric Analysis and Expectation Maximization* [53], Rozita Rastghalam propone un algoritmo de detección de melanoma basado en la fusión a nivel de decisión y un Modelo Oculto de Markov (HMM, por sus siglas en inglés), cuyos parámetros se optimizan mediante la Maximización de Expectativas (EM, por sus siglas en inglés) y el análisis asimétrico; la heterogeneidad de textura de las muestras se determina mediante análisis asimétrico, y además se presenta un clasificador HMM basado en fusión entrenado con EM, para esto una novedosa característica de textura es extraída basada en dos patrones binarios locales, denominada patrón de diferencia local (LDP, por sus siglas en inglés) y características de histograma estadístico de la imagen microscópica. El framework propuesto por Muhammad Attique Khan en *Multi-Class Skin Lesion Detection and Classification via Teledermatology* [54] consta de dos módulos, que incluyen la localización/segmentación de las lesiones cutáneas y la clasificación, para el módulo de localización, propone una estrategia híbrida

que fusiona las imágenes binarias generadas a partir del modelo de red neuronal convolucional de 16 capas diseñado y una segmentación de prominencia basada en la transformación de contraste de alta dimensión (HDCT, por sus siglas en inglés) mejorada, y para utilizar la máxima información extraída de las imágenes binarias, se propone un método de máxima información mutua, que devuelve la imagen de lesión RGB segmentada; mientras en el módulo de clasificación, se vuelve a entrenar un modelo DenseNet201 previamente entrenado en las imágenes de lesiones segmentadas mediante el aprendizaje por transferencia; posteriormente, las características extraídas de las dos capas completamente conectadas se muestrean hacia abajo utilizando el método de incrustación de vecinos estocásticos de distribución t (t-SNE), finalmente las características resultantes se fusionan utilizando un enfoque de correlación canónica múltiple (MCCA) y se pasan a un clasificador ELM de clases múltiples.

5 Metodología de Investigación Desarrollada

La metodología planteada para la correcta ejecución del proyecto se divide, como se muestra en la Figura 4, en dos partes: una referente a la forma como se evaluarán los resultados obtenidos, que será cualitativa y cuantitativamente; y la otra que se encarga de describir los aspectos clave o hitos que se tendrán en cuenta para el desarrollo del proyecto.

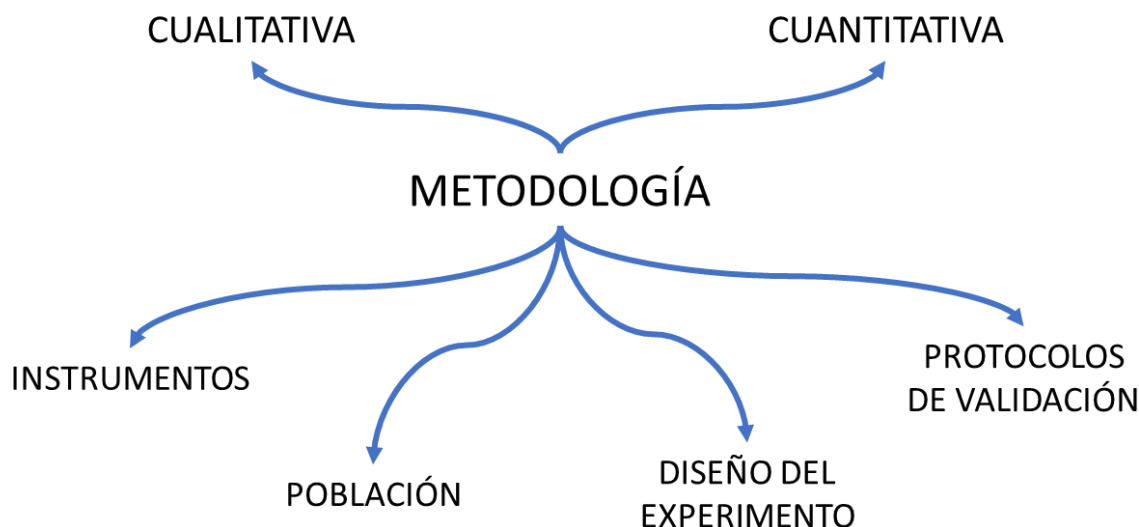


Figura 4. Desglose de la metodología propuesta.

Para este proyecto se propone desarrollar una metodología cualitativa en cuanto que se evaluarán características físicas en las imágenes o fotografías capturadas, es decir, un sistema que se basa en una técnica de observación e interpretación del aspecto que tiene la lesión, o lunar en la piel del sujeto. De igual manera manejará una metodología cuantitativa, ya que a partir de los resultados obtenidos se obtendrán estadísticas que ayudarán a soportar el concepto dado por el/los médicos que podrían usar esta herramienta y validar su correcto funcionamiento. Por lo expuesto anteriormente, se puede decir que también este es un proyecto del tipo experimental dado que se realizará una toma de datos, los cuales se analizarán a partir de una matriz de confusión para finalmente poder tener un concepto claro sobre el funcionamiento y efectividad del sistema.

Por otro lado, para poder desarrollar este proyecto exitosamente es necesario conocer de antemano trabajos de investigación sobre la detección de melanomas o que tengan alguna relación con el tema, de esta manera saber que herramientas y métodos han sido desarrollados y en qué sentido se puede mejorar tanto en el hardware como los algoritmos ya trabajados; los datos (bases de datos de imágenes) utilizados, la forma como fueron tratados y como se podría validar la efectividad del algoritmo que se desarrollará para solucionar los inconvenientes que se estén presentando actualmente en el diagnóstico de melanomas de piel. Para ello se deben tener en cuenta los siguientes ítems:

Instrumentos de la investigación: Para el correcto desarrollo y cumplimiento del proyecto descrito en este documento será necesario tener acceso principalmente a la literatura referente a herramientas de análisis como aprendizaje de máquina, deep learning y redes neuronales

convolucionales, como se menciona en el estado del arte, además de diferentes plataformas informáticas.

La literatura consultada para la realización de este proyecto debe ser referente al aprendizaje de máquina y reconocimiento de características en una imagen, permitiendo plantear la arquitectura adecuada que debe tener la red neuronal convolucional para resolver el problema en cuestión.

En cuanto a las plataformas informáticas, se trabajará con una base de datos que contenga una cantidad suficiente de imágenes o fotografías de lesiones de piel, o lunares para poder realizar las correspondientes pruebas al algoritmo desarrollado en Python, software con licencia de código abierto, además de Excel para realizar una recolección y análisis de los resultados para conocer la efectividad de la técnica propuesta.

Población a intervenir: Los sujetos que a largo plazo recibirán un beneficio directo del desarrollo de este proyecto serán todas aquellas personas que por condiciones naturales están predispuestas a la aparición de lunares u otro tipo de manchas en su piel, como es el caso de los individuos con piel clara conocida como fototipo I y II (Figura 5) que se queman con mucha facilidad, y tienen un riesgo de desarrollar cáncer de piel 10 veces mayor que las personas de raza negra (fototipo IV en adelante) [55] - [56]. Sin embargo, el impacto directo recaerá sobre los médicos generales, quienes en atención primaria podrán tener acceso a una herramienta con la cual realizar un método de tamizaje más preciso, evitando a futuro una biopsia innecesaria, que a su vez es una herramienta que facilitará el diagnóstico del médico funcionando como una segunda opinión para este último. Sin embargo, para iniciar y realizar las pruebas correspondientes se tendrá en cuenta como población muestra un data set que contenga imágenes de lesiones, o lunares en pacientes que tengan este tipo de lunares en la piel.

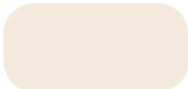
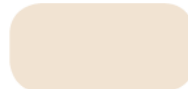
					
Fototipo I	Fototipo II	Fototipo III	Fototipo IV	Fototipo V	Fototipo VI
Siempre se quema y nunca se broncea.	Casi siempre se quema y a veces se broncea.	A veces se quema y generalmente se broncea.	Raro que se queme y siempre se broncea.	Se quema raramente. Bronceado muy intenso.	Casi nunca se quema.
Ojos y pelo claros. Piel muy blanca.	Ojos azules o pardos. Pelo rubio o pelirrojo. Piel blanca.	Tipo nórdico alpino.	Ojos y pelo oscuros. Piel blanca o un poco tostada.	Piel con un tono marrón.	Razas negras.

Figura 5. Fototipos de piel [55].

Diseño de la experimentación: Para la valoración de los resultados de este proyecto se tendrán en cuenta diferentes métricas que permitirán evaluar el funcionamiento de los modelos entrenados y, si estos son confiables. Dado que la técnica que se desarrolla en este proyecto está dirigida y finalmente será aplicada sobre el sujeto, por cuestiones bioéticas, llegado el momento de deberá realizar un documento de autorización o consentimiento informado para la recolección y tratamiento de fotografías de lunares de personas con este tipo de lesiones en la piel. Sin embargo, las pruebas para validar el correcto funcionamiento de la técnica propuesta se realizarán

inicialmente con base en un data set de fotografías de lunares, pudiendo acceder de forma directa a las imágenes y obtener resultados sin necesidad de intervenir directamente con una persona y solicitar su permiso. Con base en los resultados entregados por la clasificación realizada por cada una de las redes entrenadas, el uso de matrices de confusión y otras métricas como el AUC y F1-Score permitirán deducir cuál es la configuración de red neuronal que brinda los mejores resultados.

Protocolos de validación de la investigación: La valoración, del sistema inteligente para la detección de melanomas de piel, se llevará acabo de forma práctica mediante un data set de imágenes con diferentes casos previamente registrados, comparando su eficiencia con los resultados obtenidos respecto alguna de las otras técnicas registradas en el estado del arte, y de ser posible también contando con el criterio de un médico experto en dermatoscopia, con la experiencia en el análisis y diagnóstico en lesiones en la piel que pueda validar y apoyar los resultados obtenidos; médico que se espera pertenezca al Instituto Nacional de Cancerología o el Centro Dermatológico Federico Lleras Acosta de la ciudad de Bogotá, centros médicos donde se enviará una carta como la que se muestra en el anexo 1 solicitando colaboración de expertos y acceso a información que pueda ser de utilidad para el desarrollo del sistema propuesto como una base de datos, sí existe.

6 Arquitectura Para Sistema Detector De Melanomas

6.1 Arquitectura Propuesta

La arquitectura general manejada en cada modelo propuesto en este proyecto se encuentra conformada por una capa de entrada la cual se encuentra alimentada por la información de la imagen, esto quiere decir que en esta primera capa las neuronas están recibiendo la información correspondiente a las 3 dimensiones: ancho, alto y profundidad (RGB, por sus siglas en inglés *Red*, *Green* and *Blue*) [57]. Seguido de la entrada, la red consta de un número exacto de capas ocultas, todas ellas convolucionales junto con pooling y dropout apiladas entre sí, dedicadas a extraer la mayor cantidad de patrones posibles de una imagen y aprender de ellas, para finalmente ser usada en la salida una capa completamente conectada, donde la clasificación es dada de acuerdo con la función de activación, en este caso sigmoide, como se muestra en la Figura 6.

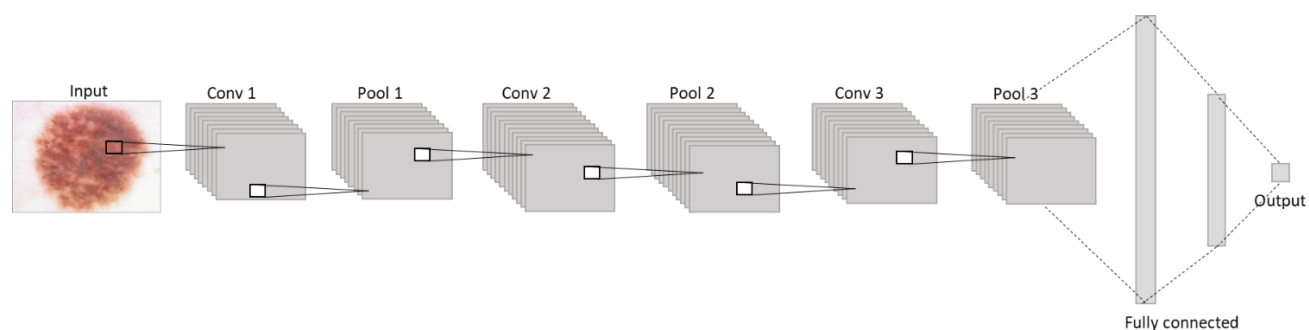


Figura 6. Arquitectura propuesta para la CNN

A la entrada de la red, esta cuenta con un set de imágenes, todas mezcladas benignas y malignas, tomadas de las fotos en la base de datos propuesta para cada fase desarrollada, donde una porción es aleatoriamente seleccionada para construir los conjuntos de imágenes de entrenamiento, validación y test, asegurando que no existan duplicados entre los diferentes sets. Estos sets deben ser pre-procesados para asegurar que todas las imágenes tengan siempre el mismo tamaño (alto y ancho). De esta manera, la CNN tendrá siempre en la entrada un set de imágenes cuyas dimensiones siempre serán entendidas como 3 dimensiones dadas por alto, ancho y la profundidad, donde la última hace referencia a los componentes RGB de la imagen, dimensiones que posteriormente se convertirán en un solo vector que alimentará la red.

Siguiendo la estructura mostrada en la Figura 6, luego de la entrada, la estructura de la CNN propuesta contiene las capas ocultas, las cuales son tres convolucionales y de pooling, 5 dropouts, una capa de flatten o aplanado y finalmente dos capas densas (dense) apiladas una seguida de la otra. De esta manera la red tiene un total de 16 capas las cuales fueron entrenadas con 33001 imágenes pertenecientes a la base de datos de ISIC [7], donde 30001 corresponden a casos benignos y 3000 son de casos melanomas. Las capas encargadas de extraer y aprender los patrones de las imágenes de entrada, principalmente lo hacen a través del uso de una operación lineal matemática entre matrices conocidas como convolución, el producto punto los pesos y una pequeña región en la entrada. Entonces el resultado obtenido es aplicado a una función de activación [58], [59], en este caso la función es Relu y es aplicada luego de cada capa de convolución.

Luego de cada capa convolucional viene una de pooling, principalmente con el fin de reducir dimensiones de la información en capas anteriores, es por esta razón que estas dos capas (convolución y pooling) comúnmente se encuentran una seguida de la otra, escogiendo el valor máximo (max pooling) en cada subregión para obtener la información más relevante que es la que pasará a la siguiente capa [59]. En este proyecto, 2x2 es el tamaño seleccionado para realizar el proceso de pooling.

Finalmente, el clasificador en la CNN este compuesto por una o más capas completamente conectadas, donde todas y cada una de las neuronas de la capa anterior están conectadas a cada una de las neuronas de la capa actual, la cual precede a la capa de salida [58], punto en el que la clasificación es dada por la función de activación. Este proyecto propone usar la función sigmoide para obtener en la salida valores de “1” o “0” para clasificar la lesión cutánea como benigna o maligna, según sea el caso.

Teniendo en cuenta como es el comportamiento general de cada componente de una CNN, será más sencillo comprender el funcionamiento de la estructura diseñada para este proyecto. Iniciando por la capa de entrada se encuentra contenida la información de las imágenes que van a ser procesadas por la red. Luego, esta información es enviada a través de 3 capas convolucionales con función de activación “Relu”, cada una seguida por una capa de pooling y un dropout luego de las dos primeras capas de pooling, obteniendo de esta manera la información más relevante del correspondiente grupo de imágenes (batch), finalmente se aplanan (flatten) la información en un vector 1D que es leído por una capa densa y la salida, donde la función de activación “sigmoide” está encargada de realizar la clasificación, construyendo de esta manera la red neuronal convolucional propuesta, como se muestra en la Figura 6.

6.2 Configuración Experimental

El conjunto de datos usado para alimentar la red diseñada fue encontrada en la página web de “The International Skin Imaging Collaboration” (ISIC) [7]. Este es un set de que inicialmente contaba con 23906 imágenes dermoscópicas de diferentes lesiones en la piel utilizadas en la fase de experimentos previos del proyecto y que actualmente contiene más de 53000 imágenes dermoscópicas, unas de ellas benignas (más de 47000 imágenes) y otras de melanomas (más de 5000). Teniendo en cuenta el limitado set de imágenes, fue necesario usar la mayoría de las imágenes del set para el entrenamiento de la red, de esta manera y dependiendo de otras variables el proyecto se dividió en 2 fases y en cada una se realizaron 5 diferentes experimentos.

La máquina usada para entrenar la red tiene las siguientes características: sistema operativo Linux Mint, procesador Intel core i7, 8 GB de Ram, 1TB HDD y una tarjeta de video NVIDIA GeForce GTX 1050 con 4 GB GDDR dedicada VRAM. Además, la arquitectura propuesta fue desarrollada usando Keras y Tensorflow, ambas librerías de fuente abierta dedicadas a procesos de aprendizaje de máquina.

Teniendo en cuenta lo anterior, el proyecto consta de 2 fases de experimentos los cuales se configuraron como sigue:

6.2.1 Fase 1: Experimentos previos

Durante esta fase de experimentos previos, se llevaron a cabo entrenamientos de diferentes modelos de la red propuesta, basados principalmente en la diferencia entre clases del dataset. En esta fase los entrenamientos realizados estuvieron enfocados mayormente en la cantidad de imágenes y la distribución de estas para los datasets de entrenamiento, validación y test. De esta manera los experimentos previos son 5 y se configuran teniendo en cuenta los valores mostrados en la Tabla 2.

Tabla 2. Rangos de los valores utilizados en las configuraciones aplicadas para los diferentes entrenamientos de los experimentos de la fase 1

Épocas	5 - 20
Tamaño de la imagen	100x100 y 200x200
Tamaño del batch	10 - 100
Pasos en Entrenamiento	100 - 1650
Pasos en Validación	10 - 1000
Número de capas convolucionales	1 - 9
Tasa de aprendizaje	1×10^{-6} - 1×10^{-2}
Función de Activación	Relu, sigmoide y softmax
Dropout	0.03 - 0.7
Peso para la clase 0 (Exp 4 y 5)	0.55891461 y 0.54999833
Peso para la clase 1 (Exp 4 y 5)	4.74342928 y 5.50016666

En los primeros 4 experimentos se usa el dataset obtenido de ISIC en el 2019 compuesto por un total de 23906 imágenes de lesiones de piel tipo melanoma y no melanoma. En el primer experimento las imágenes se encuentran distribuidas 70% para el entrenamiento (13562 de lesiones benignas y 1598 de lesiones malignas), 30% para validación (5781 lesiones benignas y 654 lesiones malignas) y 60 fotos (distribuidas 30-30) aleatoriamente seleccionadas y destinadas para testear el funcionamiento de las redes entrenadas; debido a la notoria disparidad en la cantidad de imágenes entre una clase y la otra, con un total de imágenes de lesiones benignas de aproximadamente el 90% del dataset, para el experimento 2 se decidió organizar un set de imágenes para entrenamiento con 1500 fotos de melanomas y 1500 de lesiones benignas, para validación 500 fotos de lesiones benignas y 500 de lesiones malignas, y de la misma manera para set de test se apartaron 282 de lesiones benignas y 282 de lesiones malignas, asegurando de esta manera una cantidad de información más equilibrada en el entrenamiento y así mismo poder tener más imágenes para realizar el test; el tercer experimento utilizó la misma distribución de las imágenes de entrenamiento que en el experimento 1, sin embargo, en este caso se usaron todas las imágenes para entrenamiento y validación, y se evalúan los resultados con el mismo set de validación, además en este caso se utilizó únicamente Sigmoide como la función de activación

tanto en las capas convolucionales como en la salida; siguiendo con la misma distribución del experimento anterior, los modelos del experimento 4 se entrenaron con la diferencia que la función de activación Sigmoide solo se utilizó a la salida mientras Relu fue implementada en las demás capas. Finalmente, el quinto experimento cuenta con la diferencia que el dataset de imágenes era más grande, lo que permite a la red tener una muestra mayor de donde obtener información acerca de las características que diferencian a un melanoma de un lunar común. Para este caso el data set cuenta con más de 53000 imágenes dermoscópicas, las cuales se distribuyeron de la siguiente manera: set de entrenamiento de 33001 imágenes (aprox. 61% del dataset) y 11500 para el set de validación (aprox. 21%), manteniendo algunas otras para evaluar el comportamiento de la CNN. Estas imágenes deben pasar por un preprocesamiento para estandarizar el tamaño a 200x200, asegurando que todas las imágenes en la entrada de la red tengan siempre el mismo tamaño, pudiendo así entrenar la red neuronal a través de la extracción de patrones con 3 (Figura 7) y 4 capas de convolución.

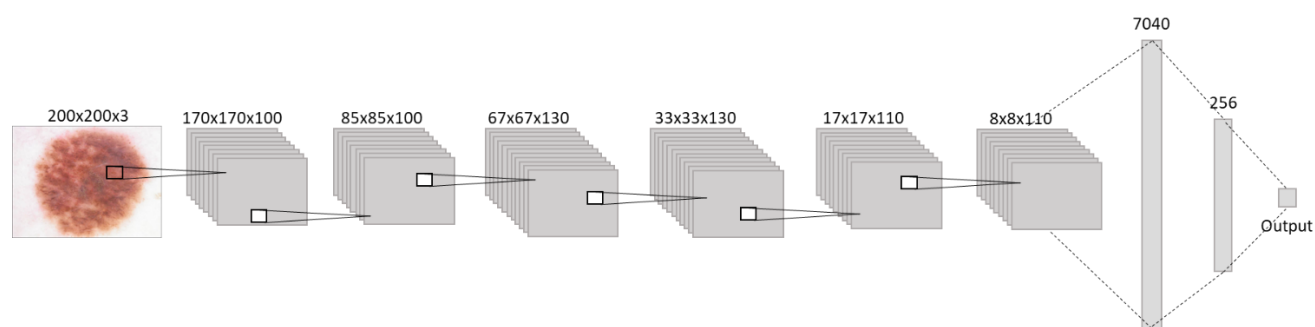


Figura 7. Arquitectura propuesta para el experimento 5

6.2.2 Fase 2

Para esta fase se igualó la cantidad de imágenes entre clases, con el fin de obtener un set de datos balanceado. De esta manera, el número total de imágenes resultante es 79015, y teniendo en cuenta una distribución 70/30, el dataset de entrenamiento queda con un total de 60015 imágenes, el de validación con 19000 y para test quedan 8809. Para esta fase se realizaron de igual manera 5 experimentos diferentes con 3 modelos entrenados en cada caso, en los cuales se variaron principalmente características de la red como lo son los tamaños de los filtros, el dropout, entre otros y respecto a la fase anterior el tamaño de las imágenes que en este caso son de 100x100.

6.2.2.1 Experimento 1

Para este caso, y en adelante, todos los modelos entrenados usaron el mismo dataset de imágenes y la misma distribución de las clases. Sin embargo, los parámetros de configuración de la red son los que cambian, y para este caso específicamente la configuración quedó como se observa en la Tabla 3.

Tabla 3. Resumen de las configuraciones aplicadas para los 3 entrenamientos del experimento 1

Épocas	10
Tamaño de la imagen	100x100

Tamaño del batch	100
Pasos en Entrenamiento	600
Pasos en Validación	190
Número de capas convolucionales	3
Tasa de aprendizaje	1×10^{-5}
Conv1	100,7,7
Conv2	130,7,7
Conv3	110,3,3
Función de Activación	Relu
	Sigmoide en capa de salida
Dropout	0.2, 0.3, 0.4

6.2.2.2 Experimento 2

En este caso, los modelos entrenados se diferencian de los anteriores por la configuración de las capas convolucionales utilizada, ya que la variación del dropout en cada caso será consistente con el experimento anterior mientras el resto de los parámetros se mantienen intactos, como se muestra en la Tabla 4.

Tabla 4. Resumen de las configuraciones aplicadas para los 3 entrenamientos del experimento 2

Épocas	10
Tamaño de la imagen	100x100
Tamaño del batch	100
Pasos en Entrenamiento	600
Pasos en Validación	190
Número de capas convolucionales	3
Tasa de aprendizaje	1×10^{-5}
Conv1	70,7,7
Conv2	140,7,7
Conv3	100,5,5
Función de Activación	Relu
	Sigmoide en capa de salida
Dropout	0.2, 0.3, 0.4

6.2.2.3 Experimento 3

En el tercer experimento se configuraron capas convolucionales más profundas respecto a los casos anteriores como se ve en la Tabla 5, e igual que en el caso anterior se mantuvo la configuración de los demás parámetros, incluyendo la variación del dropout para los 3 modelos entrenados.

Tabla 5. Resumen de las configuraciones aplicadas para los 3 entrenamientos del experimento 3

Épocas	10
Tamaño de la imagen	100x100
Tamaño del batch	100
Pasos en Entrenamiento	600
Pasos en Validación	190
Número de capas convolucionales	3
Tasa de aprendizaje	1×10^{-5}
Conv1	200,7,7
Conv2	150,7,7
Conv3	100,3,3
Función de Activación	Relu
	Sigmoide en capa de salida
Dropout	0.2, 0.3, 0.4

6.2.2.4 Experimento 4

Igual que en el experimento 3, el cambio base en la configuración de la red estuvo en el tamaño de los kernels, en este caso fueron más grandes (Tabla 6) en la búsqueda de aprender características más grandes y luego reducir rápidamente las dimensiones espaciales.

Tabla 6. Resumen de las configuraciones aplicadas para los 3 entrenamientos del experimento 4

Épocas	10
Tamaño de la imagen	100x100
Tamaño del batch	100
Pasos en Entrenamiento	600
Pasos en Validación	190

Número de capas convolucionales	3
Tasa de aprendizaje	1×10^{-5}
Conv1	200,9,9
Conv2	150,9,9
Conv3	100,7,7
Función de Activación	Relu
	Sigmoide en capa de salida
Dropout	0.2, 0.3, 0.4

6.2.2.5 Experimento 5

Basado en los resultados de los entrenamientos anteriores, este experimento se diseñó a partir de los modelos con mayor porcentaje de “val_acc” encontrados en los experimentos anteriores y a diferencia de estos, en este caso se realizaron los entrenamientos con cambiando el tamaño del pooling, como se muestra en la Tabla 7.

Tabla 7. Resumen de las configuraciones aplicadas para los 3 entrenamientos del experimento 5

Épocas	10
Tamaño de la imagen	100x100
Tamaño del batch	100
Pasos en Entrenamiento	600
Pasos en Validación	190
Número de capas convolucionales	3
Tasa de aprendizaje	1×10^{-5}
Conv1	100,7,7 - 200,7,7 - 200,9,9
Pool1	3x3 – 3x3 – 3x3
Conv2	130,7,7 - 150,7,7 - 150,9,9
Pool2	3x3 - 3x3 – 2x2
Conv3	110,3,3 - 100,7,7 - 100,7,7
Pool3	3x3 – 2x2 – 2x2
Función de Activación	Relu
	Sigmoide en capa de salida
Dropout	0.3,0.2, 0.3

7 Proceso de Evaluación (Resultados)

7.1 Fase 1: Experimentos previos

7.1.1.1 Experimento 1

Este caso se evaluó mediante el uso de la matriz de confusión y principalmente respecto de la precisión (Accuracy, Acc) de la red al momento de realizar la clasificación. Para medir el éxito de las redes entrenadas se tuvo un set de prueba con 60 imágenes, de las cuales 30 eran de casos benignos y las otras 30 de casos malignos, esto con el fin de tener un set de imágenes equilibrado, sin embargo, los resultados obtenidos no fueron 100% confiables debido al limitado número de imágenes en el set de pruebas, lo cual dejó como resultado un Acc máximo de 55% en los aciertos, los cuales en su mayoría eran casos benignos y clasificando más del 80% de los casos malignos como falsos benignos, o viceversa.

Para poder tener un mejor resultado del funcionamiento de las redes entrenadas, se realizó la prueba con el mismo set de validación obteniendo un porcentaje de aciertos de hasta 89%, sin embargo, este porcentaje de aciertos es debido a la inmensa disparidad entre clases.

Teniendo en cuenta los resultados obtenidos se puede deducir que el porcentaje de aciertos obtenido está relacionado directamente con la diferencia de imágenes que hay entre las clases, lo cual lleva a concluir que no es una red confiable ya que la cantidad de imágenes de melanomas es aproximadamente el 10% del total que se usaron para evaluar los modelos entrenados.

7.1.1.2 Experimento 2

Al igual que en el caso anterior, el funcionamiento de las redes entrenadas bajo los parámetros presentados en la sección 6.2.1, se evaluó mediante una matriz de confusión teniendo como base un set de 564 imágenes distribuidas 50/50 y seleccionadas de manera aleatoria, de esta manera se aseguró un set de pruebas más grande. Siendo así, las redes entrenadas para este experimento, se obtuvo un porcentaje de aciertos aproximado de 65.6%. Posteriormente, igual que con el experimento 1, se realizó la prueba con el set de validación que en este caso tiene un total de 1000 imágenes distribuidas 50/50, lo que entregó un Acc de hasta 73.6%. Finalmente, con base en los resultados de las pruebas, se encontró que se redujo en gran medida el número de falsos benignos, sin embargo, el número de casos mal clasificados sigue siendo bastante grande como se puede observar en la matriz de confusión de la Tabla 8.

Tabla 8. Matriz de confusión del modelo entrenado con mejor respuesta (Modelo 10).

Modelo 10	Maligno	Benigno	Real
Maligno	425	75	500
Benigno	189	311	500
Acc	73.60%		1000

7.1.1.3 Experimento 3

Este experimento a diferencia de los anteriores usó una sola función de activación en toda la red neuronal sin importar como varían los demás parámetros, sin embargo, los resultados obtenidos fueron completamente negativos, ya que se observa en la Figura 8 la red no aprendió lo necesario o nada durante el entrenamiento y a consecuencia de esto las pruebas mostraron resultados donde siempre se clasificaba 100% de los casos benignos.

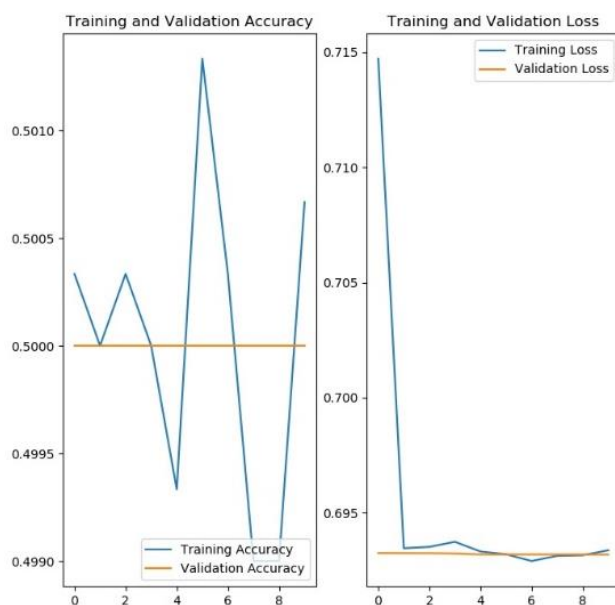


Figura 8. Entrenamiento del modelo 104. La gráfica de la izquierda muestra la curva de aprendizaje del modelo, mientras que la segunda cómo la pérdida decrece con cada época.

7.1.1.4 Experimento 4

En este caso para combatir las limitaciones del dataset se añadieron nuevos métodos para evaluar el funcionamiento de los modelos. El Área Bajo la Curva (AUC por sus siglas en inglés) y F1-Score cuyos resultados no se ven directamente afectados por diferencia en la cantidad de imágenes entre clases, afianzando los resultados obtenidos como se muestran en la Tabla 9 de la matriz de confusión correspondiente al modelo 401.

Tabla 9. Matriz de confusión del modelo entrenado con mejor respuesta (Modelo 401).

Modelo 401	Maligno	Benigno	Real	Precisión	Exhaustividad
Maligno	502	152	654	0.209	0.768
Benigno	1895	3886	5781	0.962	0.672
Acc	68.19%		6435	0.586	0.720
AUC	0.72				
F1-Score	0.646				

7.1.1.5 Experimento 5

Basados en los experimentos realizados anteriormente con las redes entrenadas y los parámetros asignados a este, las redes se evaluaron, al igual que en el experimento 4, no solo con base en la matriz de confusión y el Acc sino también con el F1-Score y el AUC.

La arquitectura propuesta fue entrenada durante 10 épocas, y con un set de imágenes más grande al usado previamente, se obtuvo una gráfica de aprendizaje más suave como se observa en la Figura 9.

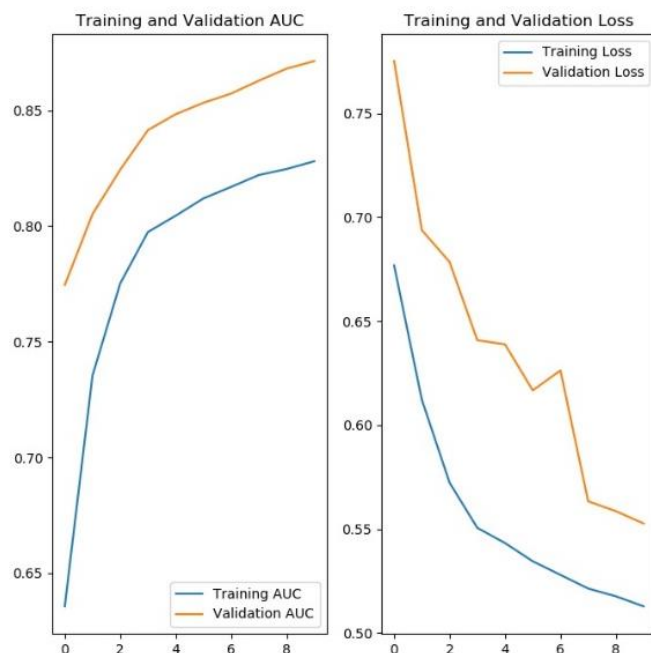


Figura 9. Entrenamiento del modelo 706. La gráfica de la izquierda muestra la curva de aprendizaje del modelo, mientras que la segunda cómo la pérdida decrece con cada época.

Evaluando el funcionamiento de los modelos entrenados con las métricas F1-Score y el AUC, usando el dataset de prueba de 8809 imágenes, se encontró que el modelo con mayor porcentaje de exactitud obtuvo un 85.46%, cómo se muestra en la siguiente Tabla 10.

Tabla 10. Matriz de confusión del modelo entrenado con mejor respuesta (Modelo 706).

Modelo 706	Maligno	Benigno	Real	Precisión	Exhaustividad
Maligno	828	298	1126	0.457	0.735
Benigno	983	6700	7683	0.957	0.872
Acc	85.46%		8809	0.707	0.804
AUC	0.834				
F1-Score	0.752				

7.2 Fase 2

7.2.1.1 Experimento 1

Con base en las configuraciones expuestas en el numeral 6.2.2.1, los modelos entrenados fueron evaluados bajo el uso de una matriz de confusión y el F1-score con ayuda del set de prueba, teniendo en cuenta que el clasificador entrega como resultado 1 o 0 a la salida igualmente interpretado melanoma o no melanoma respectivamente. Los resultados obtenidos para los 3 modelos entrenados durante este experimento se exponen en la Tabla 11.

Tabla 11. Matrices de confusión de los 3 modelos del experimento 1. El modelo 1 que fue entrenado con un dropout de 0.2, el modelo 2 con 0.3 y el modelo 3 con 0.4

Modelo 1		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	661	465	0,938	87,97%	0,923	0,587
7683	BEN	595	7088				
		VPP	VPN	Exactitud	lr+	lr-	f1
		93,84%	52,63%	87,97%	2,234	0,132	0,930

Modelo 2		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	787	339	0,952	84,90%	0,871	0,699
7683	BEN	991	6692				
		VPP	VPN	Exactitud	lr+	lr-	f1
		95,18%	44,26%	84,90%	2,893	0,185	0,910

Modelo 3		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	748	378	0,946	84,35%	0,870	0,664
7683	BEN	1001	6682				
		VPP	VPN	Exactitud	lr+	lr-	f1
		94,65%	42,77%	84,35%	2,591	0,196	0,906

En las tablas de la Tabla 11, se muestra como en cada modelo los resultados son menos variantes a diferencia de la fase 1, con valores de exactitud entre 84% y 88% y valores de exhaustividad de hasta 92% que son superiores a la sensibilidad que se estima posee la clasificación realizada por un médico general durante el tamizaje inicial.

7.2.1.2 Experimento 2

En este experimento configurado como se expone en el numeral 6.2.2.2, los modelos entrenados se pusieron a prueba con el set de prueba de 8809 imágenes y su funcionamiento se evaluó mediante el uso de matrices de confusión (Tabla 12), incluyendo la exactitud de la predicción que comparando con el experimento anterior los resultados son aparentemente superiores.

Tabla 12. Matrices de confusión de los 3 modelos del experimento 2. El modelo 4 que fue entrenado con un dropout de 0.4, el modelo 5 con 0.3 y el modelo 6 con 0.2

Modelo 4		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	483	643	0,919	88,66%	0,954	0,429
7683	BEN	356	7327				
		VPP	VPN	Exactitud	lr+	lr-	f1
		91,93%	57,57%	88,66%	1,670	0,108	0,936

Modelo 5		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	402	724	0,911	88,75%	0,965	0,357
7683	BEN	267	7416				
		VPP	VPN	Exactitud	lr+	lr-	f1
		91,11%	60,09%	88,75%	1,501	0,097	0,937

Modelo 6		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	568	558	0,928	87,91%	0,934	0,504
7683	BEN	507	7176				
		VPP	VPN	Exactitud	lr+	lr-	f1
		92,79%	52,84%	87,91%	1,885	0,131	0,931

Aparentemente los resultados obtenidos durante las pruebas de estos modelos son superiores al experimento anterior, dado que los valores de exhaustividad obtenidos son mayores, sin embargo, la especificidad solo en uno de los casos alcanza a superar el 50% y, esto implica que el modelo es capaz de acertar aproximadamente a la mitad de los casos malignos clasificándolos como lo que son realmente.

7.2.1.3 Experimento 3

Siguiendo la metodología aplicada anteriormente, los resultados obtenidos (Tabla 13) muestran una exactitud menor, aunque son más altas la *precisión* que se traduce como la probabilidad de que una predicción benigna lo sea realmente y la *especificidad* que se refiere a la probabilidad de que los resultados negativos realmente son lesiones malignas, en este caso el modelo con mayor especificidad es el 7 aunque este valor le cueste fuerza a la sensibilidad.

Tabla 13. Matrices de confusión de los 3 modelos del experimento 3. El modelo 7 que fue entrenado con un dropout de 0.2, el modelo 8 con 0.3 y el modelo 9 con 0.4

Modelo 7		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	793	333	0,952	84,36%	0,864	0,704
7683	BEN	1045	6638				
		VPP	VPN	Exactitud	lr+	lr-	f1
		95,22%	43,14%	84,36%	2,921	0,193	0,906

Modelo 8		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	622	504	0,933	87,22%	0,919	0,552
7683	BEN	622	7061				

VPP	VPN	Exactitud	lr+	lr-	f1
93,34%	50,00%	87,22%	2,053	0,147	0,926

Modelo 9		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	593	533	0,930	87,30%	0,924	0,527
7683	BEN	586	7097				
		VPP	VPN	Exactitud	lr+	lr-	f1
		93,01%	50,30%	87,30%	1,951	0,145	0,927

7.2.1.4 Experimento 4

En la Tabla 14, al igual que en los experimentos anteriores, además de los valores previamente mencionados, se tuvo en cuenta también la razón de verosimilitud (likelihood ratio, en inglés) como método para evaluar la viabilidad de los modelos entrenados y de esta manera obtener la razón entre la posibilidad de observar un resultado positivo en casos de lesiones benignas versus la posibilidad de obtener el mismo resultado en un melanoma.

Tabla 14. Matrices de confusión de los 3 modelos del experimento 4. El modelo 10 que fue entrenado con un dropout de 0.4, el modelo 11 con 0.3 y el modelo 12 con 0.2

Modelo 10		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	663	463	0,938	86,58%	0,906	0,589
7683	BEN	719	6964				
		VPP	VPN	Exactitud	lr+	lr-	f1
		93,77%	47,97%	86,58%	2,204	0,159	0,922

Modelo 11		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	725	401	0,945	86,41%	0,896	0,644
7683	BEN	796	6887				
		VPP	VPN	Exactitud	lr+	lr-	f1
		94,50%	47,67%	86,41%	2,517	0,161	0,920

Modelo 12		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	712	414	0,944	87,82%	0,914	0,632
7683	BEN	659	7024				
		VPP	VPN	Exactitud	lr+	lr-	f1
		94,43%	51,93%	87,82%	2,487	0,136	0,929

Siguiendo las comparaciones que se venían haciendo previamente se puede decir que el modelo con mejores resultados hasta ahora es el 12, esto con base en que posee una exhaustividad superior al 90% y así mismo su especificidad logra un porcentaje de 63.2%. Por otro lado, los valores obtenidos para lr+ y lr- aunque no son los mejores comparados con otros obtenidos hasta ahora, se puede rescatar que ambos son factores que refuerzan los resultados y en términos médicos ayudan al profesional a dar una mayor seguridad al resultado obtenido en el tamizaje.

7.2.1.5 Experimento 5

En la Tabla 15, los resultados de los modelos del experimento no solo muestran porcentajes de exactitud superiores al 85%, también valores de f1-score superiores a 0,9, valores que hacen referencia al rendimiento combinado entre precisión y sensibilidad.

Tabla 15. Matrices de confusión de los 3 modelos del experimento 5. En los modelos 13, 14 y 15 la diferencia respecto de los casos anteriores se encuentra en el pooling utilizado, este fue inicialmente seleccionado de 3x3 en todas las capas, sin embargo, debido a las limitaciones del dispositivo utilizado se combinó con pooling de 2x2

Modelo 13		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	459	667	0,916	87,97%	0,949	0,408
7683	BEN	393	7290				
		VPP	VPN	Exactitud	lr+	lr-	f1
		91,62%	53,87%	87,97%	1,602	0,125	0,932

Modelo 14		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	500	626	0,921	88,05%	0,944	0,444
7683	BEN	427	7256				
		VPP	VPN	Exactitud	lr+	lr-	f1
		92,06%	53,94%	88,05%	1,699	0,125	0,932

Modelo 15		MAL	BEN	Precisión	Exactitud	Exhaustividad	Especificidad
1126	MAL	732	394	0,946	86,05%	0,891	0,650
7683	BEN	835	6848				
		VPP	VPN	Exactitud	lr+	lr-	f1
		94,56%	46,71%	86,05%	2,547	0,167	0,918

8 Discusión

Respecto de los resultados obtenidos en cada uno de los experimentos realizados en este proyecto, comparando la fase 1 de experimentos previos con la fase 2, se hace notorio el errático comportamiento y la inestabilidad de los resultados al jugar con la distribución del dataset de entrenamiento para intentar controlar la inequidad entre las clases. Si se recogen y se analizan los datos de los mejores resultados de los modelos entrenados en cada experimento de la fase 1, Tabla 16, se puede observar que el modelo 33, del experimento 1, fue testeado con el set de validación con un total de 6435 imágenes, debido al limitado número de imágenes tuvo un porcentaje de Acc de 81.23% y un valor calculado de F1 de 0.59; el modelo 10, del experimento 2, diferente al experimento 1 se usó un set de validación más reducido para evaluar el funcionamiento con tan solo 1000 imágenes, ya que para este caso se decidió tener el mismo número de imágenes de melanomas y lesiones benignas tanto para entrenamiento como para el set de validación, obteniendo un Acc de 73.60% y un F1 de 0.74; en el modelo 104, del experimento 3, a diferencia de los demás experimentos se usó como función de activación “sigmoide”, dando como resultado en todos los casos un Acc de 89.84% equivalente a la totalidad de los casos benignos, esto sucedió ya que el modelo no aprendió a diferenciar y todos los casos los clasificó como benignos; para el modelo 401, del experimento 4, a diferencia del modelo 33 en la capa de salida se utilizó la función de activación sigmoide en lugar de softmax lo que dio como resultado un porcentaje de Acc menor 68.19%, sin embargo, se obtuvo un F1 de 0.65 que es mayor y asimismo indica una mayor confiabilidad de la predicción puesto que esta medida da resultados más precisos cuando se manejan clases desbalanceadas como es el caso, además para este caso se calculó el área bajo la curva AUC que dio un valor de 0.72 reafirmando la confiabilidad de la predicción del modelo; finalmente el modelo 706, del experimento 5, se tuvo un set e imágenes mayor a los casos anteriores y asimismo un set de pruebas más grande con un total de 8809 imágenes con el que se obtuvieron un Acc de 85.46%, F1 de 0.75 y AUC de 0.83, siendo estos en general los resultados más altos y que indican que se trata del modelo con predicción más exacta y confiable.

Tabla 16. Resumen de los mejores modelos entrenados en cada experimento.

Experimento	Modelo	Set de imágenes utilizado (Malignos / Benignos)	Función de activación	ACC	F1	AUC
1	33	Validación (654 / 5781)	Relu	81.23%	0.59	-
2	10	Validación (500 / 500)	Relu	73.60%	0.74	-
3	104	Validación (654 / 5781)	Sigmoide	89.84%	-	-
4	401	Validación (654 / 5781)	Relu	68.19%	0.65	0.72
5	706	Pruebas (1126 / 7683)	Relu	85.46%	0.75	0.83

De lo anterior, se puede deducir que algunos de los factores claves, para entrenar una CNN con mayor precisión en su predicción, son la cantidad de imágenes que se tienen para entrenar el modelo, y la correcta implementación de las funciones de activación para extraer y clasificar la información de dichas imágenes. Además, el aporte del uso de los diferentes métodos de evaluación utilizados para analizar el desempeño de las redes entrenadas, sobre todo en casos como

este donde la cantidad de casos de lesiones benignas es tan alta comparada con la cantidad de casos de melanoma que se tienen tanto para entrenamiento, validación y test.

Por otro lado, en la fase 2 a diferencia de la fase 1, el dataset que se utilizó para el entrenamiento fue modificado de manera que se tuvo un set de datos con una distribución 50-50, asegurando el mismo número de imágenes de lesiones benignas como de melanomas, y de esta manera fue posible mantener una cantidad considerable de imágenes en el set de test para poder realizar una prueba con resultados suficientes para poderlos analizar. De esta manera los experimentos diseñados fueron más claros y los resultados obtenidos más estables como se observó en la sección 7.2. Sin embargo, luego de analizar los resultados obtenidos para cada modelo, de los 15 entrenamientos realizados, en la Tabla 17 se muestra el resumen de algunos casos seleccionados, donde resaltaron 3 debido a los valores obtenidos respecto a la razón de verosimilitud ($lr+$ y $lr-$) puesto que la variación de estos es más notoria que en las demás métricas. Ese es el caso del modelo 5 con el $lr-$ más bajo con valor 0,097 considerado un valor altamente relevante y así mismo también el $lr+$ más bajo lo que implica una utilidad mala; en el otro extremo, el modelo 7 con un $lr+$ igual a 2,92 que da una utilidad regular y un $lr-$ igual a 0,19 indicando una utilidad buena; y finalmente un punto medio es el modelo 12 donde un $lr+$ de 2,48 y un $lr-$ de 0,13 proporcionan una mejor idea de la calidad en la predicción del modelo, frente a los otros modelos.

Tabla 17. Resumen de los mejores modelos entrenados en cada experimento de la fase 2

Experimento	Modelo	ACC	F1
1	1	87.97%	0.930
2	5	88.75%	0.937
3	7	84.36%	0.906
4	12	87.82%	0.929
5	14	88.05%	0.932

9 Conclusiones

Durante todo el desarrollo del proyecto, desde su propuesta hasta la finalización de los experimentos, se probaron diferentes arquitecturas de redes convolucionales, unas más profundas, otras con filtros más grandes o pequeños, con diferentes funciones de activación, unas con más imágenes que otras, entre otras combinaciones de los diferentes parámetros que les dan forma a las redes neuronales convolucionales. Cada una de estas combinaciones que se diseñaron, tuvieron tiempos de alrededor de 2 días de procesamiento en los que se realizaba el entrenamiento y validación del aprendizaje de la red bajo la modalidad de “aprendizaje supervisado” basado en 2 tipos de lesiones que, benignas o malignas (melanomas), ya vienen por defecto clasificadas en el dataset utilizado. Basados en los porcentajes aproximados de sensibilidad y especificidad atribuidos a la clasificación realizada por un especialista con experiencia con ayuda de un dermatoscopio, la red neuronal entrenada (modelo 12 del experimento 4 de la fase 2) mostró tener una mayor sensibilidad en la predicción de casos dando un porcentaje poco más de 2 puntos porcentuales más alto, siendo esto un 91.40%, mientras que la especificidad solo llegó a 63.2% en el modelo seleccionado, sin embargo, además de esto las otras métricas utilizadas para evaluar el funcionamiento de la red dieron valores que ayudan a soportar la confiabilidad de la predicción realizada, como es el caso del F1-score con un valor de 0.929, el $lr+$ 2.487 y $lr-$ 0.136.

Teniendo en cuenta los experimentos realizados y viendo los resultados obtenidos, es preciso afirmar que la mejor combinación de funciones de activación, para extraer los parámetros de las imágenes y poder obtener la mejor predicción, es usar Relu en las capas convolucionales para obtener una mejor y más rápida extracción de información que le permita tener un aprendizaje más rápido, y softmax a la salida para así obtener la probabilidad que tiene cada caso de pertenecer a una clase o la otra.

Con base en la alta similitud entre los casos que son y los que no son melanoma, el contraste de las lesiones con la piel que las rodea, la calidad de los colores en las imágenes, que no es posible evaluar el tamaño y evolución de las lesiones en la piel, las pruebas realizadas y los resultados obtenidos tanto en la fase 1 como la fase 2 del proyecto, se puede afirmar que la cantidad de imágenes disponibles, la distribución designada para realizar el entrenamiento y test de la red juega un papel fundamental para obtener predicciones más precisas y estables, reduciendo la cantidad de falsos positivos y falsos negativos.

Por otro lado, comparando con otros proyectos se observó que, fuera de la limitada cantidad de data que se tiene para poder entrenar la red, las características de la máquina utilizada, en términos de procesamiento, no permiten realizar entrenamientos de redes muy profundas dado que el consumo de máquina y tiempo son bastante altos y en caso de querer entrenar redes más profundas probablemente la máquina no va a poder responder y el entrenamiento no podrá finalizar, o con un mayor número de imágenes los tiempos de duración del entrenamiento van a ser mayores. Además, por esta misma razón otros parámetros, como es el caso del número de épocas utilizado para entrenar la red o el tamaño de las imágenes, se podrían ver reducidos con el fin de conseguir una mejor respuesta del equipo, lo que finalmente podría causar que se realice un entrenamiento incompleto y/o una posible pérdida en la calidad de los resultados por disminución en la calidad en la data de entrenamiento.

Finalmente, se observó que la cantidad de parámetros que se deben manejar para el diseño y entrenamiento de las redes es bastante alta, es por esta razón que se vio la necesidad de realizar un par de ajustes como igualar las clases y un preprocesamiento a las imágenes utilizadas para entrenar los modelos, este procesamiento consistió en estandarizar el tamaño con el fin de evitar que los diferentes tamaños de las imágenes se pudieran convertir en un impedimento para el entrenamiento, debido a que dicha gran variedad de tamaños se convertiría en un parámetro más a tener en cuenta por la red durante el entrenamiento y podría entonces desviar la atención de otras características que dan más información a la CNN sobre las diferencias que permitirán identificar a los melanomas.

10 Trabajos Futuros

Con base en los resultados obtenidos con este proyecto y comparándolos con los conseguidos con otras arquitecturas descritas en el estado del arte, se propone a futuro un mayor preprocesamiento de las imágenes, como por ejemplo la eliminación de artefactos, brillo y la detección del contorno para de esta manera eliminar información que pueda estar distrayendo la atención de la red, además de rotar algunos grados o invertir la dirección de las imágenes de manera aleatoria para tener un set de datos de entrenamiento más grande.

Por otro lado, se propone analizar la metadata de las imágenes de la base de datos para usarla como información extra para darle mayor confiabilidad a los resultados, ya que existen otros factores que para los especialistas son de gran ayuda para determinar la clase de lesión frente a la que se encuentran partiendo de las probabilidades según el grupo en el que se encuentre el paciente, teniendo como base registros de casos detectados e incidencias, como por ejemplo el rango de edad en que se encuentra la persona o si se trata de un hombre o una mujer.

11 Referencias

- [1] F. L. Marqués, «Clinic cloud,» [En línea]. Available: <https://clinic-cloud.com/blog/aplicaciones-de-la-tecnologia-en-la-salud/>. [Último acceso: 15 04 2022].
- [2] C. TI, «Canales TI,» 10 06 2020. [En línea]. Available: <https://itcomunicacion.com.mx/el-uso-de-la-tecnologia-en-la-medicina/>. [Último acceso: 15 04 2022].
- [3] S. Lee, H. Huang y M. Zelen, «Early detection of disease and scheduling of screening examinations,» *Statistical Methods in Medical Research*, vol. 13, nº 6, pp. 443-456, 2004.
- [4] O. T. Jones, C. K. Ranmuthu, P. N. Hall, G. Funston y F. M. Walter, «Recognising Skin Cancer in Primary Care,» *Advances in Therapy*, vol. 37, nº 1, p. 603-616, 2020.
- [5] C. Garbe, K. Peris, A. Hauschild, P. Saiag, M. Middleton, L. Bastholt, J.-J. Grob, J. Malvehy, J. Newton-Bishop, A. J. Stratigos, H. Pehamberger y A. M. Eggermont, «Diagnosis and treatment of melanoma. European consensus-based interdisciplinary guideline – Update 2016,» *European Journal of Cancer*, vol. 63, pp. 201-217, 2016.
- [6] I. P. Santos, R. van Doorn y P. Caspers, «Improving clinical diagnosis of early-stage cutaneous melanoma based on Raman spectroscopy,» *British Journal of Cancer*, vol. 119, p. 1339-1346, 2018.
- [7] The International Skin Imaging Collaboration, «ISIC,» [En línea]. Available: <https://www.isic-archive.com/#!/topWithHeader/onlyHeaderTop/gallery>. [Último acceso: January 2020].
- [8] Biblioteca Nacional de Medicina de EE. UU. , «Lunar | Nevos | MedlinePlus en español,» [En línea]. Available: <https://medlineplus.gov/spanish/moles.html>. [Último acceso: 08 09 2021].
- [9] American Cancer Society, «Estadísticas importantes sobre el cáncer de piel tipo melanoma,» [En línea]. Available: <https://www.cancer.org/es/cancer/cancer-de-piel-tipo-melanoma/acerca/estadisticas-clave.html>. [Último acceso: 09 09 2021].
- [10] Biblioteca Nacional de Medicina de EE. UU. , «Melanoma: MedlinePlus en español,» [En línea]. Available: <https://medlineplus.gov/spanish/melanoma.html>. [Último acceso: 08 09 2021].
- [11] American Cancer Society, «Etapas del cáncer de piel tipo melanoma,» [En línea]. Available: <https://www.cancer.org/es/cancer/cancer-de-piel-tipo-melanoma/deteccion-diagnostico-clasificacion-por-etapas/clasificacion-por-etapas-el-cancer-de-piel-tipo-melanoma.html>. [Último acceso: 09 09 2021].
- [12] J. A. Avilés-Izquierdo, C. Ciudad-Blanco, A. Sánchez-Herrero, A. Mateos-Mayo, L. M. Nieto-Benito y E. Rodríguez-Lomba, «Dermoscopy of cutaneous melanoma metastases: A color-based pattern classification,» *Journal of Dermatology*, vol. 46, nº 7, pp. 564-569, 2019.
- [13] V. Teigens, Inteligencia Artificial General, Cambridge Stanford Books, 2000.
- [14] V. Advany, «What is Artificial Intelligence? How does AI work, Types and Future of it?,» 19 Octubre 2021. [En línea]. Available: <https://www.mygreatlearning.com/blog/what-is-artificial-intelligence/#comments>. [Último acceso: 25 Noviembre 2021].

- [15] S. Khan, H. Rahmani, S. A. A. Shah y M. Bennamoun, *A Guide to Convolutional Neural Networks for Computer Vision*, Morgan & Claypool, 2018.
- [16] A. Hintze, «government technology,» 14 11 2016. [En línea]. Available: <https://www.govtech.com/computing/understanding-the-four-types-of-artificial-intelligence.html#:~:text=2.,Limited%20memory,and%20monitoring%20them%20over%20time..> [Último acceso: 10 04 2022].
- [17] Google, «Google Trends,» [En línea]. Available: <https://trends.google.es/trends/explore?date=all&q=Convolutional%20neural%20network,deep%20learning,machine%20learning>. [Último acceso: Agosto 2021].
- [18] R. Wason, «Deep learning: Evolution and expansion,» *Cognitive Systems Research*, vol. 52, pp. 701-708, 2018.
- [19] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu y F. Alsaadi, «A survey of deep neural network architectures and their applications,» *Neurocomputing*, vol. 234, pp. 11-26, 2017.
- [20] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu y M. S. Lew, «Deep learning for visual understanding: A review,» *Neurocomputing*, vol. 187, pp. 27-48, 2016.
- [21] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang y R. X. Gao, «Deep learning and its applications to machine health monitoring,» *Mechanical Systems and Signal Processing*, vol. 115, pp. 213-237, 2019.
- [22] F. Ercal, A. Chawla, W. Stoecker, H.-C. Lee y R. Moss, «Neural Network Diagnosis of Malignant Melanoma From Color Images,» *IEEE Transactions on Biomedical Engineering*, vol. 41, n° 9, pp. 837-845, 1994.
- [23] P. Vannoorenberghe, O. Colot y D. De Brucq, «Dempster-Shafer's theory as an aid to color information processing. Application to melanoma detection in dermatology,» de *Proceedings 10th International Conference on Image Analysis and Processing*, Venice, Italy, 1999.
- [24] L. Yu, H. Chen, Q. Dou, J. Qin y P.-A. Heng, «Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks,» *IEEE TRANSACTIONS ON MEDICAL IMAGING*, vol. 36, n° 4, pp. 994-1004, 2017.
- [25] Y. Yuan, M. Chao y Y.-C. Lo, «Automatic Skin Lesion Segmentation Using Deep Fully Convolutional Networks with Jaccard Distance,» *IEEE TRANSACTIONS ON MEDICAL IMAGING*, vol. 36, pp. 1876-1886, 2017.
- [26] N. Codella, Q.-B. Nguyen, S. Pankanti, D. Gutman, B. Helba, A. Halpern y J. R. Smith, «Deep learning ensembles for melanoma recognition in dermoscopy images,» *IBM Journal of Research and Development*, vol. 21, pp. 5:1-5:15, 2017.
- [27] I. González-Díaz, «DermaKNet: Incorporating the knowledge of dermatologists to Convolutional Neural Networks for skin lesion diagnosis,» *IEEE Journal of Biomedical and Health Informatics*, vol. 23, n° 2, pp. 547-559, 2018.
- [28] E. Nasr-Esfahani, S. Samavi, N. Karimi y S. Soroushmehr, «Melanoma Detection by Analysis of Clinical Images Using Convolutional Neural Network,» de *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, 2016.

- [29] M. u. Rehman, S. H. Khan, S. M. D. Rizvi, Z. Abbas y A. Zafar, «Classification of Skin Lesion by Interference of Segmentation and Convolution Neural Network,» de *2018 2nd International Conference on Engineering Innovation (ICEI)*, Bangkok, 2018.
- [30] Z. Yu, D. Ni, S. Chen, J. Qin, S. Li, T. Wang y B. Lei, «Hybrid dermoscopy image classification framework based on deep convolutional neural network and Fisher vector,» de *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, Melbourne, 2017.
- [31] S. Sigurdsson, P. A. Philipsen, L. K. Hansen, J. Larsen, M. Gniadecka y H. C. Wulf, «Detection of skin cancer by classification of Raman spectra,» *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, vol. 51, n° 10, pp. 1784 - 1793, 2004.
- [32] I. Pirnóg, R. O. Preda, C. Oprea y C. Paleologu, «Automatic lesion segmentation for melanoma diagnostics in macroscopic images,» de *23rd European Signal Processing Conference (EUSIPCO)*, Nice, 2015.
- [33] R. Kasmi y K. Mokrani, «Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule,» *IET Image Processing*, vol. 10, n° 6, pp. 448-455, 2016.
- [34] J. R. Varatharaj, N. J. P. R y A. K. L, «Detection of melanoma skin cancer using digital camera images,» *ARPJ Journal of Engineering and Applied Sciences*, vol. 10, pp. 3082-3085, 2015.
- [35] K. Shimizu, H. Iyatomi, . M. E. Celebi, K.-A. Norton y M. Tanaka, «Four-Class Classification of Skin Lesions With Task Decomposition Strategy,» *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, vol. 62, n° 1, pp. 274 - 283, 2014.
- [36] M. Sadeghi, T. K. Lee, D. McLean, H. Lui y M. S. Atkins, «Detection and Analysis of Irregular Streaks in Dermoscopic Images of Skin Lesions,» *Detection and Analysis of Irregular Streaks in Dermoscopic Images of Skin Lesions*, vol. 32, n° 5, pp. 849 - 861, 2013.
- [37] Y. Gu, Z. Ge, C. P. Bonnington y J. Zhou, «Progressive Transfer Learning and Adversarial Domain Adaptation for Cross-Domain Skin Disease Classification,» *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFOMATICS*, vol. 24, n° 5, pp. 1379 - 1393, 2020.
- [38] Q. Zhou, Y. Shi, Z. Xu, R. Qu y G. Xu, «Classifying Melanoma Skin Lesions Using Convolutional Spiking Neural Networks With Unsupervised STDP Learning Rule,» *IEEE Access*, vol. 8, pp. 101309 - 101319, 2020.
- [39] L. Wei, K. Ding y H. Hu, «Automatic Skin Cancer Detection in Dermoscopy Images based on Ensemble Lightweight Deep Learning Network,» *IEEE Access*, vol. 8, pp. 99633 - 99647, 2020.
- [40] Q. Zhou, C. Ren y S. Qi, «An Imbalanced R-STDP Learning Rule in Spiking Neural Networks for Medical Image Classification,» *IEEE Access*, vol. 8, pp. 224162 - 224177, 2020.
- [41] A. A. Adegun y S. Viriri, «Deep Learning-Based System for Automatic Melanoma Detection,» *IEEE Access*, vol. 8, pp. 7160 - 7172, 2019.
- [42] S. Albahli, N. Nida, A. Irtaza, M. H. Yousaf y M. T. Mahmood, «Melanoma Lesion Detection and Segmentation Using YOLOv4-DarkNet and Active Contour,» *IEEE Access*, vol. 8, pp. 198403 - 198414, 2020.

- [43] Q. Wang, L. Sun, Y. Wang, M. Zhou, M. Hu, J. Chen, Y. Wen y Q. Li, «Identification of Melanoma From Hyperspectral Pathology Image Using 3D Convolutional Networks,» *IEEE Transactions on Medical Imaging*, vol. 40, n° 1, pp. 218 - 227, 2021.
- [44] K. Thurnhofer-Hemsi, E. López-Rubio, E. Domínguez y D. A. Elizondo, «Skin Lesion Classification by Ensembles of Deep Convolutional Networks and Regularly Spaced Shifting,» *IEEE Access*, vol. 9, pp. 112193 - 112205, 2021.
- [45] M. Goyal, A. Oakley, P. Bansal, D. Dancey y M. H. Yap, «Skin Lesion Segmentation in Dermoscopic Images With Ensemble Deep Learning Methods,» *IEEE Access*, vol. 8, pp. 4171-4181, 2020.
- [46] L. Ichim y D. Popescu, «Melanoma Detection Using an Objective System Based on Multiple Connected Neural Networks,» *IEEE Access*, vol. 8, pp. 179189-179202, 2020.
- [47] L. Talavera-Martínez, P. Bibiloni y M. González-Hidalgo, «Hair Segmentation and Removal in Dermoscopic Images Using Deep Learning,» *IEEE Access*, vol. 9, pp. 2694-2704, 2021.
- [48] A. Gong, X. Yao y W. Lin, «Dermoscopy Image Classification Based on StyleGANs and Decision Fusion,» *IEEE Access*, vol. 8, pp. 70640-70650, 2020.
- [49] T.-C. Pham, A. Doucet, C.-M. Luong, C.-T. Tran y V.-D. Hoang, «Improving Skin-Disease Classification Based on Customized Loss Function Combined With Balanced Mini-Batch Logic and Real-Time Image Augmentation,» *IEEE Access*, vol. 8, pp. 150725-150737, 2020.
- [50] B. Zhang, Z. Wang, J. Gao, C. Rutjes, K. Nufer, D. Tao, D. D. Feng y S. W. Menzies, «Short-Term Lesion Change Detection for Melanoma Screening With Novel Siamese Neural Network,» *IEEE Transactions on Medical Imaging*, vol. 40, n° 3, pp. 840-851, 2021.
- [51] H. Arab, L. Chioukh, M. Dashti Ardakani, S. Dufour y S. O. Tatu, «Early-Stage Detection of Melanoma Skin Cancer Using Contactless Millimeter-Wave Sensors,» *IEEE Sensors Journal*, vol. 20, n° 13, pp. 7310-7317, 2020.
- [52] L. Song, J. Lin, Z. J. Wang y H. Wang, «An End-to-End Multi-Task Deep Learning Framework for Skin Lesion Analysis,» *IEEE Journal of Biomedical and Health Informatics*, vol. 24, n° 10, pp. 2912-2921, 2020.
- [53] R. Rastghalam, H. Danyali, M. S. Helfroush, M. E. Celebi y M. Mokhtari, «Skin Melanoma Detection in Microscopic Images Using HMM-Based Asymmetric Analysis and Expectation Maximization,» *IEEE Journal of Biomedical and Health Informatics*, vol. 25, n° 9, pp. 3486-3497, 2021.
- [54] M. A. Khan, K. Muhammad, M. Sharif, T. Akram y V. H. C. d. Albuquerque, «Multi-Class Skin Lesion Detection and Classification via Teledermatology,» *IEEE Journal of Biomedical and Health Informatics*, vol. 25, n° 12, pp. 4267-4275, 2021.
- [55] novaderma, «novaderma CLINICA DERMATOLOGICA,» [En línea]. Available: <https://www.clinicanovaderma.com/la-dermatoscopia-digital/>. [Último acceso: 08 2021].
- [56] E. Galán y D. Puerto, Manual para la detección temprana del cancer de piel y recomendaciones para la disminución de exposición a radiación ultravioleta, Bogotá: Instituto Nacional de Cancerología ESE, 2015.

- [57] A. Karpathy , F.-F. Li, J. Johnson y S. Yeung, «CS231n: Convolutional Neural Networks for Visual Recognition.,» Stanford CS class, 2016-2019. [En línea]. Available: <http://cs231n.github.io/>. [Último acceso: January 2020].
- [58] T. Guo, J. Dong, H. Li y Y. Gao, «Simple convolutional neural network on image classification,» de *IEEE 2nd International Conference on Big Data Analysis (ICBDA)*(, Beijing, 2017.
- [59] S. ALBAWI, T. A. MOHAMMED y S. AL-ZAWI, «Understanding of a convolutional neural network,» de *International Conference on Engineering and Technology (ICET)*, Antalya, 2017.
- [60] R. Sarkar, C. C. Chatterjee y A. Hazra, «Diagnosis of melanoma from dermoscopic images using a deep depthwise separable residual convolutional network,» *IET Image Processing*, vol. 13, 2019.
- [61] M. A. Khan, M. Sharif, M. Raza, A. Anjum, T. Saba y S. A. Shad, «Skin lesion segmentation and classification: A unified framework of deep neural network features fusion and selection,» *Expert Systems*, 2019.
- [62] Z. Yu, X. Jiang, F. Zhou, J. Qin, D. Ni, S. Chen, B. Lei y T. Wang, «Melanoma Recognition in Dermoscopy Images via Aggregated Deep Convolutional Features,» *IEEE Transactions on Biomedical Engineering*, vol. 66, nº 4, pp. 1006-1016, 2019.
- [63] K. Jayapriya y I. J. Jacob, «Hybrid fully convolutional networks-based skin lesion segmentation and melanoma detection using deep feature,» *International Journal of Imaging Systems and Technology*, 2019.
- [64] A. Namozov y Y. I. Cho, «Convolutional Neural Network Algorithm with Parameterized Activation Function for Melanoma Classification,» de *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, Korea (South), 2018.
- [65] J. Yang, F. Xie, H. Fan, Z. Jiang y J. Liu, «Classification for Dermoscopy Images Using Convolutional Neural Networks Based on Region Average Pooling,» *IEEE Access*, vol. 6, pp. 65130-65138, 2018.
- [66] Y. Li y L. Shen, «Skin Lesion Analysis towards Melanoma Detection Using Deep Learning Network,» *Sensors*, vol. 18, nº 2, p. 556, 2018.
- [67] P. Tang, Q. Liang, X. Yan, S. Xiang y D. Zhang, «GP-CNN-DTEL: Global-Part CNN Model With Data-Transformed Ensemble Learning for Skin Lesion Classification,» *IEEE Journal of Biomedical and Health Informatics*, vol. 24, nº 10, pp. 2870 - 2882, 2020.
- [68] R. Ashraf, S. Afzal, A. U. Rehman, S. Gul, J. Baber, M. Bakhtyar, I. Mehmood, O.-Y. Song y M. Maqsood, «Region-of-Interest Based Transfer Learning Assisted Framework for Skin Cancer Detection,» *IEEE Access*, vol. 8, pp. 147858 - 147871, 2020.
- [69] H. Wu, J. Pan, Z. Li, Z. Wen y J. Qin, «Automated Skin Lesion Segmentation Via an Adaptive Dual Attention Module,» *IEEE Transactions on Medical Imaging*, vol. 40, nº 1, pp. 357-370, 2021.
- [70] Y. Jiang, S. Cao, S. Tao y H. Zhang, «Skin Lesion Segmentation Based on Multi-Scale Attention Convolutional Neural Network,» *IEEE Access*, vol. 8, pp. 122811-122825, 2020.

- [71] A. Gong, X. Yao y W. Lin, «Classification for Dermoscopy Images Using Convolutional Neural Networks Based on the Ensemble of Individual Advantage and Group Decision,» *IEEE Access*, vol. 8, p. 2020, 2020.
- [72] J. Zhang, Y. Xie, Y. Xia y C. Shen, «Attention Residual Learning for Skin Lesion Classification,» *IEEE Transactions on Medical Imaging*, vol. 38, n° 9, pp. 2092-2103, 2019.

12 Anexos

Dataset	Ref	Clasificador	Acc	AUC	F1	Procesador	Ram	GPU	Imagenes
ISBI 2016	[25]	Fully convolutional neural networks with Jaccard Distance	95.5%	-	-	Dell XPS 8900 desktop with Intel(R) i7-6700 3.4 GHz CPU	-	GPU of Nvidia GeForce GTX 1060 with 6GB GDDR5 memory	1279
ISBI 2016	[24]	Very Deep Residual Networks	85.5%	0.78	-	-	-	NVIDIA TITAN X GPU	1250
CNNs	[28]	Pre-trained CNN	81,00%	-	-	Intel Core i7-4790K	32 GB	NVIDIA GeForce GTX Titan X GPU cards with scalable link interface (SLI)	6120
PH2	[60]	Deep depth wise separable residual convolutional algorithm	96.7%	0.9949	0,9948	2.8 GHz	16 BG	GPU support of 4 GB	-
ISBI 2016	[61]	Optimized colour feature (OCF) of lesion segmentation & DCNN	92.1%	0.97	-	-	-	The proposed model is implemented in on MATLAB 2017b to perform experiments using high-processing GPU systems	-
ISBI 2016	[62]	DCNN-FV (fusion)	86.81%	0.852	-	Inter Xeon E5-2680 @ 2.70 GHz	128GB	NVIDIA Quadro K4000	1279
ISBI 2017	[63]	FCNs based on VGG-16 and googleNet	85.3%	-	-	-	-	NVIDIA TITAN X GPU	1250
ISBI 2017	[64]	Transfer learning	95,86%	-	-	Intel(R) Core(TM) i7-7700HQ CPU 2.80 GHz	-	Nvidia GTX 1050 Ti GPU	-
ISBI 2017	[61]	Optimized colour feature (OCF) of lesion segmentation & DCNN	96.5%	0.99	-	-	-	The proposed model is implemented in on MATLAB 2017b to perform experiments using high-processing GPU systems	-
ISBI 2017	[65]	CNN + RAPooling +RankOpt	83,00%	0.842	-	-	-	NVIDIA GTX1080 GPU with 8GB of memory	-
ISBI 2017	[66]	Two (FCNR) & LICU	85.7%	0.912	-	-	-	GeForce GTX TITAN X, 12 GB RAM	-
ISBI 2018	[61]	Optimized colour feature (OCF) of lesion segmentation & DCNN	85.1%	0.92	-	-	-	The proposed model is implemented in on MATLAB 2017b to perform experiments using high-processing GPU systems	-
ISIC Archive	[60]	Deep depth wise separable residual convolutional algorithm	99.5%	0.994	-	-	-	GPU support of 4 GB	-
Med-Node	[60]	Deep depth wise separable residual convolutional algorithm	95.2%	0.944	-	-	-	GPU support of 4 GB	-
ISCI 2016	[67]	GP-CNN and a data-transformed ensemble learning strategy	86.3%	0,861	0,893	Intel(R) Core (TM) i7-7820X CPU @ 3.60GHz	-	RTX 2080Ti GPU	1279
ISIC 2017	[67]	GP-CNN and a data-transformed ensemble learning strategy	85,00%	0,891		Intel(R) Core (TM) i7-7820X CPU @ 3.60GHz	-	RTX 2080Ti GPU	2600

DermIS	[68]	Convolutional Neural Network (CNN) based transfer learning model with data augmentation for ROI images of DermIS and DermaQuest datasets	97.9% AUG - 81,30%	-	0,9770	-	-	6GB NVIDIA GPU card	2779
DermQUEST	[68]	Convolutional Neural Network (CNN) based transfer learning model with data augmentation for ROI images of DermIS and DermaQuest datasets	97.4% AUG - 74,2%	-	0,9790	-	-	6GB NVIDIA GPU card	959
IS-BI2017	[69]	Adaptive dual attention module (ADAM) and a dual encoder architecture (DEA)	95,70%	-	-	Intel i7-7700k CPU	32 GB DDR4 RAM	Nvidia GeForce GTX 1080Ti (11 GB)	2750
ISIC2018	[69]	Adaptive dual attention module (ADAM) and a dual encoder architecture (DEA)	94,70%	-	-	Intel i7-7700k CPU	32 GB DDR4 RAM	Nvidia GeForce GTX 1080Ti (11 GB)	2594
ISIC 2017	[70]	CSARM-CNN (Channel & Spatial Attention Residual Module) model	94,96%	-	-	Intel Xeon (R) CPU E5-2620 v3 2.40GHz	-	NVIDIA Tesla K80 (12G)GPU,	2750
PH2	[70]	CSARM-CNN (Channel & Spatial Attention Residual Module) model	95,23%	-	-	Intel Xeon (R) CPU E5-2620 v3 2.40GHz	-	NVIDIA Tesla K80 (12G)GPU,	200
ISIC 2018	[71]	Classification method based on the ensembled of individual advantage and group decision in dermoscopy images	92,60%	0,918	0,491	-	-	Four NVIDIA Tesla P100 GPUs with 16GB of memory	10015
ISIC 2019	[71]	Classification method based on the ensembled of individual advantage and group decision in dermoscopy images	92,40%	0,919	0,488	-	-	Four NVIDIA Tesla P100 GPUs with 16GB of memory	25331
ISBI 2016	[39]	Ensemble Lightweight deep learnign network	87,60%	0,854	-	-	-	NVIDIA Tesla K80 GPU (12G)	1279
ISIC-skin 2017	[72]	Attention residual learning convolutional neural network (ARL-CNN)	85,00%	0,875	-	-	-	ARL-CNN50 model with one NVIDIA GTX Titan XP GPU	2750