

Universidad de La Habana  
Facultad de Matemática y Computación



## Título de la tesis

Autor:

**Deborah Famadas Rodríguez**

Tutores:

**Dr. Reinaldo Rodríguez Ramos, Universidad de la  
Habana**

**Dr. Yudivian Almeida, Universidad de la Habana**

Trabajo de Diploma  
presentado en opción al título de  
Licenciado en Ciencia de la Computación

Fecha

[github.com/deborahfam/Thesis](https://github.com/deborahfam/Thesis)

Dedicación

# Agradecimientos

Agradecimientos

# Opinión del tutor

Opiniones de los tutores

# Resumen

El Machine Learning (ML) o Aprendizaje Automático (AM) es un subcampo de la Inteligencia Artificial (IA), que se define como la capacidad de una máquina para aprender, mejorando su rendimiento. Los avances en las tecnologías de IA y ML prometen. El potencial de estas herramientas para generar conocimiento a partir de cantidades masivas de datos es enorme, pudiendo así ayudar a tomar decisiones, que incluirán intervenciones, y tratamientos de precisión contra el cáncer. La detección de cáncer de piel utilizando Redes Neuronales Convolucionales (CNN) es un área prometedora en el campo de la inteligencia artificial y el aprendizaje automático. Las CNN son especialmente adecuadas para el procesamiento de imágenes y han demostrado un alto rendimiento en tareas de clasificación y detección de objetos. Este trabajo de diploma presenta una serie de implementaciones, las que incluyen CNN, el modelo EfficientNetB1 y el uso de Learning Rate Adjustment a un Dataset de imágenes para lograr una detección efectiva de diferentes tipos de cáncer de piel a partir del entrenamiento de una IA. Estos resultados pueden sentar las bases para futuras investigaciones y mejoras en el campo de la detección temprana del cáncer de piel utilizando técnicas de aprendizaje automático y redes neuronales convolucionales.

# Abstract

Machine Learning (ML) or Machine Learning (ML) is a subfield of Artificial Intelligence (AI), which is defined as the ability of a machine to learn, improving its performance. Advances in AI and ML technologies hold promise. The potential of these tools to generate knowledge from massive amounts of data is enormous, thus being able to help make decisions, which will include interventions, and precision in cancer treatments. The detection of skin cancer using Convolutional Neural Networks (CNN) is a promising area in the field of artificial intelligence and machine learning. CNNs are particularly well suited for image suitable for image processing and have demonstrated high performance in object classification and object detection tasks. This diploma work presents a number of implementations, which include CNNs, the EfficientNetB1 model, and the use of Learning Rate Adjustment to an image dataset to achieve effective detection of different types of skin cancer by training an from the training of an AI. These results can lay the foundation for future research and improvements in the field of early skin cancer detection using machine learning techniques.

# Índice general

<b>Introducción</b>	<b>1</b>
<b>1. Estado del Arte</b>	<b>2</b>
<b>2. Propuesta</b>	<b>4</b>
2.1. Metodología . . . . .	4
2.2. Preparación y Carga de Datos . . . . .	4
2.2.1. Transformación de Datos . . . . .	4
2.2.2. Modelación y división del conjunto de datos . . . . .	5
2.3. Generadores de Datos y Preprocesamiento . . . . .	5
2.4. Diseño y Entrenamiento del Modelo . . . . .	6
2.4.1. Capas Adicionales y Regularización . . . . .	6
2.4.2. Optimización . . . . .	7
2.4.3. Ajuste Dinámico de la Tasa de Aprendizaje . . . . .	7
<b>3. Detalles de Implementación y Experimentos</b>	<b>8</b>
3.1. Resultados . . . . .	8
3.1.1. Estadísticas básicas . . . . .	8
3.1.2. Estadísticas de eficacia . . . . .	9
3.1.3. Estadísticas de aprendizaje . . . . .	11
3.2. Ventajas y Desventajas . . . . .	11
<b>Conclusiones</b>	<b>12</b>
<b>Recomendaciones</b>	<b>13</b>
<b>Bibliografía</b>	<b>14</b>

# Índice de figuras

3.1. Estadísticas básicas de algunas iteraciones del modelo. . . . .	9
3.2. Estadísticas de eficacia del modelo al estimar los resultados en el conjunto de pruebas. . . . .	10
3.3. Estadísticas de aprendizaje a lo largo del proceso de entrenamiento (en rojo el proceso de entrenamiento en el conjunto de entrenamiento, en verde en el conjunto de validación). . . . .	11



## Ejemplos de código

# Introducción

# Capítulo 1

## Estado del Arte

El cáncer de piel es el más frecuente del mundo *annadir cita* . A diferencia de muchos otros tipos de cáncer que se desarrollan internamente, el cáncer de piel se forma externamente y suele ser visible, lo que lo convierte en el cáncer que se ve". Esta visibilidad subraya la importancia de los exámenes de la piel, tanto realizados por uno mismo como por dermatólogos. El diagnóstico y el tratamiento oportunos son primordiales. La mayoría de los casos de cáncer de piel son tratables si se detectan precozmente, lo que subraya la necesidad de identificar la enfermedad con prontitud. La detección precoz no sólo salva vidas, sino que también evita que la enfermedad se vuelva peligrosa, desfigurante o incluso mortal.

## Teledermatología

Las imágenes dermatológicas se utilizan para detectar el cáncer de piel, mediante el análisis de las lesiones cutáneas y la pigmentación. Esta es una técnica utilizada para examinar las lesiones cutáneas mediante un instrumento manual denominado dermatoscopio, que amplía la piel hasta 10 veces. Con el avance de la tecnología, la atención sanitaria ha experimentado una transformación. Se ha producido un aumento de dispositivos de diagnóstico y se ha desplazado hacia el desarrollo de nuevas competencias en estadística y psicología de la toma de decisiones médicas.

La teledermatología comenzó a ganar terreno en la década de 2000 como una manera de proporcionar consulta dermatológica a distancia. Un trabajo pionero es el de Whited et al Whited y col. 2002, que exploró la eficacia de la teledermatología en 2002. El estudio constató que la teledermatología reducía significativamente el tiempo de intervención de los pacientes, con una mediana de 5 días para las derivaciones teledermatológicas frente a 28 días para las derivaciones tradicionales. El estudio también constató que la teledermatología era especialmente eficaz para pacientes con afecciones urgentes o semi-urgentes. El artículo destaca las posibles ventajas de la

tele dermatología para mejorar el acceso a la asistencia y reducir los tiempos de espera de los pacientes.

# Capítulo 2

## Propuesta

### 2.1. Metodología

En este capítulo se propone una método para la detección de cáncer de piel mediante el aprendizaje automático. Este enfoque de solución se basa en la implementación de una red neuronal profunda para analizar y categorizar automáticamente imágenes, con el objetivo de asistir en la detección de posibles casos malignos y mejorar la eficiencia del proceso diagnóstico.

A continuación, se introducen algunos de los conceptos básicos de los métodos o técnicas que se emplearan a lo largo del capítulo.

### 2.2. Preparación y Carga de Datos

Los datos utilizados como medio de aprendizaje para este proyecto son: Imágenes y Metadatos. Los datos son extraídos del Dataset: HAM1000-segmentation-and-classification *Conjunto de datos HAM10000* 2023. El mismo contiene imágenes tomadas de varios tipos de cáncer de piel y contiene además un archivo Excel formato csv de Metadatos. Los metadatos contienen información relacionada con cada imagen con un formato de *one hot encoding*.

#### 2.2.1. Transformación de Datos

Para optimizar la eficiencia del algoritmo, se procesan las imágenes realizando diversas modificaciones, que incluyen: eliminación de cabello, luces y sombras, división de canales y aplicación de un leve desenfoque.

Los metadatos, asociados a la clasificación, están etiquetados utilizando el método de *one hot encoding*. Fue necesario convertirlos a un formato *categorico* para su

procesamiento..

El *one hot encoding* es una técnica de procesamiento de datos en la cual cada valor categórico se representa mediante un vector binario cuyo tamaño corresponde al número de categorías posibles. En dicho vector, todos los elementos son cero, salvo el correspondiente a la categoría del valor, que es uno *One-hot encoding* 2023.

Se define *categorico* como el formato en el que a cada elemento se le asigna el nombre de una categoría como propiedad.

### 2.2.2. Modelación y división del conjunto de datos

Los datos se modelan a partir de un Dataframe de Pandas *Dataframes de Pandas* 2023. Los mismo son divididos luego en 3 conjuntos: Entrenamiento, Validación y Prueba. Estas divisiones son necesarias para que el modelo pueda aprender y ser evaluado correctamente.

Aquí, se define que el 95% de los datos se utilizará para entrenamiento y la cantidad restante se considera en un conjunto dummy, el cual luego es dividido en validación y prueba. En esta última división al conjunto de validación le corresponde una cantidad del 50% del conjunto dummy (lo que es equivalente al 2.5% del df original) y el conjunto de prueba el 50% restante (también equivalente al 2.5% del df original). Además se mezclan aleatoriamente los datos y se utiliza una variable fija para garantizar que la división sea reproducible.

En el contexto del Machine Learning (Aprendizaje de Máquinas), un DataFrame de Pandas es una estructura de datos bidimensional, similar a una tabla, proporcionada por la biblioteca Python *Pandas*. Facilita la manipulación y análisis de grandes conjuntos de datos, permitiendo una organización clara en forma de filas y columnas.

## 2.3. Generadores de Datos y Preprocesamiento

Se detectó que el conjunto de datos estaba altamente desequilibrado. Por lo tanto, limitamos el número máximo de muestras por clase a 300 para equilibrar el conjunto de datos.

Se empleó *ImageDataGenerator* *ImageDataGenerator* 2023 para cargar, transformar y generar lotes de imágenes durante el entrenamiento. El *ImageDataGenerator* es una herramienta ofrecida por la biblioteca *Keras*, diseñada específicamente para el procesamiento y manejo de imágenes en tareas de Machine Learning. Se inicializa con varias transformaciones (como rotación, desplazamiento y zoom) para aplicar automáticamente estas transformaciones a las imágenes a medida que se cargan devolver lotes de imágenes listos para el entrenamiento o la evaluación del modelo, lo cual permite alimentar la red neuronal de manera eficiente *Aumentación de datos* 2023 .

Además, implementamos *Class Weighting*. Este método se refiere a la asignación de pesos diferenciados a cada clase durante el proceso de entrenamiento del modelo debido conjuntos de datos desequilibrados, con el objetivo de reforzar la señal de entrenamiento de las clases menos representadas, o sea, las clases/categorías con menos cantidad de datos.

## 2.4. Diseño y Entrenamiento del Modelo

Se utiliza la arquitectura de red neural denominada *EfficientNetB1* *EfficientNetB1* s.f. . Esta arquitectura forma parte de la familia EfficientNet, que está diseñada para proporcionar alta precisión mientras se mantiene un tamaño de modelo y una complejidad computacional eficientes.

Para aprovechar los conocimientos previos y acelerar el entrenamiento, se carga el modelo *EfficientNetB1* pre-entrenado con los pesos obtenidos de entrenar en el dataset *ImageNet*, que es una amplia base de datos de imágenes utilizada comúnmente para entrenamiento y *benchmarking* en tareas de visión por computadora. Utilizar un modelo pre-entrenado permite aprovechar las características que ya ha aprendido de este amplio conjunto de datos, facilitando su adaptación al dataset.

Se utiliza el *EfficientNetB1* sin incluir su capa superior, con el objetivo de añadir y personalizar capas adicionales. Tras obtener la salida del modelo base, se aplica una serie de transformaciones, incluyendo normalización por lotes, una capa densa con regularizaciones, una técnica de *Dropout* para prevenir el sobre-ajuste y una capa de optimización.

### 2.4.1. Capas Adicionales y Regularización

Aunque el modelo EfficientNetB1 es poderoso por sí solo, no está perfectamente adaptado a nuestro problema específico. Para ajustarlo a nuestras necesidades, se añaden capas adicionales:

1. Normalización por Lotes: Esta capa busca estandarizar las activaciones del modelo para cada lote de entrenamiento. La misma regulariza el modelo y, al mismo tiempo, suele acelerar el entrenamiento ya que permite utilizar tasas de aprendizaje más altas. [*Regularización* 2023]
2. Densa: Se introduce una capa completamente conectada con 256 neuronas. Esta capa tiene una activación ReLU y utiliza regularización L1 y L2. La regularización L1 y L2 penaliza los pesos grandes en la red, ayudando a evitar el sobre-ajuste y garantizando que el modelo se generalice bien a nuevos datos. [*Capas densas* 2023]

3. Dropout: Esta técnica de regularización desactiva aleatoriamente una fracción (en este caso, el 45 %) de las neuronas durante el entrenamiento. Esta aleatoriedad ayuda a evitar la dependencia excesiva en cualquier neurona individual, lo que a su vez ayuda a prevenir el sobre-ajuste. *Dropout* 2023
4. Capa de Salida: Finalmente, se agrega una capa densa que tiene un número de neuronas igual al número de clases en nuestro problema. La activación *softmax* se utiliza aquí para transformar las salidas en probabilidades de clase.

### 2.4.2. Optimización

Para el proceso de entrenamiento, se utiliza el optimizador *Adamax* *Adamax* 2023. Este es una variante del conocido optimizador Adam, que combina las ventajas de los métodos adaptativos de tasa de aprendizaje con una implementación más robusta en presencia de momentos espurios.

Se configura el modelo para minimizar la categorical crossentropy, que es una medida común de error para problemas de clasificación, y se rastrea la precisión como métrica principal.

### 2.4.3. Ajuste Dinámico de la Tasa de Aprendizaje

Una de las características más innovadoras de este enfoque es la inclusión de un "Callback" personalizado para ajustar dinámicamente la tasa de aprendizaje durante el entrenamiento. Este ajuste se basa en la precisión del entrenamiento y la pérdida de validación. En resumen, si el modelo no mejora lo suficientemente rápido (según ciertos criterios preestablecidos), la tasa de aprendizaje se reduce, con la esperanza de mejorar la convergencia del modelo *Tasa de aprendizaje* 2023.

Este enfoque puede ser especialmente útil cuando nos enfrentamos a superficies de pérdida complejas con múltiples mínimos locales; al variar la tasa de aprendizaje, es más probable que el modelo salga de mínimos locales subóptimos y encuentre una solución mejor.



## Capítulo 3

# Detalles de Implementación y Experimentos

---

### 3.1. Resultados

Tras la implementación del modelo, los resultados obtenidos de este modelo de aprendizaje automático son muy útiles y eficientes en el campo de la detección de melanomas. Con una eficacia de validación cercana al 78,9%, el modelo es capaz de detectar la presencia de melanoma con una precisión razonable. Los resultados son un paso significativo hacia la automatización del diagnóstico del cáncer de piel, que tradicionalmente ha dependido de la inspección visual por parte de expertos humanos.

La eficacia del modelo también puede mejorarse aún más optimizando su consumo de tiempo y memoria. Esta optimización será esencial para convertir los resultados del modelo en una solución a gran escala que pueda ser utilizada por profesionales sanitarios y pacientes de todo el mundo. Es importante señalar que el entrenamiento inicial del modelo es un proceso que se realiza una sola vez y no necesita ejecutarse continuamente. Por lo tanto, si se optimiza el proceso de entrenamiento, la eficacia del modelo puede mejorar considerablemente.

#### 3.1.1. Estadísticas básicas

Estos resultados de la tabla siguiente corresponden a la evaluación del modelo a lo largo de 40 epochs (o iteraciones) de entrenamiento. Cada fila representa una época y se presentan las siguientes métricas:

**loss:** Es una medida de la diferencia entre las predicciones del modelo y los valores

reales de los datos de entrenamiento en esa época. El objetivo del entrenamiento es minimizar esta métrica.

**acc (accuracy):** Es la proporción de predicciones correctas realizadas por el modelo sobre el conjunto de datos de entrenamiento en esa época.

**v loss (vl) (pérdida de validación):** Es la pérdida en el conjunto de datos de validación, que es un conjunto de datos independiente que no se utiliza para entrenar el modelo, sino para evaluar su generalizabilidad.

**v acc (precisión de validación):** Es la precisión en el conjunto de datos de validación.

E	Loss	Acc	V loss	V acc	LR	M	Batch
1	9.535	37.735	9.396	38.4	$10^{-2}$	Acc	1337.97
2	7.791	66.306	7.832	61.6	$10^{-2}$	Acc	414.44
3	6.852	79.624	7.042	66.8	$10^{-2}$	Acc	291.03
4	6.191	85.544	6.398	67.6	$10^{-2}$	Acc	281.46
5	5.594	90.894	5.772	72.8	$10^{-2}$	Acc	288.00
6	5.111	92.943	5.311	74.4	$10^{-2}$	Vl	280.35
7	4.653	95.788	4.916	74.4	$10^{-2}$	Vl	274.80
8	4.268	95.788	4.631	73.6	$10^{-2}$	Vl	280.88
9	3.888	97.097	4.164	76.8	$10^{-2}$	Vl	302.48
10	3.547	97.553	3.840	76.4	$10^{-2}$	Vl	275.76
11	3.233	98.349	3.622	75.2	$10^{-2}$	Vl	281.52
...	...	...	...	...	...	...	...
31	0.592	99.488	1.133	77.2	.00025	Vl	420.26

Figura 3.1: Estadísticas básicas de algunas iteraciones del modelo.

En general, los resultados muestran que el modelo mejora con el tiempo, ya que la pérdida disminuye y la precisión aumenta tanto en el conjunto de datos de entrenamiento como en el de validación.

### 3.1.2. Estadísticas de eficacia

La matriz de confusión proporciona información valiosa sobre el rendimiento del modelo en relación de Actual/Predicho, en términos de su capacidad para clasificar correctamente cada una de las siete clases de cáncer de piel. La diagonal principal de la matriz representa los verdaderos positivos (TP), el número de casos en los que

el modelo ha predicho correctamente la clase correspondiente. Los valores fuera de la diagonal principal indican errores de clasificación. Estos errores pueden ser falsos positivos (FP), falsos negativos (FN) o verdaderos negativos (TN), dependiendo de la posición en la matriz. AKIEC, BCC, BKL, DF, MEL, NV y VASC son los acrónimos de los distintos tipos de lesiones cutáneas utilizados en la tarea de clasificación multiclase.

Algunos patrones que pueden observarse en esta matriz son:

- . El modelo parece tener dificultades para clasificar las muestras de la clase AKIEC, ya que sólo predice correctamente 5 de 7 casos. Además, en la fila AKIEC hay un falso positivo (el modelo predijo la clase BKL en un caso que en realidad era AKIEC) y un falso negativo (el modelo predijo la clase MEL en un caso que en realidad era AKIEC).
- . En la clase BCC, el modelo predice correctamente todos los casos.
- . Las muestras de las clases BKL y MEL parecen ser las más difíciles de clasificar, ya que hay varios falsos positivos y falsos negativos en esos rangos.
- . La mayoría de las predicciones se clasifican correctamente como NV, que es la clase más común en los datos.

Actual	AKIEC	5	0	1	0	1	0	0
	BCC	0	10	0	0	0	0	0
	BKL	4	1	14	1	5	5	0
	DF	0	0	0	3	0	0	0
	MEL	3	1	1	1	21	7	1
	NV	3	3	2	1	10	141	3
	VASC	0	0	0	0	0	0	3
		AKIEC	BCC	BKL	DF	MEL	NV	VASC
		Predicted						

Figura 3.2: Estadísticas de eficacia del modelo al estimar los resultados en el conjunto de pruebas.

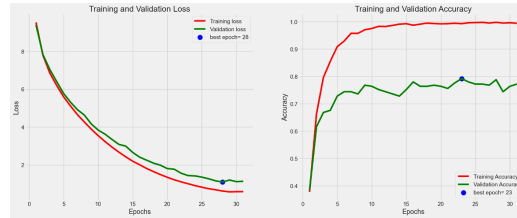


Figura 3.3: Estadísticas de aprendizaje a lo largo del proceso de entrenamiento (en rojo el proceso de entrenamiento en el conjunto de entrenamiento, en verde en el conjunto de validación).

### 3.1.3. Estadísticas de aprendizaje

La imagen anterior muestra un gráfico del curso temporal de las épocas realizadas, mostrando cómo el algoritmo iba obteniendo resultados más precisos y disminuyendo el error.

## 3.2. Ventajas y Desventajas

Las ventajas y desventajas del proceso son las siguientes. Ventajas:

1. Selección y representación del problema y de los datos. Este proceso ayuda a comprender mejor el problema en cuestión y a identificar los datos relevantes necesarios para entrenar el modelo.
2. Preprocesamiento de datos e imágenes. Eliminar el ruido y los artefactos de las imágenes mejora la calidad de los datos y puede mejorar el rendimiento del modelo. La selección de canales adecuados y la reducción del ruido también pueden mejorar el rendimiento del modelo.

Desventajas:

1. Modificación y procesamiento de datos. Si se utilizan conjuntos de datos inadecuados o se normalizan incorrectamente, el modelo puede no ser preciso o útil.

# Conclusiones

Los resultados obtenidos muestran una alta eficacia (79% aproximadamente) que puede estar sujeta a futuras mejoras del modelo y del algoritmo. En general, los modelos de detección de cáncer de piel pueden tener una precisión del 85% al 95% o superior, dependiendo de la complejidad del problema, la calidad de los datos y la selección del algoritmo utilizado. Sin embargo, el 78% es un porcentaje elevado para los algoritmos de aprendizaje automático, y con algunas mejoras en el modelo puede incrementarse. También se desprende de este resultado la intención de convertir este proyecto en un producto a gran escala para ser utilizado, inicialmente, por los profesionales sanitarios de atención primaria, como prueba complementaria de alta efectividad a la hora de derivar a un paciente con una lesión cutánea al área de oncología. Con la premisa de convertir los resultados obtenidos en un producto al servicio de la sociedad, el equipo investigador pretende seguir investigando y desarrollando una solución mejor.

# Recomendaciones

Recomendaciones

# Bibliografía

- Adamax. (2023). (Vid. pág. 7).
- Aumentación de datos. (2023). (Vid. pág. 5).
- Capas densas. (2023). (Vid. pág. 6).
- Conjunto de datos HAM10000. (2023). (Vid. pág. 4).
- Dataframes de Pandas. (2023). (Vid. pág. 5).
- Dropout. (2023). (Vid. pág. 7).
- EfficientNetB1. (s.f.). (Vid. pág. 6).
- ImageDataGenerator. (2023). (Vid. pág. 5).
- One-hot encoding. (2023). (Vid. pág. 5).
- Regularización. (2023). (Vid. pág. 6).
- Tasa de aprendizaje. (2023). (Vid. pág. 7).
- Whited, J. D., Hall, R. P., Foy, M. E., Marbrey, L. E., Grambow, S. C., Dudley, T. K., Datta, S., Simel, D. L., & Oddone, E. Z. (2002). Teledermatology's impact on time to intervention among referrals to a dermatology consult service. *Telemedicine Journal and e-Health*, 8(3), 313-321 (vid. pág. 2).