



Universidad Autónoma del Estado de México

CENTRO UNIVERSITARIO UAEM TEXCOCO

Detección de Lesiones en la Piel Usando Visión Artificial

*Para obtener el grado de Maestro en Ciencias de la Computación
presenta:*

Erick García Espinosa

Director de tesis:

Dr. José Sergio Ruiz Castilla

Octubre 2024

INTRODUCCIÓN

1. PROBLEMÁTICA	2
1.1 JUSTIFICACIÓN	5
1.2. OBJETIVOS	6
1.3. HIPÓTESIS	6
1.4. ANTECEDENTES	7
2. MARCO TEÓRICO	7
2.1 INTELIGENCIA ARTIFICIAL	7
2.2 ENFERMEDADES DE LA PIEL	7
2.2.1 MELANOMA	8
2.2.2 HERPES ZOSTER	9
2.2.3 SARAMPIÓN	9
2.2.4 SARNA	9
2.2.5 MONKEY POX	10
2.2.6 LUPUS ERITEMATOSO	10
2.2.7 VARICELA	10
2.3 TECNOLOGÍAS PARA LA CLASIFICACIÓN DE IMÁGENES	11
2.3.1 AZURE CUSTOM VISION	11
2.3.2 RED NEURONAL CONVOLUCIONAL	11
2.3.3 FUNCIÓN DE ACTIVACIÓN MISH	11
2.3.4 ORANGE DATA MINING	12
2.4 TÉCNICAS DE PREPROCESAMIENTO	12
2.4.1 ESCALA DE GRISES	12
2.4.2 ESPACIO DE COLOR HSL	13
2.4.3 SEGMENTACIÓN DE IMÁGENES	14
2.4.4 AUMENTO DE DATOS	15
2.4.5 RASPBERRY PI 4	17
2.4.6 VIT TRANSFORMER	18
3. ESTADO DEL ARTE	18
5. METODOLOGÍA	27
4.1. DATASET	27
4.2. PREPROCESAMIENTO	28
4.2.1. SEGMENTACIÓN	28
4.2.2 AUMENTO DE DATOS	29
4.3. AZURE CUSTOM VISION	30
4.4. ORANGE DATA MINING	31
4.5. CNN CREADA EN GOOGLE COLAB	37
4.6. DISPOSITIVO DE DETECCIÓN MÓVIL	38
4.6.1. HARDWARE	38
4.6.2. SOFTWARE	39

4.6.3. DATASET	40
4.6.4. ENTRENAMIENTO	40
5.1. DISCUSIÓN DE LOS RESULTADOS	41
CONCLUSIONES	44
<hr/>	
REFERENCIAS	47

Introducción

Detectar enfermedades dermatológicas de manera oportuna es crucial para un tratamiento efectivo y la prevención de complicaciones graves. En este sentido, la Inteligencia Artificial y la Visión por Computadora se presentan como herramientas clave para asistir a los especialistas en salud en esta tarea.

En esta tesis se evaluarán las capacidades y resultados de diversas herramientas, como Orange Datamining, Custom Vision de Microsoft Azure y redes neuronales convolucionales, aprovechando los recursos proporcionados por Google Colab. Se analizarán enfermedades como melanomas, varicela, sarampión, lupus, herpes, sarna y monkeypox.

El propósito principal de este trabajo es desarrollar un modelo basado en IA (Inteligencia Artificial) que pueda ayudar al personal médico en clínicas de nivel 1 en México a realizar un prediagnóstico de enfermedades dermatológicas. Este modelo tiene como finalidad identificar lesiones cutáneas con precisión, lo que permitiría a los profesionales de la salud tomar decisiones más informadas sobre la necesidad de referir a los pacientes a centros médicos de mayor nivel para un diagnóstico y tratamiento más especializado.

Para ello, es importante comprender cómo un dermatólogo analiza las lesiones en la piel. Este evaluará el color de la lesión, si es uniforme o si presenta variaciones en tono. Además, se analizará la forma y tamaño de la lesión, así como su evolución en el tiempo.

Se espera que este modelo contribuya significativamente a mejorar la detección temprana de enfermedades dermatológicas en clínicas de nivel más bajo, lo que podría resultar en una mejor atención para los pacientes y en la prevención de complicaciones graves.

Este trabajo se compone de 5 capítulos. El Capítulo 1. presenta los antecedentes, la justificación y los objetivos de este proyecto, destacando la importancia de esta iniciativa en el contexto de la salud pública. El Capítulo 2. abordará la metodología necesaria para comprender el proyecto, desde los conceptos básicos de IA hasta el funcionamiento de las Redes Neuronales Convolucionales (CNN), que constituyen la base del modelo propuesto. En el Capítulo 3. se detallará la metodología específica empleada para el desarrollo del proyecto, incluyendo la recolección y preparación de datos, el diseño y entrenamiento de los modelos, y la evaluación de su desempeño. Además, se presentarán tres aproximaciones distintas para el desarrollo del modelo: una basada en CNN, otra utilizando Orange Datamining y una tercera empleando

Custom Vision de Microsoft Azure. En el Capítulo 4. se discutirán los resultados obtenidos de cada una de las aproximaciones realizadas, comparando los resultados de precisión, tanto en el entrenamiento como en la validación de los modelos desarrollados. Finalmente, en el Capítulo 5. se presentarán las conclusiones del proyecto, resumiendo las principales contribuciones, limitaciones y posibles direcciones futuras de investigación, así como el impacto potencial del modelo desarrollado.

1. Problemática

Las enfermedades de la piel representan un amplio espectro de condiciones médicas que afectan la integridad y función de este órgano vital. Desde el cáncer de piel, con sus variantes como melanomas y no melanomas, hasta enfermedades infecciosas como el herpes zóster y la sarna, estas afecciones pueden tener impactos significativos en la salud y calidad de vida de las personas. Además, factores como la exposición a la radiación ultravioleta, la genética, y las condiciones socioeconómicas pueden influir en su desarrollo y curso. Comprender la naturaleza y el manejo de estas enfermedades es crucial para brindar atención médica efectiva y mejorar la salud de quienes las padecen.

El cáncer es una enfermedad caracterizada por el crecimiento anormal y acelerado de células cancerígenas que tienden a esparcirse a otros órganos del cuerpo. Es la segunda causa de muerte en el mundo después de las enfermedades cardiovasculares. El cáncer de piel se divide en melanomas y no melanomas, siendo los primeros más agresivos y los segundos menos. El cáncer de piel es una de las enfermedades oncológicas más comunes en el mundo. El melanoma es el decimonoveno cáncer más común en hombres y mujeres, con alrededor de 300,000 nuevos casos alrededor del mundo en el año 2018. El no melanoma es el quinto cáncer más común en hombres y mujeres con más de un millón de diagnósticos alrededor del mundo en el año 2018 (World Cancer Research Fund International, 2021).

Los hematomas son síntomas de diversos tipos de cáncer como la leucemia. En México, el cáncer de piel tiene una incidencia de aproximadamente 1000 casos anuales. El riesgo de contraer cáncer en la piel depende de múltiples factores como la edad, el sexo, la raza, el consumo de tabaco y alcohol, y la exposición a la

radiación UV (Ultravioleta). A pesar del aumento del IDH (Índice de Desarrollo Humano) logrado por México en el año 2018, solo cuenta con un aproximado de 1-3 máquinas de radioterapia por cada millón de habitantes. La leucemia es la principal causa de muerte en México en la población con menos de 15 años y en los jóvenes de 15 a 29 años (INEGI, 2021).

Para valorar un hematoma o melanoma se requiere de estudios como radiografías, ecografías, análisis de sangre y resonancias magnéticas. Esto puede traducirse en

Detección de lesiones en la piel usando visión artificial

tiempo y dinero para la persona que los padece, además de dificultar el acceso a una atención médica de calidad en zonas de alta marginación. El problema que se pretende resolver es determinar si una inteligencia artificial puede realizar un prediagnóstico en base a fotografías de lesiones en la piel otorgadas por el usuario y si este prediagnóstico es realizado con eficacia para poder apoyar a un especialista al momento de realizar un diagnóstico.

Por otro lado, la varicela es una enfermedad común en la población infantil de los países desarrollados, siendo la más frecuente después de la inmunización sistemática contra el sarampión y la rubéola. Aunque suele ser benigna, puede tener complicaciones graves, especialmente en personas inmunodeficientes, con neumopatías, embarazadas, y en niños menores de un año, adolescentes y adultos. Los costos directos e indirectos de la varicela incluyen medicamentos, hospitalización, absentismo escolar y laboral, y afectan significativamente a la sociedad y la economía. La edad de aparición de la varicela está disminuyendo, y la mayoría de los adultos mayores de 25 años son inmunes. Las complicaciones más comunes en niños sanos son infecciones bacterianas secundarias, seguidas de complicaciones neurológicas y pulmonares. En personas inmunodeprimidas y adultos, las complicaciones más frecuentes son debido al virus varicela-zoster, incluyendo varicela diseminada, con riesgo de afectación pulmonar, hepática y neurológica. La varicela puede ser fatal, especialmente en adultos. La vacunación contra la varicela ha demostrado ser eficaz, segura y reduce la incidencia de complicaciones y la carga económica y social asociada con la enfermedad. (Moraga & Campins, 2000).

Martínez-Fernández y colaboradores (2023) proporcionan una visión general de la infección por virus del monkeypox, un virus zoonótico que pertenece al género Orthopoxvirus. A pesar de su benignidad habitual, el monkeypox puede tener complicaciones graves, especialmente en poblaciones de riesgo como personas inmunodeficientes, con neumopatías, embarazadas y niños menores de un año, adolescentes y adultos. Desde la erradicación de la viruela, el monkeypox se ha identificado como el Orthopoxvirus más común que causa enfermedad en humanos. Aunque endémico en ciertas regiones de África, el brote actual de monkeypox, iniciado en 2022, se ha extendido a numerosos países, generando preocupación global con más de 87,545 casos y 141 muertes reportadas hasta mayo de 2023, la

mayoría en países no endémicos, principalmente debido a la transmisión de persona a persona. La revisión aborda aspectos clave como reservorios animales, modos de transmisión, epidemiología, características clínicas e inmunológicas, diagnóstico, tratamientos, vacunas y estrategias de prevención, ofreciendo información valiosa para futuras investigaciones y el desarrollo de intervenciones efectivas.

Algunas de las enfermedades de la piel han tenido rebrotes como lo expone Suvvari

Detección de lesiones en la piel usando visión artificial

et al. (2023) discuten la reemergencia del sarampión como una amenaza global inminente debido al declive en las tasas de vacunación durante la pandemia de COVID-19, con un enfoque especial en el reciente brote en India. La interrupción de los procedimientos de vacunación rutinaria durante la pandemia ha llevado a brotes esporádicos de sarampión entre los niños. Se estima que más de 23 millones de niños han perdido sus dosis de vacunación de rutina, lo que ha afectado la distribución de la vacuna contra el sarampión. El sarampión, causado por el virus del sarampión (MeV), es altamente contagioso, con un número de reproducción básico (R0) de 12-18, y se transmite por contacto directo y a través del aire. Durante la pandemia de COVID-19, se redujo significativamente el número de dosis de vacunas administradas en comparación con la era pre-pandémica. La vacilación ante la vacunación, las condiciones de vida precarias, los servicios de salud inadecuados y la falta de nutrición adecuada han contribuido a los brotes de sarampión. Desde el comienzo de 2022, Mumbai ha registrado más de 20 brotes de sarampión, con la cepa D8 del virus asociada con el brote más reciente. A pesar de que la estructura del virus no ha cambiado, los brotes recientes se deben principalmente a lagunas en la inmunización. India, junto con muchos otros países del sur de Asia, se ha comprometido a eliminar el sarampión y controlar la rubéola para 2020. Sin embargo, muchos países han postergado o dado menos prioridad a las campañas de vacunación contra el sarampión durante la pandemia de COVID-19, lo que ha aumentado el riesgo de brotes de sarampión en todo el mundo. Se han reportado alrededor de 13,000 casos de sarampión en India en 2022, considerado el brote más grande del año, lo que amenaza el objetivo del país de eliminar esta enfermedad para su fecha límite de 2023. La Organización Mundial de la Salud (OMS) y los Centros para el Control y la Prevención de Enfermedades (CDC) han advertido que el sarampión es ahora una amenaza inminente en todo el mundo debido a la disminución de las tasas de vacunación y la vigilancia debilitada del virus del sarampión durante la pandemia de COVID-19. Es fundamental intensificar las campañas de vacunación contra el sarampión a gran escala en India y a nivel mundial para proteger la salud y la vida de los niños.

El estudio realizado por J.A. Avilés y colaboradores (2013) analizó a 101 pacientes diagnosticados con lupus eritematoso cutáneo en un hospital de tercer nivel entre 2007 y 2011. Se encontró que la edad media al diagnóstico fue de 40 años, con un predominio del 84% de mujeres, representando el 94% en las formas subagudas. Se observó que el 79% de las formas crónicas discoides se localizaron exclusivamente en la cabeza y el cuello, mientras que el 97% de las formas subagudas se localizaron en el tronco.

Se identificaron un total de 8017 casos de herpes zóster confirmados por registros médicos. La edad promedio de los pacientes con herpes zóster fue de 51.8 años (DE, 23.7), con un rango de edad de 0.3 a 101.7 años. En general, el 58.7% de los casos fueron mujeres. La población del condado de Olmsted creció de aproximadamente

Detección de lesiones en la piel usando visión artificial

44,000 en 1945 a 140,000 en 2007, con un aumento modesto en la mezcla racial y étnica, de menos del 1% de no blancos en 1945 a 7.3% de no blancos en 2007 y menos del 1% de hispanos en 1945 a 3.8% en 2007. Aproximadamente el 6.6% de los pacientes con herpes zóster eran inmunocomprometidos y el 35.4% recibió terapia antiviral (Kawai, Yawn, Wollan, & Harpaz, 2016). El porcentaje de pacientes con estado inmunocomprometido con herpes zóster varió entre 1.4% en 1945-1949, 8.8% en 1990-1999 y 6.0% en 2000-2007. El uso de terapia antiviral aumentó del 11.8% en 1980-1989 al 46.8% en 2000-2007 (Kawai et al., 2016).

En el trabajo de Welch y colaboradores (2021), se destaca que la sarna, reconocida por la OMS como una de las enfermedades tropicales desatendidas más importantes desde 2017, no solo tiene consecuencias significativas para la salud, sino también efectos psicosociales profundos. Los esfuerzos recientes se han centrado en establecer la carga global de la sarna, desarrollar criterios de diagnóstico e introducir intervenciones de salud pública.

Se menciona que la sarna es causada por ácaros microscópicos transmitidos a través del contacto cercano, y prospera en comunidades con recursos escasos, con hacinamiento y pobreza. La reinfestación rápida es común en áreas endémicas debido al difícil acceso al tratamiento y al hacinamiento doméstico. La carga económica en estas comunidades incluye costos directos relacionados con medicamentos, días de enfermedad y hospitalización.

Aunque la sarna ocurre en todo el mundo, la mayor prevalencia se documenta en regiones con climas cálidos y tropicales, como las islas del Pacífico, América Central y el norte de Australia. Grupos vulnerables como los niños pequeños y los ancianos están afectados de manera desproporcionada. Por ejemplo, en Fiji se informó que más de un tercio de los residentes tenían sarna y más de la mitad de los niños estaban afectados. Además, se destaca que las consultas médicas por sarna representaron el 58% de las consultas de enfermedades infecciosas/dermatología entre los migrantes que llegaban a Italia y el 56.5% entre las poblaciones sin hogar de París.

1.1 Justificación

La detección precisa y oportuna de enfermedades dermatológicas es fundamental para garantizar un prediagnóstico y la canalización adecuada a especialistas, lo que puede prevenir complicaciones graves y mejorar la calidad de vida de los pacientes. Sin embargo, la evaluación manual de las lesiones cutáneas puede ser subjetiva y propensa a errores, lo que resalta la necesidad de utilizar enfoques apoyados en visión por computadora para mejorar la precisión y eficiencia del diagnóstico dermatológico.

En este contexto, la Inteligencia Artificial y la Visión por Computadora ofrecen un enfoque prometedor para mejorar la detección de enfermedades dermatológicas a través

Detección de lesiones en la piel usando visión artificial

del análisis automatizado de imágenes digitales de lesiones cutáneas. Estas tecnologías permiten a los profesionales de la salud aprovechar algoritmos avanzados y modelos de aprendizaje automático para identificar patrones y características específicas asociadas con diversas enfermedades de la piel.

La utilización de herramientas de inteligencia artificial en el campo de la dermatología podría ser de gran ayuda en el prediagnóstico de enfermedades cutáneas. Estas herramientas podrían procesar rápidamente grandes cantidades de imágenes de lesiones cutáneas, identificando patrones y características que podrían escapar al ojo humano. Al proporcionar análisis objetivos y precisos, estas herramientas podrían servir como una herramienta complementaria para los dermatólogos, ayudándoles a tomar decisiones más informadas y precisas en el proceso de diagnóstico. Además, al permitir la detección temprana de enfermedades, estas herramientas podrían mejorar significativamente los resultados clínicos y la calidad de vida de los pacientes.

1.2. Objetivos

Objetivo General: Clasificar lesiones cutáneas usando Redes neuronales artificiales a partir de imágenes para identificar enfermedades tales como: cáncer de piel, la sarna, el sarampión, entre otras y apoyar en el prediagnóstico de enfermedades de la piel.

Objetivos específicos:

1. Evaluar la efectividad de las redes neuronales artificiales en el prediagnóstico de diversas enfermedades de la piel
2. Construir un dataset robusto que abarque enfermedades relevantes y de alta recurrencia en la población en general
3. Comparar varios modelos y plataformas que ofrecen sistemas o modelos de clasificación además de sus resultados obtenidos
4. Elaborar un sistema móvil con un modelo entrenado para la detención de lesiones

1.3. Hipótesis

¿Es posible detectar una variedad de lesiones en la piel utilizando inteligencia artificial y modelos de visión artificial? Se plantea que sí, dado que un dermatólogo experto realiza la detección de lesiones cutáneas benignas o malignas basándose en la observación de características como la asimetría, los bordes irregulares, el color, el diámetro y la evolución de la lesión. Esta habilidad humana puede ser emulada por algoritmos de inteligencia artificial, como redes neuronales convolucionales, que pueden analizar imágenes de lesiones cutáneas y clasificarlas con alta precisión, mejorando así el diagnóstico dermatológico y facilitando la detección temprana de enfermedades.

1.4. Antecedentes

Para abordar el desafío planteado en la sección 1.1, es esencial explorar dos áreas clave: sistemas de visión y medicina. En este apartado, se presentan los fundamentos del problema tratado en esta tesis, así como las técnicas y tecnologías de vanguardia existentes para su resolución. Nos enfocaremos en el estudio de la visión artificial, la segmentación de imágenes y los algoritmos y plataformas que nos permiten clasificar imágenes de manera efectiva.

2. Marco Teórico

2.1 Inteligencia Artificial

La inteligencia artificial (IA) se refiere al esfuerzo por hacer que los computadores realicen tareas que, cuando son realizadas por personas, requieren de inteligencia. En un sentido amplio, se trata de desarrollar máquinas con mentes, capaces de realizar funciones cognitivas como la toma de decisiones, la resolución de problemas y el aprendizaje. La IA también se ocupa del estudio de cómo lograr que los computadores realicen tareas de manera racional, es decir, de manera que se considere "correcta" en función del conocimiento que poseen (Russell & Norvig, 2021).

Alan Turing, en su artículo "Computing machinery and intelligence", propuso la Prueba de Turing como un método para evaluar la inteligencia de una máquina. La prueba consiste en que un evaluador humano realice preguntas a una máquina y a un humano sin verlos, y debe determinar cuál es cuál basándose en las respuestas. Si el evaluador no puede distinguir entre la máquina y el humano, se considera que la máquina tiene inteligencia (Turing, 1950).

2.2 Enfermedades de la piel

En este apartado describiremos brevemente las enfermedades de estudio las cuales se pueden ver en la figura 1, las enfermedades incluidas abarcan una variedad de condiciones dermatológicas, cada una con características distintivas en términos de apariencia y comportamiento clínico. Estas enfermedades pueden afectar diferentes capas de la piel, manifestándose en forma de lesiones, erupciones, manchas o protuberancias. Es fundamental reconocer y diferenciar estas enfermedades, ya que algunas pueden ser benignas, mientras que otras pueden tener implicaciones más graves para la salud.

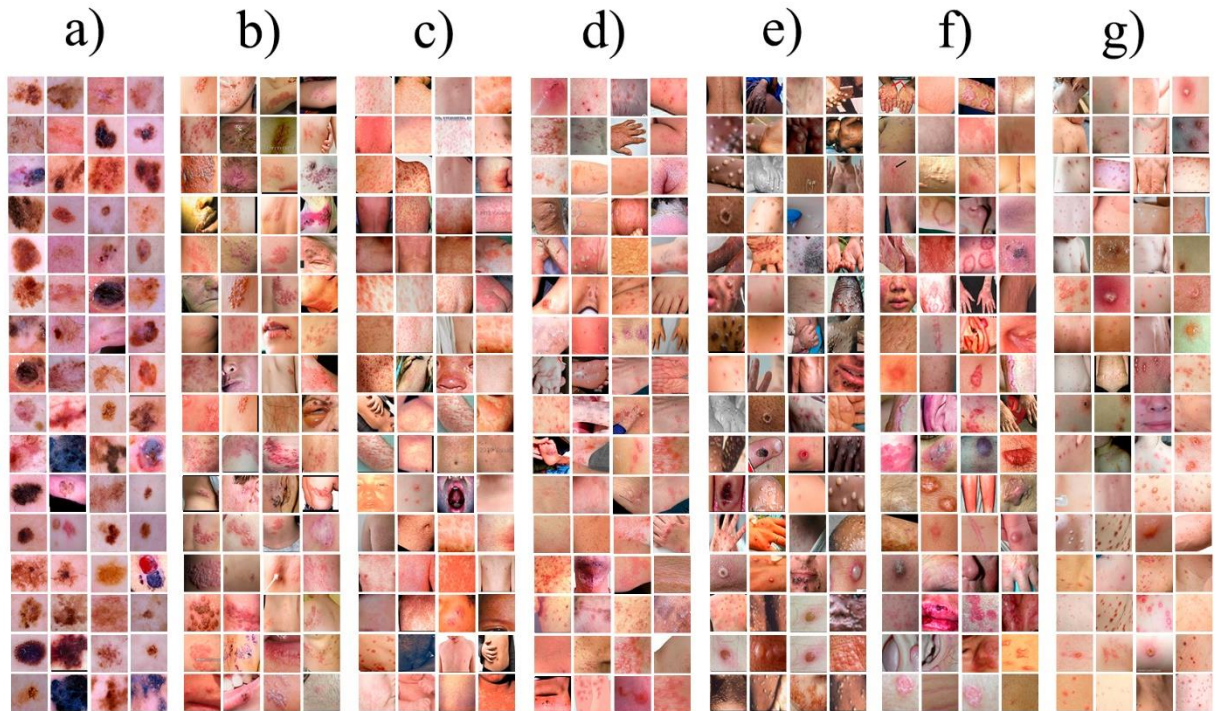


Figura. 1 Enfermedades evaluadas en el estudio: a) Melanoma, b) Herpes, c) Sarampión, d) Sarna, e) Monkeypox, f) Lupus y g) Varicela

2.2.1 Melanoma

Los melanomas son (véase Figura 1 letra a) son un tipo de cáncer de piel que se desarrolla a partir de los melanocitos, las células que producen el pigmento de la piel llamado melanina. Estos cánceres generalmente se presentan como lunares o manchas pigmentadas en la piel que han cambiado de tamaño, forma o color con el tiempo. El melanoma puede diseminarse rápidamente a otros órganos si no se trata tempranamente, por lo que es importante detectarlo y tratarlo lo antes posible.

El cáncer es una enfermedad caracterizada por la formación de tumores malignos, es decir, que se diseminan a otras partes del cuerpo. Los tumores malignos son resultado de la división anormal de las células cancerosas estas pueden invadir y destruir tejido sano de otros órganos del cuerpo sin importar cuales sean, el cáncer también puede diseminarse a través del torrente sanguíneo o el sistema linfático. El instituto Nacional del Cáncer lo define como:

El cáncer es una enfermedad por la que algunas células del cuerpo se multiplican sin control y se diseminan a otras partes del cuerpo. Es posible que el cáncer comience en cualquier parte del cuerpo humano, formado por billones de células. En condiciones normales, las células humanas se forman y se multiplican (mediante un proceso que se llama división celular) para formar células nuevas a medida que el cuerpo las necesita. Cuando las células envejecen o se dañan, mueren y las células nuevas las reemplazan. [...] Estas células tal vez formen tumores, que son bultos de tejido. Los tumores son cancerosos (malignos) o no cancerosos (benignos). (Instituto Nacional del Cáncer, 2021)

2.2.2 Herpes Zoster

Herpes zóster (culebrilla) es una enfermedad causada por el virus varicela-zóster como resultado de una disminución en la inmunidad celular específica que ocurre fisiológicamente con el proceso de envejecimiento o debido a la inmunosupresión y condiciones que involucran inmunodeficiencia. Las lesiones de herpes causan ampollas dolorosas en un lado del cuerpo (Véase Figura 1 letra b). Para el herpes oral (tipo 1), aparecen en los labios y forman costras. Para el herpes genital (tipo 2), aparecen en los genitales. Las lesiones son contagiosas y pueden reaparecer durante períodos de estrés o cuando el sistema inmunológico está debilitado. (Gross, Schofer, Wassilew, Friese & Timm, 2003)

2.2.3 Sarampión

El sarampión, también conocido como rubeola, es una enfermedad viral aguda febril altamente contagiosa y prevenible las lesiones destacan por ser pequeñas lesiones en la piel en forma de granos como se puede ver en la figura 1 letra c. A pesar de la disponibilidad de una vacuna efectiva, sigue siendo una causa importante de mortalidad y morbilidad a nivel mundial, especialmente en las regiones de África y el sudeste asiático. Aunque los funcionarios de salud pública declararon la eliminación del sarampión en los Estados Unidos en 2000 y en la región de las Américas en 2016, los brotes siguen ocurriendo a través de la enfermedad importada y la transmisión entre grupos de niños no vacunados en la comunidad. Según los Centros para el Control y la Prevención de Enfermedades (CDC), hubo 372 casos en 2018 y 764 casos hasta mayo de 2019. Esta actividad proporciona una visión general de la prevención, transmisión y diagnóstico del sarampión, y destaca el papel del equipo interprofesional en ayudar a sus pacientes a prevenir esta enfermedad (Kondamudi & Waymack, 2023).

2.2.4 Sarna

La sarna es una infestación común que se presenta de forma variable en diferentes situaciones clínicas (Véase Figura 1 letra d). En entornos desarrollados, los brotes en instituciones y residencias desafían a los servicios de salud y sociales. Por otro lado, en áreas con recursos limitados, las bacteriemias estafilocócicas y estreptocócicas, provocadas por el rascado, tienen un impacto significativo en la salud comunitaria a largo plazo. En los últimos años, la sarna ha sido reconocida como una "enfermedad tropical desatendida" por la Organización Mundial de la Salud. (Chandler & Fuller, 2019).

2.2.5 Monkey Pox

Es enfermedad viral zoonótica causada por el virus del monkeypox se puede observar en la Figura 1 letra e, ha emergido como un desafío significativo para la salud pública mundial. Hasta mayo de 2022, el virus era endémico en países de África occidental y central, con casos esporádicos fuera de este mismo. Sin embargo, desde mayo de 2022, se han identificado numerosos casos en países no endémicos sin historial de viajes desde áreas endémicas, lo que indica una posible transmisión a largo plazo en ausencia de vínculos epidemiológicos conocidos. Este cambio es preocupante, ya que plantea interrogantes sobre la evolución del virus y su capacidad para la transmisión no detectada.

Se han reportado siete casos de monkeypox relacionados con viajes fuera del continente africano entre septiembre de 2018 y noviembre de 2021, incluidos casos en Israel, Singapur y Estados Unidos. Este cambio en la distribución geográfica y la aparición de casos en lugares sin conexiones epidemiológicas conocidas plantean desafíos adicionales para el control y la prevención de la enfermedad. La situación actual subraya la importancia de la vigilancia continua y la implementación de medidas preventivas efectivas para contener la propagación del virus del monkeypox (Ferdous et al., 2023).

2.2.6 Lupus eritematoso

El lupus eritematoso cutáneo (LEC) se manifiesta en diversas formas clínicas, incluyendo el LEC agudo, subagudo y crónico. La versión aguda de este puede aparecer como un sarpullido en la cara, mientras que el subagudo es altamente sensible a la luz solar y suele afectar la parte superior de la espalda, hombros, cuello y pecho como puede observarse en la Figura 1 letra f. Por otro lado, el crónico se caracteriza por placas escamosas en el cuero cabelludo, la cara y las orejas, que pueden dejar cicatrices y cambios en la pigmentación de la piel. Además de estos subtipos, existen otras formas menos comunes de LEC. Esta enfermedad también puede causar otros problemas en la piel, como vasculitis, cambios en la coloración de la piel, pérdida de cabello, lesiones en los dedos y ampollas. (Walling & Sontheimer, 2009)

2.2.7 Varicela

La varicela es una enfermedad infecciosa altamente contagiosa causada por el virus de la varicela-zóster que afecta principalmente a niños y se caracteriza por la aparición de un sarpullido característico en todo el cuerpo como puede verse en la Figura 1 letra g. Aunque generalmente es una enfermedad leve y autolimitada, puede causar complicaciones graves, especialmente en adultos y personas con sistemas inmunológicos debilitados. En un estudio reciente sobre la varicela, se investigaron genes de susceptibilidad a la enfermedad en diferentes continentes, lo que sugiere que

factores genéticos pueden influir en la susceptibilidad y la gravedad de la varicela en poblaciones diversas (Irham et al., 2023).

2.3 Tecnologías para la clasificación de imágenes

En la actualidad, el análisis de imágenes se ha convertido en una herramienta fundamental en diversas áreas, como la medicina, la industria, la agricultura y la investigación científica. Para llevar a cabo una clasificación precisa y eficiente de imágenes, se han desarrollado diversas tecnologías basadas en inteligencia artificial y aprendizaje automático. Entre estas tecnologías destacan las redes neuronales convolucionales, que han demostrado ser altamente efectivas en el procesamiento de imágenes. Plataformas como Azure de Microsoft ofrecen herramientas avanzadas basadas en estas tecnologías, que permiten realizar tareas de clasificación de imágenes de manera automatizada y precisa. En este capítulo, se explorarán algunas de las tecnologías más relevantes en este campo, así como su aplicación en diferentes contextos y sus beneficios en términos de eficiencia y precisión en la clasificación de imágenes.

2.3.1 Azure Custom Vision

Microsoft Azure's Custom Vision es un servicio de reconocimiento de imágenes que permite a los usuarios crear, implementar y mejorar sus propios modelos de identificación de imágenes. A diferencia del servicio Azure AI Vision, Custom Vision permite a los usuarios definir sus propias etiquetas y entrenar modelos personalizados para detectarlas (Microsoft, 2023).

2.3.2 Red Neuronal Convolucional

Las Redes Neuronales Convolucionales (CNNs) tienen varias capas para procesar datos. La idea de este modelo comenzó en la década de 1950 cuando los investigadores Hubel y Wiesel estudiaron cómo funcionan los sistemas visuales de los animales (Carrión, 2023).

Una CNN consta de una capa de entrada, una o más capas ocultas y una capa de salida. Las CNN se utilizan principalmente para tareas como la clasificación y la

visión por computadora. Son excelentes para clasificar imágenes porque pueden identificar patrones importantes en las imágenes, lo que facilita su reconocimiento y clasificación.

2.3.3 Función de Activación MISH

MISH es una alternativa prometedora a ReLU en algoritmos de clasificación de imágenes. Una de sus principales ventajas es su capacidad para abordar el problema de la desaparición del gradiente. Mientras que ReLU puede causar gradientes

MISH es más suave y ayuda a aliviar estos problemas (Liu y Furber, 2016) . Esta función es una función de activación neuronal suave y no monótona definida como:

$$f(x)=x * \tanh(\ln(1 + e^x)), \text{ donde } \zeta(x) = \ln(1 + e^x) \quad (1)$$

2.3.4 Orange Data Mining

Orange Data Mining es un software de código abierto utilizado para visualización de datos, aprendizaje automático, minería de datos y análisis de datos (Orange Visual Programming, 2023). Proporciona una interfaz de programación visual que permite a los usuarios explorar datos cualitativos y visualizarlos de forma interactiva. Además, incluye el modelo Inception v3 como un modelo CNN preentrenado para tareas de clasificación de imágenes. Este modelo de aprendizaje profundo, desarrollado por Google, es conocido por su alta precisión en la clasificación de imágenes. Es una variación de la arquitectura Inception que utiliza una combinación de capas convolucionales, capas de agrupamiento y capas completamente conectadas para aprender características de las imágenes de entrada. (Sharma & Lohchab, 2021)

2.4 Técnicas de preprocesamiento

En visión artificial, una técnica de preprocesamiento se refiere a un conjunto de operaciones o transformaciones aplicadas a una imagen digital antes de que sea analizada por un algoritmo de visión por computadora. Estas técnicas tienen como objetivo mejorar la calidad de la imagen, resaltar características importantes, reducir el ruido o simplificar la información para facilitar la tarea de análisis o reconocimiento de patrones por parte de los algoritmos.

2.4.1 Escala de grises

En el ámbito de la visión artificial, la conversión de una imagen a escala de grises es un paso fundamental que permite simplificar la información visual al eliminar la componente de color, preservando únicamente la información de luminosidad. Una de las técnicas más comunes para lograr esta conversión es mediante la fórmula de luminancia ponderada, la cual considera la sensibilidad del ojo humano a diferentes colores. Esta fórmula asigna un valor de intensidad único a cada píxel de la imagen, combinando los componentes de rojo, verde y azul en proporciones específicas (0.299, 0.587 y 0.114 respectivamente) de la siguiente manera:

$$\text{gris}=0.299 \times \text{rojo} + 0.587 \times \text{verde} + 0.114 \times \text{azul}$$

De esta manera, se obtiene una representación en escala de grises que conserva las

características de luminosidad de la imagen original, pero simplificada y lista para ser procesada por algoritmos de visión por computadora.

2.4.2 Espacio de color HSL

El espacio de color RGB es conocido por su uso en pantallas LCD a color, mientras que HSL se emplea en procesamiento de imágenes. Cada espacio tiene sus propias características: RGB es hardware-orientado, mientras que HSL está orientado al humano (Wu & Sun, 2013). HSL, similar a RGB, se compone de tres componentes: tono, saturación y luminancia. El tono abarca todas las mezclas posibles de los colores primarios y se representa en grados de 0 a 360. La saturación indica la pureza del color, y la luminancia, la cantidad de luz reflejada (Aptoula & Lefèvre, 2009).

Las ecuaciones de conversión de tono y saturación son ampliamente utilizadas, pero algunas requieren operaciones trigonométricas, lo que las hace menos adecuadas para implementar en arquitectura digital (Nnolim, 2015).

$$L = \frac{R_{max} + R_{min}}{2} \quad (2)$$

$$S = \frac{R_{max} - R_{min}}{R_{max} + R_{min}} \quad \text{si } L < 0.5$$

$$S = \frac{R_{max} - R_{min}}{2.0 - (R_{max} + R_{min})} \quad \text{si } L \geq 0.5$$

$$H = \frac{G - B}{R_{max} - R_{min}} \quad \text{si } R_{max} = R$$

$$H = \frac{B - R}{R_{max} - R_{min}} \quad \text{si } R_{max} = G$$

$$H = \frac{R - G}{R_{max} - R_{min}} \quad \text{si } R_{max} = B$$

En esta fórmula R_{max} y R_{min} son los valores máximo y mínimo de los componentes RGB del píxel respectivamente, y L, S, H representan los valores de luminosidad, saturación y matiz del píxel en el espacio de color HSL. Esta representación facilita el análisis y manipulación de colores en aplicaciones de procesamiento de imágenes.

2.4.3 Segmentación de imágenes

Para la labor de la clasificación de imágenes es indispensable la segmentación esta es la separación de los elementos de una imagen, a través de un proceso de codificación y decodificación (Véase Figura 2), esta se encarga de dividir las imágenes en pequeños segmentos u objetos (Szeliski, 2010). Este proceso es crucial en áreas como el análisis de imágenes del ámbito médico, vehículos inteligentes, video vigilancia y la realidad aumentada (Forsyth & Ponce, 2002). A lo largo del tiempo, se han desarrollado diferentes algoritmos para segmentar imágenes, desde métodos clásicos como el umbralizado (Otsu, 1979) y el crecimiento o eliminación de elementos como la dilatación y erosión hasta algoritmos más avanzados como puede ser el clustering k-means o el Algoritmo Watershed (Najman & Schmitt, 1994).

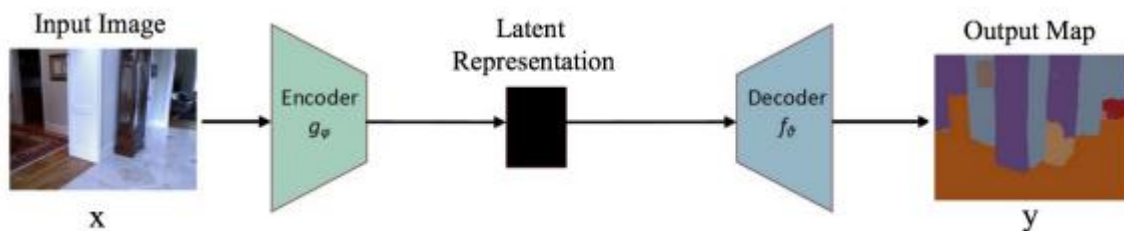


Figura. 2 Ejemplo del proceso de segmentación (Minaee,2020)

La segmentación de imágenes puede ser tratada como una tarea de asignación de clases a píxeles basada en su contenido semántico, o como la división de objetos individuales en una imagen. (Plath, Toussaint, & Nakajima, 2009). La segmentación semántica se encarga de que para pixel le sea asignada una etiqueta, por otro lado, la segmentación de instancias va más allá al delinear y detectar cada objeto de relevante en la imagen (Starck, Elad, & Donoho, 2005).

En su artículo "Image Segmentation Using Deep Learning: A Survey", Minaee et al. (2020) proporcionan una revisión exhaustiva de los últimos avances en segmentación de imágenes utilizando aprendizaje profundo, destacando la importancia y las aplicaciones de este campo en constante evolución.

La segmentación de imágenes se fundamenta en tres principios básicos que guían el proceso de división de la imagen en segmentos significativos:

- Similitud: Este principio se refiere a que los píxeles que pertenecen a un mismo objeto tienen valores similares con respecto a alguna propiedad específica. En el contexto de la segmentación, se estudia la regularidad de los valores de nivel de gris presentes en la imagen. Es decir, los píxeles que forman parte de un objeto tienden a tener propiedades similares, como intensidades de color cercanas en imágenes en color o niveles de gris similares

- **Discontinuidad:** Los objetos y las regiones que se desean segmentar suelen destacarse del entorno y presentar contornos o bordes más definidos que los objetos que los rodean. Este principio se basa en la observación de que los bordes entre objetos representan áreas donde hay cambios significativos en las propiedades de los píxeles, como cambios abruptos en el color, la textura o la intensidad.
- **Conectividad:** Este principio se refiere a que los píxeles que pertenecen a un mismo objeto tienden a estar adyacentes, es decir, a agruparse y formar regiones homogéneas en la imagen. La conectividad se utiliza para distinguir entre objetos separados y asegurar que los segmentos resultantes sean coherentes y representen entidades individuales en la escena.

2.4.4 Aumento de Datos

Las técnicas tradicionales de aumento de datos son fundamentales en el procesamiento de imágenes y la mejora del rendimiento de los modelos de aprendizaje automático (Vease Figuras 3-7). Aquí tienes un resumen de las principales técnicas de aumento de datos y cómo se aplican:

Transformaciones geométricas: Estas operaciones básicas incluyen voltear, recortar y rotar imágenes para aumentar la diversidad de datos.

- **Volteo (Flipping):** Refiere al proceso de invertir imágenes horizontal o verticalmente, lo que ayuda a evitar problemas de clasificación debido a la dirección en las imágenes. Por ejemplo, en un conjunto de datos de perros y gatos, algunas imágenes de perros pueden estar mirando hacia la izquierda,
- lo que podría confundir al modelo. El volteo horizontal puede resolver este problema.



Figura. 3 Ejemplo función de volteo en el aumento

- **Recorte (Cropping):** Esta técnica recorta una parte aleatoria de la imagen y la redimensiona a un tamaño específico, lo que ayuda a estandarizar el tamaño de las imágenes de entrenamiento.



Figura. 4 Ejemplo función de corte en el aumento

- Rotación (Rotation): Las imágenes se rotan por un ángulo especificado para aumentar la variabilidad de los datos. Sin embargo, la rotación puede causar pérdida de información en los bordes de las imágenes.
- Transformaciones fotométricas: Estas técnicas alteran los valores de los píxeles en lugar de su posición.



Figura. 5 Ejemplo función de transformación fotométrica en el aumento

- Cambios de color (Colour Jittering): Se manipulan los canales de color de una imagen, como el brillo, el contraste o los colores, para generar variaciones en los datos.
- Kernel/Filtro: Los kernels juegan un papel importante en el aprendizaje profundo al extraer características de las imágenes como filtro.

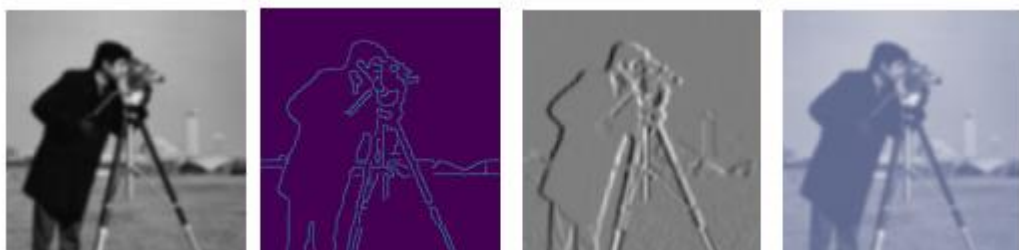


Figura. 6 Ejemplo de cambio de Kernel en el aumento

- Detección de bordes (Edge Detection): Se utilizan filtros como Sobel o Canny para detectar bordes en las imágenes.
- Transformaciones de ruido (Noise Transformations): Estas técnicas agregan ruido a las imágenes para aumentar la resistencia del modelo al ruido en los datos reales.



Figura. 7 Ejemplo del uso de transformaciones de ruido en el aumento

- Ruido Gaussiano (Gaussian Noise): Se agrega ruido gaussiano a las imágenes para simular imperfecciones.

Para concluir este apartado, podemos citar el artículo "Data Augmentation in Classification and Segmentation: A Survey and New Strategies" de Khaled Alomar, Halil Ibrahim Aysel y Xiaohao Cai. En este artículo, los autores exploran diversas estrategias de aumento de datos en clasificación y segmentación de imágenes, lo que destaca la importancia de estas técnicas en el contexto de las redes neuronales convolucionales y otras aplicaciones de visión por computadora. Algunas de las estrategias discutidas incluyen transformaciones geométricas, transformaciones fotométricas, el uso de kernels y filtros, así como transformaciones de ruido. Este estudio resalta la necesidad de abordar los desafíos asociados con el entrenamiento de modelos de redes neuronales profundos, como los requisitos de datos a gran escala, el tiempo de entrenamiento prolongado, el sobreajuste y la disminución del rendimiento con la escasez de datos. Estos resultados subrayan la importancia de investigar y desarrollar nuevas estrategias de aumento de datos para mejorar la generalización y la efectividad de los modelos de visión por computadora basados en redes neuronales convolucionales.

2.4.5 Raspberry Pi 4

La Raspberry Pi 4 es una microcomputadora de bajo costo y alto rendimiento desarrollada por la Fundación Raspberry Pi. Diseñada para ser una herramienta educativa y accesible, esta pequeña pero potente computadora puede ejecutar una variedad de sistemas operativos, incluyendo su propio Raspberry Pi OS basado en Linux, y aplicaciones que normalmente se ejecutan en computadoras de escritorio más grandes y costosas.

La Raspberry Pi 4 es notable por su avance significativo en velocidad y rendimiento respecto a modelos anteriores. Por primera vez, se ha creado una experiencia de escritorio completa. Ya sea que estés editando documentos, navegando por la web con varias pestañas abiertas, manejando hojas de cálculo o elaborando una presentación, encontrarás que la experiencia es fluida y muy reconocible. Todo esto en una máquina que es más pequeña, más eficiente energéticamente y mucho más rentable (Raspberry,2019).

2.4.6 ViT Transformer

Un Transformador de Visión (ViT, por sus siglas en inglés) funciona de manera similar al Transformer utilizado en el procesamiento de lenguaje natural (NLP, por sus siglas en inglés) en el sentido de que utiliza atención para asignar pesos a diferentes partes de la secuencia de entrada, permitiendo que el modelo se enfoque en partes específicas según su relevancia para la tarea. Sin embargo, ViT tiene adaptaciones específicas para el procesamiento de imágenes. Aquí tienes un resumen de cómo funciona:

- **Incrustaciones de Parches:** La imagen de entrada se divide en parches (por ejemplo, 16x16 píxeles) y se convierte en vectores usando una capa lineal. Cada uno de estos vectores representa un "token" en la secuencia.
- **Incrustaciones Posicionales:** Se añaden incrustaciones posicionales a estos vectores de parches para indicar la posición relativa de cada parche en la imagen.
- **Clasificación:** Se añade un token especial [CLS] al principio de la secuencia de vectores. Este token se utiliza para la tarea final de clasificación.
- **Codificador Transformer:** La secuencia de vectores (parches + [CLS]) se introduce en un codificador Transformer. El Transformer consiste en múltiples capas de atención y redes neuronales completamente conectadas. La atención se utiliza para capturar las relaciones entre diferentes parches, mientras que las capas completamente conectadas se utilizan para aprender representaciones más abstractas de la imagen.
- **Capa de Clasificación:** Después de que la secuencia ha pasado por el codificador Transformer, se toma la representación correspondiente al token [CLS] y se introduce en una capa final de clasificación (por ejemplo, una capa lineal seguida de una función softmax) para realizar la clasificación de la imagen.

3. Estado del Arte

En el trabajo "Segmentación y Clasificación de Lesiones de la Piel para el Diagnóstico de Enfermedades", se propone un enfoque novedoso para la segmentación automática y clasificación de lesiones cutáneas. Se filtran las imágenes de la piel para eliminar vellos no deseados y ruido, y luego se lleva a cabo el proceso de segmentación para extraer áreas de lesiones. Posteriormente, las áreas de lesiones extraídas se representan mediante características de color y textura. Se utilizan clasificadores SVM y k-NN junto con su fusión para la clasificación utilizando las características extraídas. El rendimiento del sistema se prueba en un conjunto de datos propio de 726 muestras de 141 imágenes que consisten en 5 clases diferentes de enfermedades. Los resultados son prometedores, con un 46.71% y un 34% de medida-F utilizando los clasificadores SVM y k-NN respectivamente, y con un 61% de medida-F para la fusión de SVM y k-NN. Para este trabajo se creó un conjunto de

Detección de lesiones en la piel usando visión artificial

datos propio descargando imágenes de recursos en línea, debido a la falta de un conjunto de datos de referencia y problemas de permisos. Las imágenes recolectadas a través de internet exhiben grandes variaciones intraclase con menos variaciones interclase, lo cual era necesario para garantizar una técnica de preprocesamiento precisa como el filtrado. Se seleccionaron cinco clases diferentes de enfermedades de la piel: Melanoma, bullae, queratosis seborreica, culebrilla y células escamosas. Sumitra et al. (2015)

En el trabajo "Clasificación de Trastornos Dermatológicos Basado en Aprendizaje Profundo", se aborda la importancia de la diagnosis médica automatizada en apoyo a los médicos. Se destaca la creciente demanda de inventar redes neuronales convolucionales (CNN) y aprendizaje profundo (DL) para analizar imágenes médicas, con la dermatología como uno de los campos recientemente objetivo de especialistas en IA para introducir nuevos algoritmos de DL o mejorar las arquitecturas de CNN. El estudio evalúa el desempeño de seis arquitecturas de CNN (VGG16, EfficientNet, InceptionV3, MobileNet, NasNet y ResNet50) para los tres principales trastornos dermatológicos que aparecen con frecuencia en el Medio Oriente. Se aplicó filtrado de imagen y eliminación de ruido para mejorar la calidad de la imagen y aumentar el rendimiento de la arquitectura. Los resultados experimentales revelaron que MobileNet alcanzó el mejor rendimiento y precisión entre las arquitecturas de CNN, pudiendo clasificar los trastornos con alta precisión (95.7%). Se propone una futura expansión del dataset para incluir más imágenes que consideren nuevos trastornos y variaciones. AlSuwaidan (2022).

En el artículo "Un Referente para la Clasificación Automática Visual de Imágenes de Enfermedades Cutáneas Clínicas", se presenta un conjunto de datos de referencia para enfermedades cutáneas clínicas con el objetivo de apoyar la diagnosis médica automatizada en el campo de la dermatología. Este conjunto de datos, que según los autores es el más grande actualmente disponible para el reconocimiento visual de enfermedades de la piel, consta de 6,584 imágenes de 198 clases diferentes, que varían en escala, color, forma y estructura. Se espera que este conjunto de datos estimule la investigación adicional en la clasificación visual de enfermedades de la piel. Además, se realiza un análisis exhaustivo de este conjunto de datos utilizando métodos de vanguardia, incluidas las redes neuronales convolucionales (CNN). Sun et al. (2016).

El estudio "Clasificación de enfermedades cutáneas basado en Deep Learning utilizando teléfonos inteligentes" aborda la detección automatizada de enfermedades de la piel mediante una aplicación móvil inteligente basada en aprendizaje profundo. Se entrenaron diferentes redes neuronales convolucionales (CNN) preentrenadas para la clasificación de imágenes de lesiones cutáneas, y la red más adecuada en términos de rendimiento y compatibilidad móvil se transformó al formato TensorFlow Lite. Se creó una aplicación móvil en la plataforma Android que utiliza la cámara del teléfono

Detección de lesiones en la piel usando visión artificial

inteligente para obtener imágenes y utiliza TensorFlow Lite para realizar predicciones. El sistema propuesto logra una precisión de clasificación del 74.27% para siete clases en un conjunto de datos combinado, obteniendo resultados comparables o mejores que la literatura. Se espera que esta aplicación ayude a los pacientes a realizar un diagnóstico preliminar de sus lesiones utilizando sus teléfonos inteligentes, lo que puede alentar a los pacientes de alto riesgo a visitar el hospital para un diagnóstico definitivo y evitar el estrés innecesario y las falsas alarmas. Oztel (2023).

En el artículo "DSCC_Net: Modelos de Aprendizaje Profundo de Clasificación Múltiple para el Diagnóstico de Cáncer de Piel Utilizando Imágenes Dermoscópicas", se propone una red de clasificación de cáncer de piel basada en aprendizaje profundo (DSCC_Net) que se basa en una red neuronal convolucional (CNN) e implementada en tres conjuntos de datos de referencia públicos (ISIC 2020, HAM10000 y DermIS). La DSCC_Net propuesta obtuvo un AUC del 99.43%, junto con una precisión del 94.17%, una recuperación del 93.76%, una

precisión del 94.28% y un puntaje F1 del 93.93% en la categorización de los cuatro tipos distintos de enfermedades de cáncer de piel. La precisión de ResNet-152, Vgg-19, MobileNet y Vgg-16, EfficientNet-B0 e Inception-V3 es del 89.68%, 92.51%, 91.46%, 89.12%, 89.46% y 91.82%, respectivamente. Los resultados mostraron que el modelo DSCC_Net propuesto funciona mejor en comparación con los modelos base, ofreciendo un apoyo significativo a los dermatólogos y expertos en salud para diagnosticar el cáncer de piel. Tahir et al. (2023).

El artículo "Clasificación de Enfermedades de la Piel Utilizando un Enfoque de Localización Basado en IA Híbrida" propone un enfoque de inteligencia artificial híbrida para la clasificación de enfermedades de la piel. Se implementa una técnica de clasificación de enfermedades de la piel utilizando un modelo de extracción de características basado en matrices de co-ocurrencia estructurales (SCM) y una red neuronal convolucional mejorada (ECNN). El proceso de clasificación se basa en la extracción de características estructurales de las imágenes y su posterior clasificación utilizando ECNN. El enfoque propuesto se evalúa en un conjunto de datos de imágenes dermoscópicas recopiladas de conjuntos de datos públicos como PH2 y ISIC obteniendo una precisión máxima de 85%. Sreekala et al. (2022).

En la investigación de Khafaga et al. (2022), se propone un enfoque novedoso para el diagnóstico automatizado de la enfermedad de viruela del mono a partir de imágenes. Dado que las enfermedades de la piel humana, como la viruela del mono, han aumentado en prevalencia en las últimas décadas, es fundamental contar con métodos de diagnóstico eficaces. La propuesta utiliza un modelo de red neuronal convolucional profunda (CNN) ajustado con el algoritmo de optimización basado en el radio de la Tierra de Al-Biruni (BER) para clasificar la enfermedad a partir de

Detección de lesiones en la piel usando visión artificial

imágenes. Se utilizó un conjunto de datos de enfermedades de la piel humana obtenido de un hospital africano para validar el enfoque propuesto. Los resultados experimentales muestran la eficacia del marco propuesto, superando a numerosos ejemplos de investigaciones previas en problemas de enfermedades de la piel. El estudio también incluye experimentos estadísticos con pruebas de Wilcoxon y análisis de varianza (ANOVA) para evaluar la efectividad y la estabilidad del enfoque propuesto que logro obtener con su mejor modelo BERSFS-CNN: un porcentaje de 0.9883.

Por otro lado, Wei et al. (2023), se presenta un modelo de clasificación de enfermedades de la piel basado en la fusión de DenseNet y ConvNeXt. Dado que las enfermedades de la piel tienen categorías intrincadas, síntomas muy similares en la etapa temprana y muestras de lesiones extremadamente desequilibradas, su clasificación es un desafío. Se propone un modelo de red neuronal convolucional para la clasificación de enfermedades de la piel basado en la fusión de modelos, que fortalece la capacidad de extracción de características del modelo mediante la fusión de características profundas y superficiales y la introducción de un módulo de atención. Además, se llevaron a cabo una serie de trabajos como pre-entrenamiento del modelo, aumento de datos y ajuste fino de parámetros para mejorar el rendimiento de clasificación del modelo. Los resultados experimentales muestran que, al trabajar en un conjunto de datos privado dominado por enfermedades de la piel similares al acné, el modelo propuesto superó a los dos modelos base de DenseNet201 y ConvNeXt_L en un 4,42% y un 3,66%, respectivamente. En el conjunto de datos público HAM10000, la precisión y el puntaje f1 del modelo propuesto fueron del 95,29% y el 89,99%, respectivamente, lo que también logró buenos resultados en comparación con otros modelos de vanguardia.

En la aproximación de Cai et al. (2022), se presenta un modelo de red neuronal novedoso para la clasificación de enfermedades de la piel que utiliza imágenes y metadatos clínicos. Se propone un multimodal Transformer, que consta de dos codificadores para imágenes y metadatos, y un decodificador para fusionar la información multimodal. Se utiliza un modelo Vision Transformer (ViT) como base para extraer características profundas de las imágenes, mientras que los metadatos se tratan como etiquetas y se utiliza un nuevo codificador de etiquetas suaves (SLE) para incrustarlos. Además, en la parte del decodificador, se propone un bloque de Atención Mutua (MA) para fusionar mejor las características de la imagen y los metadatos. El modelo se evalúa en un conjunto de datos privado de enfermedades de la piel y en el conjunto de datos de referencia ISIC 2018, mostrando un rendimiento superior en comparación con métodos de vanguardia teniendo un porcentaje bastante alto en el dataset de ISIC con una tasa de precisión de 0.9932%.

En la propuesta de Thieme et al. se presenta un algoritmo de aprendizaje profundo para clasificar lesiones cutáneas causadas por el virus monkeypox. El algoritmo,

Detección de lesiones en la piel usando visión artificial

llamado MPXV-CNN, se desarrolló para identificar las lesiones características de la infección por monkeypox mediante redes neuronales convolucionales profundas. Utilizaron un conjunto de datos de 139,198 imágenes de lesiones cutáneas,

divididas en cohortes de entrenamiento/validación y prueba, que incluían imágenes no-monkeypox y monkeypox. En las cohortes de validación y prueba, la sensibilidad del MPXV-CNN fue de 0.83 y 0.91, respectivamente, la especificidad fue de 0.965 y 0.898, y el área bajo la curva fue de 0.967 y 0.966, respectivamente. En la cohorte prospectiva, la sensibilidad fue de 0.89. El rendimiento de clasificación del MPXV-CNN fue robusto en diferentes tonos de piel y regiones corporales. Se desarrolló una aplicación web para facilitar el acceso al algoritmo, lo que podría ayudar en la mitigación de brotes de monkeypox. El conjunto de datos utilizado incluyó imágenes de lesiones cutáneas monkeypox y no-monkeypox de diversas fuentes públicas y datos institucionales.

En la investigación de "Clasificación de lesiones cutáneas de imágenes demoscópicas utilizando aprendizaje automático y redes neuronales convolucionales." se utilizó el dataset HAM10000, que consta de 10015 imágenes de lesiones cutáneas pigmentadas. Se realizó un trabajo de clasificación de lesiones cutáneas utilizando técnicas de aprendizaje automático y redes neuronales convolucionales (CNN). Se implementó un modelo con aumento de datos que tiende a aprender características y rasgos más distintivos en comparación con un modelo sin aumento de datos. Se utilizó la técnica de validación cruzada k-fold para hacer que el modelo sea robusto. Se analizó la precisión de la clasificación de algoritmos de aprendizaje automático y modelos de CNN, concluyendo que la CNN proporciona una mejor precisión en comparación con otros algoritmos de aprendizaje automático. El trabajo propuesto logró una precisión máxima del 95.18% con el modelo CNN. Este trabajo ayuda a la identificación temprana de siete clases de enfermedades de la piel y puede validarse y tratarse adecuadamente por los médicos. Shetty et al. (2022).

En el artículo de "Clasificación de enfermedades de la piel utilizando redes neuronales profundas con MobileNet V2 y LSTM." se propuso un proceso computarizado para clasificar enfermedades de la piel mediante el uso de MobileNet V2 y Long Short Term Memory (LSTM) basados en aprendizaje profundo. El modelo MobileNet V2 demostró ser eficiente con una mayor precisión que puede funcionar en dispositivos informáticos ligeros. El modelo propuesto es eficiente en mantener información estatal para predicciones precisas. Se utilizó una matriz de co-ocurrencia de nivel de gris para evaluar el progreso del crecimiento de las enfermedades. La precisión se comparó con otros modelos de vanguardia como Fine-Tuned Neural Networks (FTNN), Convolutional Neural Network (CNN), Very

Deep Convolutional Networks for Large-Scale Image Recognition desarrollados por

Detección de lesiones en la piel usando visión artificial

Visual Geometry Group (VGG) y arquitectura de red neuronal convolucional que se amplió con algunos cambios. El conjunto de datos HAM10000 se utilizó y el método propuesto superó a otros métodos con más del 85% de precisión. Su robustez en el reconocimiento de la región afectada mucho más rápido con casi 2 veces menos cálculos que el modelo MobileNet convencional resulta en un esfuerzo computacional mínimo. Además, se diseñó una aplicación móvil para una acción instantánea y adecuada. Ayuda al paciente y a los dermatólogos a identificar el tipo de enfermedad a partir de la imagen de la región afectada en la etapa inicial de la enfermedad de la piel. Estos hallazgos sugieren que el sistema propuesto puede ayudar a los médicos de atención primaria a diagnosticar eficientemente y de manera efectiva afecciones cutáneas, reduciendo así las complicaciones y la morbilidad posteriores. Srinivasu et al. (2021).

En la investigación liderada por In Hoang et al., se desarrolló un modelo de aprendizaje profundo que utiliza una técnica de segmentación y la arquitectura Wide- ShuffleNet para clasificar lesiones cutáneas. Este enfoque se evaluó en dos conjuntos de datos ampliamente reconocidos en el campo médico: HAM10000 e ISIC2019. Los resultados mostraron un nivel de precisión del 86.33% en la segunda fase experimental, que se centró principalmente en el conjunto de datos HAM10000. Asimismo, se logró una precisión del 82.56% al aplicar este modelo al conjunto de datos ISIC.

Por otro lado, Aldhyani et al. propusieron un modelo de clasificación de lesiones cutáneas que se destaca por ser ligero y eficiente. Este modelo integra núcleos de tamaño variable en sus capas y emplea tanto la función de activación ReLU como la función de activación leaky ReLU. Sus pruebas en el conjunto de datos HAM10000 arrojaron una impresionante precisión general del 97.85%.

En contraste, Ozkan et al. se enfocaron en clasificar las lesiones cutáneas en tres grupos distintos utilizando cuatro métodos de aprendizaje automático diferentes: ANN, SVM, KNN y Árbol de decisión. Este enfoque resultó en una tasa de clasificación del 90.45%, demostrando así su efectividad en la tarea de clasificación de lesiones cutáneas.

Por su parte, Barros et al. construyeron un modelo de clasificación que abarca 12 tipos diferentes de lesiones cutáneas. Para lograrlo, generaron un conjunto de datos extenso utilizando algoritmos de aumento de datos y emplearon una arquitectura basada en ResNet-152. Este modelo logró una precisión del 94.50%, lo que muestra su capacidad para manejar una amplia variedad de tipos de lesiones cutáneas.

En la investigación de Hamed et al. introdujeron un algoritmo de clasificación Multi-Clase Multi-Nivel (MCML). Este algoritmo fue evaluado en un conjunto de 3672 imágenes recopiladas de diversas fuentes y logró una precisión del 96.47%, demostrando su eficacia en la clasificación de lesiones cutáneas en múltiples niveles y clases.

En el trabajo de Duggani et al. (2023), se presenta un enfoque para la clasificación automatizada de lesiones cutáneas en benignas o melanoma mediante dos modelos híbridos de redes neuronales convolucionales (CNN) con un clasificador SVM en la capa de salida. Estos modelos lograron una precisión del 88.02% y 87.43% respectivamente, superando a los modelos CNN tradicionales. Utilizaron el conjunto de datos ISBI 2016, logrando mejores resultados que los modelos CNN del estado del arte. La combinación de DenseNet-201 y MobileNet obtuvo la mayor precisión. La detección automática de melanoma es crucial para aumentar la tasa de supervivencia, y este enfoque podría mejorar significativamente la precisión en la clasificación de imágenes de dermatoscopia.

En la investigación de Ali et al. (2023), se presenta un enfoque para la clasificación de cáncer de piel en múltiples clases utilizando EfficientNets como un primer paso hacia la prevención del cáncer de piel. Se utilizaron EfficientNets B0-B7 para el aprendizaje por transferencia y el ajuste fino de las CNN para el conjunto de datos HAM10000. Se evaluaron la precisión, la recuperación, la exactitud, la puntuación F1, el área bajo la curva y las matrices de confusión para evaluar el rendimiento de

EfficientNets B0-B7 en esta tarea de clasificación multiclase desequilibrada. El mejor modelo, EfficientNet B4, logró un F1 Score del 87% y una precisión del 87.91%. Los hallazgos indican el rendimiento superior de EfficientNets B0-B7 para la clasificación de cáncer de piel en múltiples clases en el conjunto de datos HAM10000.

Por otro lado, Aljohani y Turki (2022), nos presentan una investigación donde se aborda la clasificación automática del cáncer de piel melanoma utilizando redes neuronales convolucionales (CNN). Se destaca la importancia de un diagnóstico preciso y temprano para detectar el melanoma, que es uno de los tipos más peligrosos de cáncer de piel. Tradicionalmente, los dermatólogos utilizan un microscopio para examinar y proporcionar un informe sobre una biopsia para el diagnóstico, un proceso que requiere experiencia y no es fácil. Por lo tanto, se propone utilizar técnicas de inteligencia artificial para facilitar el proceso de diagnóstico sin comprometer la precisión.

En el trabajo de "Diagnóstico asistido por computadora de enfermedades de la piel utilizando redes neuronales profundas." se exploró el potencial del aprendizaje profundo (Deep Learning) para clasificar cientos de enfermedades de la piel. Se utilizaron dos de los conjuntos de datos de imágenes de piel más grandes disponibles públicamente, DermNet y ISIC Archive, y se empleó la taxonomía de enfermedades para mejorar el rendimiento de clasificación de los modelos. En DermNet se estableció un nuevo estado del arte con un 80% de precisión y un 98% de Área Bajo la Curva (AUC) para la clasificación de 23 enfermedades, y se logró una precisión del 67% y un AUC del 98% para clasificar las 622 subclases únicas en este conjunto

Detección de lesiones en la piel usando visión artificial

de datos. En ISIC Archive se clasificaron las 7 enfermedades con un 93% de precisión promedio y un AUC del 99%. Las redes neuronales profundas mostraron un gran potencial para clasificar una amplia gama de enfermedades de la piel con una precisión cercana a la humana y una reproducibilidad mucho mejor. Bajwa et al. (2020).

Por otro lado, el trabajo de Kousis (2022), "Métodos de aprendizaje profundo para el reconocimiento preciso del cáncer de piel y la aplicación móvil.", se aborda el reconocimiento de cáncer de piel a partir de imágenes médicas como un área activa de investigación con el objetivo de lograr resultados más precisos. Se presenta un resumen de los conocimientos médicos necesarios sobre el cáncer de piel, seguido de un resumen extenso de los trabajos relacionados más recientes. Se presentan 11 arquitecturas de redes neuronales convolucionales (CNN) candidatas únicas. Se entrenan y prueban esas 11 arquitecturas CNN, utilizando el conjunto de datos HAM10000, que abarca siete clases de lesiones en la piel. Para abordar el problema del desequilibrio y la alta similitud entre las imágenes de algunas lesiones en la piel, se aplica aumento de datos (durante el entrenamiento), transferencia de aprendizaje y ajuste fino. De las 11 configuraciones de arquitectura CNN, DenseNet169 produjo los mejores resultados. Logró una precisión del 92.25%, una sensibilidad (recall) del 93.59% y un puntaje F1 del 93.27%, superando los esfuerzos actuales del estado del arte. Se utilizó una versión ligera de DenseNet169 en la construcción de una aplicación móvil para Android, que se mapeó como un modelo de dos clases (benigno o maligno). Se toma una foto a través de la cámara del dispositivo móvil y, después de recortarla manualmente, se clasifica como benigna o maligna. La aplicación también puede informar al usuario sobre el tiempo de exposición al sol permitido según el grado actual de radiación UV, el fototipo de piel del usuario y el grado del protector solar utilizado. En conclusión, se lograron resultados de vanguardia en el reconocimiento de cáncer de piel basados en un solo modelo de aprendizaje profundo relativamente ligero, que también se utilizó en una aplicación móvil.

Alwakid et al. (2023) nos expone en el trabajo " Detección de melanoma utilizando clasificaciones basadas en aprendizaje profundo.", la propuesta de el uso de algoritmos de aprendizaje profundo (DL) para extraer una zona de lesión con precisión en el contexto del cáncer de piel. Primero, se mejora la calidad de la imagen utilizando Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN). Luego, se utiliza la segmentación para segmentar las Regiones de Interés (ROI) de la imagen completa. Se emplea la aumentación de datos para corregir la disparidad de datos. La imagen se analiza luego con una red neuronal convolucional (CNN) y una versión modificada de Resnet-50 para clasificar las lesiones en la piel. Este análisis utilizó una muestra desigual de siete tipos de cáncer de piel del conjunto de datos HAM10000. Con una precisión del 0.86, una precisión del 0.84, una sensibilidad del 0.86 y un puntaje F de 0.86, el modelo propuesto basado en CNN superó significativamente los resultados de estudios anteriores. El estudio culmina

Detección de lesiones en la piel usando visión artificial
 con un método automatizado mejorado para diagnosticar el cáncer de piel que beneficia a los profesionales médicos y a los pacientes.

Tabla 1 Tabla recopilatoria del estado del arte

Autor	Dataset	Enfermedades Analizadas	Modelo	Precisión
R.Sumitra et al. (2015).	Conjunto de datos propio	Melanoma, bullae, queratosis seborreica y otros.	SVM y k-NN	0.61
AlSuwaidan (2022).	Eczema y psoriasis	Eczema, dermatitis atópica y psoriasis.	CNN	0.95
Sun et al. (2016).	DermQuest	198 enfermedades	CNN CaffeNET VGGNET	0.52
Oztel & Sahin (2023).	HAM10000	Queratosis actínica carcinoma de células basales, melanoma y otras	CNN	0.74
Tahir et al. (2023).	ISIC 2020, HAM10000 y DermIS.	Melanoma, carcinoma de células basales y otras	DSCC_Net	0.94
Sreekala et al. (2022).	PH2 y ISIC	Melanoma	ECNN	0.97
Khafaga et al. (2022).	Viruela del mono y Sarampión	Monkeypox	BERSFS-CNN	0.98
Wei et al. (2023).	HAM10000 y un conjunto de datos privado	Melanoma, carcinoma de células basales, carcinoma de células escamosas y otras	DenseNet201 y ConvNeXt_L	y privado: 96,49% HAM10000: 95,29%
Cai et al. (2022).	ISIC 2018 y un conjunto de datos privado	Melanoma, carcinoma de células basales, carcinoma de células escamosas y otras	Vision Transformer (ViT)	0.93
Thieme et al. (2023)	139,198 imágenes de lesiones cutáneas	Monkeypox	CNN	Sensibilidad 0.91; Especificidad :0.965
Shetty et al. (2022).	HAM10000	Melanoma, carcinoma de células basales y otras	CNN,RF	RF:0.87 CNN:0.94
Srinivasu et al. (2021).	HAM10000	Melanoma, carcinoma de células basales, carcinoma de células escamosas y otras	MobileNet V2-LSTM	0.90
Hoang et al. (2022).	ISIC 2019	Melanoma, carcinoma de células basales, carcinoma de células escamosas y otras	EW-FCM + EfficientNet-B0	0.84
Aldhyani et al. (2022).	HAM10000	Melanoma, carcinoma de células basales, carcinoma de células escamosas y otras	DKCNN model	0.98
Ozkan & Koklu (2017).	PH2	Melanoma	Neuronal Network	0.92
Barros & Silva (2018).	MED-NODE, Edinburgh 11K, Dataset, DermIs y otros	Actinic Keratosis, Carcinoma Basocelular, Nevo Melanocítico (lunar), y otras	ResNet-152	0.94
Hameed et al. (2020).	ISIC, PH2	Melanoma	AlexNet MCML	0.96
Duggani et al. (2023).	ISBI 2016	Melanoma Nevus	DenseNet-201 y MobileNet	0.88
Ali et al. (2023).	HAM10000	Melanoma, carcinoma de células basales, carcinoma de células escamosas y nevus melanocítico.	EfficientNet B4	0.87
Aljohani & Turki (2022).	ISIC	Melanoma	CNN	0.74
Bajwa et al. (2020).	ISIC Dermnet	Acné y Rosácea Queratosis Actínica, y otras	CNN	ISIC:0.93 Dermnet:0.80
Kousis et al. (2022).	HAM10000	Melanoma	DenseNet169	0.92
Alwakid et al. (2022).	HAM10000	Melanoma Nevus	Resnet-50	0.86

5. Metodología

En este capítulo, se detalla con profundidad el proceso metodológico empleado para este trabajo de tesis, el cual está dividido en dos principales vertientes, cada una crucial para el desarrollo y resultado de la investigación.

Preprocesamiento: La primera etapa, el preprocesamiento, resulta fundamental para garantizar la calidad de los datos de entrada al sistema de visión. Aquí, se describen técnicas como el aumento y la segmentación y la importancia de las mismas al momento de preparar el dataset para este estudio. El aumento de datos es un proceso crucial que implica la generación de nuevas imágenes a partir de las originales mediante transformaciones como rotaciones, desplazamientos y cambios en la escala. Por otro lado, la segmentación permite dividir una imagen en regiones más pequeñas para un análisis más detallado, lo que resulta esencial para la detección precisa de lesiones en la piel.

Comparación de enfoques: La segunda etapa implica la comparación de tres enfoques diferentes para la detección de lesiones en la piel. El primero utiliza redes neuronales convolucionales (CNN), conocidas por su eficacia en tareas de visión por computadora, especialmente en el reconocimiento de patrones en imágenes. El segundo enfoque se basa en un modelo creado con Orange Data Mining, una herramienta de minería de datos que permite crear y evaluar modelos de aprendizaje automático. Este enfoque ofrece una perspectiva alternativa y complementaria a las redes neuronales convolucionales, brindando una mayor comprensión sobre la viabilidad de otras técnicas en el campo de la detección de lesiones en la piel. El tercer enfoque utiliza Custom Vision de Microsoft Azure, una plataforma de visión por computadora que facilita el entrenamiento y la implementación de modelos personalizados en la nube. Este enfoque destaca por su accesibilidad y facilidad de uso, lo que lo convierte en una opción atractiva para aquellos sin experiencia previa en el desarrollo de modelos de visión por computadora.

4.1. Dataset

Recopilamos un conjunto de datos de imágenes de enfermedades cutáneas de diversas fuentes para crear nuestro propio dataset. Estas fuentes incluyeron el desafío HAM10000 2019, Kaggle, Google Images, Dermnet NZ, Bing Images, Yandex, Hellenic Atlas y Dermatological Atlas. Para compilar una amplia gama de imágenes representativas, empleamos técnicas de web scraping utilizando la biblioteca Selenium. Esta técnica facilitó la extracción sistemática de imágenes relevantes de las fuentes mencionadas, asegurando la diversidad y cantidad necesarias para construir el conjunto de datos. Este conjunto de datos está disponible en Kaggle[67-

68] en dos conjuntos diferentes: el "Conjunto de datos original" y el "Conjunto de datos de lesiones cutáneas utilizando segmentación".

Posteriormente, el conjunto de datos se dividió en siete categorías que representan diferentes enfermedades de la piel. Estas categorías incluyeron viruela, varicela, viruela simia, herpes, lupus, melanoma, sarampión y sarna. Para garantizar una evaluación precisa del modelo, los datos se dividieron estratégicamente en 80/20. Esto significaba que el 80% de las imágenes se asignaron al conjunto de entrenamiento, mientras que el 20% restante se reservó para el conjunto de prueba. Esta división proporcionó una cantidad adecuada de datos para entrenar el modelo y, al mismo tiempo, una muestra representativa para evaluar su rendimiento en datos previamente no vistos.

Tabla 2 Composición del dataset

Nombre	ID	Cantidad de imágenes	Resolución	Formato
Viruela	CH	12204	224x224	JPG, PNG
Herpes	HZ	11918	224x224	JPG, PNG
Lupus	LP	12096	224x224	JPG, PNG
Melanoma	Mel	12738	224x224	JPG, PNG
Sarna	SC	11901	224x224	JPG, PNG
Monkey Pox	MP	12040	224x224	JPG, PNG
Sarampión	MS	11901	224x224	JPEG, PNG

4.2. Preprocesamiento

4.2.1. Segmentación

Para segmentar las imágenes, seguimos los siguientes pasos (ver Figura 8): comenzamos convirtiéndolas al espacio de color HSV y extrayendo sus canales de tono (H), saturación (S) y valor (V). Luego, aplicamos segmentación por color utilizando umbrales definidos en el canal de tono (H) dentro del rango [0, 60], una saturación (S) mayor a 90 y un valor (V) entre 10 y 200, lo que nos permite identificar las regiones de interés. Después de este paso, la imagen se convierte a escala de grises y se aplica umbral adaptativo con un tamaño de bloque de 11 y una constante de 2 para resaltar los bordes. Posteriormente, se invierten los valores de píxel en la imagen umbralizada, seguido por operaciones de erosión con un núcleo de 4x4 y luego operaciones de dilatación con un núcleo de 3x3 para refinar la segmentación.

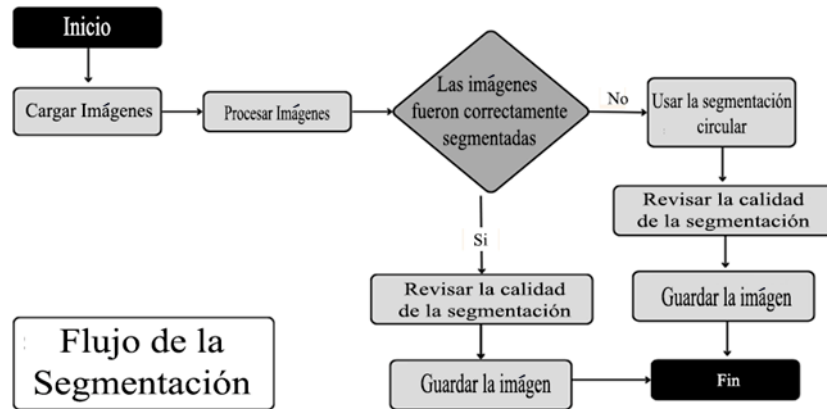


Figura. 8 Flujo de segmentación

Es importante señalar que el algoritmo no segmenta correctamente todas las imágenes del conjunto de datos debido a factores como la proximidad de las lesiones, la iluminación ambiente, etc. Para abordar esto, se tomó la decisión de crear un marco circular alrededor de los bordes de la imagen para resaltar mejor las lesiones (ver Figura 9)

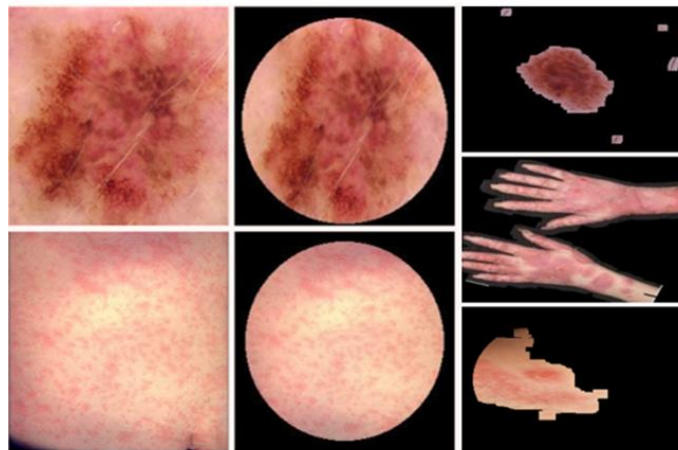


Figura. 9 Resultados de ambos tipos de segmentación

4.2.2 Aumento de Datos

Posteriormente, realizamos un aumento de datos para hacer que el conjunto de datos sea más robusto (ver Figura 10). Para este fin, se definen tres funciones de ampliación:

volteo horizontal, rotación aleatoria de ángulo y cambio aleatorio de factor de escala. A continuación, procesamos las imágenes del directorio de origen y creamos copias de cada imagen en el directorio de destino. Posteriormente, para cada copia, se aplica una combinación aleatoria de estos filtros, generando múltiples versiones ampliadas de cada imagen.

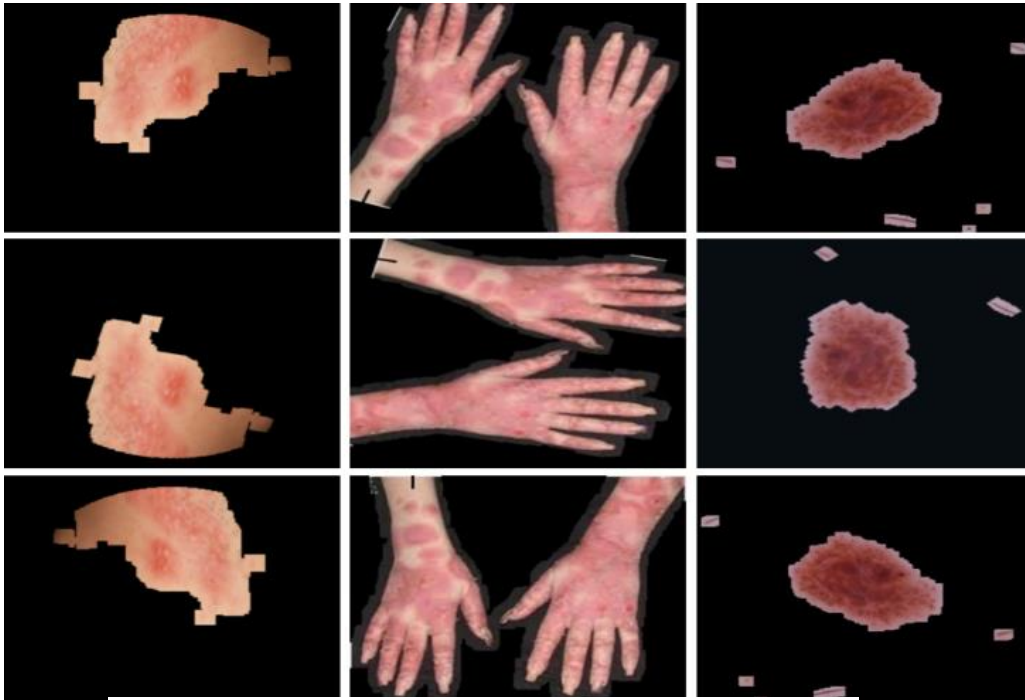


Figura. 10 Ejemplo de aumento en 3 imágenes diferentes

4.3. Azure Custom Vision

Azure Custom Vision logró tasas de precisión del 88,6% en el conjunto de entrenamiento y del 62,7% en el conjunto de validación utilizando los siguientes parámetros: Umbral de probabilidad del 60% y el modelo General (compacto). Es importante destacar que el modelo fue entrenado con solo 52,576 imágenes porque determinó que algunas imágenes eran similares (ver Tabla 3).

El modelo General (compacto) de Azure Custom Vision es una versión simplificada del modelo General, diseñado para aplicaciones donde se prioriza la velocidad y la eficiencia computacional sin sacrificar demasiado la precisión. Este modelo es adecuado para escenarios donde el tiempo de inferencia y los recursos computacionales son limitados, pero aún se requiere una precisión aceptable en las predicciones.

Tabla 3 Resultados Azure Custom Vision a detalle

Enfermedad	Entrenamiento	Cantidad de imágenes	Validación	Numero de imágenes para validación
Melanoma	99.3%	8334	100.0%	100
Sarampión	90.3%	6863	38.0%	100
MonkeyPox	89.7%	7105	49.0%	100
Sarna	86.3%	6918	78.0%	100
Herpes	84.5%	7069	70.0%	100
Varicela	83.8%	6851	73.0%	100
Lupus	83.6%	7436	95.0%	100

Detección de lesiones en la piel usando visión artificial

Los resultados de Azure Custom Vision para la precisión en el entrenamiento y la validación de diferentes enfermedades de la piel fueron los siguientes:

- Melanoma: Precisión del 99.3% en el conjunto de entrenamiento con 8334 imágenes, y del 100.0% en el conjunto de validación con 100 imágenes.
- Sarampión: Precisión del 90.3% en el conjunto de entrenamiento con 6863 imágenes, y del 38.0% en el conjunto de validación con 100 imágenes.
- Viruela del mono: Precisión del 89.7% en el conjunto de entrenamiento con 7105 imágenes, y del 49.0% en el conjunto de validación con 100 imágenes.
- Sarna: Precisión del 86.3% en el conjunto de entrenamiento con 6918 imágenes, y del 78.0% en el conjunto de validación con 100 imágenes.
- Herpes: Precisión del 84.5% en el conjunto de entrenamiento con 7069 imágenes, y del 70.0% en el conjunto de validación con 100 imágenes.
- Varicela: Precisión del 83.8% en el conjunto de entrenamiento con 6851 imágenes, y del 73.0% en el conjunto de validación con 100 imágenes.
- Lupus: Precisión del 83.6% en el conjunto de entrenamiento con 7436 imágenes, y del 95.0% en el conjunto de validación con 100 imágenes.

Estos resultados indican que el modelo tuvo un buen rendimiento en el conjunto de entrenamiento estos pueden verse mejor en la Figura 11, pero algunas enfermedades como el sarampión y la varicela tuvieron una precisión más baja en el conjunto de validación, lo que podría sugerir áreas donde se podría mejorar el modelo.

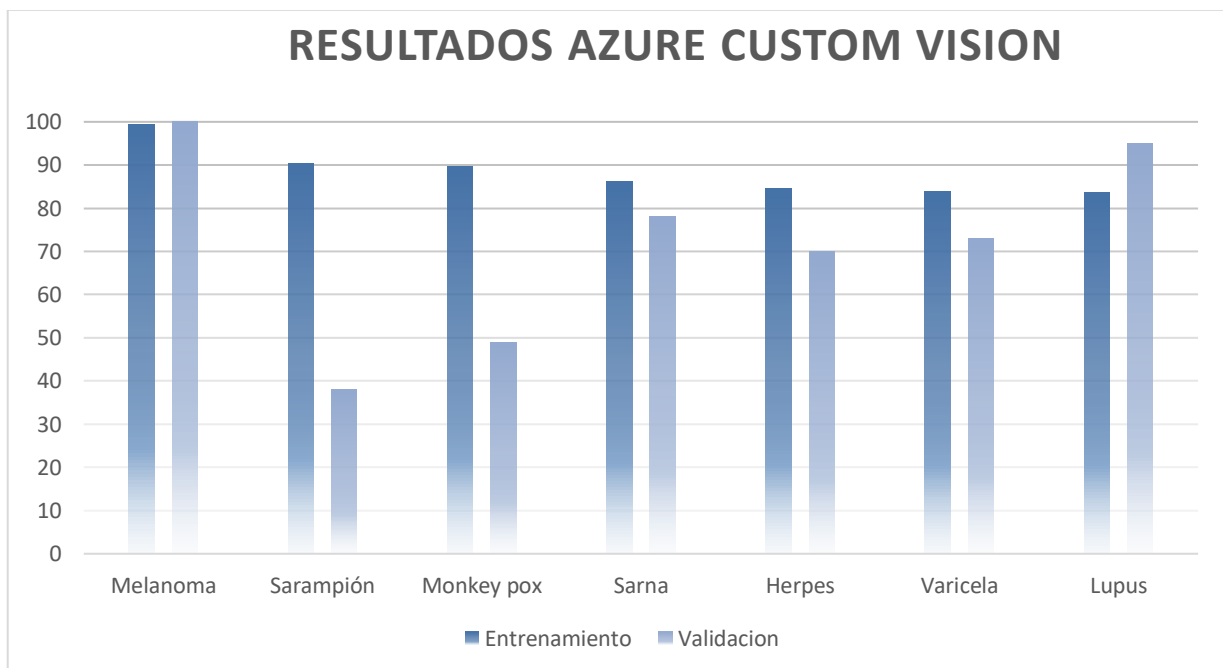


Figura. 11 Grafica de Resultados de Azure Custom Vision

4.4. Orange Data Mining

Por otro lado, Orange Data Mining se destaca como el método más efectivo según

Detección de lesiones en la piel usando visión artificial

los resultados obtenidos, tal como se muestra en la Tabla 4. Logra valores altos de AUC y precisión tanto en los conjuntos de entrenamiento como en los de validación. Además, es fácil de implementar y visualizar, como se indica en las figuras 4-5. No obstante, puede resultar un tanto desafiante comprender el funcionamiento interno del programa al analizar el código de la biblioteca paso a paso.

Este programa funciona a través de generar flujos los cuales nos permiten realizar funciones como entrenar un modelo utilizando diferentes algoritmos (Véase Figura 12) y la evaluación del mismo a través de un test, comenzamos nuestro flujo importando las imágenes y realizando un fine tuning al modelo inception V3 posteriormente estudiamos los resultados del modelo en la ventana “Test and Score” y en las ventanas que nos proporciona el programa en el widget “Confusion Matrix” para posteriormente guardar el modelo en formato .pckls

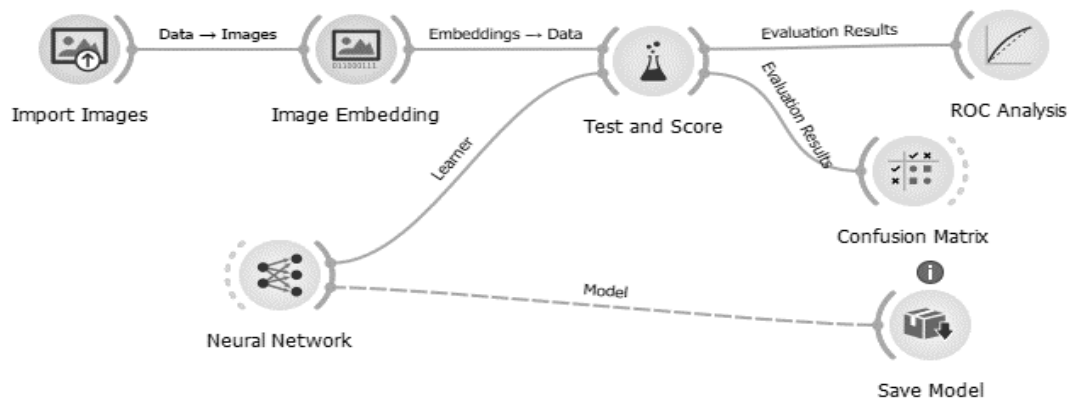


Figura. 12 Flujo de Entrenamiento de Orange Datamining

Por último, volvemos a cargar este modelo para evaluarlo con las imágenes del conjunto de validación y realizar predicciones con el modelo ya entrenado (Véase Figura 13)

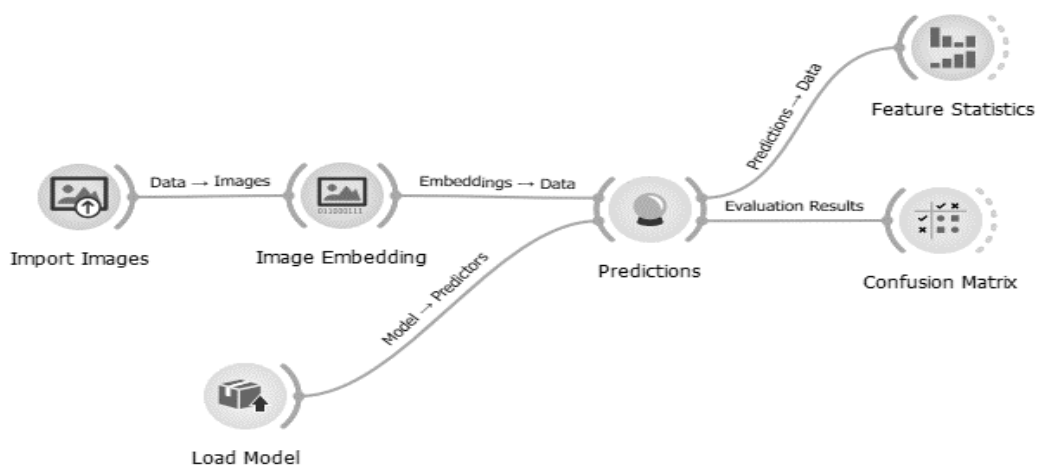


Figura. 13 Flujo de Validación de Orange Datamining

Los resultados de Orange Data Mining para la precisión en el entrenamiento y la validación de diferentes enfermedades de la piel son los siguientes:

- Melanoma: Precisión del 99.6% en el conjunto de entrenamiento con 12738 imágenes, y del 100.0% en el conjunto de validación con 100 imágenes.
- Sarampión: Precisión del 99.1% en el conjunto de entrenamiento con 11901 imágenes, y del 100.0% en el conjunto de validación con 100 imágenes.
- Viruela del mono: Precisión del 98.4% en el conjunto de entrenamiento con 12040 imágenes, y del 97.0% en el conjunto de validación con 100 imágenes.
- Sarna: Precisión del 97.7% en el conjunto de entrenamiento con 11901 imágenes, y del 99.0% en el conjunto de validación con 100 imágenes.
- Herpes: Precisión del 97.5% en el conjunto de entrenamiento con 11918 imágenes, y del 100.0% en el conjunto de validación con 100 imágenes.
- Varicela: Precisión del 97.9% en el conjunto de entrenamiento con 12204 imágenes, y del 99.0% en el conjunto de validación con 100 imágenes.
- Lupus: Precisión del 97.5% en el conjunto de entrenamiento con 12096 imágenes, y del 100.0% en el conjunto de validación con 100 imágenes.

Tabla 4 Resultados de Orange Data Mining a detalle

Enfermedad	Entrenamiento	Cantidad de imágenes para entrenamiento	Validación	Cantidad de imágenes para validación
Melanoma	99.6%	12738	100.0%	100
Sarampión	99.1%	11901	100.0%	100
MonkeyPox	98.4%	12040	97.0%	100
Sarna	97.7%	11901	99.0%	100
Herpes	97.5%	11918	100.0%	100
Varicela	97.9%	12204	99.0%	100
Lupus	97.5%	12096	100.0%	100

Estos resultados demuestran un alto nivel de precisión en el entrenamiento y la validación de las diferentes enfermedades de la piel utilizando Orange Data Mining como puede verse en la Figura 14.

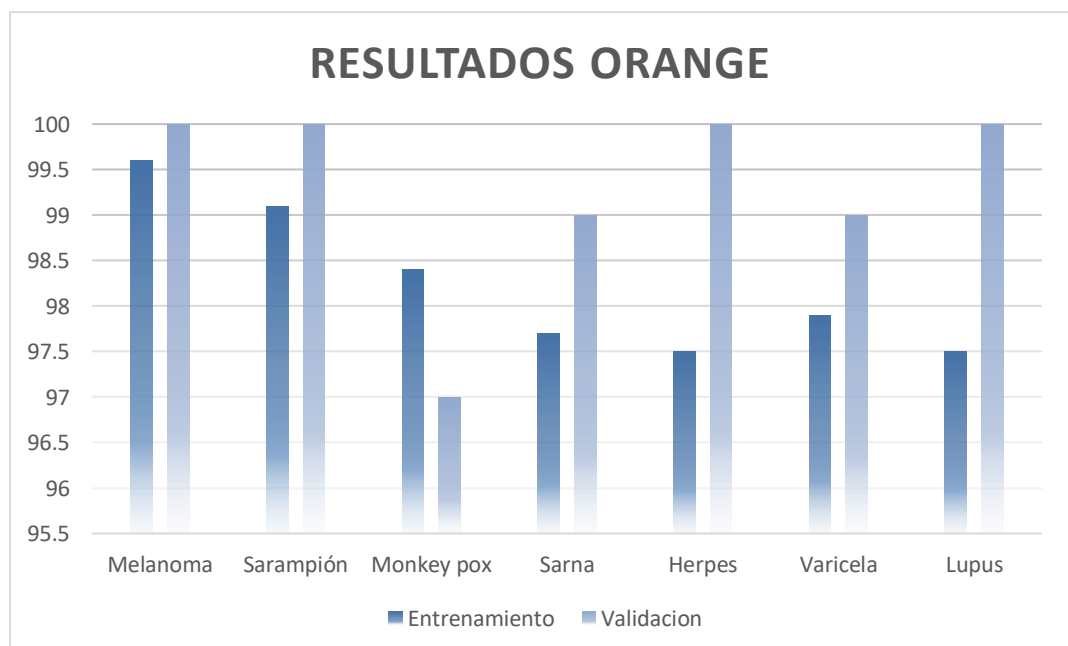


Figura. 14 Resultados de Orange Data Mining

Para la extracción de características (ver Figura 15), Orange principalmente muestra tamaño (a), ancho (b) y alto (c), junto con una variedad de características que no están explícitamente (d) especificadas en términos de lo que extraen de la imagen, como se ve en la Figura 6. Sin embargo, puedes usar un mapa de calor para identificar qué partes de la imagen son relevantes para el modelo InceptionV3, lo

que te permite comprender qué extrae de cada imagen, como se muestra en la Figura 7. Esto ayuda a proporcionar información sobre la selección de características del modelo y qué encuentra significativo en las imágenes.

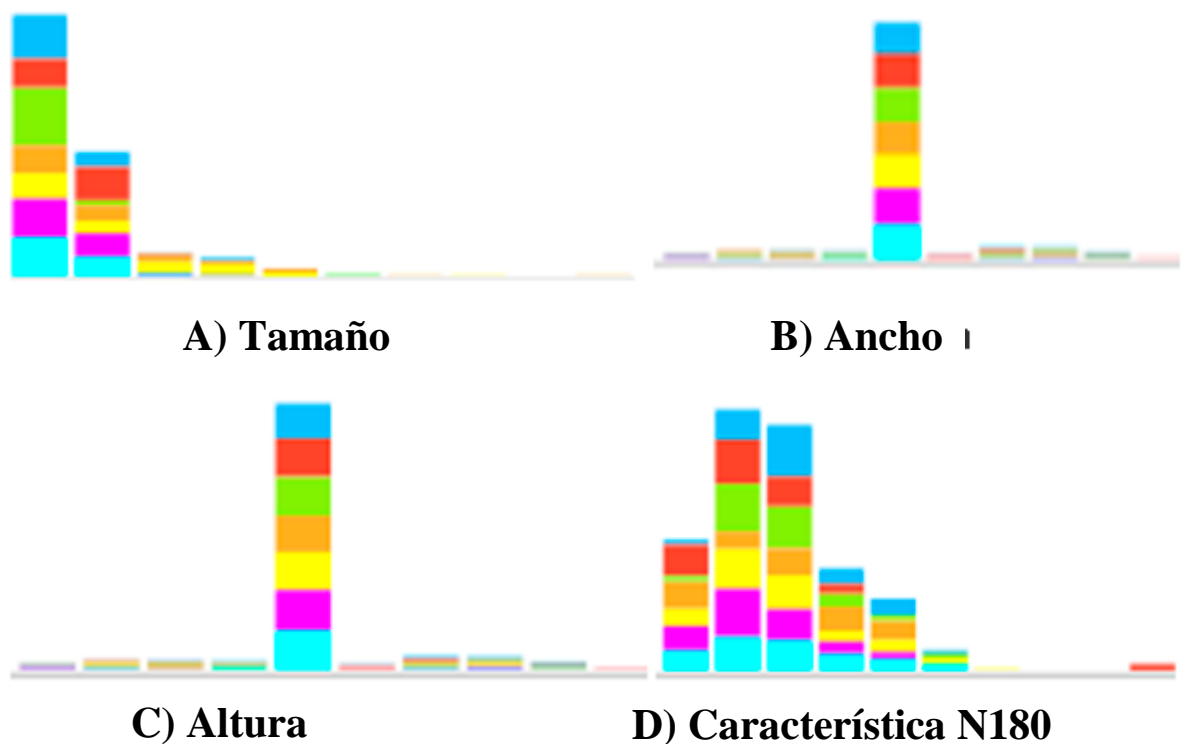


Figura. 15 Extracción de características por Orange

La tabla muestra las estadísticas resumidas de las características extraídas de las imágenes utilizando Orange Data Mining.

- **Tamaño:** representa el tamaño total de las imágenes en píxeles cuadrados, con una media de 21144.53 y una variabilidad baja (dispersión de 0.86). El tamaño mínimo de la imagen es de 1203 píxeles cuadrados y el máximo es de 159351 píxeles cuadrados.
- **Ancho y altura:** representan el ancho y alto de las imágenes respectivamente, con valores medios de 225.85. Ambos tienen una baja variabilidad (dispersión de 0.14), con valores mínimos y máximos de 113 y 334 respectivamente.

- N180: es una característica no especificada explícitamente en cuanto a su significado, pero se observa que tiene una media de 1.43446, una moda de 5.435 y una mediana de 1.32143. Su dispersión es moderada (0.5641), y los valores oscilan entre 0.081473 y 5.435. No hay valores faltantes en ninguna de las características.

Estas estadísticas proporcionan una visión general de las características extraídas de las imágenes, incluyendo su tamaño, dimensiones y otras características que podrían haber sido relevantes para el modelo utilizado en el análisis véase Figura 16.

Tabla 5 Características extraídas por Orange Data Mining

Característica	Media	Moda	Mediana	Dispersión	Min	Max	Perdida
Size	21144.53	1411	16347	0.86	1203	159351	0%
Width	225.85	224	224	0.14	113	334	0%
Height	225.85	224	224	0.14	113	334	0%
N180	1.43446	5.435	1.32143	0.5641	0.081473	5.435	0%

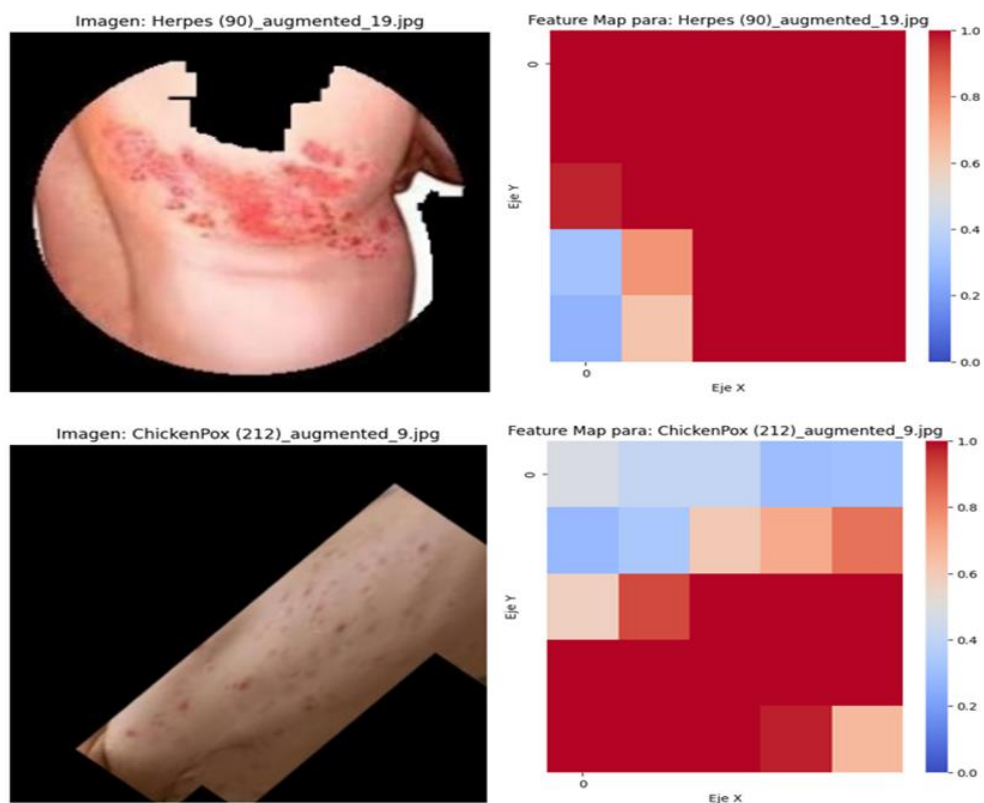


Figura. 16 Mapas de calor obtenidos de Inception v3

Para crear los mapas de calor se utilizó el siguiente código, en el cual carga el modelo preentrenado InceptionV3 y se crea un modelo intermedio que produce la salida de la penúltima capa. Este modelo se utiliza para extraer características de las imágenes de un directorio de entrenamiento. El código selecciona aleatoriamente un número específico de imágenes por subcarpeta, las procesa utilizando el modelo intermedio y muestra las imágenes junto con sus respectivos mapas de calor. Estos mapas resaltan las áreas de interés identificadas por el modelo, lo que puede ser útil

para comprender qué partes de la imagen son más relevantes para la red neuronal en tareas de reconocimiento visual.

```

1.  import os
2.  import numpy as np
3.  import random
4.  import matplotlib.pyplot as plt
5.  from tensorflow.keras.preprocessing import image
6.  from tensorflow.keras.applications.inception_v3 import preprocess_input
7.  from tensorflow.keras.models import Model
8.  import seaborn as sns
9.
10.
11.  train_data_dir = '/content/drive/MyDrive/Train'
12.
13.
14.  base_model = InceptionV3(weights='imagenet', include_top=False)
15.
16.  intermediate_layer_model = Model(inputs=base_model.input,
outputs=base_model.layers[-2].output)
17.
18.
19.  def load_and_process_image(img_path):
20.      img = image.load_img(img_path, target_size=(224, 224))
21.
22.      img = image.img_to_array(img)
23.      img = np.expand_dims(img, axis=0)
24.      img = preprocess_input(img)
25.      return img
26.
27.
28.  num_images_per_subfolder = 3
29.
30.  subfolders = os.listdir(train_data_dir)
31.
32.
33.  for subfolder in subfolders:
34.      subfolder_path = os.path.join(train_data_dir, subfolder)
35.      images = os.listdir(subfolder_path)
36.      random.shuffle(images)
37.
38.      for img_name in images[:num_images_per_subfolder]:
39.          img_path = os.path.join(subfolder_path, img_name)
40.          img = load_and_process_image(img_path)
41.          features = intermediate_layer_model.predict(img)
42.
43.
44.          plt.figure(figsize=(6, 6))
45.          plt.imshow(image.load_img(img_path))
46.          plt.title(f'Imagen: {img_name}')
47.          plt.axis('off')

```

```

48.         plt.show()
49.
50.
51.         plt.figure(figsize=(6, 6))
52.         ax = sns.heatmap(features[0, :, :, 0], cmap='coolwarm',
xticklabels=8, yticklabels=8, vmin=0, vmax=1)
53.         ax.set_xlabel('Eje X')
54.         ax.set_ylabel('Eje Y')
55.         plt.title(f'Feature Map para: {img_name}')
56.         plt.show()
57.

```

4.5. CNN creada en Google Colab

El modelo de CNN creado en Colab usando como función de activación MISH, logró un puntaje AUC decente de 0.9742 como se puede ver en la Tabla 6. Sin embargo, sus tasas de precisión fueron comparativamente más bajas, con un 86.4% (sin desglosar precisión por enfermedad) en el conjunto de entrenamiento y solo un 28% en el conjunto de validación, como se puede ver en la Tabla 6. Estos resultados sugieren que el rendimiento de la CNN creada en Colab podría ser subóptimo para la tarea dada, especialmente en términos de precisión (Véase Figura 17).

Tabla 6 Resultados de Validación de la CNN

Enfermedad	Total de Imágenes	Validación	Numero de Imágenes
Melanoma	12738	0%	100
Sarampion	11901	96%	100
MonkeyPox	12040	8%	100
Sarna	11901	1%	100
Herpes	11918	5%	100
Varicela	12204	87%	100
Lupus	12096	0%	100

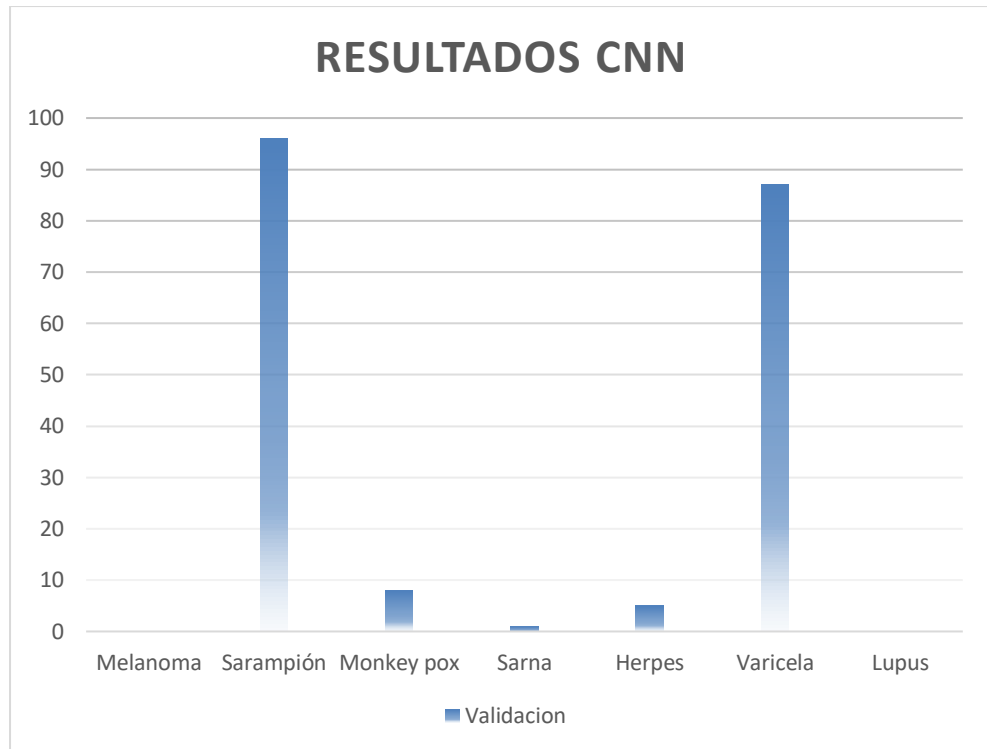


Figura. 16 Resultados de la CNN elaborada con Colab

Destacando los siguientes puntos de los resultados obtenidos:

- La enfermedad con la tasa de precisión más alta en el conjunto de entrenamiento fue el Sarampión, con un 96% , seguido por la varicela con un 87% de precisión en validación .
- Hay varias enfermedades en las que la precisión es 0%

4.6. Dispositivo de Detección Móvil

En este capítulo se describe el desarrollo de un dispositivo móvil para la detección simplificada de lesiones en la piel. Nuestra meta fue crear un dispositivo llamado D.A.N.N. (Dermatologic Analysis with Neural Networks por sus siglas en ingles), que fuera fácil de transportar, inalámbrico y útil en clínicas de nivel 1 (Véase Figura 18).

4.6.1. Hardware

La base para este dispositivo es la Raspberry Pi 4 Modelo B con 8 GB de RAM. Esta Raspberry Pi cuenta con un sistema en un chip Broadcom BCM2711, un procesador de cuatro núcleos a 1.5 GHz con ARM Cortex-A72 y una GPU VideoCore VI.

En cuanto a conectividad, el dispositivo incluye Wi-Fi 802.11ac, Bluetooth 5.0 y

Detección de lesiones en la piel usando visión artificial

Ethernet Gigabit. Para la salida de video y audio, cuenta con dos puertos micro-HDMI que admiten pantallas 4K a 60 Hz a través de HDMI 2.0, un puerto de pantalla MIPI DSI, un puerto de cámara MIPI CSI, salida estéreo de 4 polos y un puerto de video compuesto. También cuenta con 2 puertos USB 3.0 y 2 puertos USB 2.0.

El dispositivo D.A.N.N funciona con una fuente de alimentación de 20000 mAh que proporciona una salida USB-C de 5V 3A y dos salidas USB de 5V 3A. Para la interacción del usuario, utilizamos una pantalla táctil sensible a la presión de 3.5 pulgadas, que se controla con un lápiz.

Para capturar imágenes, incorporamos el Módulo de Cámara Raspberry Pi 3, que ofrece una resolución fija de 11.9 megapíxeles con un sensor de 4608 x 2592 píxeles. Este módulo puede grabar en varios modos de video, como 2304 x 1296p a 56 fps, 2304 x 1296p a 30 fps en HDR y 1536 x 864p a 120 fps. Además, para garantizar la calidad de las fotografías, incluimos una pequeña lámpara USB de 1.2W (Ver Figura 1).

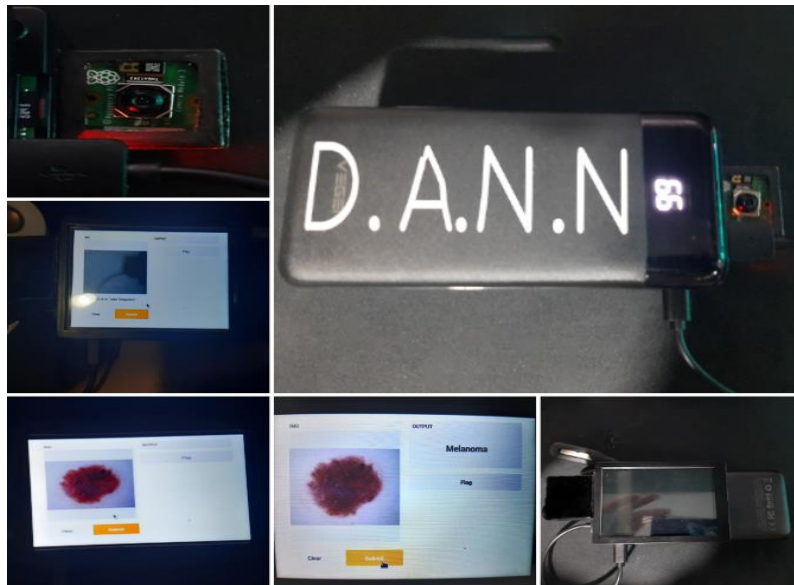


Figura. 18 Dispositivo D.A.N.N

4.6.2. Software

El dispositivo fue instalado con el sistema operativo Debian Bookworm Full de 64 bits, junto con las bibliotecas Transformers, PyTorch, Torchvision, Pillow y TIMM para obtener el modelo pre-entrenado, así como la versión 2.0.0 de Gradio para crear interfaces. Esta configuración permite la interacción con los modelos y la captura de fotografías utilizando la cámara Raspberry Pi 3. La interfaz de Gradio fue configurada en el navegador predeterminado, que en este caso es Firefox, y se utiliza para mostrar los resultados de predicción.

4.6.3. Dataset

Para este experimento, buscamos generar un modelo ligero pero eficiente, ya que debe cargarse en la primera clasificación cada vez que el dispositivo se reinicia. Para lograr esto, creamos un conjunto de datos más liviano cuyas características pueden observarse en la Tabla 7, el cual también está disponible en la plataforma de Kaggle[70] pero centrándonos únicamente en 5 enfermedades como puede observarse en la Figura 19.

Tabla. 7 Composición del dataset ligero

Nombre	ID	Cantidad de imágenes	Resolución	Formato
Viruela	CH	322	224x224	JPG, PNG
Herpes	HZ	386	224x224	JPG, PNG
Melanoma	Mel	450	224x224	JPG, PNG
Monkey Pox	MP	386	224x224	JPG, PNG
Sarampión	MS	293	224x224	JPEG, PNG

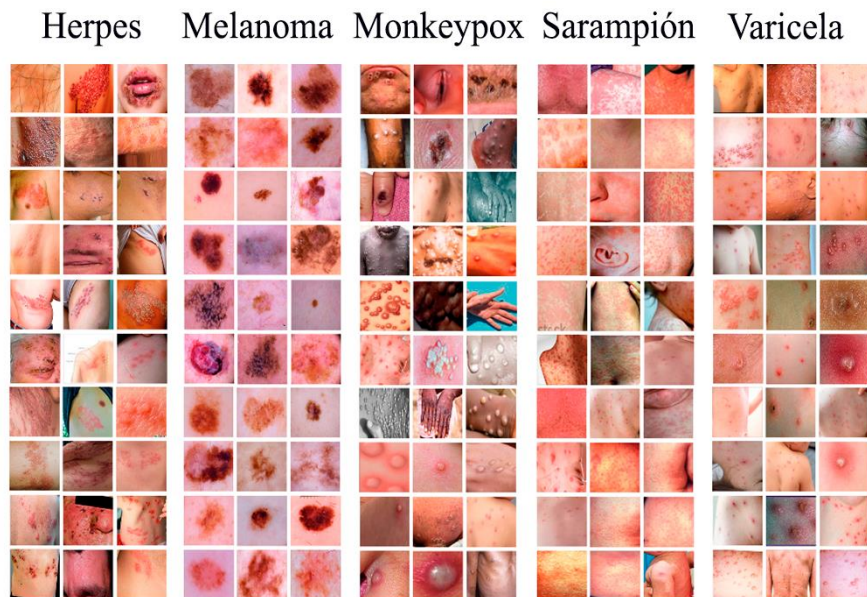


Figura. 19 Galería del Dataset

4.6.4. Entrenamiento

El modelo utilizado para este dispositivo es el modelo ViT-Base-Patch16-224 desarrollado por Google. Este modelo fue preentrenado en una gran colección de imágenes de manera supervisada como se aprecia en la Figura 20. Se preentrenó en ImageNet-21k, que consta de 14 millones de imágenes y 21,843 clases, con una resolución de 224x224 píxeles. Después de este preentrenamiento, el modelo fue afinado en ImageNet 2012, un conjunto de datos que comprende 1 millón de imágenes y 1,000 clases, también con una resolución de 224x224 (Dosovitskiy,2020). Más tarde, entrenamos esta última capa congelando los pesos

Detección de lesiones en la piel usando visión artificial

de las capas anteriores para poder clasificar nuestros objetos de estudio utilizando transferencia de aprendizaje.

El proceso de entrenamiento implicó probar dos conjuntos de datos, uno segmentado y otro no segmentado, ambos utilizando los siguientes hiperparámetros: siete épocas, una función de pérdida de entropía cruzada y un optimizador SGD con una tasa de aprendizaje de $1e-4$, un momento de 0.9 y un decaimiento de peso de $1e-4$.

Durante el entrenamiento, iteramos sobre el DataLoader que contiene las imágenes de entrenamiento, calculando la pérdida en cada iteración y ajustando los pesos del modelo para minimizar esta pérdida. La pérdida y la precisión del modelo se evaluaron por época. El modelo se guardó después de cada iteración para su posterior evaluación.

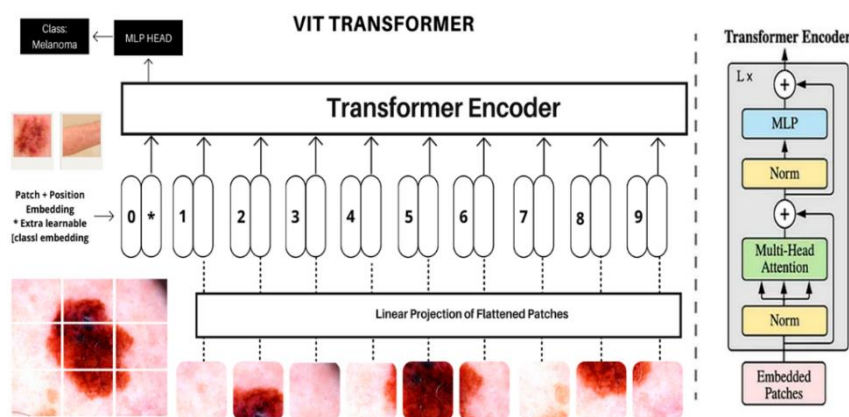


Figura. 20 Modelo propuesto para el Dispositivo D.A.N.N basado en el Modelo de Aprendizaje por Transferencia VIT-Base-Patch16-224.

5.1. Discusión de los resultados

En resumen, el análisis de los resultados revela que Azure Custom Vision demostró una precisión moderada, alcanzando un 88.6% en el conjunto de entrenamiento y un 62.7% en el conjunto de validación. En contraste, Orange Data Mining exhibió consistentemente una alta precisión en ambos conjuntos, con algunas enfermedades mostrando resultados notablemente sólidos. Por otro lado, la CNN mostró una precisión más baja, especialmente en el conjunto de validación, donde solo alcanzó un 28% (ver Tabla 8).

Al comparar nuestros resultados con estudios previos, observamos una amplia variabilidad en la precisión de los modelos de detección de enfermedades de la piel. Mientras que algunos modelos alcanzaron precisiones superiores al 0.95, como en el caso de AlSuwaidan (2022) y la importancia de la obtención de imágenes dermatológicas que hayan sido tomadas en los escenarios óptimos como es el caso del dataset HAM10000 en el cual se recopilaron imágenes de una excelente calidad. Nuestro modelo de CNN en Colab mostró una precisión razonable en el conjunto de entrenamiento (0.864), pero una caída en la labor de validación (0.2814), sugiriendo

Detección de lesiones en la piel usando visión artificial

margen de mejora, especialmente en la generalización de la validación. Estos hallazgos resaltan la importancia continua de mejorar los modelos de aprendizaje automático para una detección precisa de enfermedades de la piel.

Tabla. 8 Comparación de los modelos y sus resultados

Modelo	Parametros	AUC	Precisión en entrenamiento	Precisión en validación
Orange Datamining	Función: Relu	0.982	98.2%	99.8%
	Train:90% Test:10% Épocas:20 Folds:10			
Azure Custom Vision	General Compact S1	-	88.6%	62.7%
CNN	Función: MISH	0.974	86.4%	28.14%
	Train:80%			
	Test:20%			
	Épocas:20 Batch size:128			

Por la parte de las pruebas en el modelo transformer del dispositivo móvil, los resultados del entrenamiento con las imágenes segmentadas mostraron un aumento significativo en la precisión del modelo a lo largo de las épocas, demostrando un gran rendimiento en la clasificación, como se muestra en la Figura 21, logrando una precisión del 99.9% en la última época, mientras que la versión no segmentada alcanzó un porcentaje del 99.4%. Ambos modelos lograron un AUC de 1.0, como se muestra en la Tabla 9.

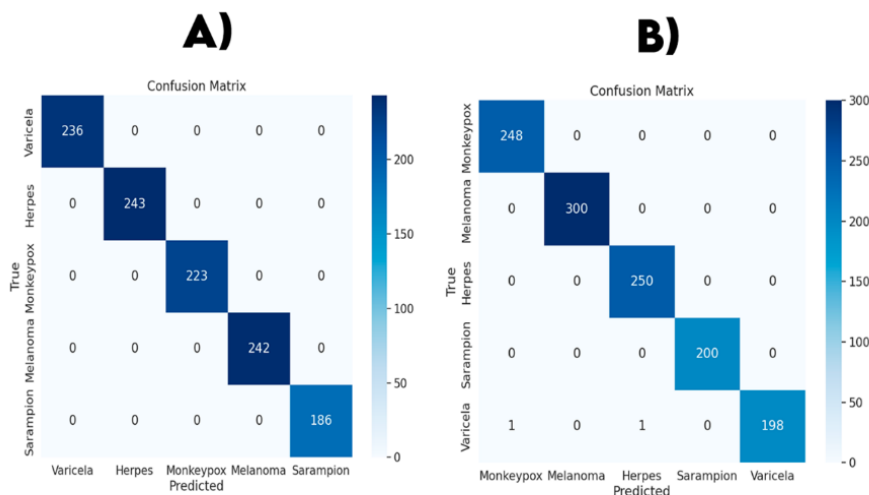


Figura. 21 Matrices de confusión durante la etapa de entrenamiento A) Segmentado B) No segmentado

Tabla. 9 Comparación de resultados entre el modelo Segmentado y Sin Segmentación

Modelo	Precisión	Loss	AUC	ACC	Recall	F1-Score
Segmentado	99.9%	0.007	1.0	1.0	1.0	1.0
No Segmentado	99.6%	0.020	1.0	1.0	1.0	1.0

Para la validación, evaluamos los últimos 3 modelos generados en cada época porque cada uno mostró un rendimiento diferente (ver Figura 22 y 23). En algunos casos, el modelo número 7

Detección de lesiones en la piel usando visión artificial

no siempre es el más efectivo para esta tarea, como se muestra en la tabla. Por lo tanto, elegimos el modelo con la mejor capacidad de validación, en este caso, el modelo número 6 sin segmentación (ver Tabla 9 y 10).

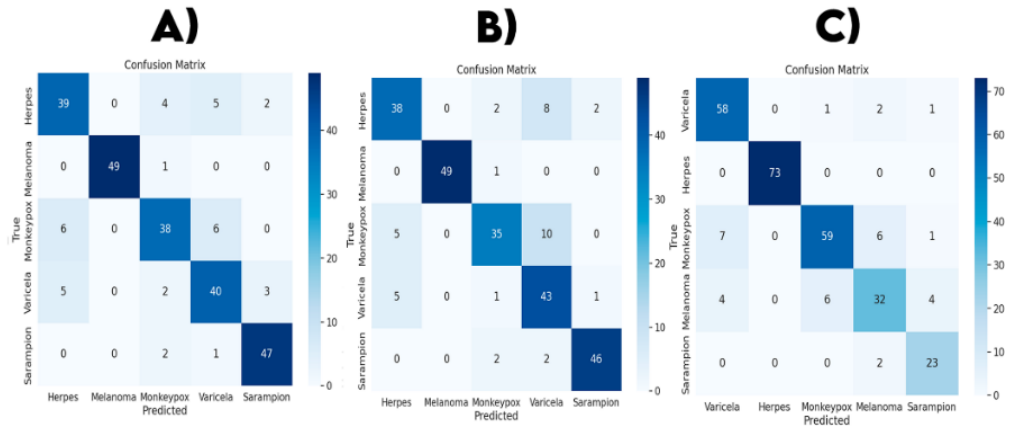


Figura. 22 Resultados de validación de los modelos segmentados A) modelo 5, B) modelo 6 y C) modelo 7

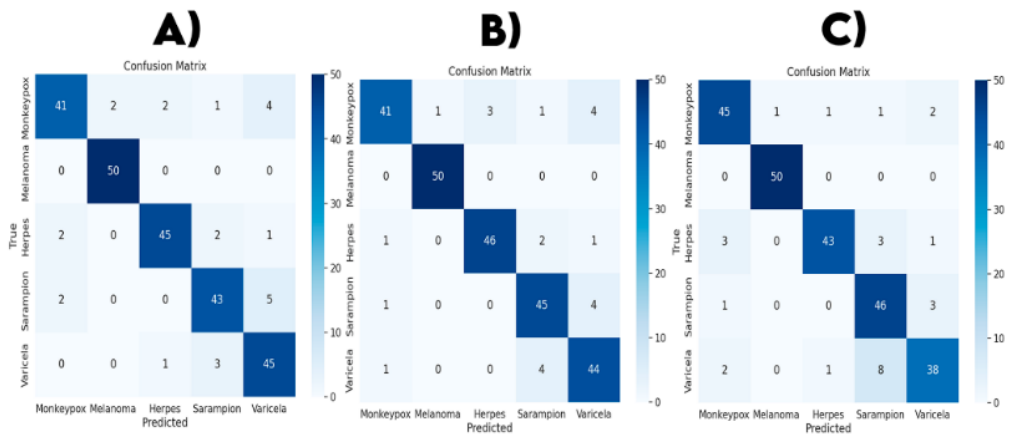


Figura. 33 Resultados de validación de los modelos no segmentados A) modelo 5, B) modelo 6 y C) modelo 7

Tabla. 9 Comparación de resultados entre el modelo Segmentado y Sin Segmentación

Modelo/Precisión	Época 5	Época 6	Época 7
Segmentado	85.20%	84.40%	87.81%
No Segmentado	89.95%	90.76%	89.15%

Tabla. 10 Resultados desglosados por enfermedad del modelo 6 sin segmentar

Enfermedad	Precisión	ACC	Recall	F1-Score
Monkeypox	89.0%	0.88	0.82	0.85
Melanoma	98.0%	0.88	0.99	0.99
Herpes	94.0%	0.88	0.92	0.92
Sarampion	79.0%	0.88	0.83	0.83
Varicela	81.0%	0.88	0.80	0.80

Para las pruebas en el dispositivo, se eligió el mejor modelo, en este caso, la sexta época sin segmentación. Se tomaron fotografías a varias distancias, concluyendo que 15 cm es la distancia óptima para tomar la fotografía. Durante las pruebas, se

Detección de lesiones en la piel usando visión artificial

utilizaron conjuntos de datos con zooms de salida del 10%, 30% y 50%, observando una disminución en la precisión y el rendimiento a medida que la distancia de captura se alejaba de los 15 cm. Los resultados muestran que, sin zoom, se logró una precisión del 90.76, con un zoom del 10% fue del 87.9, con un zoom del 30% fue del 85.9, y con un zoom del 50% fue del 72.2, como se puede ver en la Figura 24.

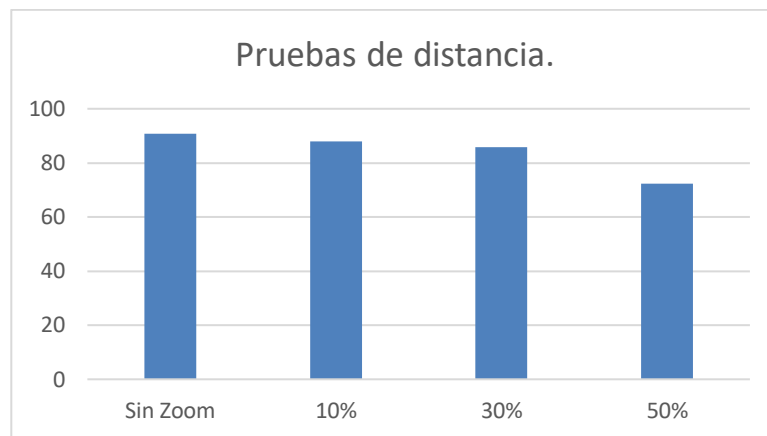


Figura. 24 Test de distancia para las fotografías

CONCLUSIONES

Este estudio comparativo examinó plataformas como Orange, Azure Custom Vision y Google Colab para la identificación de lesiones cutáneas, con el objetivo de obtener una comprensión general de estas herramientas y sus resultados. Se observó que la diversidad de imágenes y el entrenamiento son aspectos cruciales en la efectividad de los modelos. Por ejemplo, en Orange se utilizó el modelo InceptionV3, entrenado con una amplia variedad de imágenes y clases, mientras que en Azure se emplearon modelos compactos que también demostraron una gran solidez en los resultados. En contraste, se comparó con una CNN creada desde cero, cuyo entrenamiento se basó únicamente en las imágenes utilizadas.

Uno de los principales desafíos fue la creación de un conjunto de datos adecuado y la segmentación de lesiones, aspectos que se abordará más adelante en esta conclusión. Además, se destacaron las limitaciones de usar modelos desarrollados en software que no permiten un control completo sobre la arquitectura, como sería el caso de un modelo programado línea por línea.

Por la parte del dispositivo móvil se destaca la agilidad al momento de realizar la predicción además de que el modelo solo pesa 327.4MB en contraste con el modelo de Orange el 1.09GB y la precisión en validación que tiene un sólido 90.76% de validación, es importante destacar que en este apartado no se realizó aumento, aunque sería bueno a futuro realizar un aumento con técnicas avanzadas como el MixUp o un CutUp

Para cerrar, se resalta la efectividad del modelo en la agilización de tiempos, lo que puede ser una herramienta valiosa para médicos especialistas y no especialistas al

Detección de lesiones en la piel usando visión artificial

buscar una segunda opinión sobre el diagnóstico de una lesión o para canalizar a un paciente hacia la atención adecuada. Se sugiere continuar el desarrollo de la investigación explorando formas de recrear el mismo modelo con recursos que permitan un mayor control y adaptación, así como un estudio más profundo del mismo.

- Dataset y Preprocesamiento

En este apartado, se destaca la complejidad de crear un dataset eficiente, donde las imágenes proporcionen las características adecuadas en el escenario correcto. Por ejemplo, en datasets como HAM10000 e ISIC, las imágenes dermatológicas son de alta calidad, a diferencia de aquellas obtenidas de otros conjuntos de datos o a través de técnicas como web scrapping, que, si bien pueden ser útiles, no garantizan la validación de las imágenes.

- Librerías

Se utilizaron diversas librerías, como OpenCV para la segmentación y TensorFlow y Keras para el entrenamiento de los modelos. Sin embargo, surgió un desafío con Orange y su librería Orange 3, debido al formato único que permite exportar los modelos. Esto dificultó la transferencia a formatos más flexibles y con mayor soporte, como el .h5, utilizado por otras librerías y entornos.

- Segmentación de imágenes

La tarea de segmentación en este proyecto resultó ser bastante complicada debido a la variedad de lesiones, las cuales presentan diferentes formas y características. Algunas de estas lesiones incluso comparten similitudes con otros elementos del cuerpo humano, como los ojos, pezones o labios. Estos bordes, protuberancias o áreas con alta coloración en la piel pueden confundir al segmentador y complicar la labor.

Sería beneficioso considerar nuevas opciones para la segmentación, como los segmentadores privados de Meta o tecnologías diferentes como YoLo. Estas alternativas podrían ofrecer métodos más efectivos para lidiar con la complejidad de las lesiones cutáneas y mejorar la precisión en la identificación y delimitación de las mismas.

Observando hacia el futuro, me gustaría continuar explorando soluciones para este problema, que es extremadamente amplio. Además, me gustaría abordar el desafío de analizar otros tipos de lesiones cutáneas, dado que la cantidad de enfermedades que presentan lesiones en la piel es abismal, aunque ello requiera diferentes modelos. También tengo interés en mejorar el modelo utilizado en la Raspberry Pi, para que este dispositivo sea aún más valioso en el campo de la dermatología.

En resumen, este estudio comparativo destacó la importancia del entrenamiento y la diversidad de imágenes en la efectividad de los modelos de identificación de lesiones cutáneas. A pesar de los desafíos encontrados, se evidenció el potencial de estas herramientas para agilizar los tiempos de diagnóstico y ofrecer segundas opiniones a médicos especialistas y no especialistas. Continuar investigando en esta área puede llevar a avances significativos en el campo de la dermatología y visión por computadora.

Referencias

- [1] World Cancer Research Fund International. (2021). Skin Cancer. Obtenido de American Institute for Cancer Research: <https://cutt.ly/sO3cxT1>
- [2] INEGI. (2021,14 de Febrero). ESTADÍSTICAS A PROPÓSITO DEL DÍA MUNDIAL CONTRA EL CÁNCER, [Comunicado de prensa]. Obtenido de <https://cutt.ly/JO3cZKb>
- [3] Moraga, F. A., & Campins, M. (2000). Vacuna de la varicela. Una inmunización del niño y del adulto. *Enfermedades Infecciosas y Microbiología Clínica*, 18(10), 516-518.
- [4] Martínez-Fernández, D. E., Fernández-Quezada, D., Casillas-Muñoz, F. A. G., Carrillo-Ballesteros, F. J., Ortega-Prieto, A. M., Jiménez-Guardeño, J. M., & Regla-Nava, J. A. (2023). Human Monkeypox: Una revisión integral de epidemiología, patogénesis, diagnóstico, tratamiento y estrategias de prevención. *Patógenos*, 12(947).
- [5] Suvvari, T. K., Kandi, V., Mohapatra, R. K., Chopra, H., Islam, M. A., & Dhama, K. (2023). The re-emergence of measles is posing an imminent global threat owing to decline in its vaccination rates amid COVID-19 pandemic: a special focus on recent outbreak in India – a call for massive vaccination drive to be enhanced at global level. *International Journal of Surgery*, 109(2), 198–200. <https://doi.org/10.1097/JS9.0000000000000228>
- [6] Avilés Izquierdo, J.A., Cano Martínez, N., & Lázaro Ochoita, P. (2013). Características epidemiológicas de los pacientes con lupus eritematoso cutáneo. *Actas Dermo-Sifiliográficas*, (5), 409-414.
- [7] Kawai, K., Yawn, B. P., Wollan, P., & Harpaz, R. (2016). Increasing incidence of herpes zoster over a 60-year period from a population-based study. *Clinical Infectious Diseases*, 63(2), 221–226. <https://doi.org/10.1093/cid/ciw296>
- [8] Welch, E., Romani, L., & Whitfeld, M. J. (2021). Recent advances in understanding and treating scabies. *Fac Rev*, 10, 28. <https://doi.org/10.12703/r/10-28>
- [9] Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach*. Pearson.
- [10] Turing, A. (1950). Computing machinery and intelligence. *Mind*, 236, 433-460.
- [11] Gross, G., Schofer, H., Wassilew, S., Friese, K., & Timm, A. (2003). Herpes zoster guideline of the German Dermatology Society. *Journal of Clinical Virology*, 26, 277-289.
- [12] Wenzel, J. (2019). Cutaneous lupus erythematosus: new insights into pathogenesis and therapeutic strategies. *Nat Rev Rheumatol*, 15(9), 519-532. <https://doi.org/10.1038/s41584-019-0272-0>
- [13] Kondamudi, N. P., & Waymack, J. R. (2023). Measles. In StatPearls. Last Update: August 12, 2023. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/31399711/>

- [14] Chandler, D. J., & Fuller, L. C. (2019). A Review of Scabies: An Infestation More than Skin Deep. *Dermatology*, 235(2), 79–90. <https://doi.org/10.1159/000495290>
- [15] Ferdous, J., Barek, M. A., Hossen, M. S., Bhowmik, K. K., & Islam, M. S. (2023). A review on monkeypox virus outbreak: New challenge for world. *Health Science Reports*, 6(1), e1007. <https://doi.org/10.1002/hsr2.1007>
- [16] Walling, H. W., & Sontheimer, R. D. (2009). Lupus eritematoso cutáneo: aspectos en diagnóstico y tratamiento. *American Journal of Clinical Dermatology*, 10(6), 365-381.
- [17] Irham, L. M., Adikusuma, W., Lolita, L., Puspitaningrum, A. N., Afief, A. R., Sarasmita, M. A., ... Chong, R. (2023). Investigation of susceptibility genes for chickenpox disease across multiple continents. *Biochemistry and Biophysics Reports*, 33, 101419. <https://doi.org/10.1016/j.bbrep.2022.101419>
- [18] Microsoft. (2023). Azure AI Custom Vision. Recuperado de <https://n9.cl/msazure>
- [19] Carrión, C. B. (2023). REDES CONVOLUCIONALES. Recuperado de <https://n9.cl/carmelobonilla>
- [20] Liu, Qian, y Steve Furber. "Noisy Softplus: a biology inspired activation function." En Conferencia Internacional sobre Procesamiento de Información Neural, pp. 405-412. Springer, Cham, 2016. Recuperado de <https://n9.cl/liu-quian>
- [21] "Orange Visual Programming," Orange3, 2023. [Online]. Recuperado de: <https://n9.cl/orange3> . [Consultado: Nov. 9, 2023].
- [22] S. Sharma and S. Lohchab, "Personal Authentication Using Finger Vein Biometric Technology with Implementation of Transfer Learning CNN Model," SSRN Papers, Dec. 28, 2021. [Online]. Available: <https://n9.cl/sharma> . [Accessed: Nov. 9, 2023].
- [23] Wu, D., and Sun, D.-W. Colour measurements by computer vision for food quality control – A review. *Trends in Food Science & Technology* 2013; 29(1), 5-20
- [24] Aptoula, E., and Lefèvre, S. On the morphological processing of hue. *Image and Vision Computing* 2009; 27(9), 1394-1401.
- [25] Nnolim, U. A. Design and implementation of novel, fast, pipelined HSI2RGB and log-hybrid RGB2HSI colour converter architectures for image enhancement. *Microprocessors and Microsystems* 2015; 39(4–5),223-236.
- [26] R. Szeliski, *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.
- [27] D. Forsyth and J. Ponce, *Computer vision: a modern approach*. Prentice Hall Professional Technical Reference, 2002.
- [28] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [29] L. Najman and M. Schmitt, "Watershed of a continuous function," *Signal Processing*, vol. 38, no. 1, pp. 99–112, 1994
- [30] N. Plath, M. Toussaint, and S. Nakajima, "Multi-class image segmentation using conditional random fields and global classification," in *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM, 2009, pp. 817–824.
- [31] J.-L. Starck, M. Elad, and D. L. Donoho, "Image decomposition via the

- combination of sparse representations and a variational approach,” *IEEE transactions on image processing*, vol. 14, no. 10, pp. 1570–1582, 2005.
- [32] Minaee, S., Boykov, Y., Porikli, F., Plaza, A., Kehtarnavaz, N., & Terzopoulos, D. (2020). Image segmentation using deep learning: A survey. *arXiv preprint arXiv:2001.05566v5* [cs.CV].
- [33] Yulistira, N.; Kurita, T. Gated spatio and temporal convolutional neural network for activity recognition: Towards gated multimodal deep learning. *Eurasip J. Image Video Process.* 2017, 2017, 1–12.
- [34] Alomar, K., Aysel, H. I., & Cai, X. (2023). Data Augmentation in Classification and Segmentation: A Survey and New Strategies. *J. Imaging*, 9(2), 46. <https://doi.org/10.3390/jimaging9020046>
- [35] Raspberry. (2019). Raspberry Pi 4 Model B. Raspberry Pi. Consultado en 2024, de <https://www.raspberrypi.com/products/raspberry-pi-4-model-b/>
- [36] Sumitra., Suhilb, M., & Guruc, D. (2015). Segmentation and Classification of Skin Lesions for Disease Diagnosis. En *International Conference on Advanced Computing Technologies and Applications (ICACTA 2015)*.
- [37] AlSuwaidan, L. (2022). Deep Learning Based Classification of Dermatological Disorders. *International Conference on Advanced Computing Technologies and Applications*.
- [38] Sun, X., Yang, J., Sun, M., & Wang, K. (2016). A Benchmark for Automatic Visual Classification of Clinical Skin Disease Images (Vol. 9910). DOI: 10.1007/978-3-319-46466-4_13
- [39] Oztel, I., Yolcu Oztel, G., & Sahin, V. H. (2023). Deep Learning-Based Skin Diseases Classification using Smartphones. *Advanced Intelligent Systems*, 5, 2300211. <https://doi.org/10.1002/aisy.202300211>
- [40] Tahir, M., Naeem, A., Malik, H., Tanveer, J., Naqvi, R. A., & Lee, S.-W. (2023). DSCC_Net: Multi-Classification Deep Learning Models for Diagnosing of Skin Cancer Using Dermoscopic Images. *Advanced Intelligent Systems*. DOI: 10.1002/aisy.202300211.
- [41] Sreekala, K., et al. (2022). Skin Diseases Classification Using Hybrid AI Based Localization Approach. *Computational Intelligence and Neuroscience*. DOI: [DOI]
- [42] Khafaga, D.S., Ibrahim, A., El-Kenawy, E.-S.M., Abdelhamid, A.A., Karim, F.K., Mirjalili, S., Khodadadi, N., Lim, W.H., Eid, M.M., Ghoneim, M.E. (2022). An Al-Biruni Earth Radius Optimization-Based Deep Convolutional Neural Network for Classifying Monkeypox Disease. *Diagnostics*, 12, 2892. <https://doi.org/10.3390/diagnostics12112892>
- [43] Wei, M., Wu, Q., Ji, H., Wang, J., Lyu, T., Liu, J., & Zhao, L. (2023). A Skin Disease Classification Model Based on DenseNet and ConvNeXt Fusion. *Electronics*, 12, 438.
- [44] Cai, G., Zhu, Y., Wu, Y., Jiang, X., Ye, J., & Yang, D. (2022). A multimodal transformer to fuse images and metadata for skin disease classification. *European*

- [45] Thieme, A. H., Zheng, Y., Machiraju, G., Sadee, C., Mittermaier, M., Gertler, M., . Gevaert, O. (n.d.). A deep-learning algorithm to classify skin lesions from mpox virus infection.
- [46] Shetty, B., Fernandes, R., Rodrigues, A. P., Chengoden, R., Bhattacharya, S., & Lakshmana, K. (2022). Skin lesion classification of dermoscopic images using machine learning and convolutional neural network. *Scientific Reports*, 12, 18134. <https://doi.org/10.1038/s41598-022-22644-9>.
- [47] Srinivasu, P. N., SivaSai, J. G., Ijaz, M. F., Bhoi, A. K., Kim, W., & Kang, J. J. (2021). Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM. *Sensors*, 21(8), 2852. <https://doi.org/10.3390/s21082852>.
- [48] L. Hoang, S.-H. Lee, E.-J. Lee, and K.-R. Kwon, "Multiclass Skin Lesion Classification Using a Novel Lightweight Deep Learning Framework for Smart Healthcare," *Appl. Sci.*, vol. 12, no. 5, pp. 2677, Mar. 2022, Available: <https://n9.cl/hoang> , doi: 10.3390/app12052677
- [49] Aldhyani THH, Verma A, Al-Adhaileh MH, Koundal D. Multi-Class Skin Lesion Classification Using a Lightweight Dynamic Kernel Deep-Learning-Based Convolutional Neural Network. *Diagnostics (Basel)*, Available: <https://n9.cl/theyazn> , 2022 Aug 24;12(9):2048. doi: 10.3390/diagnostics12092048. PMID: 36140447; PMCID: PMC9497471
- [50] I. A. Ozkan and M. Koklu, "Skin Lesion Classification using Machine Learning Algorithms," *Int. J. Intell. Syst. Appl. Eng.*, vol. 5, no. 4, pp. 285-289, Dec. 2017, Available: <https://n9.cl/ozkan> , doi: 10.18201/ijisae.2017534420.
- [51] D. B. Barros and N. C. D. Silva, "Skin Lesions Classification Using Convolutional Neural Networks in Clinical Images," *arXiv preprint arXiv:1812.02316*, Dec. 2018. [Online]. Available: <https://n9.cl/barros> . [Accessed: Nov. 10, 2023].
- [52] N. Hameed, A. M. Shabut, M. K. Ghosh, and M. A. Hossain, "Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques," *Expert Syst. Appl.*, vol. 141, Available: <https://n9.cl/hameed> ,pp. 112961, Mar. 2020, doi: 10.1016/j.eswa.2019.112961
- [53] Duggani, K., Venugopal, V., Nath, M. K., & Mishra, M. (2023). Hybrid convolutional neural networks with SVM classifier for classification of skin cancer. *Biomedical Engineering Advances*, 5(2023), 100069.
- [54] Ali, K., Shaikh, Z. A., Khan, A. A., & Laghari, A. A. (2023). Multiclass skin cancer classification using EfficientNets – a first step towards preventing skin cancer. *Neuroscience Informatics*.
- [55] Aljohani, K., & Turki, T. (2022). Automatic Classification of Melanoma Skin Cancer with Deep Convolutional Neural Networks. *AI*, 3(2), 512–525. <https://doi.org/10.3390/ai3020029>
- [56] Bajwa, M. N., Muta, K., Malik, M. I., Siddiqui, S. A., Braun, S. A., Homey, B.,

- Dengel, A., & Ahmed, S. (2020). Computer-Aided Diagnosis of Skin Diseases using Deep Neural Networks. *Applied Sciences*, 10(7), 2488. DOI: 10.3390/app10072488
- [57] Kousis, I., Perikos, I., Hatzilygeroudis, I., & Virvou, M. (2022). Deep Learning Methods for Accurate Skin Cancer Recognition and Mobile Application. *Electronics*, 11(9), 1294. <https://doi.org/10.3390/electronics11091294>.
- [58] Alwakid, G., Gouda, W., Humayun, M., & Sama, N. U. (2022). Melanoma Detection Using Deep Learning-Based Classifications. *Healthcare*, 10(10), 2481. <https://doi.org/10.3390/healthcare10102481>.
- [59] "Dermnet," Kaggle, Jun. 24, 2020. [Online]. Available: <https://n9.cl/dermnet>
- [60] "Monkeypox Skin Lesion Dataset," Kaggle, Jul. 5, 2022. [Online]. Available: <https://n9.cl/monkeypox> . [Accessed: Nov. 10, 2023].
- [61] "POX DATASET," Kaggle, Aug. 31, 2022. [Online]. Available: <https://n9.cl/balasubramaniamv> . [Accessed: Nov. 10, 2023].
- [62] D. Bala and M. S. Hossain, "Monkeypox Skin Images Dataset (MSID)," Mendeley Data, Feb. 23, 2023. [Online]. Available: <https://n9.cl/mendeley-pox> . [Accessed: Nov. 10, 2023].
- [63] U. Shahab, "Dermnet Skin Disease Images," Kaggle, May 21, 2021. [Online]. Available: <https://n9.cl/dermnet2> . [Accessed: Nov. 10, 2023].
- [64] "ISIC Challenge Datasets," ISIC Archive, Nov. 8, 2023. [Online]. Available: <https://n9.cl/isic-challenge> . [Accessed: Nov. 10, 2023].
- [65] R. P. Usatine and B. D. Madden, "Interactive Dermatology Atlas," Dermatlas.net. [Online]. Available: <https://n9.cl/dermatlas> . [Accessed: Nov. 10, 2023].
- [66] Hellenic Dermatological Atlas, "Home | Hellenic Dermatological Atlas," <https://n9.cl/hellenic> . [Accessed: Nov. 10, 2023].
- [67] "Skin Lesion Dataset Using Segmentation," Kaggle, Jul. 5, 2022. [Online]. Available: <https://n9.cl/bigdataset> . [Accessed: Nov. 10, 2023].
- [68] "Skin Lesion Dataset Using Segmentation," Kaggle, Jul. 5, 2022. [Online]. Available: <https://n9.cl/rawdataset> . [Accessed: Nov. 10, 2023].
- [69] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," arXiv:2010.11929 [cs.CV], Oct. 2020.
- [70] DEVDOPE. (2024, May 1). Skin Disease Lightweight Dataset. Kaggle. <https://www.kaggle.com/datasets/devdope/skin-disease-lightweight-dataset>

