

## Inteligencia artificial en la detección de melanomas

## Artificial intelligence in the detection of melanomas

Cristian Borja Borja<sup>1</sup> , Geovanny Euclides Silva Peñafiel<sup>2</sup> , Víctor Alfonso Cusco Vinuesa<sup>2</sup> , Paco Jovanni Vasquez Carrera<sup>2</sup> 

<sup>1</sup>Universidad Internacional de la Rioja, Logroño – España

<sup>2</sup>Universidad Técnica de Cotopaxi extensión La Maná, La Maná – Ecuador

Correo de correspondencia: cristianborjab@gmail.com, geovanny.silva1764@utc.edu.ec, victor.cusco7756@utc.edu.ec, paco.vasquez@utc.edu.ec.

### Información del artículo

**Tipo de artículo:**  
Artículo original

**Recibido:**  
19/03/2022

**Aceptado:**  
10/05/2022

**Publicado:**  
15/06/2022

**Revista:**  
DATEH

OPEN ACCESS



### Resumen

En vista que se ha abierto el campo de la inteligencia artificial en la detección de melanoma, este trabajo de investigación se realiza con el objetivo de intentar aportar un grano de arena a este ámbito, cabe mencionar que el cáncer de piel tipo melanoma es de los más agresivos en cuanto a mortalidad y la mejor forma de combatirlo es mediante su detección temprana. Se implementaron cinco modelos predictivos, cuatro basados en redes neuronales convolucionales CNN, y uno bajo la arquitectura Vision Transformer ViT, todos estos modelos se entrenaron con datasets armados a partir de imágenes dermatoscópicas obtenidas del sitio web oficial de la plataforma ISIC que promueve la implementación de la inteligencia artificial para la lucha contra el cáncer de piel. Los resultados obtenidos no fueron los esperados, sin embargo, se exponen las posibles causas y se proyecta una línea de trabajo futuro.

**Palabras clave:** *inteligencia artificial, melanoma, cáncer de piel, redes neuronales, datasets.*

### Abstract

Given that the field of artificial intelligence in the detection of melanoma has been opened, this research work is carried out with the aim of trying to contribute a grain of sand to this field, it is worth mentioning that melanoma skin cancer is one of the most more aggressive in terms of mortality and the best way to combat it is through its early detection. Five predictive models were implemented, four based on CNN convolutional neural networks, and one under the Vision Transformer ViT architecture, all these models were trained with datasets built from dermoscopic images obtained from the official website of the ISIC platform that promotes the implementation of artificial intelligence for the fight against skin cancer. The results obtained were not as expected, however, the possible causes are exposed and a line of future work is projected.

**Keywords:** *artificial intelligence, melanoma, skin cancer, neural networks, datasets*

**Forma sugerida de citar (APA):** López-Rodríguez, C. E., Sotelo-Muñoz, J. K., Muñoz-Venegas, I. J. y López-Aguas, N. F. (2024). Análisis de la multidimensionalidad del brand equity para el sector bancario: un estudio en la generación Z. Retos Revista de Ciencias de la Administración y Economía, 14(27), 9-20. <https://doi.org/10.17163/ret.n27.2024.01>.

### INTRODUCCIÓN

estos días se ha constatado el apogeo de la informática y electrónica que ha dado lugar a la automatización de tareas y procesos llevados a cabo enteramente por intervención humana, con el avance tecnológico en forma exponencial que existe en el presente, se han ido desarrollando computadoras con mayor y mejor capacidad de procesamiento de datos, haciéndolas más autónomas y, llevándose a una analogía con los seres humanos, haciendo que estas computadoras lleguen a ser “inteligentes”. En consecuencia, ha surgido la rama de la ingeniería denominada como “Inteligencia Artificial”, cuya principal finalidad es la de servir a la humanidad en

la ejecución de tareas que resultan complejas para los seres humanos y brindar asistencia en procedimientos críticos o delicados como es el caso de la detección de cáncer de piel tipo melanoma dentro del campo de la medicina y que es objeto de este trabajo de investigación.

Según Mezquita [1] el melanoma es una de las afecciones más agresivas en lo que a cáncer de piel se refiere ya que produce alrededor del 75% de las muertes por cáncer de piel en España según revela el Registro Nacional de Melanoma Cutáneo; hasta el año 2002 la OMS (Organización Mundial de la Salud) estimaba que 20.000 personas menores de 55 años en los países miembros de la

Unión Europea (UE) padecían esta enfermedad. Para la American Cancer Society [2] el melanoma, a pesar de ser un tipo de cáncer poco frecuente, tiene una alta tasa de mortalidad ya que es mucho más probable que se propague a otras partes del cuerpo si no se combate a tiempo. El procedimiento normal para la detección temprana de melanoma se realiza a través de imágenes dermatoscópicas que capturan estructuras cutáneas invisibles a simple vista o con fotografías convencionales, y se lleva a cabo solamente por dermatólogos especializados.

En consecuencia, Marín et al. [3] manifiesta que se han implementado numerosos algoritmos y modelos de inteligencia artificial para replicar de forma automática y eficiente el procedimiento de detección de melanoma. Adegun & Viriri [4] dice que la utilización de la inteligencia artificial en la detección de melanoma aún debe enfrentar varios desafíos, por tal motivo, se alienta a investigadores de todo el mundo a contribuir con esta temática. Tal afirmación ha sido la motivación para desarrollar este trabajo investigativo.

Las redes neuronales convolucionales CNN han jugado un papel preponderante en el desarrollo de modelos predictivos de imágenes debido a su estupenda capacidad de procesarlas, por tal motivo, este trabajo se centrará en implementar un modelo CNN y a partir de él, realizar modificaciones y aplicar otras técnicas, tecnologías y/o herramientas con el anhelo de mejorar los resultados obtenidos hasta el momento. La predicción de la que se habla en un modelo predictivo, consiste en acercarse en la mayor medida posible a la realidad en función de lo ocurrido en el pasado, en tal virtud, el modelo realiza un entrenamiento con datos conocidos y aprende a identificar ciertos aspectos que establecen que dichos datos pertenezcan a una u otra categoría; concretamente en este proyecto, el modelo aprende a reconocer factores que determinen el hecho de que una lesión cutánea pertenezca a la categoría de melanoma para luego, en base a este aprendizaje, poder predecir con la mayor precisión posible si una futura lesión cutánea puede catalogarse como melanoma. El entrenamiento se realiza con un dataset debidamente procesado cuyas imágenes se obtuvieron gratuitamente de la plataforma de ISIC dedicada a almacenar imágenes dermatoscópicas para este tipo de proyectos.

Los modelos construidos en este trabajo de investigación se comparan con los del estado del arte y lamentablemente no se consiguieron los resultados esperados, la causa probable para ello, fue el hecho contar con recursos computacionales limitados, principalmente la memoria RAM que no permitió trabajar con un tamaño lo suficientemente grande en las imágenes de entrada.

### Antecedentes

En la actualidad, en sintonía con los avances científicos y tecnológicos, se ha vuelto común la utilización de herramientas computacionales para llevar a cabo tareas de la medicina, entre ellas el diagnóstico de cáncer de piel mediante el análisis de lesiones cutáneas que eventualmente podrían transformarse en melanomas; el empleo de estas metodologías modernas y automáticas, además suponen un avance en el accionar de los profesionales médicos ya que representan una optimización en el tiempo y esfuerzo dedicado comparadas con las metodologías manuales convencionales.

En consecuencia, han surgido numerosos modelos de deep learning para el tratamiento de imágenes dermatoscópicas que han ofrecido magníficos resultados pero que aún enfrentan desafíos y dificultades debido a la unicidad y complejidad de cada paciente.

La aplicación de modelos de inteligencia artificial para la detección de melanoma ha visto la luz en proyectos realizados, por ejemplo, por Rangel-Cortes et al. [5], donde efectúan la clasificación de melanomas a través de CNN empleando tres técnicas: la primera mediante la añadidura de metadatos sobre edad, género y tamaño del melanoma comparada con un análisis sin añadir dichos metadatos, la segunda con imágenes netamente del melanoma eliminando la piel de alrededor y la tercera con una entrada de 2.000 imágenes sin pre-procesamiento de 300x300 píxeles. La precisión predictiva real alcanzada por el modelo fue: 77.50% sin metadatos, 77.50% con metadatos, 67.50% sin la piel que rodea el melanoma y 80.00% sin un pre-procesamiento en las imágenes entrantes.

Abhinav & Dheeba [6] desarrollaron un clasificador mediante la implementación de CNN y basado en una arquitectura de Transfer Learning con la finalidad de trabajar sobre los avances logrados por otros modelos construidos previamente, tales modelos fueron: Inception v3, InceptionResNet v2 y ResNet 152; la base de datos empleada en el entrenamiento y posterior evaluación de este modelo fue la proporcionada por ISIC; el modelo alcanzó un 93.50% de precisión predictiva.

Yu et al. [7] aplicaron una CNN en 724 imágenes dermatoscópicas de melanomas acrales y nevos benignos obtenidas de manos y pies de los pacientes, dichas imágenes estuvieron clasificadas en 350 imágenes de melanomas de 81 pacientes y 374 imágenes de nevos benignos de 194 pacientes; la precisión predictiva lograda por el modelo fue del 83.51% en verdaderos positivos y 80.23% en verdaderos negativos.

Nasiri et al. [8] propusieron un enfoque de clasificación de lesiones cutáneas para la detección de melanoma bajo un sistema de razonamiento basado en casos CBR (Case-Based Reasoning) con la finalidad de recuperar nuevas imágenes de entrada desde la base de datos original y obtener un nuevo conjunto de entrenamiento más organizado y fácil de entender para el modelo que fue construido con 16 capas convolucionales, este modelo se entrenó utilizando imágenes del dataset de ISIC utilizándose 1.346 imágenes para el entrenamiento y 450 para la evaluación; la precisión de predicción alcanzada fue del 75.00%.

Mahecha et al. [9] desarrollaron una GPU (Graphic Processing Unit) para el análisis de imágenes clínicas de la piel cuyo funcionamiento se basa en un sistema de redes neuronales con distintos módulos encargados de tareas específicas tales como la detección de la silueta del melanoma y el estudio del contorno del mismo para determinar el diagnóstico médico final; los resultados obtenidos en este proyecto son de una precisión predictiva del 76.67% en la detección de melanomas y un 74.07% en lesiones cutáneas benignas.

Adegun & Viriri [4] enlistan una serie de proyectos exitosos que se detallan a continuación:

Alqudah et al. [10] desarrollaron un modelo empleando la técnica de transfer learning de las arquitecturas GoogleNet y AlexNet, el optimizador utilizado fue ADAM; el modelo se probó con el dataset de ISIC con imágenes segmentadas y no segmentadas, para el primer caso se alcanzó una predicción de clasificación del 92.20% y para el segundo caso la predicción llegó al 89.80%.

El-Khatib et al. [11] construyeron un sistema donde integraron las arquitecturas de GoogleNet, ResNet-101 y NasNet-Large pre-entrenadas en extensas bases de datos como ImageNet y Places365, también utilizaron un SVM (Support Vector Machine) para la detección de características obteniendo altos valores de precisión en la diferenciación entre melanomas y nevos benignos.

Almaraz-Damian et al. [12] implementaron un sistema donde el pre-procesamiento de imágenes se llevó a cabo mediante la regla ABCD, la extracción de características se realizó a través de una CNN, el sistema fue cotejado con otras técnicas de clasificación como Regresión Lineal, Support Vector Machine (SVM) y Relevant Vector Machines (RVM); al evaluarse con el dataset de ISIC el modelo logró buenos resultados en la detección de melanomas.

Yuan & Tavildar [13] elaboraron un modelo llamado SegAN que consiste en dos partes: un segmentador y la

red como tal, el segmentador genera un mapa de probabilidades en cada imagen de entrada de ser o no ser una lesión cutánea (que potencialmente pudiere convertirse en melanoma), mientras tanto, la red separa o clasifica los dos tipos de entradas; este modelo alcanzó una precisión predictiva del 93.70%.

Hao et al. [14] construyeron un modelo con una arquitectura de red tomada del modelo Deep-lab, pre-entrenada mediante PASCAL VOC-2012 y empleando métodos de bagging sobre modelos VGG16, U-net, DenseNet y Inception v3; obtuvo una precisión predictiva del 94.50% entrenándose con 20.000 iteraciones.

Molina-Moreno et al. [15] diseñaron una red híbrida entre FCN (Fully Convolutional Network) y una RPN elíptica (Region Proposal Network) para la segmentación automática de lesiones cutáneas (que pudieren convertirse en melanomas); el modelo FCN provee una segmentación acertada píxel por píxel de la imagen mientras que el modelo RPN extrae una segmentación a baja resolución de las zonas elípticas del potencial melanoma; este proyecto contó con el 94.70% de precisión predictiva.

Zhou et al. [16] efectuaron un ensemble de algunos modelos de deep learning del estado del arte como: DenseNet121, Se-ResNext50, Se-ResNext101, EfficientNet-B2, EfficientNet-B3 y EfficientNet-B4 aplicando la técnica del transfer learning; obtuvieron una precisión predictiva del 78.00%.

Pacheco et al. [17] realizaron un ensemble de varios clasificadores CNN, los cuales son: SENet, PNASNet, InceptionV4, ResNet-50, ResNet-101, ResNet-152, DenseNet-121, DenseNet-169, DenseNet-201, MobileNetV2, GoogleNet, VGG-16 y VGG-19; cada uno de estos clasificadores se adaptaron para tareas específicas; el modelo fue pre-entrenado con ImageNet y logró una precisión predictiva del 89.20%.

Chouhan [18] empleó el modelo DenseNet-161 pre-entrenado con el dataset ImageNet para clasificar lesiones cutáneas; la precisión de predicción obtenida fue del 87.00%.

Dat et al. [19] emplearon los modelos CNN EfficientNet e Inception ResNet del estado del arte para solucionar inconvenientes de desbalance de los datos de entrada, previnieron el overfitting con técnicas de muestreo ascendente y una función novel loss; alcanzaron una precisión predictiva del 87.20%.

Zhang [20] formuló un modelo de deep learning llamado MelaNet usando como base una capa densa de 169 neuronas, empleó una función softmax y una función

sigmoid para establecer la no linealidad requerida, este modelo fue concebido a partir de una arquitectura Fully Connected con una capa de salida de 8 unidades para construir clasificadores multi-clase; la precisión de predicción alcanzada fue del 89.70%.

Xing et al. [21] propone un ensemble de modelos con la misma estructura y entrenados con los mismos hiperparámetros, tal ensemble alcanzó una precisión predictiva del 75.40%.

Yousef & Motahari [22] emplearon 7 modelos con la base de datos ImageNet para desarrollar un clasificador, los modelos usados fueron: DenseNet121, InceptionV3, InceptionResNetV2, Xception, EfficientNetB1, EfficientNetB2 y EfficientNetB3; la precisión predictiva del clasificador fue del 88.50%.

Cohen & Shimoni [23] utilizaron técnicas de ensemble de modelos y generaron otros modelos desde diferentes puntos de vista, partieron con la construcción de un algoritmo que obtiene relaciones entre los atributos de entrada y los atributos objetivo, luego crearon un Diversity Generator para generar diversos modelos y un Combiner para combinar los resultados de todos los modelos, finalmente implementaron CNN para entrenar mediante loss functions y test-time augmentation; la capacidad predictiva que obtuvieron fue del 88.40%.

## MATERIAL Y MÉTODOS

El desarrollo de este trabajo de investigación se basa en una comparativa de soluciones al problema planteado de utilizar la inteligencia artificial para detectar melanomas, partiendo de modelos basados en redes neuronales convolucionales CNN; de modo que, se pretende evaluar la precisión, efectividad y confiabilidad de los modelos existentes, con los modelos que se implementan en este trabajo investigativo. Finalmente, se buscan aportar nuevas técnicas, tecnologías y/o herramientas que pudieran mejorar la efectividad actual con la que cuenta el campo de la detección automática de melanomas.

## CONTRIBUCIÓN

La implementación de la inteligencia artificial en el campo de la medicina, concretamente, en la detección de melanoma ha abierto un mar de conocimiento por explorar; en virtud de ello, este proyecto ha pretendido adentrarse en este nuevo mundo para intentar aportar un grano de arena en el entendimiento de este conocimiento que se ha develado.

Mediante la utilización de los recursos computacionales que estuvieron al alcance, se constató que el procesamiento de imágenes dermatoscópicas para detectar melanomas requiere una gran capacidad en memoria RAM, puesto que las características que definen una

lesión cutánea como melanoma o no, son muy finas, en consecuencia, el modelo predictivo debe contar con muchas capas convolucionales en el caso de ser una CNN, pero sobretodo, las imágenes de entrada al modelo deben ser lo suficientemente grandes para lograr encontrar estas diferencias entre “melanoma” y “no melanoma” que pueden darse en un solo píxel.

Se implementó también un modelo basado en la arquitectura Vision Transformer que ha venido dando buenos resultados en la clasificación de imágenes, sin embargo, no se obtuvieron los resultados deseados a causa de la limitación de memoria RAM, por tal motivo, se detalla una línea de trabajo futuro en la sección VII enfatizando la consecución de más recursos informáticos.

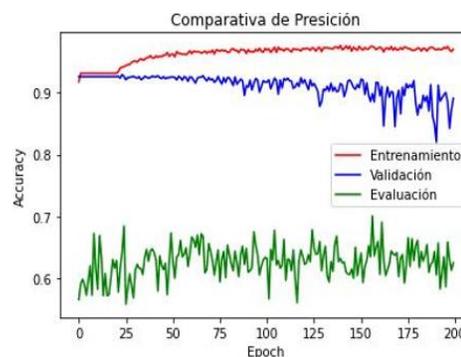
## RESULTADOS Y DISCUSIÓN

A continuación, se describen los resultados obtenidos en cada uno de los 5 modelos implementados en este proyecto.

### Modelo inicial

Este modelo es el fruto de experimentar con múltiples hiperparámetros, luego de un considerable número de intentos y ensayos, este modelo resultó ser el más óptimo de este proyecto. Por limitaciones de memoria no se pudo establecer un tamaño de entrada de imágenes mayor a 64x64 píxeles. Este modelo alcanzó una precisión del 62.54%.

En la Figura 1 se realiza una comparativa de precisión “accuracy” entre las etapas de entrenamiento, validación y evaluación del modelo inicial

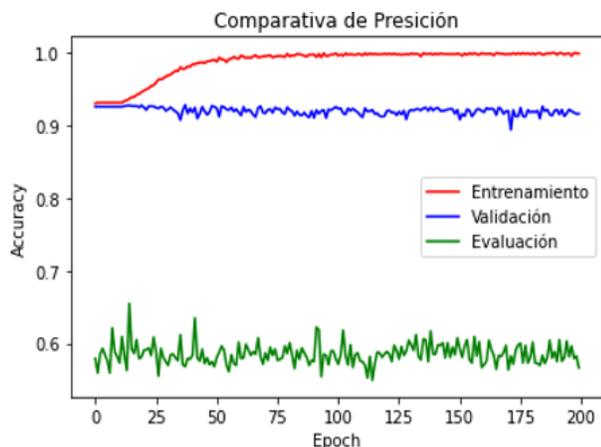


**Figura 1.** Curvas que describen la métrica de “accuracy” en el modelo inicial durante el entrenamiento (color rojo), validación (color azul) y evaluación (color verde) del modelo inicial en el transcurso de las 200 “epochs” establecidas.

### Modelo 1: Data augmentation

La diferencia de este modelo respecto al modelo inicial, valga la redundancia, es que se aplicó la técnica de “data augmentation”. La precisión alcanzada por este modelo fue del 56.62%.

En la Figura 2 se realiza una comparativa de precisión “accuracy” entre las etapas de entrenamiento, validación y evaluación del modelo 1

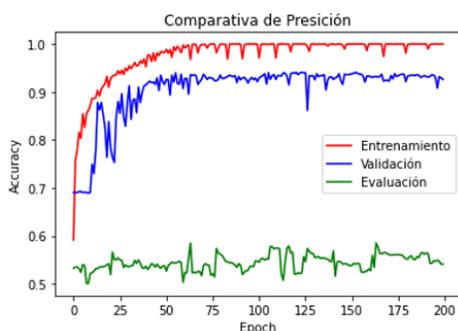


**Figura 2.** Curvas que describen la métrica de “accuracy” en el modelo 1 durante el entrenamiento (color rojo), validación (color azul) y evaluación (color verde) del modelo inicial en el transcurso de las 200 “epochs” establecidas.

#### Modelo 2: Menor dataset

La primera diferencia de este modelo respecto al modelo inicial, valga la redundancia, es que se empleó un dataset con un menor número de imágenes, 5.170 para entrenar y 1.164 para evaluar. La precisión alcanzada por este modelo fue del 54.12%.

En la Figura 3 se realiza una comparativa de precisión “accuracy” entre las etapas de entrenamiento, validación y evaluación del modelo 2.



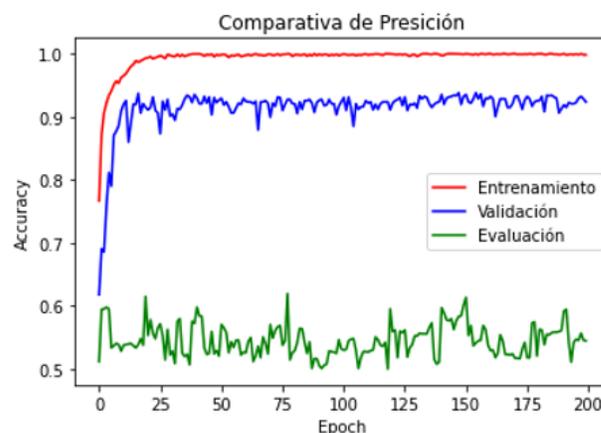
**Figura 3.** Curvas que describen la métrica de “accuracy” en el modelo 2 durante el entrenamiento (color rojo), validación (color azul) y evaluación (color verde) del modelo inicial en el transcurso de las 200 “epochs” establecidas.

#### Modelo 3: Menor dataset y mayor tamaño de imágenes

La primera diferencia de este modelo respecto al modelo inicial, valga la redundancia, es que se incrementó el tamaño de las imágenes de entrada de 64x64 a 128x128 píxeles, que fue el valor máximo permitido por la memoria RAM otorgada por Google Colaboratory. La segunda

diferencia es que se utilizó el dataset de imágenes usado por el modelo 2, es decir, 5.170 imágenes para entrenar y 1.164 para evaluar. La precisión alcanzada por este modelo fue del 54.47%.

En la Figura 4 se realiza una comparativa de precisión “accuracy” entre las etapas de entrenamiento, validación y evaluación del modelo 3.

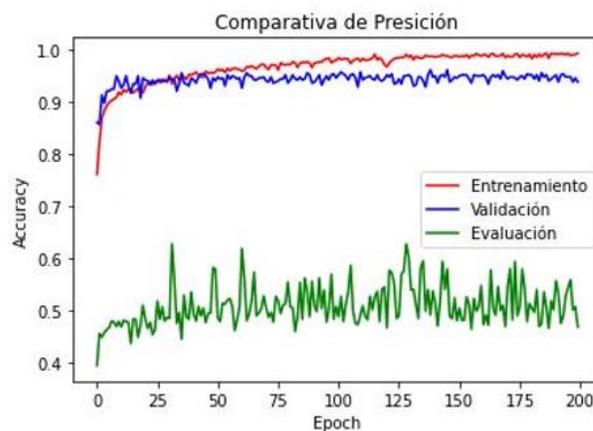


**Figura 4.** Curvas que describen la métrica de “accuracy” en el modelo 3 durante el entrenamiento (color rojo), validación (color azul) y evaluación (color verde) del modelo inicial en el transcurso de las 200 “epochs” establecidas.

#### Modelo 4: Vision Transformer ViT

Este modelo se construyó reemplazando las redes neuronales convolucionales por la arquitectura Vision Transformer, se empleó el mismo dataset que el modelo 2 y 3, es decir, 5.170 im para entrenar y 1.164 para evaluar. La precisión alcanzada por este modelo fue del 50.00%.

En la Figura 5 se realiza una comparativa de precisión “accuracy” entre las etapas de entrenamiento, validación y evaluación del modelo 4



**Figura 5.** Curvas que describen la métrica de “accuracy” en el modelo 4 durante el entrenamiento (color rojo), validación (color azul) y evaluación (color verde) del modelo inicial en el transcurso de las 200 “epochs” establecidas.

## DISCUSIÓN

Se pudo constatar que las redes neuronales convolucionales CNN son efectivas en la clasificación de imágenes, sin embargo, en este proyecto las imágenes tienen bastantes similitudes entre sus categorías, dicho de otro modo, las imágenes catalogadas como “melanoma” se parecen en muchos aspectos a las imágenes catalogadas como “no melanoma”; esto conllevó a que sea bastante difícil para los modelos, poder diferenciar adecuadamente los factores que definen a una imagen como “melanoma” y como “no melanoma”, en consecuencia, no se obtuvieron buenas clasificaciones y por lo tanto, no se consiguió el objetivo central de construir modelos para la detección de melanomas a partir de imágenes dermatoscópicas puesto que, para que un detector sea considerado como tal, debe contar con una precisión bastante alta, quizá en el orden del 95% o superior, mientras que la máxima precisión obtenida en este proyecto fue del 62.54%.

La capacidad predictiva de redes neuronales convolucionales y modelos de inteligencia artificial en general, obtiene mejores resultados conforme existe un mayor número de datos para el entrenamiento, por tal motivo, se armó un dataset con 22.457 imágenes para entrenar y 1.164 imágenes para evaluar, que son valores relativamente altos, tales imágenes tienen tamaños variables que llegan hasta 6708x4419 píxeles; los modelos construidos en este proyecto admiten un tamaño estandarizado constante para todas las imágenes, por consiguiente, se redujo el tamaño a 64x64 píxeles que fue el máximo permitido por los recursos computacionales con los que se trabajó; dichos recursos fueron proporcionados gratuitamente por la plataforma Google Colaboratory y extraoficialmente se sabe que otorga alrededor de 12 GB de memoria RAM para ejecutar notebooks; en consecuencia, las imágenes con las que se entrenaron los modelos, resultaron demasiado pequeñas considerando que existen bastantes similitudes entre la categoría “melanoma” y “no melanoma”; además, dada esta limitación de recursos, no se pudo implementar un mayor número de capas convolucionales, que en teoría, ayudaría a mejorar el rendimiento del modelo. Por ello, el modelo inicial que trabajó con este dataset obtuvo una precisión del 62.54%. De manera que, se afirma con un alto valor de confianza que el rendimiento de los modelos construidos en este proyecto aumentaría si se utiliza una mayor capacidad de memoria RAM ya que esto permite un mayor tamaño de las imágenes de entrada y así el modelo podrá identificar con mayor detalle aquellas características que establecen que una lesión cutánea sea catalogada como “melanoma” y como “no melanoma”; además, se podría contar con más capas convolucionales.

A fin de reducir costos computacionales, se construyó un dataset menor con 5.170 imágenes para entrenar y 1.164 imágenes para evaluar y se realizaron ensayos con el mismo; el resultado fue el modelo 2 en el que se comprobó que con 4 capas convolucionales el modelo alcanza su mejor rendimiento, es decir, al incorporar más capas convolutivas, el modelo no adquiere mejoras sustanciales. La precisión predictiva alcanzada por este modelo fue del 54.12%, lo cual era esperado puesto que se entrenó con un menor número de imágenes que el modelo inicial que logró una precisión del 62.54%.

El dataset menor pudo emplearse en imágenes de mayor tamaño dando lugar al modelo 3 que trabajó con imágenes de 128x128 píxeles a diferencia de los modelos anteriores que trabajaron con imágenes de 64x64 píxeles. La precisión predictiva lograda mejoró un poco respecto al modelo 2, subió al 54.47%; esto confirmaría la afirmación de que, con un mayor tamaño en las imágenes de entrada, se pueden obtener mejores resultados.

En vista del inconveniente de la limitación de recursos informáticos, principalmente de la memoria RAM, se buscaron herramientas que traten de solucionar esta problemática, esto llevó a implementar el modelo 4 aplicando la técnica de “Vision Transformer ViT” en lugar de redes neuronales convolucionales CNN, dicho sea de paso, no se encontró en el estado del arte trabajos previos que hayan utilizado esta técnica para la detección de melanomas. Los resultados obtenidos fueron bastante parecidos a los resultados del modelo 3, tomando en cuenta que se utilizaron exactamente el mismo dataset, es decir, 5.170 imágenes para entrenar, 1.164 imágenes para evaluar y un tamaño para las imágenes de entrada de 128x128 píxeles, la precisión alcanzada por el modelo 4 fue del 51.20%; esto significa que la arquitectura “Vision Transformer ViT” para clasificar imágenes podría convertirse en un buen competidor para las CNN. Sin embargo, no ayudó a mejorar los resultados del presente trabajo de investigación, se podría afirmar que nuevamente se debió al hecho de trabajar con un tamaño demasiado pequeño en las imágenes de entrada.

En el caso del modelo 1 se agregó una etapa de “data augmentation” al modelo inicial pero no se consiguieron buenos resultados sino al contrario, esto podría explicarse tal vez porque las modificaciones que hace esta técnica a las imágenes dermatoscópicas no contribuyen en una mejor identificación y diferenciación de características por parte del modelo.

## CONCLUSIONES

Se implementaron un total de cinco modelos de inteligencia artificial para la detección de melanomas, para todos ellos se utilizó un dataset compuesto enteramente de

imágenes dermatoscópicas debido a su capacidad de capturar estructuras cutáneas imperceptibles a simple vista o con fotografías convencionales, dichas imágenes se obtuvieron del sitio web de ISIC y tienen tamaños de hasta 6708x4419 píxeles.

Se realizó una búsqueda en el estado del arte en la temática de la inteligencia artificial en la detección de melanomas donde se encontraron varios trabajos exitosos que llegan a obtener una precisión predictiva de hasta el 93.50%. Además, se hallaron varias técnicas complementarias a las redes neuronales convolucionales que han contribuido exitosamente a la predicción de melanoma alcanzando precisiones de hasta el 94.70%

En total se construyeron cuatro modelos basados en redes neuronales convolucionales, el que mejores resultados produjo es el modelo inicial con una capacidad predictiva del 62.54%. Además, se implementó un modelo bajo la arquitectura Vision Transformer ViT como sustituto de las CNN y alcanzó una predicción del 51.20%.

El principal problema encontrado en el desarrollo de este trabajo de investigación fue el de las limitaciones de recursos computacionales, especialmente de memoria RAM, que al trabajar en la plataforma de Google Colaboratory, se limita su uso con la finalidad de proveer sus servicios a usuarios de todo el mundo de forma gratuita.

La limitación de memoria RAM conllevó a que se trabajó con un tamaño excesivamente pequeño para las imágenes de entrada, de 64x64 píxeles para el modelo inicial y modelo 1, y de 128x128 para el modelo 2, modelo 3 y modelo 4. Con un alto grado de confianza se puede decir que esto provocó que no se obtengan buenos resultados puesto que las imágenes catalogadas como “melanoma” y “no melanoma” guardan bastantes similitudes, en consecuencia, a los modelos predictivos se les dificulta diferenciar aquellas características que definen una u otra categoría.

La limitación de memoria RAM llevó también a que no se puedan implementar un mayor número de capas convolucionales en el modelo inicial, esto podría haber ayudado a conseguir mejores resultados. Sin embargo, en el modelo 2 se implementaron varias capas convolutivas y a partir de la cuarta no se obtuvieron mejoras significativas; por lo tanto, se vuelve a afirmar el hecho de que las imágenes de entrada fueron demasiado pequeñas. El modelo 4 que se implementó a través de la herramienta Vision Transformer ViT demostró ser tan eficiente como las CNN puesto que obtuvo resultados parecidos al modelo 3; ambos modelos trabajaron exactamente con el mismo dataset.

Una ventaja de la arquitectura Vision Transformer ViT respecto a las CNN es su menor consumo de recursos computacionales, sin embargo, no se pudo incrementar el tamaño de las imágenes de entrada y permaneció en 128x128 píxeles.

Debido a las limitaciones de memoria RAM no se pudo combinar varios modelos y probar resultados.

## REFERENCIAS BIBLIOGRÁFICAS

- [1]E. Mezquita, “Detección de melanomas con inteligencia artificial,” *Diario Médico*, p. 12, 13 Enero 2014.
- [2]American Cancer Society, “¿Qué avances hay en las investigaciones sobre el cáncer de piel tipo melanoma?,” 14 Agosto 2019. [En línea]. Available: [https://www.cancer.org/es/cancer/cancer-de-piel-tipo-melanoma/acerca/nuevas-investigaciones.html#escrito\\_por](https://www.cancer.org/es/cancer/cancer-de-piel-tipo-melanoma/acerca/nuevas-investigaciones.html#escrito_por).
- [3]C. Marín, G. H. Alférez, J. Córdova y V. González, “Detection of melanoma through image recognition and artificial neural networks,” *IFMBE Proceedings*, vol. 51, pp. 832-835, 7-12 Junio 2015.
- [4]A. Adegun y S. Viriri, “Deep learning techniques for skin lesion analysis and melanoma cancer detection: a survey of state-of-the-art,” *Artificial Intelligence Review*, 27 Junio 2020.
- [5]J. J. Rangel-Cortes, J. S. Ruiz-Castilla, F. García-Lamont y J. Cervantes-Canales, “Redes Neuronales Convolucionales en la identificación de melanomas benignos y malignos,” *Revista Ibérica de Sistemas e Tecnologías de Informação*, nº E23, pp. 15-27, Octubre 2019.
- [6]S. Abhinav y J. Dheeba, “Convolutional Neural Network for Classifying Melanoma Images,” 2020.
- [7]C. Yu, S. Yang, W. Kim, J. Jung, K.-Y. Chung, S. W. Lee y B. Oh, “Acral melanoma detection using a convolutional neural network for dermoscopy images,” *PloS ONE*, vol. 13, nº 3, p. e0193321, 24 Abril 2018.
- [8]S. Nasiri, J. Helsper, M. Jung y M. Fathi, “DePicT Melanoma Deep-CLASS: a deep convolutional neural networks approach to classify skin lesion images,” *BMC Bioinformatics*, vol. 21, nº 84, 11 Marzo 2020.
- [9]M. S. S. Mahecha, O. J. S. Parra y J. B. Velandia, “Design of a System for Melanoma Detection Through the Processing of Clinical Images Using Artificial Neural Networks,” *Challenges and Opportunities in the Digital Era*, vol. 11195, pp. 605-616, 12 Octubre 2018.

- [10]A. M. Alqudah, H. Alquraan y I. A. Qasmieh, "Segmented and non-segmented skin lesions classification using transfer learning and adaptative moment learning rate technique using pretrained convolutional neural network," *Journal of Biomimetics, Biomaterials and Biomedical Engineering*, vol. 42, pp. 67-78, 2019.
- [11]H. El-Khatib, D. Popescu y L. Ichim, "Deep learning-based methods for automatic diagnosis of skin lesions," *Sensors*, vol. 20, n° 6, p. 1753, 2020.
- [12]J. A. Almaraz-Damian, V. Ponomaryov, S. Sadovnychiy y H. Castillejos-Fernandez, "Melanoma and nevus skin lesion classification using handcraft and deep learning feature fusion via mutual information measures," *Entropy*, vol. 22, n° 4, p. 484, 2020.
- [13]Q. Yuan y S. Tavildar, "An open solution to ISIC 2018 classification and segmentation challenges," 2018.
- [14]D. Hao, J. Y. Seok, D. Ng, N. K. Yuan y F. M, "ISIC Challenge 2018," 2018.
- [15]M. Molina-Moreno, I. González-Díaz y F. Díaz-de-María, "An elliptical shape-regularized convolutional neural network for skin lesion segmentation," 2018.
- [16]S. Zhou, Y. Zhuang y R. Meng, "Multi-category skin lesion diagnosis using dermoscopy images and deep CNN ensembles," 2019.
- [17]A. G. C. Pacheco, A. R. Alib y T. Trappenber, "Skin cancer detection based on deep learning and entropy to detect outlier samples," 2019.
- [18]V. Chouhan, "Skin lesion analysis towards melanoma detection with deep convolutional neural network," 2019.
- [19]T. Dat, D. T. Lan, T. T. H. Nguyen, T. T. N. Nguyen, H. P. Nguyen, L. Phuong y T. Z. Nguyen, "Ensembled skin cancer classification (ISIC 2019 challenge submission)," 2019.
- [20]P. Zhang, "MelaNet: a deep dense attention network for melanoma detection in dermoscopy images," 2019.
- [21]J. Xing, C. Zeng, H. Yangwen, W. Tao, Y. Mao, S. Wang, Y. Zheng y R. Wang, "Open-set recognition of dermoscopic images with ensemble of deep convolutional networks," 2019.
- [22]Z. M. Yousef y H. Motahari, "Skin lesion analysis towards melanoma detection using softmax ensemble model and sigmoid ensemble model," 2019.
- [23]S. Cohen y N. Shimoni, "TTA meta learning for anomaly detection on skin lesion," 2019.