

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

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

Introducción	1
1. Estado del Arte	2
2. Propuesta	11
2.1. Metodología	11
2.2. Preparación y carga de datos	11
2.2.1. Transformación de datos	11
2.2.2. Modelación y división del conjunto de datos	12
2.3. Generadores de datos y preprocesamiento	12
2.4. Diseño y entrenamiento del modelo	13
2.4.1. Capas adicionales y regularización	13
2.4.2. Optimización	14
2.4.3. Ajusted dinámico de la tasa de aprendizaje	14
3. Detalles de Implementación y Experimentos	15
3.1. Resultados	15
3.1.1. Estadísticas básicas	15
3.1.2. Estadísticas de eficacia	16
3.1.3. Estadísticas de aprendizaje	18
3.2. Ventajas y Desventajas	18
Conclusiones	19
Recomendaciones	20
Bibliografía	21

Índice de figuras

3.1.	Estadísticas básicas de algunas iteraciones del modelo.	16
3.2.	Estadísticas de eficacia del modelo al estimar los resultados en el conjunto de pruebas.	17
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).	18

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

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

Técnicas Tempranas de Clasificación de Imágenes

Antes del auge de las redes neuronales y del aprendizaje profundo, había varias técnicas tradicionales utilizadas para la clasificación de imágenes. Estas técnicas se basaban en la extracción manual de características y en métodos de aprendizaje automático clásicos. La clave de la clasificación tradicional de imágenes radica en la extracción efectiva de características. Estas características deberían ser invariantes a pequeñas transformaciones y variaciones en las imágenes. Principalmente se tenía en cuenta:

1. El Histograma de Gradientes Orientados (HOG): Esta técnica captura la distribución de direcciones de gradientes (orientaciones) en una imagen. *add reference*
2. Scale-Invariant Feature Transform (SIFT): Es utilizado para la detección y descripción de características locales en imágenes. El descriptor SIFT es invariante a cambios de escala, rotación y parcialmente invariante a cambios de punto de vista y afín. *add reference*
3. Binary Robust Independent Elementary Features (BRIF): Es un descriptor de características que utiliza un conjunto de pruebas binarias para describir una región de interés en una imagen. *add reference*
4. Color Histograms: describen la distribución de colores en una imagen y son útiles para tareas en las que el color es un indicador importante. *add reference*

En 1974, se publicó un artículo en *IEEE Transactions on Automatic Control* relacionado con el reconocimiento de patrones y el aprendizaje automático, lo que indica que ya en la década de 1970 se estaban explorando estas técnicas para el reconocimiento de patrones en datos.

Con el tiempo, los investigadores comenzaron a aplicar algoritmos de aprendizaje automático tradicionales como Máquinas de Soporte Vectorial (SVM) y Árboles de Decisión para clasificar imágenes basándose en características extraídas manualmente. Esta ha logrado grandes avances en las últimas décadas en las siguientes tres áreas: (1) desarrollo y uso de algoritmos de clasificación avanzados, como algoritmos de clasificación de subpíxeles, por campo y basados en el conocimiento; (2) uso de múltiples funciones de teledetección, incluida información espectral, espacial, multi-temporal y multisensor; y (3) incorporación de datos auxiliares en los procedimientos de clasificación, incluidos datos como topografía, suelo, carreteras y datos censales. La

evaluación de la precisión es una parte integral de un procedimiento de clasificación de imágenes. La evaluación de la precisión basada en la matriz de errores es el enfoque más comúnmente empleado para evaluar la clasificación por píxel, mientras que los enfoques difusos están ganando atención para evaluar los resultados de la clasificación difusa.

Otro algoritmo de machine learning utilizado en las primeras etapas de la clasificación de imágenes es el árbol de decisión. Los árboles de decisión son un tipo de algoritmo de aprendizaje supervisado que se utiliza para la clasificación y la regresión. En el contexto de la clasificación de imágenes, un árbol de decisión podría usarse para tomar decisiones basadas en ciertos criterios, como el color, la forma, el tamaño, etc., para clasificar una imagen en una categoría específica.

Con el tiempo, se introdujeron otros algoritmos de aprendizaje automático como los Clasificadores de Bosques Aleatorios, y Clasificadores de Bayes Ingenuos, que también se aplicaron a la clasificación de imágenes utilizando características manuales. A medida que se avanzaba, se desarrollaron técnicas más avanzadas como las Redes Neuronales Artificiales y, eventualmente, las Redes Neuronales Convolucionales (CNN), que mostraron una capacidad significativa para aprender características directamente de los datos, reduciendo así la necesidad de extracción manual de características.

Introducción y Desarrollo de Redes Neuronales

Las primeras aplicaciones de redes neuronales en imágenes médicas se orientaron hacia el análisis y clasificación de dichas imágenes para apoyar en diagnósticos y tratamientos.

Análisis de imágenes médicas

El análisis de imágenes médicas mediante redes neuronales se ha enfocado en campos como ultrasonido, resonancia magnética, medicina nuclear y radiología, permitiendo la identificación y clasificación de patologías o condiciones específicas. Además, estas tecnologías han encontrado aplicaciones en áreas como la oftalmología, para el diagnóstico de enfermedades oculares a partir de imágenes de retina, y en la cardiología, para la evaluación de imágenes de ecocardiogramas.

Ultrasonido : La aplicación de redes neuronales en el ultrasonido ha mejorado significativamente la detección y caracterización de lesiones. Mediante el análisis automatizado de imágenes, se pueden identificar patrones sutiles que podrían pasarse por alto en la evaluación humana.

Resonancia Magnética (RM) : En el ámbito de la RM, las redes neuronales han permitido un avance en la segmentación de imágenes, mejorando la precisión

en la identificación de estructuras anatómicas y patológicas. Esto es especialmente relevante en la neurología, donde se utilizan para detectar y monitorizar enfermedades como la esclerosis múltiple.

Medicina Nuclear : La aplicación de redes neuronales en medicina nuclear, como en la tomografía por emisión de positrones (PET), ha mejorado la interpretación de los estudios, permitiendo una detección más temprana y precisa de diversas patologías, incluyendo el cáncer.

Radiología : En radiología, las redes neuronales han permitido el desarrollo de sistemas de ayuda al diagnóstico, especialmente en la detección de patologías pulmonares y mamarias. Estos sistemas pueden identificar anomalías en las radiografías y mamografías con un nivel de precisión que se acerca al de los especialistas humanos.

Clasificación y Reconocimiento de patrones en imágenes médicas

Cuando se trata de la clasificación de imágenes, es esencial considerar que nuestro sistema visual humano (SVH) primero recibe las ondas electromagnéticas que pertenecen al espectro visible y luego las interpreta en el cerebro. Sin embargo, en el campo de la visión artificial, cuando introducimos una imagen en un ordenador, lo que se interpreta es una matriz de números generalmente en el rango de $[0,255]$ y con tres dimensiones en caso de que sea una imagen a color (RGB). Como resultado, se puede notar una gran brecha entre el significado semántico de la clase asociada a una imagen y los valores de píxeles que la componen, lo que hace que la tarea de clasificación sea compleja para sistemas artificiales.

En el ámbito de la medicina, se ha observado que la mayoría de los sistemas artificiales de diagnóstico, en un punto, toman decisiones que están cada vez menos relacionadas con la apariencia física de la imagen tal como la vería un radiólogo. En su lugar, estos sistemas se basan en los detalles del patrón matemático de las características físicas individuales de la imagen, que son extraídas por un sistema de visión artificial o un radiólogo, para tomar su decisión final. Estos patrones matemáticos han sido objeto de estudio durante décadas por científicos que han utilizado diversos métodos analíticos, incluyendo las redes neuronales.

Para abordar esta brecha entre la representación de la imagen y su significado, se han desarrollado varios tipos de algoritmos de clasificación. Algunos de estos algoritmos se basan en la detección de bordes, como el algoritmo de Canny Sin embargo, estos algoritmos son robustos cuando se trata de identificar una clase específica, pero si se desea clasificar una clase diferente, es necesario crear un nuevo modelo desde cero.

Otros algoritmos, como el k-Nearest Neighbors (KNN), se basan en medir la diferencia entre los valores de píxeles o distancias para comparar la similitud entre imágenes. Este enfoque es simple, pero requiere una configuración óptima de los hiperparámetros para obtener buenos resultados, lo que hace que estos algoritmos sean dependientes del problema que se está abordando.

Sin embargo, enfrentaron una serie de limitaciones significativas que restringieron su capacidad para abordar tareas complejas y desafiantes. La escasez de datos adecuados, la falta de recursos computacionales avanzados, arquitecturas simples, dificultades en el entrenamiento, generalización limitada, problemas de gradiente, falta de interpretabilidad y largos tiempos de entrenamiento fueron obstáculos clave en su desarrollo inicial.

Estas limitaciones condujeron a la necesidad de innovación y evolución en el campo de las redes neuronales. Esta necesidad llevó al surgimiento de un enfoque revolucionario en la forma de las Redes Neuronales Convolucionales (CNN), que abordaron muchas de estas limitaciones de manera efectiva.

Advenimiento de las Redes Neuronales Convolucionales (CNN)

Mediante el uso de redes neuronales convolucionales, para facilitar la identificación de características relevantes en las imágenes médicas que pudieran ser indicativas de alguna condición médica particular las redes neuronales convolucionales han probado ser muy eficientes. Estas aprovecharon estructuras específicas de datos, como imágenes, para capturar patrones con mayor precisión, redujeron la carga computacional y mejoraron la generalización y la interpretabilidad. Gracias a estas mejoras, las CNN han sido fundamentales en el avance de la visión por computadora y han demostrado ser herramientas poderosas en una amplia variedad de aplicaciones, desde reconocimiento de objetos hasta diagnóstico médico.

Una red neuronal convolucional está formada por diferentes capas, entre ellas las principales son las capas convolucionales, las capas de max-pooling, y las capas completamente conectadas. La capa convolucional tiene como objetivo realizar la convolución a la imagen de entrada, para extraer sus características. Realizar una convolución a una imagen, consiste en filtrar dicha imagen utilizando una máscara o ventana. La máscara se va desplazando por toda la imagen, multiplicándose de forma matricial.

Una revisión sistemática de 2018 destacó el uso de CNNs para clasificar lesiones cutáneas. En este estudio, se revisaron 13 artículos que utilizaban CNNs para la clasificación de lesiones cutáneas, y se encontró que los métodos más comunes eran aquellos que usaban una CNN ya entrenada con otro gran conjunto de datos y luego optimizaban sus parámetros para la clasificación de lesiones cutáneas. Las CNNs mos-

traron un alto rendimiento como clasificadores de lesiones cutáneas, aunque se señaló la dificultad de comparar diferentes métodos de clasificación debido a la utilización de conjuntos de datos no públicos en algunos casos.

Sin embargo, las técnicas de aprendizaje automático, incluidos los modelos de aprendizaje profundo, han surgido como herramientas prometedoras para el análisis de imágenes médicas, especialmente imágenes de lesiones cutáneas, para la detección del cáncer de piel melanoma. ^{.A} Deep Learning Approach to Skin Cancer Detection in Dermoscopy Images.^{es} un proyecto que sugiere un modelo de aprendizaje profundo para identificar cáncer de piel usando imágenes de lesiones en la piel. En esta investigación, se utilizaron 3400 imágenes de la colección de fotos dermatoscópicas HAM10000, que incluyen lesiones tanto melanoma como no melanoma. Se armó una red neuronal convolucional profunda para separar las imágenes en benignas y malignas. Se aplicó una técnica de aprendizaje transferido usando AlexNet como el modelo ya entrenado. El modelo sugerido toma la imagen tal cual es como entrada y descubre por su cuenta características valiosas en la imagen para clasificarla. Así, se salta los procesos complicados de segmentación de lesiones y extracción de características.

Un notable estudio publicado el 28 de octubre de 2022 publicado en Nature "Skin lesion classification of dermoscopic images using machine learning and convolutional neural network" se utilizó un subconjunto del conjunto de datos HAM10000 para clasificar lesiones cutáneas de imágenes dermatoscópicas y demostró la eficacia del uso de aprendizaje automático y redes neuronales convolucionales (CNN) en la clasificación de lesiones cutáneas a partir de imágenes dermatoscópicas. La metodología propuesta mostró resultados prometedores a la hora de distinguir entre lesiones malignas y benignas. El estudio compara la precisión de clasificación de distintos algoritmos de aprendizaje automático y modelos CNN. Se concluye que los modelos CNN proporcionan una mayor precisión en comparación con otros algoritmos de aprendizaje automático. En el sistema propuesto, alcanzan una precisión del 95,18% con el modelo CNN.

Un artículo publicado en PubMed ^{.A} An efficient deep learning-based skin cancer classifier for an imbalanced dataset" propuso un eficaz clasificador de cáncer de piel basado en el aprendizaje profundo empleando el conjunto de datos HAM10000. Se utilizaron varios modelos de aprendizaje profundo como AlexNet, InceptionV3 y RegNetY-320 para clasificar el cáncer de piel. El rendimiento del marco propuesto fue mejor que el de los métodos convencionales. La precisión, la puntuación F1 y el valor de la curva ROC obtenidos con el marco propuesto fueron del 91 %, el 88,1 % y el 0,95.

Técnicas Modernas y Avances Recientes

Uso de arquitecturas CNN avanzadas (ej. ResNet, Inception, etc.). Técnicas de transferencia de aprendizaje. Implementaciones de detección en tiempo real y aplicaciones móviles. Integración de métodos de inteligencia artificial con otras modalidades de diagnóstico.

El avance en la detección de cáncer mediante el uso de la inteligencia artificial (IA) y el aprendizaje automático (ML) ha sido significativo en las últimas décadas, el mismo, ha demostrado un rendimiento excepcional en tareas de reconocimiento de imágenes, que es fundamental en la detección del cáncer de piel. Se llevó a cabo una revisión exhaustiva para evaluar el impacto de las técnicas de aprendizaje profundo en la detección precoz, en la que se analizaron diversos resultados de investigación y se presentaron mediante herramientas, gráficos, tablas y marcos para comprender mejor las técnicas predominantes en este campo

Además, se introdujo un enfoque innovador de combinación de técnicas de aprendizaje automático y aprendizaje profundo para abordar el problema de la detección del cáncer de piel. El modelo de aprendizaje profundo en este estudio empleó redes neuronales de última generación para extraer características de las imágenes, mientras que el modelo de aprendizaje automático procesó estas características de la imagen para clasificar las lesiones cutáneas.

En otro artículo publicado en PubMed "Design and validation of a new machine-learning-based diagnostic tool for the differentiation of dermatoscopic skin cancer images" Tadjerian y col. 2023 se presentó un enfoque metodológico para mejorar el diagnóstico de las lesiones cutáneas pigmentadas utilizando imágenes dermatoscópicas del conjunto de datos HAM10000. Este conjunto de datos, una colección de 10015 imágenes dermatoscópicas recogidas a lo largo de 20 años, se utilizó para analizar lesiones cutáneas pigmentadas. El modelo obtiene los mejores resultados en la detección de lesiones de nevos melanocíticos, con una puntuación F1 de 0,93. La puntuación F1 para queratosis actínica, carcinoma basocelular, queratosis benigna, dermatofibroma, melanoma y lesiones vasculares fue consecutivamente de 0,63, 0,72, 0,70, 0,54, 0,58 y 0,80.

Retos y Limitaciones Actuales

Dificultades en la Interpretación de Resultados y la Falta de Explicabilidad de los Modelos

Uno de los desafíos más significativos en el uso de algoritmos avanzados de aprendizaje automático y aprendizaje profundo es la interpretación y explicabilidad de sus resultados. Aunque estos modelos pueden lograr un alto grado de precisión, a

menudo operan como cajas negras", lo que significa que sus procesos internos y la forma en que llegan a una conclusión específica no son transparentes ni fácilmente comprensibles para los humanos. Esta falta de transparencia puede ser un obstáculo importante en el ámbito clínico, donde los profesionales de la salud necesitan comprender el razonamiento detrás de un diagnóstico para confiar y actuar según los resultados proporcionados por estos sistemas.

Problemas Relacionados con el Sobreajuste y la Generalización de los Modelos

Otro reto importante es el sobre-ajuste y la generalización de los modelos. El sobre-ajuste ocurre cuando un modelo se ajusta demasiado a los datos de entrenamiento, perdiendo la capacidad de generalizar a nuevos datos. Este problema es particularmente prevalente en situaciones donde los conjuntos de datos de entrenamiento son limitados o no representativos de la variabilidad real de casos clínicos. Por ejemplo, la falta de diversidad en los tipos de piel y las características de las lesiones en los conjuntos de datos puede llevar a modelos que no se desempeñan bien en poblaciones no representadas durante el entrenamiento.

Limitaciones de los Métodos Tradicionales de Aprendizaje Automático en Diagnóstico Clínico

Aunque los métodos tradicionales de aprendizaje automático, como las Máquinas de Vectores de Soporte (SVM), XGBoost y árboles de decisión, han sido útiles en la clasificación del cáncer de piel, enfrentan limitaciones significativas en el contexto clínico. Estos métodos requieren la extracción manual de características de las imágenes de enfermedades de la piel, un proceso que puede ser subjetivo y limitante. Además, la selección restringida de características puede impedir que estos algoritmos capturen la complejidad y variabilidad de las lesiones cutáneas, limitando su capacidad para generalizar a un espectro más amplio de tipos de cáncer de piel.

Escasez de Datos y Variedad Limitada en Categorías de Cáncer de Piel

La escasez de datos y la limitada variedad en las categorías de cáncer de piel presentes en los conjuntos de datos disponibles son problemas centrales en el desarrollo de algoritmos efectivos para la clasificación del melanoma. Esta limitación en los datos disponibles afecta negativamente la capacidad de los modelos para aprender y reconocer una amplia gama de manifestaciones del cáncer de piel, lo que puede

resultar en un rendimiento deficiente al enfrentarse a casos menos comunes o atípicos en la práctica clínica real.

Conclusión del Estado del Arte

Los estudios revisados reflejan un claro progreso en la aplicación de CNNs y otras técnicas de aprendizaje profundo, no solo en términos de eficacia sino también en la reducción de tiempos de espera y mejora en el acceso a diagnósticos. Es notable cómo estos avances han permitido el desarrollo de sistemas de clasificación más robustos y precisos, capaces de distinguir entre lesiones cutáneas malignas y benignas con altos niveles de precisión.

En el contexto de estos avances, esta investigación aporta un valor distintivo en varios aspectos. Primero, la implementación de una arquitectura EfficientNetB5 para la clasificación de cáncer de piel, una técnica relativamente reciente y menos explorada en la literatura comparada con modelos más establecidos como AlexNet o InceptionV3. Esta elección representa un intento de equilibrar la eficiencia y precisión en un campo donde la carga computacional y la exactitud son críticas.

Además, el enfoque de esta investigación hacia el tratamiento de conjuntos de datos desequilibrados aborda una limitación significativa que ha sido un desafío persistente en estudios anteriores. Al proponer y validar métodos que manejan de manera efectiva la desproporción en las categorías de datos, se está contribuyendo a un área de necesidad crítica, mejorando potencialmente la capacidad del modelo para generalizar y funcionar eficazmente en escenarios clínicos reales.

Finalmente, esta investigación propone un análisis detallado de la interpretabilidad y la explicabilidad del modelo propuesto, un área que ha sido tradicionalmente un desafío en el campo de la inteligencia artificial. Entender cómo y por qué el modelo toma ciertas decisiones es crucial no solo para la confianza de los médicos y pacientes en estas tecnologías, sino también para identificar áreas de mejora y refinamiento en futuros trabajos.

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 esta 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 aleato-

riedad 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. Ajusted 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.

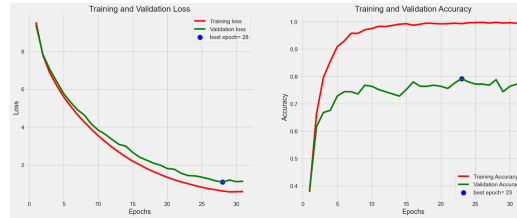


Figura 3.3: Estadísticas de aprendizaje a lo largo del proceso de entrenamiento (en rojo el proceso de entrenamiento en el conjunto de entrenamiento, en verde en el conjunto de validación).

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. 14).
- Aumentación de datos. (2023). (Vid. pág. 12).
- Capas densas. (2023). (Vid. pág. 13).
- Conjunto de datos HAM10000. (2023). (Vid. pág. 11).
- Dataframes de Pandas. (2023). (Vid. pág. 12).
- Dropout. (2023). (Vid. pág. 14).
- EfficientNetB1. (s.f.). (Vid. pág. 13).
- ImageDataGenerator. (2023). (Vid. pág. 12).
- One-hot encoding. (2023). (Vid. pág. 12).
- Regularización. (2023). (Vid. pág. 13).
- Tajerian, A., Kazemian, M., Tajerian, M., & Akhavan Malayeri, A. (2023). Design and validation of a new machine-learning-based diagnostic tool for the differentiation of dermoscopic skin cancer images. *Plos one*, 18(4), e0284437 (vid. pág. 8).
- Tasa de aprendizaje. (2023). (Vid. pág. 14).
- 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).