

Sistema de detección de cáncer de piel con aprendizaje de máquina para dispositivo de bajo consumo

Carlos Vicente Niño Rondón

Nota de Aceptación

Certificamos que el presente Trabajo de Grado Satisface, en alcances y calidad, todos los requisitos que demanda un Trabajo de Grado de Maestría.



Manuel Guillermo Forero Vargas
Director

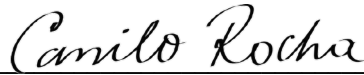


Eugenio Tamura Morimitsu
Jurado



Mario Julian Mora Cardona
Jurado

Aprobado en cumplimiento de los requisitos exigidos por la Pontificia Universidad Javeriana Cali, para optar el título de Magister en Ciencia de Datos



HERNÁN CAMILO ROCHA NIÑO Ph. D.
Decano Facultad de Ingeniería y Ciencias



JUAN CARLOS MARTÍNEZ ARIAS
Director Posgrados de Ingeniería y Ciencias

Cali, 26 de abril de 2024.



Acta de Correcciones al Documento de Trabajo de Grado

Santiago de Cali, 30 de abril de 2024

Autor: Carlos Vicente Niño Rondón

Título del Trabajo de Grado: Sistema de detección de cáncer de piel con aprendizaje de máquina para dispositivo de bajo consumo

Director: PhD. Manuel Guillermo Forero Vargas

Como indica el artículo 2.13 de las Directrices para Trabajo de Grado de Maestría, he verificado que el estudiante indicado arriba ha implementado todas las correcciones que los Jurados del Proyecto de Trabajo de Grado definieron que se efectuaran, como consta en el Acta de Evaluación correspondiente.

PhD. Manuel Guillermo Forero Vargas
Director del Proyecto de grado

Santiago de Cali, 01 de abril de 2024

Señores:

Maestría en Ciencia de Datos
Pontificia Universidad Javeriana – Cali

Respetuosamente, damos conocimiento a ustedes que autorizamos la sustentación del trabajo de grado de maestría titulado “**Sistema de detección de cáncer de piel con aprendizaje de máquina para dispositivo de bajo consumo**”, desarrollado por el estudiante **Carlos Vicente Niño Rondón**, dirigido por **Manuel Guillermo Forero Vargas PhD**, y codirigido por **Sergio Alexander Castro Casadiego MSc**. La presente se realiza dada la respetuosa solicitud de sustentación anticipada.

En constancia firman el director y el codirector del proyecto de grado.



Manuel Guillermo Forero Vargas
Director del proyecto de grado
manuel.forero@unibague.edu.co



Sergio Alexander Castro Casadiego
Codirector del proyecto de grado
sergio.castroc@ufps.edu.co

FICHA RESUMEN

TÍTULO: Sistema de detección de cáncer de piel con aprendizaje de máquina para dispositivo de bajo consumo

1. **ÁREA DE TRABAJO:** Procesamiento de imágenes, modelos computacionales.
2. **TIPO DE PROYECTO:** Aplicado
3. **ESTUDIANTE:** Carlos Vicente Niño Rondón
4. **CORREO ELECTRÓNICO:** carlosvicenteninor@javerianacali.edu.co
5. **DIRECCIÓN Y TELEFONO:** Calle 9N #2-59 Barrio El Bosque, Cúcuta, Norte de Santander, 3209674979
6. **DIRECTOR:** Manuel Guillermo Forero
7. **VINCULACIÓN DEL DIRECTOR:** Universidad de Ibagué
8. **CORREO ELECTRÓNICO DEL DIRECTOR:** manuel.forero@unibague.edu.co
9. **CO-DIRECTOR:** Sergio Alexander Castro Casadiego
10. **GRUPO O EMPRESA QUE LO AVALA:** Universidad Francisco de Paula Santander
11. **OTROS GRUPOS O EMPRESAS:** Grupo de Investigación y Desarrollo en Electrónica y Telecomunicaciones, GIDET, Universidad Francisco de Paula Santander
12. **PALABRAS CLAVE (al menos 5):** Cáncer de piel, diagnóstico asistido, visión computacional, microprocesadores, análisis de datos.
13. **FECHA DE INICIO:** 30 de junio de 2023
14. **DURACIÓN:** 8 meses
15. **RESUMEN:**

El cáncer de piel, que representa aproximadamente el 10% de los casos de cáncer a nivel mundial, se enfrenta a desafíos significativos en términos de diagnóstico preciso. Los expertos en la materia han señalado que tres de cada cuatro herramientas de diagnóstico asistido para el cáncer de piel clasifican incorrectamente alrededor del 30% de los casos de melanoma. Estas herramientas, caracterizadas por su alto costo computacional, inciden directamente en los gastos de implementación, afectando así el acceso a los servicios de salud.

En el presente trabajo se presentan un sistema de diagnóstico de cáncer de piel diseñado específicamente para dispositivos de bajo consumo. Se inicia con el análisis del conjunto de datos HAM10000, seguido por el desarrollo de un modelo de aprendizaje profundo destinado a la clasificación de esta enfermedad. Posteriormente, se planteó la creación de una herramienta de diagnóstico asistido, adaptada para su uso en dispositivos de bajo consumo. La evaluación de su rendimiento a nivel hardware y software completa el proceso.



Pontificia Universidad
JAVERIANA
Cali

**Sistema de detección de cáncer de piel con aprendizaje de máquina para
dispositivo de bajo consumo**

Carlos Vicente Niño Rondón
Código 8980110

*Proyecto Aplicado para optar al título de
Magister en Ciencia de Datos*

Director

PhD. Manuel Guillermo Forero Vargas

Codirector

MSc. Sergio Alexander Castro Casadiego

FACULTAD DE INGENIERÍA Y CIENCIAS
MAESTRIA EN CIENCIA DE DATOS
SANTIAGO DE CALI
2024

Agradecimientos

Agradezco en primer lugar a Dios por permitirme cumplir uno a unos mis sueños, y llegar a este momento, el de haber culminado mi tesis y mis estudios de maestría. Quiero agradecer a mis padres por el apoyo incondicional brindado durante este proceso académico. Igualmente, quiero agradecer a mi tío y a mis tías por verme como otro hijo y apoyarme en todo momento. A mi novia, por ser mi mano derecha y mi motivadora en este caminar de la investigación.

También agradezco al ingeniero Manuel Forero, director de mi proyecto de grado, por todas las enseñanzas, por toda la disposición y por su asesoramiento durante el desarrollo del proyecto. Extiendo también mis más sinceros agradecimientos a mi codirector, el ingeniero Sergio Castro Casadiego, quien desde el pregrado me orienta, me forma y a quien tengo la fortuna de poder llamar amigo. Al ingeniero Byron Medina Delgado, por su apoyo incondicional y su contribución a formarme como profesional integral.

Finalmente, quiero agradecer a los grupos de investigación GIDET y GITecInSo de la Universidad Francisco de Paula Santander por brindarme todo el apoyo necesario para el desarrollo del proyecto. Asimismo, al programa de Jóvenes Investigadores e Innovadores de MinCiencias bajo la convocatoria 891, por financiar parte de los recursos requeridos para el estudio de la maestría.

Dedicatoria

Quiero dedicar este logro académico a mis padres Carlos Arturo y Celina, quienes lo son todo en mi vida y quienes con su ejemplo día a día me forman como un ciudadano integral. A mis tías Leonor y Alcira, quienes constituyen un bastón indispensable y a quienes respeto y admiro. A Gabriela, quien desde el año 2015 es mi compañera de vida, consejera y motivadora.

Quiero cerrar esta dedicatoria con un mensaje al cielo, donde estoy seguro que mi tío Vicente está contento por verme cumplir otra meta. Su legado, su inspiración y su fortaleza me seguirán acompañando durante el resto de mi vida.

¡Disfruten este logro, es por y para ustedes!

FICHA RESUMEN

TÍTULO: Sistema de detección de cáncer de piel con aprendizaje de máquina para dispositivo de bajo consumo

1. **ÁREA DE TRABAJO:** Procesamiento de imágenes, modelos computacionales.
2. **TIPO DE PROYECTO:** Aplicado
3. **ESTUDIANTE:** Carlos Vicente Niño Rondón
4. **CORREO ELECTRÓNICO:** carlosvicenteninor@javerianacali.edu.co
5. **DIRECCIÓN Y TELEFONO:** Calle 9N #2-59 Barrio El Bosque, Cúcuta, Norte de Santander, 3209674979
6. **DIRECTOR:** Manuel Guillermo Forero
7. **VINCULACIÓN DEL DIRECTOR:** Universidad de Ibagué
8. **CORREO ELECTRÓNICO DEL DIRECTOR:** manuel.forero@unibague.edu.co
9. **CO-DIRECTOR:** Sergio Alexander Castro Casadiego
10. **GRUPO O EMPRESA QUE LO AVALA:** Universidad Francisco de Paula Santander
11. **OTROS GRUPOS O EMPRESAS:** Grupo de Investigación y Desarrollo en Electrónica y Telecomunicaciones, GIDET, Universidad Francisco de Paula Santander
12. **PALABRAS CLAVE (al menos 5):** Cáncer de piel, diagnóstico asistido, visión computacional, microprocesadores, análisis de datos.
13. **FECHA DE INICIO:** 30 de junio de 2023
14. **DURACIÓN:** 8 meses
15. **RESUMEN:**

El cáncer de piel, que representa aproximadamente el 10% de los casos de cáncer a nivel mundial, se enfrenta a desafíos significativos en términos de diagnóstico preciso. Los expertos en la materia han señalado que tres de cada cuatro herramientas de diagnóstico asistido para el cáncer de piel clasifican incorrectamente alrededor del 30% de los casos de melanoma. Estas herramientas, caracterizadas por su alto costo computacional, inciden directamente en los gastos de implementación, afectando así el acceso a los servicios de salud.

En el presente trabajo se presentan un sistema de diagnóstico de cáncer de piel diseñado específicamente para dispositivos de bajo consumo. Se inicia con el análisis del conjunto de datos HAM10000, seguido por el desarrollo de un modelo de aprendizaje profundo destinado a la clasificación de esta enfermedad. Posteriormente, se planteó la creación de una herramienta de diagnóstico asistido, adaptada para su uso en dispositivos de bajo consumo. La evaluación de su rendimiento a nivel hardware y software completa el proceso.

TABLA DE CONTENIDO

	Pág.
INTRODUCCIÓN	11
1. DEFINICIÓN DEL PROBLEMA	13
1.1. PLANTEAMIENTO DEL PROBLEMA	13
1.2. FORMULACIÓN DEL PROBLEMA	14
2. OBJETIVOS DEL PROYECTO	15
2.1. OBJETIVO GENERAL.....	15
2.2. OBJETIVOS ESPECÍFICOS.....	15
2.3. RESULTADOS ESPERADOS.....	15
3. MARCO TEORÍCO Y ANTECEDENTES.....	16
3.1. MARCO TEÓRICO.....	16
3.1.1. Cáncer de piel	16
3.1.2. Aprendizaje de máquina.....	16
3.1.3. Balanceo de datos en el ámbito médico	16
3.1.4. Red neuronal convolucional	16
3.1.5. Eficiencia de modelos computacionales.....	17
3.1.6. Herramientas de diagnóstico asistido por computador	17
3.1.7. Código abierto	18
3.1.8. Dispositivos de bajo consumo	18
3.2. ANTECEDENTES	18
4. METODOLOGÍA.....	22
5. ANÁLISIS DEL CONJUNTO DE IMÁGENES DERMOSCÓPICAS.....	25
6. MODELO DE APRENDIZAJE DE MÁQUINA PARA LA CLASIFICACIÓN DEL CÁNCER DE PIEL 31	

7. HERRAMIENTA DE DIAGNÓSTICO ASISTIDO EN DISPOSITIVO DE BAJO CONSUMO	44
8. DESEMPEÑO DEL SISTEMA DE CLASIFICACIÓN EN DISPOSITIVO DE BAJO CONSUMO.	50
9. CONCLUSIONES	58
10. TRABAJOS FUTUROS	61
11. REFERENCIAS BIBLIOGRÁFICAS.....	62

LISTA DE FIGURAS

	Pág.
Figura 1. Metodología propuesta.	22
Figura 2. Imágenes de lesiones cutáneas del HAM10000	28
Figura 3. Comparación de la distribución de imágenes de las lesiones cutáneas.	29
Figura 4. Gráfico de dispersión bidimensional para el conjunto de entrenamiento original y aumentado.	29
Figura 5. Curva KDE de la primera componente principal del conjunto de datos original y el conjunto de datos aumentados.	30
Figura 6. Arquitectura de CNN propuesta para la clasificación del cáncer de piel.	32
Figura 7. Evolución del entrenamiento de la CNN.	36
Figura 8. Matriz de confusión porcentual del modelo computacional desarrollado.	37
Figura 9. Curva de ROC con valores de AUC para cada clase de lesiones cutáneas.	38
Figura 10. Desempeño general del modelo basado en Random forest.	40
Figura 11. Desempeño general del modelo basado en Random forest.	41
Figura 12. Interfaz de inicio de sesión al sistema.	44
Figura 13. Recuperación de credenciales para ingresar al sistema.	45
Figura 14. Interfaz gráfica desarrollada para el diagnóstico de la lesión cutánea.	47
Figura 15. Interfaz gráfica desarrollada con funcionalidad.	49
Figura 16. Implementación en sistema embebido.	54
Figura 17. Desempeño del sistema de clasificación respecto al tiempo de respuesta.	55
Figura 18. Desempeño del sistema de clasificación en cuanto a niveles de CPU y RAM requerido.	56

Figura 19. Desempeño del sistema de clasificación en cuanto a la temperatura del procesador del sistema embebido. 57

LISTA DE TABLAS

	Pág.
Tabla 1. Descripción de las lesiones cutáneas del HAM10000 dataset.....	25
Tabla 2. Distribución de las lesiones cutáneas del HAM10000 dataset.	26
Tabla 3. Muestra de descriptores de imágenes.....	27
Tabla 4. Resumen capa a capa del modelo desarrollado.....	33
Tabla 5. Evolución del proceso de entrenamiento de la CNN propuesta.	34
Tabla 6. Desempeño del modelo de clasificación propuesto.	37
Tabla 7. Desempeño de la técnica de Random Forest en la clasificación.	39
Tabla 8. Desempeño de la técnica de XGBoost en la clasificación.	40
Tabla 9. Comparación de rendimiento entre el modelo propuesto y trabajos del estado del arte publicados en revistas Q1 y Q2 de Scopus.....	43
Tabla 10. Librerías utilizadas para el desarrollo de la interfaz gráfica de usuario.	45
Tabla 11. Pseudocódigo con funcionalidades básicas de la interfaz de usuario desarrollada.	46
Tabla 12. Casos de uso de la interfaz gráfica desarrollada para el diagnóstico.....	48
Tabla 13. Peso de los parámetros para la selección del dispositivo de bajo consumo.	50
Tabla 14. Comparación de las herramientas de hardware respecto a la velocidad de procesamiento.....	51
Tabla 15. Comparación de las herramientas de hardware respecto a la disponibilidad de interfaces de entrada y salida.....	51

Tabla 16. Comparación de las herramientas de hardware respecto al consumo de energía.	52
Tabla 17. Comparación de las herramientas de hardware respecto a la capacidad de disipación de calor	52
Tabla 18. Comparación de las herramientas de hardware respecto al precio y la disponibilidad en el mercado.	53
Tabla 19. Selección de herramienta de hardware de bajo consumo.	53

LISTA DE ANEXOS

Anexo I. Manual de usuario de la herramienta de diagnóstico asistido por computador ... 67

INTRODUCCIÓN

El cáncer es una enfermedad que se produce en las células y que se multiplican a las demás partes del cuerpo. En condiciones normales, las células se multiplican mediante división celular, creando así nuevas células de manera progresiva, y cuando estas envejecen y mueren, las células nuevas las reemplazan [1]. No obstante, existen casos en que las células anormales se forman y se multiplican, dando cabida a posibles tumores cancerosos (malignos) que se diseminan e invaden a los tejidos cercanos dando lugar a múltiples tumores malignos. El cáncer de piel es el crecimiento anormal de células de piel y suele desarrollarse en secciones expuestas al sol. Asimismo, afecta a todos los tonos de piel incluidas personas de tez oscura, siendo estos casos en donde se presenta la afección en las zonas que no están expuestas al sol. Los tres tipos principales de cáncer de piel corresponden al melanoma, y carcinomas de células escamosas y basocelular [2]. La detección temprana es crucial para mejorar las tasas de supervivencia y reducir la morbimortalidad asociada. La inteligencia artificial (IA) ha emergido como un recurso valioso en el campo de la medicina, específicamente en el ámbito de la clasificación de imágenes médicas. En el caso del cáncer de piel, la capacidad de análisis y detección de patrones complejos por parte de modelos computacionales ofrece una nueva perspectiva para mejorar la eficacia de los diagnósticos [3].

El diagnóstico tradicional del cáncer de piel implica la evaluación visual de lesiones dermatológicas por parte de dermatólogos. Este enfoque, aunque ha sido efectivo durante mucho tiempo, tiene limitaciones en términos de subjetividad, variabilidad Inter observador y la necesidad de expertos altamente especializados. Las herramientas de diagnóstico asistido por computador se presentan como apoyo a los profesionales de la salud en el proceso de diagnóstico de enfermedades, así como para la toma de decisiones [4]. Las principales aplicaciones de las herramientas de diagnóstico asistidas por computador se enfocan en la detección de anomalías, clasificación de imágenes, apoyo en toma de decisiones, seguimiento de pacientes y optimización de recursos.

Los dispositivos de bajo consumo tienen relevancia en el campo de la medicina ya que, permite entre otros, la portabilidad y análisis de datos en tiempo real, el registro de parámetros de salud para decisiones médicas acertadas, la reducción de costos de implementación y la innovación en el desarrollo de tecnologías portátiles para el diagnóstico de enfermedades [5]. La introducción de modelos computacionales en estos dispositivos busca abordar estos desafíos, proporcionando un apoyo a los profesionales de la salud en el proceso de diagnóstico [6].

En el presente trabajo se desarrolló un sistema para la clasificación del cáncer de piel para ejecución sobre dispositivos de bajo consumo, basado en un modelo de aprendizaje de máquina. En primer lugar, se analizó el conjunto de datos HAM10000, empleando la modificación HMNIST 28x28x3 para el equilibrio de clases por sobremuestreo empleando

inducción de ruido normal. Para el entrenamiento del modelo se diseñó una estructura convolucional propia de 22 capas que fue comparada considerando métricas de software como precisión, recall, f1-score, especificidad y exactitud, y métricas de rendimiento en hardware como tiempo de entrenamiento, y requerimientos de CPU y RAM respecto a 2 los modelos (optimizados por búsqueda de hiperparámetros mediante la técnica de GridSearch) de aprendizaje automático Random Forest y XGBoost, y respecto a documentos del estado del arte donde se emplearon arquitecturas basadas en aprendizaje profundo para la clasificación de lesiones cutáneas. En ambos escenarios, la red neuronal convolucional propuesta presentó mejor desempeño. Además, se diseñó una herramienta de diagnóstico asistido para su ejecución en dispositivos de bajo consumo, diseñada con QtDesigner y programada en Python. Finalmente, se evaluó el desempeño y el rendimiento de la herramienta de diagnóstico asistido en un sistema embebido.

1. DEFINICIÓN DEL PROBLEMA

1.1. PLANTEAMIENTO DEL PROBLEMA

El cáncer tiene un impacto significativo en la salud mundial. La Organización Mundial de la Salud indica que el cáncer es la segunda causa principal de muertes a nivel global, siendo superada solamente por las muertes derivadas de enfermedades cardiovasculares. El cáncer de piel ha tenido un aumento importante en los reportes de casos en los últimos años, siendo diagnosticados alrededor de 3 millones de casos de tipo melanoma, y más de doscientos millones de casos de diversos tipos de cáncer de piel [7].

Uno de los factores preocupantes respecto al cáncer de piel es la carga económica, dado que los costos que se asocian al diagnóstico, así como al tratamiento y cuidado de la enfermedad son sustanciales, además de todas aquellas consecuencias que tiene en la calidad de vida de los pacientes y en su producción laboral. Asimismo, algunas de las herramientas tecnológicas utilizadas como apoyo en la detección del cáncer de piel presentan errores en la detección y clasificación de la enfermedad, donde incluso los especialistas en la materia exponen que 3 de cada 4 aplicaciones clasifican erróneamente más del 30 % de los melanomas como lesiones menores en la piel. La precisión limitada de las herramientas de diagnóstico requiere del juicio e interpretación de los expertos en el área, y la variabilidad de las imágenes afecta la capacidad de los algoritmos y modelos computacionales para la obtención de resultados fiables [8], [9]. Sumado a esto, las herramientas de diagnóstico asistido por computador empleadas para la detección y clasificación de cáncer de piel tienen como limitantes las barreras de implementación, ya que, requieren de infraestructuras especiales de tecnologías de la información, lo que se ve reflejado directamente en los costos de adquisición, mantenimiento y capacitación para su uso e implementación.

Los dispositivos de bajo consumo están diseñados para funcionar eficientemente con una baja cantidad de energía. Estos dispositivos se utilizan en aplicaciones médicas y de inteligencia computacional donde es crucial maximizar la vida útil de la batería, reducir los costos de operación y minimizar el impacto ambiental. Para lograr un bajo consumo, estos dispositivos requieren entre otros, del uso de tecnologías que requieren niveles bajos de energía, microcontroladores de baja potencia y la implementación de algoritmos que permitan el ahorro energético. Por esta razón, las alternativas basadas en el uso de teléfonos celulares no se consideran apropiadas para implementación de aplicaciones de baja energía debido a factores tales como su alto consumo de energía, ya que, generalmente, estos dispositivos están equipados con múltiples componentes y características que impiden el ahorro energético. Además, otros factores considerados importantes en las aplicaciones médicas como la fiabilidad y la precisión de los dispositivos limitan el uso de los dispositivos celulares, donde la estabilidad de las mediciones, precisión del sensor y consistencia de resultados no están garantizadas para su uso en el ámbito médico [10].

De igual forma, la calidad de la información presente en los conjuntos de datos de imágenes dermoscópicas, así como la calidad del etiquetado de las imágenes es otro factor que afecta considerablemente a las tasas de aciertos en la clasificación del tipo de cáncer de piel [11]. Sumado a esto, la falta de diversidad en los datos y el desbalanceo entre los mismos, también contribuye a que se presenten errores en la detección y clasificación.

Considerando esto, surgió la necesidad de desarrollar un sistema de diagnóstico de cáncer de piel que logre la identificación del tipo de lesión cutánea, y que permita su ejecución en dispositivos de bajo consumo, como sistemas embebidos, de fácil uso y que esté basado en código abierto con el objetivo de mejorar el acceso a la atención médica y la detección temprana de esta enfermedad.

1.2. FORMULACIÓN DEL PROBLEMA

Dicho lo anterior, surge la siguiente pregunta de investigación:

¿Cómo desarrollar un sistema de diagnóstico de cáncer de piel que se ejecute sobre sistemas de bajo consumo?

De igual forma, surgen las siguientes preguntas de sistematización:

- ¿Cómo preparar el conjunto de datos derivado de un dataset de imágenes dermoscópicas?
- ¿Cómo desarrollar un modelo de aprendizaje de máquina para la clasificación de cáncer de piel?
- ¿Cómo diseñar una herramienta de diagnóstico asistido por computador que se adapte a dispositivos de bajo consumo?
- ¿Cómo evaluar el desempeño del sistema de diagnóstico para la clasificación de cáncer de piel ejecutable en dispositivos de bajo consumo?

2. OBJETIVOS DEL PROYECTO

2.1. OBJETIVO GENERAL

Desarrollar un sistema de detección de cáncer de piel para ejecución sobre dispositivos de bajo consumo.

2.2. OBJETIVOS ESPECÍFICOS

- Analizar el conjunto de imágenes dermoscópicas para lesiones benignas y lesiones malignas para obtener información de la naturaleza de las mismas.
- Desarrollar un modelo de aprendizaje de máquina para identificar y clasificar el cáncer de piel.
- Diseñar la herramienta de diagnóstico asistido por computador ejecutable sobre dispositivos de bajo consumo.
- Evaluar el desempeño del sistema de diagnóstico y clasificación de cáncer de piel.

2.3. RESULTADOS ESPERADOS

- Dataset con información de imágenes dermoscópicas de lesiones cutáneas.
- Modelos computacionales para la detección y clasificación del cáncer de piel.
- Herramienta de apoyo en el diagnóstico de cáncer de piel ejecutable en dispositivos de bajo consumo.
- Documento final con descripción de la herramienta desarrollada para el diagnóstico de cáncer de piel.

3. MARCO TEORÍCO Y ANTECEDENTES

3.1. MARCO TEÓRICO

A continuación, se describen los principales conceptos considerados para contextualizar la investigación desarrollada.

3.1.1. Cáncer de piel

El cáncer de piel refiere a una enfermedad en la que se forman células malignas o cancerosas en los tejidos de la piel [12]. Principalmente, se clasifica en tres tipos: melanomas, carcinomas de células basales y carcinomas de células escamosas. Su origen radica generalmente por la exposición excesiva a rayos ultravioleta del sol o fuentes artificiales de radiación ultravioleta. Los principales factores de riesgo para el desarrollo de cáncer de piel incluyen la exposición excesiva al sol, antecedentes familiares de cáncer de piel, tener piel clara, quemaduras solares frecuentes y el uso de camas de bronceado.

3.1.2. Aprendizaje de máquina

Es el área de la inteligencia artificial enfocada en el desarrollo de algoritmos y modelos con capacidades para aprender y tomar decisiones a partir de datos. Los sistemas de aprendizaje de máquina se basan en patrones y ejemplos para la generación de predicciones, identificación de tendencias y toma de decisiones basadas en información. Estos modelos se alimentan con información de entrada y mediante algoritmos, se ajustan para el reconocimiento de patrones y relaciones significativas entre los datos [13]. El aprendizaje de máquina tiene aplicación directa en el campo de la medicina, la industria, bienes y servicios, entre otros.

3.1.3. Balanceo de datos en el ámbito médico

En el ámbito médico, los modelos deben ser capaces de servir a todos los pacientes, independientemente de su raza, etnia, género, edad o condición médica. Un conjunto de datos desequilibrado puede no reflejar la diversidad de la población, lo que puede llevar a que los modelos sean menos precisos o incluso sesgados [14]. El balanceo de datos en el ámbito médico se refiere a asegurarse de que los conjuntos de datos utilizados en investigaciones, análisis y modelos en el campo de la salud sean representativos y equilibrados. Esto implica garantizar que haya una distribución adecuada de casos positivos y negativos, así como una diversidad suficiente en términos de características demográficas, clínicas y otras variables relevantes.

3.1.4. Red neuronal convolucional

Las redes neuronales convolucionales (CNN) son un tipo de red neuronal artificial que se utiliza para procesar y analizar datos, así como imágenes y videos. Las CNN son especialmente efectivas para aprender patrones espaciales y jerárquicos presentes en datos

visuales. Las redes neuronales convolucionales han demostrado ser herramientas poderosas en el campo médico, especialmente en el diagnóstico de enfermedades a partir de imágenes médicas [15]. Algunos aspectos clave de su importancia incluyen la extracción de características, segmentación, mejora de la precisión, aprendizaje automático, reducción de tiempos de diagnóstico.

3.1.5. Eficiencia de modelos computacionales

Refiere a la capacidad de un modelo para lograr sus objetivos de manera efectiva y con recursos optimizados. Generalmente se evalúa en términos de la relación entre el rendimiento óptimo y los recursos utilizados, como tiempo de ejecución, memoria y capacidad de procesamiento. A continuación, se relacionan algunas consideraciones para determinar la eficiencia de un modelo computacional.

- **Tiempo de ejecución:** La eficiencia se relaciona directamente con la rapidez con la que un modelo realiza su tarea. Modelos eficientes completan sus operaciones en un tiempo razonable, fundamental en aplicaciones en tiempo real o en entornos donde el rendimiento rápido es esencial, como en el campo médico [16].
- **Uso de recursos:** Un modelo eficiente utiliza de manera óptima los recursos disponibles, como la memoria y la capacidad de procesamiento. La eficiencia se puede medir evaluando cuántos recursos se necesitan para lograr un resultado específico. Modelos con niveles de rendimiento con recursos limitados son considerados eficientes [17].
- **Escalabilidad:** La eficiencia también se relaciona con la capacidad de un modelo para manejar volúmenes crecientes de datos o tareas más complejas sin un aumento proporcional en los recursos requeridos. Un modelo escalable puede adaptarse y mantener su rendimiento incluso cuando se enfrenta a desafíos adicionales [18].
- **Consumo de energía:** En aplicaciones móviles o embebidas, la eficiencia se evalúa a menudo en términos de consumo de energía. Modelos que pueden realizar tareas de manera eficiente en términos energéticos son preferidos, especialmente en dispositivos con recursos limitados [19].

3.1.6. Herramientas de diagnóstico asistido por computador

Las herramientas de diagnóstico asistido corresponden a sistemas diseñados como apoyo a los especialistas médicos en el proceso de diagnóstico. Se basan principalmente en aprendizaje de máquina y, utilizan algoritmos y modelos entrenados con grandes volúmenes de imágenes y datos para identificar patrones, anomalías y características relevantes en los datos [20]. Su principal incidencia se encuentra en el análisis de imágenes de radiografías, tomografías computarizadas, resonancias magnéticas y mamografías para la detección temprana de signos de enfermedad o anomalía [21]. Específicamente en el cáncer de piel, las herramientas de diagnóstico asistido tienen funcionalidad en el análisis de imágenes dermoscópicas, el seguimiento a la evolución de las lesiones, diagnóstico diferencial entre

tipos de cáncer de piel, evaluación de factores de riesgo personalizados, y apoyo en la toma de decisiones para profesionales de la salud

3.1.7. Código abierto

El código abierto se refiere a un modelo de desarrollo de software en el cual el código fuente del programa está disponible públicamente y puede ser modificado y distribuido por cualquier persona. Este enfoque fomenta la colaboración y transparencia, permitiendo a la comunidad de desarrolladores acceder, mejorar y compartir el software de manera abierta. En el campo de la medicina, el uso de software de código abierto ha ganado relevancia debido a su flexibilidad, transparencia y accesibilidad [22]. Los sistemas de gestión de información médica, herramientas de análisis de datos clínicos y plataformas para el intercambio de información sanitaria a menudo aprovechan el enfoque de código abierto. Esto permite a los profesionales de la salud personalizar y adaptar las soluciones a sus necesidades específicas, facilitando la interoperabilidad y el desarrollo colaborativo en el ámbito médico.

3.1.8. Dispositivos de bajo consumo

Son componentes diseñados para consumir la menor cantidad de energía durante su funcionamiento. Estos dispositivos son ideales en aplicaciones de uso prolongado con energía disponible limitada. Además, son utilizados con frecuencia en sistemas embebidos, dispositivos portátiles, dispositivos de internet de las cosas y sensores inalámbricos. De igual forma, facilitan la integración de tecnologías de menor tamaño y mayor portabilidad [23]. Los dispositivos de bajo consumo tienen varias aplicaciones en el ámbito del diagnóstico de enfermedades, aportando beneficios en términos de movilidad, accesibilidad y eficiencia. La utilización de dispositivos de bajo consumo en el diagnóstico de enfermedades médicas no solo contribuye a la eficiencia energética, sino que también amplía las posibilidades de acceso a la atención médica, especialmente en entornos con recursos limitados o en situaciones donde la movilidad y la autonomía son cruciales.

3.2. ANTECEDENTES

Asimismo, se presentan trabajos desarrollados enfocados a la clasificación de cáncer de piel empleando modelos computacionales.

- **Skin Cancer Detection Based on Extreme Learning Machine and a Developed Version of Thermal Exchange Optimization [24].**

Los autores proponen un estudio de detección de melanoma sobre imágenes dermoscópicas empleando el dataset SIIM-ISIC Melanoma. El método empleado se basa en etapas de segmentación de regiones de interés mediante el método de Otsu, extracción de características y una red optimizada para clasificar los casos de melanoma. Para la optimización se basan en un algoritmo de optimización de intercambio térmico dTEO. Previo a la segmentación de la región de interés de la lesión, emplean técnicas de

preprocesamiento para reducción de ruido gaussiano mediante realce de contrastes. El algoritmo utilizado se basa en la temperatura de los objetos y posición variante generando así unas posiciones actualizadas de los mismos. Los autores, además, consideran las funciones de activación de las redes para el realce de las características propias. El sistema se desarrolló en el entorno Matlab 2018b en un dispositivo de hardware de 16GB de RAM, y procesador Core i7-4720HQ de 1.6GHz de frecuencia. Para determinar la eficiencia de la herramienta desarrollada compararon parámetros como eficiencia, sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo respecto a trabajos basados en análisis fractal, redes neuronales convolucionales, método de lado a lado, máquinas de vectores de soporte y algoritmos genéticos. Las pruebas arrojan que el valor de sensibilidad se encuentra superior al 95 % para 30 pruebas de ejecución por técnica. Este trabajo se limita la detección de melanoma ejecutable sobre un dispositivo de alto costo. Asimismo, el trabajo no presenta el desempeño en hardware del sistema desarrollado.

- **DermlA: Machine Learning to Improve Skin Cancer Screening [25]**

Este trabajo presenta una herramienta orientada al seguimiento del cáncer de piel. La herramienta se diseñó en una aplicación móvil y usa la cámara del dispositivo móvil para escanear lesiones sospechosas. Es una aplicación desarrollada por GeniaLabs, y está disponible en Google Play Store, con una versión demo gratuita, y una versión premium por suscripción de 6 dolares mensuales. Se lanzó al mercado en mayo de 2019 y ha tenido una aceptación media de 4 (siendo 5 el máximo valor). En óptimas condiciones, la herramienta presenta una eficiencia de hasta 95 % para la versión gratuita y de 98 % para la versión por suscripción. La principal limitante de esta herramienta son las variaciones en los niveles de luminosidad y distancia a la lesión al momento de la captura de la imagen, lo que incide directamente en la eficiencia de la herramienta.

- **Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning [26]**

En este trabajo se emplea el dataset ISIC 2018. La arquitectura se basa en una estrategia de “Transfer DL”, empleando en primer lugar una arquitectura ESRGAN para lograr una edición mejorada de las lesiones cutáneas, empleando bloques “resilient-in-residual” en lugar de una red residual básica. Además, estos modelos no cuentan con capas de normalización por lotes con el objetivo de filtrar la imagen, por lo que los bordes de la lesión se mejoran mediante ESRGAN. Igualmente, emplearon aumentado de datos con variaciones geométricas que incluían rotación, reflexión, brillo y cambio de tamaño. De esta forma, se pasó de 1440 imágenes benignas a 1760, y de 1197 imágenes de lesiones malignas a 1773. El desbalance de los datos se trató mediante un sobremuestreo en las imágenes malignas. Se probaron las arquitecturas ResNet50, Inception V3 y Inception ResNet. Los resultados muestran que el método propuesto elevó a 83.1 % la tasa de acierto en la clasificación de la lesión al emplearse una CNN convencional, a 83.6 % al emplear la Resnet50, 84.1 % con la Restnet50-Inception y 85.7 % con la arquitectura InceptionV3. Este es un trabajo de clasificación binaria y que solamente está validado por la métrica de exactitud (accuracy),

no considerando entre otras, métricas como el F1-Score, que enfoca la precisión y la recuperación, el desequilibrio de clases y la dominancia de clases. Además, no se cuenta con una prueba de los datos aumentados respecto a los datos originales

- **Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering [27]**

Este trabajo presenta un enfoque hacia la detección de melanoma en fases tempranas empleando una arquitectura de redes neuronales convolucionales basadas en regiones rápidas (RCNN). En primer lugar, se preprocesan las imágenes para mitigar los efectos de ruido e iluminación antes de utilizar la RCNN con el objetivo de encontrar un vector de características con longitud fija. Además, se emplea una matriz de neuronas con valores de atributos para segmentar la sección de piel con melanoma. El método se probó con los dataset ISBI-2016, ISIC-2017, y PH2. En Promedio, el método que proponen los autores tiene una precisión media de 95.4 %, 93.1 % y 95.6 % al probarse con los dataset ISIC-2016, ISIC-2017, y PH2 respectivamente. Además, presenta robustez en la segmentación de las lesiones.

- **On the Automatic Detection and Classification of Skin Cancer Using Deep Transfer Learning [28]**

En este documento, se desarrolla un sistema de clasificación automática de cáncer de piel empleando “transfer Learning”. La contribución de este documento es el desarrollo de un sistema de escaneo de lesiones para identificar entre melanomas y no melanomas, así como el uso de 13 modelos y la definición de métricas para evaluar el modelo de clasificación. El trabajo se desarrolló en MATLABR2021 y se ejecutó sobre un dispositivo HP OMEN 30L con 64GB de RAM, GPU NVIDIA GeForce RTX 3080 GPU, con un procesador Intel Core i7, 3.8GHz de frecuencia y 1TB de disco duro de estado sólido. El trabajo muestra un valor de F1-Score variante entre 51.5 % y 64.8 %, la precisión media cercana al 65 %, especificidad superior al 93.5 % y exactitud fluctuante entre 71.6 % y 76.7 %. La limitante de este trabajo se encuentra en la baja tasa de valores de exactitud en clasificación entre clases con valores variantes entre 43.1 % y 92.5 %. Sumado a esto, el hecho de requerir un dispositivo especializado de hardware para su desarrollo e implementación se convierte en una limitante para su adopción a escala, por el elevado costo de implementación.

- **Melanoma Skin Cancer Detection Using Recent Deep Learning Models [29]**

Para este trabajo se utilizó un agregado de imágenes de Kaggle SIIM-ISIC 2020 para una clasificación binaria (lesiones benignas y lesiones malignas) con lesiones redimensionadas a 1024x1024. El dataset generado contiene 32542 imágenes de lesiones benignas y 3960 de lesiones malignas posterior a una etapa de aumento de datos con transformaciones entre 5 y 10 grados. Se probaron diversos modelos entre los que destacan las arquitecturas VGG16, VGG19, DenseNet121, MobileNetv2 y NasNetLarge. Para la detección de melanoma

se presentan valores superiores al 95 % de exactitud en todas las arquitecturas de red neuronal convolucional probadas. No obstante, los resultados pueden estar sesgados debido al desbalance entre clases, teniendo la clase de lesiones benignas aproximadamente 10 veces mayor cantidad de imágenes, siendo esta la principal limitante de dicho trabajo. Además, no se considera el rendimiento en hardware del sistema de clasificación

- **Skin lesion classification of dermoscopic images using machine learning and convolutional neural network[30].**

Para este trabajo se emplea el HAM10000 dataset, del cual, se toma un subconjunto de datos y se emplea un aumentado de datos, considerando que los modelos con aumentado tienden a mejorar el aprendizaje con características distintivas. Asimismo, con el objetivo de mejorar la robustez del sistema se emplea prueba cruzada. Para la clasificación emplean una red neuronal convolucional que presentó una exactitud de 95.18 %, que mejora a algoritmos clásicos como Inception V3 en casi 5 %, y VGG16+GoogLeNet en casi 14 %. La limitante de este trabajo está en la herramienta de hardware empleada, correspondiente a un computador con procesador Intel Core i5 Y 8GB de RAM, ya que, se muestra como una herramienta de alto costo para la ejecución del sistema de clasificación de la lesión.

4. METODOLOGÍA

La metodología propuesta para la clasificación de lesiones cutáneas se resume en la figura 1. Destacan 4 etapas básicas: el análisis del conjunto de imágenes dermoscópicas, el desarrollo del modelo convolucional para la clasificación, el diseño de la herramienta de diagnóstico asistido y la evaluación del desempeño del sistema de clasificación.

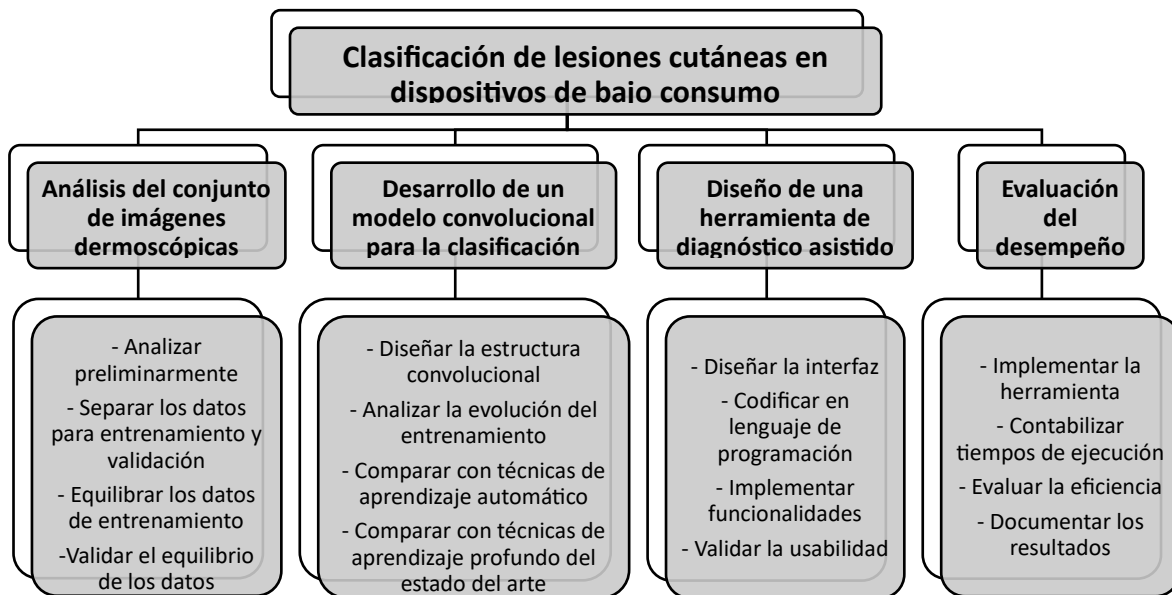


Figura 1. Metodología propuesta.

La forma en que se solucionaron cada uno de los 4 objetivos específicos propuestos se muestra a continuación. En la primera etapa de la metodología, se realizó un análisis del conjunto de imágenes dermoscópicas HAM10000, equilibrando las clases mediante la modificación HMNIST 28x28x3 y validando la calidad de la modificación a través de métricas específicas. La segunda etapa se centró en el desarrollo de un modelo de aprendizaje de máquina, utilizando una estructura convolucional diseñada y comparándola con arquitecturas de aprendizaje automático preentrenadas, con énfasis en parámetros de hardware y software. La tercera etapa de la metodología se enfocó en el diseño de una herramienta de diagnóstico asistido con una interfaz intuitiva, programada en Python y validada mediante pruebas de funcionalidad. Finalmente, en la última etapa, se evaluó el desempeño del sistema implementado en dispositivos de bajo consumo, midiendo el tiempo de ejecución y los recursos de CPU y RAM utilizados, con la documentación detallada de los resultados obtenidos durante la investigación.

- **Objetivo 1:** Analizar el conjunto de imágenes dermoscópicas para lesiones benignas y lesiones malignas para obtener información de la naturaleza de las mismas.

Actividades:

1. Realizar un análisis exploratorio de datos en el conjunto HAM10000 para comprender la distribución y características de las lesiones cutáneas benignas y malignas.
 2. Implementar la modificación HMNIST 28x28x3 para equilibrar las clases mediante sobremuestreo y la inducción de ruido normal, asegurando una representación equitativa de ambos tipos de lesiones.
 3. Validar la calidad de la modificación del conjunto de datos mediante métricas específicas, como la proporción de clases equilibradas y la preservación de características esenciales.
 4. Generar visualizaciones y estadísticas descriptivas que proporcionen información detallada sobre las características distintivas de las lesiones benignas y malignas presentes en el conjunto de datos.
- Objetivo 2: Desarrollar un modelo de aprendizaje de máquina para identificar y clasificar el cáncer de piel.

Actividades:

1. Diseñar y entrenar una estructura convolucional para el modelo de aprendizaje de máquina, considerando la complejidad y la capacidad de representación de características.
 2. Obtener información relevante de la evolución del proceso de entrenamiento de la red neuronal convolucional diseñada.
 3. Comparar el rendimiento de la estructura convolucional diseñada con modelos de aprendizaje automático considerando parámetros de hardware y software.
 4. Revisar la literatura del estado del arte en arquitecturas basadas en aprendizaje profundo para la clasificación de lesiones cutáneas y comparar los resultados obtenidos con la estructura convolucional diseñada.
- Objetivo 3: Diseñar la herramienta de diagnóstico asistido por computador ejecutable sobre dispositivos de bajo consumo.

Actividades:

1. Utilizar QtDesigner para diseñar una interfaz de usuario intuitiva y amigable para la herramienta de diagnóstico asistido por computador.
2. Programar la herramienta en Python, asegurando la compatibilidad y eficiencia en dispositivos de bajo consumo.
3. Implementar funcionalidades de carga de imágenes, clasificación y presentación de resultados en la interfaz gráfica de usuario.
4. Validar la usabilidad de la herramienta mediante pruebas de funcionalidad.

- Objetivo 4: Evaluar el desempeño del sistema de diagnóstico y clasificación de cáncer de piel.

Actividades.

1. Implementar la herramienta de diagnóstico asistido en dos sistemas embebidos para evaluar su desempeño en entornos reales.
2. Medir el tiempo de ejecución en los sistemas embebidos y compararlo con las métricas obtenidas durante el entrenamiento del modelo.
3. Evaluar la eficiencia del sistema en términos de recursos de CPU y RAM utilizados durante la ejecución en dispositivos de bajo consumo.
4. Documentar los resultados de la investigación.

5. ANÁLISIS DEL CONJUNTO DE IMÁGENES DERMOSCÓPICAS

Se aplica un análisis del conjunto de imágenes dermoscópicas con el objetivo de distinguir entre lesiones benignas y malignas. En primer lugar, se identifica la naturaleza del conjunto de datos utilizado, identificando sus características fundamentales. De igual forma, se analiza la calidad y relevancia de las imágenes que conforman el dataset para asegurar la fiabilidad de los resultados. Posteriormente, se definen criterios de balanceo de datos para las muestras de imágenes. Finalmente, se consolidan los datos para el desarrollo del modelo computacional.

Para la clasificación de cáncer de piel se emplea generalmente el HAM10000 dataset, un conjunto de datos que incluye imágenes dermatoscópicas de lesiones cutáneas, que proporciona información para la detección temprana y precisa de diversas afecciones de la piel. Las imágenes son diversas ya que abarca una amplia gama de categorías de lesiones, incluyendo tanto lesiones benignas como malignas. Cada imagen está acompañada de metadatos detallados que proporcionan información adicional sobre la lesión, como el diagnóstico clínico, el diagnóstico histopatológico y la ubicación en el cuerpo. Contiene un total de 10,015 imágenes dermoscópicas de alta resolución, distribuidas en siete categorías principales de lesiones cutáneas que son descritas a continuación.

Tabla 1. Descripción de las lesiones cutáneas del HAM10000 dataset.

Código	Lesión	Descripción
MEL	Melanoma	Una forma de cáncer de piel que surge a partir de la proliferación descontrolada de células de la piel conocidas como melanocitos.
NV	Nevus	Estados no cancerosos de melanocitos.
BCC	Carcinoma de células basales	Un tipo de cáncer que afecta a las células basales, responsables de reemplazar las células envejecidas de la piel con nuevas.
AKIEC	Queratosis actínica	Condiciones resultantes de una exposición prolongada al sol, con un riesgo muy bajo de desarrollar cáncer.
BKL	Queratosis benigna	Protuberancias cutáneas inofensivas que generalmente aparecen en personas mayores.
DF	Dermatofibroma	Enfermedad cutánea benigna caracterizada por lesiones firmes con apariencia fibrosa en varios tamaños.
VASC	Lesión vascular	Ocurre debido a problemas vasculares, donde los capilares superficiales de la piel se agrandan o se vuelven más frecuentes de lo normal.

Originalmente, el tamaño de las imágenes es de 600 x 600 píxeles, con algunas imágenes significativamente más grandes. Por otro lado, la profundidad de color está dada por los canales rojo, verde y azul, y las imágenes muestran características distintivas que reflejan diferentes etapas de la lesión. Algunas imágenes contienen artefactos como marcas de instrumentos médicos, reflejos de luz y sombras.

Considerando lo anterior, en la tabla 2 se relaciona el tipo de lesión, con su clasificación como benigna o malignas y la cantidad de imágenes del conjunto de datos HAM10000. Se identifican tres categorías malignas, específicamente Melanoma (MEL) con 1113 imágenes, Carcinoma de Células Basales (BCC) con 514 imágenes y Queratosis Actínica (AKIEC) con 327 imágenes. Por otro lado, las lesiones benignas incluyen Nevus (NV) con 6705 imágenes, Queratosis Benigna (BKL) con 1099 imágenes, Dermatofibroma (DF) con 115 imágenes y Lesión Vasculosa (VASC) con 142 imágenes. Se observa además un notable desbalance en la distribución de imágenes entre las categorías malignas y benignas. La cantidad total de imágenes benignas asciende a 8061, mientras que las malignas suman 1954. Esta disparidad en el número de imágenes tiene implicaciones en la capacidad del modelo para aprender y generalizar efectivamente, con tendencia a favorecer la clasificación de las clases mayoritarias.

Tabla 2. Distribución de las lesiones cutáneas del HAM10000 dataset.

Código de lesión	Tipo	Cantidad de imágenes
MEL	Malignas	1113
BCC		514
AKIEC		327
NV	Benignas	6705
BKL		1099
DF		115
VASC		142
Cantidad de imágenes benignas		8061
Cantidad de imágenes malignas		1954
Cantidad de imágenes		10015

Se optó por emplear un enfoque que se basa en el conjunto de datos HMNIST, una versión adaptada del HAM10000. A diferencia de la versión original que utiliza imágenes de 600 x 600 píxeles en tres canales, HMNIST emplea descriptores de imágenes redimensionados a 28x28x3 píxeles. El conjunto de datos contiene 2352 columnas que representan la intensidad de color en la ubicación específica de cada una de las 10015 imágenes en tamaño 28x28. En la tabla 3 se presenta una muestra de las últimas 5 imágenes expresadas en forma de descriptores de imágenes, donde se muestran los 5 primeros y los 5 últimos descriptores.

Tabla 3. Muestra de descriptores de imágenes.

Imagen	Descriptores de imágenes									
	0001	0002	0003	0004	0005	2348	2349	2350	2351	2352
10011	183	165	181	182	165	186	186	206	187	189
10012	2	3	1	38	33	23	21	3	4	1
10013	132	118	118	167	149	159	153	172	151	145
10014	160	124	146	164	131	147	166	185	162	172
10015	175	142	121	181	150	137	125	175	139	126

Con esta modificación en la resolución y representación de las imágenes se obtuvo un conjunto de datos más manejable, permitiendo una exploración eficiente de las características fundamentales para la clasificación de lesiones cutáneas. Además, se destaca que el diseño de ejecución se centra en la eficiencia para dispositivos de bajo consumo, buscando implementaciones prácticas y accesibles que no comprometan significativamente la calidad del diagnóstico dermatológico. Esta orientación hacia la eficiencia en dispositivos de bajo consumo es crucial para facilitar la aplicación práctica y la accesibilidad generalizada del modelo de clasificación. Previo al balanceo de las clases, se hizo la división de los descriptores de imagen de modo que el 80 % de los descriptores se emplearon para entrenamiento y el 20 % de los descriptores se emplearon para pruebas.

Por esto, el balanceo se aplica mediante sobremuestreo a las clases minoritarias por inducción de ruido normal a los descriptores de imágenes que conforman el conjunto de entrenamiento. El propósito fue generar pequeñas fluctuaciones que podrían manifestarse en distintas instancias de la misma clase, mejorando la capacidad del modelo para generalizar a datos nuevos y disminuyendo el riesgo de sobreajuste. Considerando X como una variable aleatoria que representa un píxel determinado en la imagen original y Y representa una variable aleatoria correspondiente al píxel aumentado, entonces, el aumentado de datos se representa como $Y = X + Z$, con Z siendo la variable aleatoria que sigue la distribución del ruido que se agrega. Considerando que Z sigue una distribución normal con media cero y desviación estándar σ , es decir $Z \sim N(0, \sigma^2)$, el aumento quedaría expresado como $Y = X + N(0, \sigma^2)$.

En términos de valor esperado, la expresión se transforma en $E[Y] = E[X] + E[N(0, \sigma^2)]$ y en términos de varianza se transforma en $Var[Y] = Var[X] + Var[N(0, \sigma^2)]$. Además, dado que la media del ruido es cero la expresión de la imagen aumentada queda reducida a $E[Y] = E[X]$. Esto implica que el incremento de datos no tiene un alto impacto en la expectativa del píxel, conservando la información central de la clase original. No obstante, la variabilidad aumenta debido al ruido. La magnitud de la desviación estándar empleada es 10, para regular la cantidad de variabilidad introducida durante el aumento.

En la figura 2 se muestra a forma de imagen los descriptores originales y los descriptores aumentados. En las imágenes originales se observa una diversidad de pigmentaciones y condiciones de piel, que reflejan la variabilidad natural en los datos. En las imágenes aumentadas se introducen variaciones en la región de interés, abarcando tanto las lesiones como la piel circundante. Este proceso de aumento de datos simula distintas manifestaciones de las lesiones y variaciones en las pigmentaciones cutáneas.



(a)



(b)

Figura 2. Imágenes de lesiones cutáneas del HAM10000. (a) Descriptores de imágenes originales, (b) Descriptores de imágenes aumentados por inducción de ruido.

Se utilizó la técnica de inducción de ruido controlado para el aumento de datos en el equilibrio de las clases con el objetivo de evitar repetir imágenes y de experimentar la incidencia de dicha técnica en la clasificación del cáncer de piel. Además, aplicar esta técnica es más eficiente a nivel computacional, ya que, los descriptores representan características abstractas y de menor dimensionalidad de las imágenes, lo que significa que operar en ellos es menos intensivo en términos de cálculos y memoria.

La relación del conjunto de datos posterior al proceso de aumento de datos se muestra a continuación en la figura 3. En la sección (a) se presenta la distribución original del conjunto de datos de entrenamiento, mientras que en la sección (b) la resultante con el aumentado. En términos de descriptores de imágenes, inicialmente, el conjunto de datos de entrenamiento constaba de 8012 filas y 2352 columnas, generando un total de 18.891.264 datos. Posteriormente, mediante el aumento de datos, el número de filas se incrementó a 37569, manteniendo la misma cantidad de columnas. Esta expansión resultó en un conjunto de entrenamiento ampliado que contiene un total de 88.362.288 datos de descriptores de imagen.

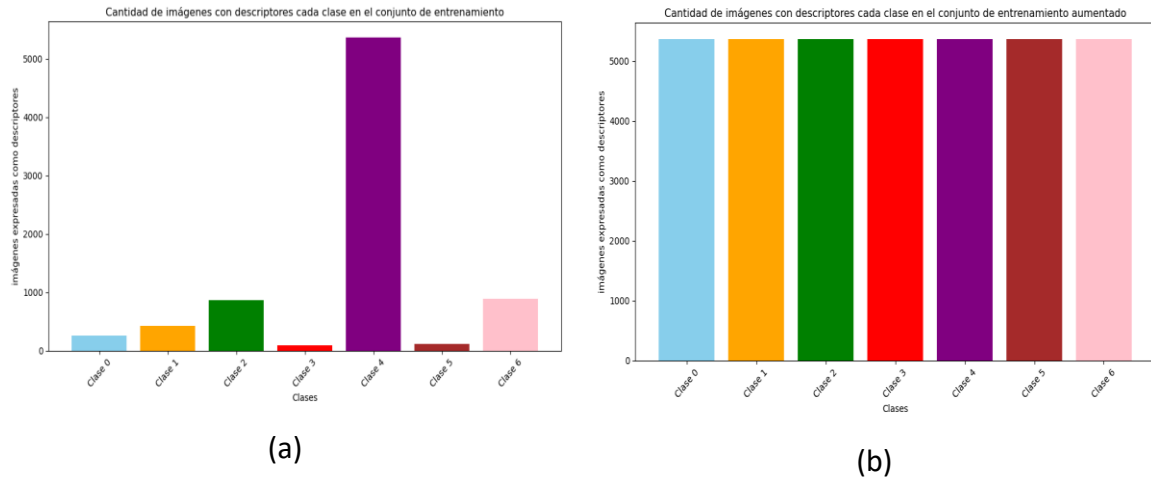


Figura 3. Comparación de la distribución de imágenes de las lesiones cutáneas.

Para validar el aumento de datos se comparan los descriptores de las imágenes originales y los descriptores de las imágenes aumentadas del conjunto de datos de entrenamiento, con el objetivo de evaluar la similitud estructural de las mismas. La comparación entre las imágenes originales y las aumentadas se realiza mediante un gráfico de dispersión bidimensional utilizando el análisis de componentes principales (PCA) mostrado en la figura 4. Los puntos resultantes en el espacio bidimensional representan las imágenes y se muestra la dispersión de los descriptores originales en color verde, y los descriptores aumentados en color rojo.

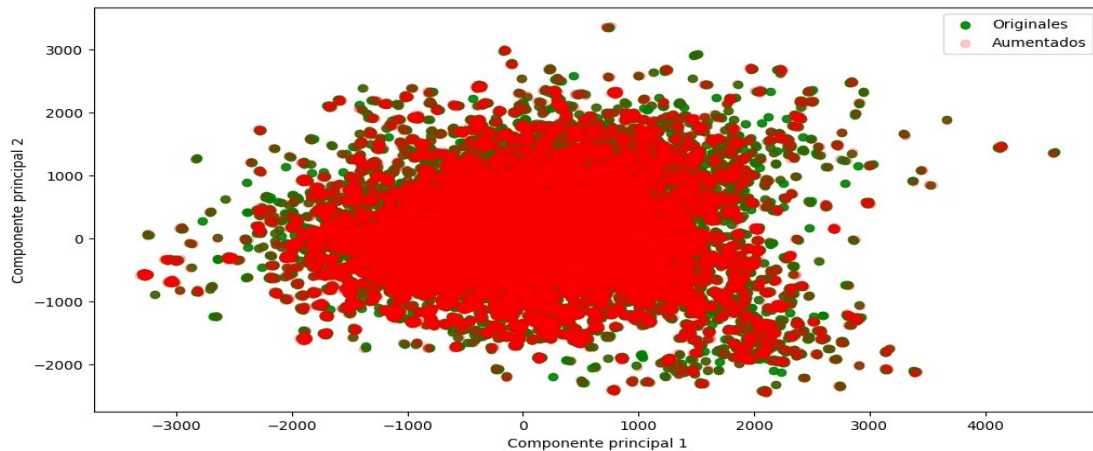


Figura 4. Gráfico de dispersión bidimensional para el conjunto de entrenamiento original y aumentado.

Adicionalmente, se genera una curva KDE (Kernel Density Estimate) donde se representan estimaciones suavizadas de las funciones de densidad de probabilidad para la primera componente principal, que se muestra en la figura 5.

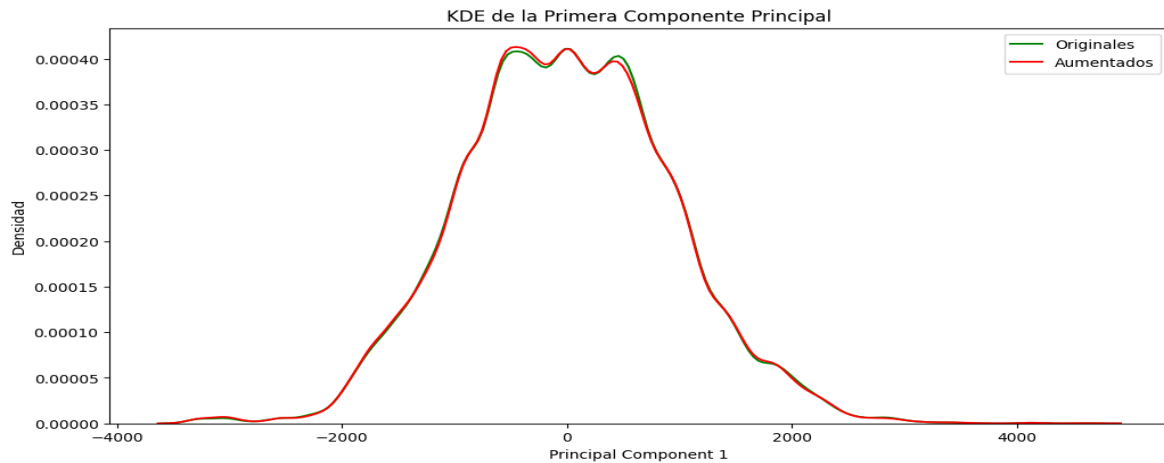


Figura 5. Curva KDE de la primera componente principal del conjunto de datos original y el conjunto de datos aumentados.

Al estar muy cerca las curvas entre sí, se infiere el proceso de aumento de datos ha mantenido la estructura subyacente de la distribución. Con esto se asegura que las variaciones introducidas no distorsionan significativamente las características fundamentales de los descriptores de imágenes originales y no se introdujeron características que generen sesgos.

6. MODELO DE APRENDIZAJE DE MÁQUINA PARA LA CLASIFICACIÓN DEL CÁNCER DE PIEL

Definida la proporción de datos para entrenamiento y prueba, se diseñó una estructura propia de red neuronal convolucional para el entrenamiento del modelo. La arquitectura combina capas convolucionales para extraer características espaciales y capas densas para la clasificación final, con técnicas como la normalización por lotes y Dropout para mejorar la generalización del modelo. A continuación, se detalla el modelado de la arquitectura propuesta para la clasificación del cáncer de piel.

- **Entrada:** La red espera una entrada con forma (28, 28, 3). Es decir, la entrada son los descriptores de una imagen a color de 28x28 píxeles.
- **Capas convolucionales:** La red tiene tres conjuntos de capas convolucionales separadas por capas de agrupación Max-pooling y normalización por lotes.
 - ❖ Primera etapa convolucional:
 - Filtros: 16
 - Tamaño del Kernel: (3,3)
 - Función de activación: ReLU
 - Padding: Same (preservando el tamaño de entrada)
 - ❖ Segunda etapa convolucional:
 - Filtros: 32
 - Tamaño del Kernel (3,3)
 - Función de activación ReLU
 - ❖ Tercera etapa convolucional:
 - Filtros: 64
 - Tamaño del Kernel (3,3)
 - Función de activación ReLU
- **Capas de agrupación Max-pooling:** Se utilizan dos capas de agrupación Max-pooling con un tamaño de ventana de (2, 2) para reducir la dimensionalidad de las salidas de las capas convolucionales y capturar características espaciales.
- **Capas de normalización por lotes:** Se utilizan capas de normalización por lotes después de cada etapa convolucional y de agrupación Max-pooling. Estas capas ayudan a estabilizar el proceso de entrenamiento y mejorar el rendimiento de la red.
- **Capas densas:** Después de las capas convolucionales, la red se aplanada para convertir la salida de la última capa convolucional en un vector unidimensional. Luego, se agregan varias capas densas completamente conectadas, de la siguiente forma:
 - ❖ Primera capa densa: 256 neuronas con activación ReLU y normalización por lotes con un 15% de dropout.

- ❖ Segunda capa densa: 128 neuronas con activación ReLU y normalización por lotes con un 15% de dropout.
- ❖ Tercera capa densa: 64 neuronas con activación ReLU y normalización por lotes con un 15% de dropout.
- ❖ Cuarta capa densa: 32 neuronas con activación ReLU y normalización por lotes con un 15% de dropout.
- **Capa de salida:** Conformada por 7 neuronas y utiliza la función de activación softmax.

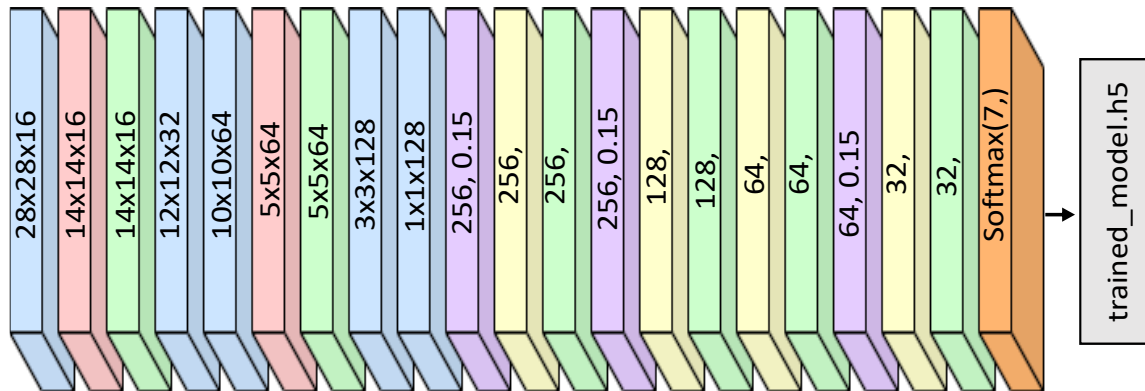


Figura 6. Arquitectura de CNN propuesta para la clasificación del cáncer de piel.

En la figura 6 se muestra la arquitectura de la CNN desarrollada. En la primera capa convolucional se aplican 16 filtros con un tamaño de Kernel de 3x3 y una función de activación ReLU a datos de entrada de 28x28 píxeles en 3 canales de color RGB. En la capa 2, se reduce la dimensión espacial de la salida de la capa anterior con un tamaño de pool de 2x2, y la tercera capa se utiliza para mejorar la estabilidad a través de la normalización. Se observa otra capa convolucional en la capa 4, pero esta vez con 32 filtros. Luego, se aplica una nueva capa convolucional con 64 filtros. Como sexta capa, se redimensiona nuevamente usando otra capa MaxPooling y, a continuación, en la capa 7, se normaliza. El proceso se repite en la capa 8, donde se aplica convolución, pero esta vez con 128 filtros, y en la capa 9, con 256 filtros. En la capa 10, la salida de las capas anteriores se convierte en un vector unidimensional antes de pasar a las capas densas. En la capa 11, y con el objetivo de prevenir el sobreajuste durante el entrenamiento, se aplica un Dropout con una tasa del 15%. La capa 12 es una capa densa completamente conectada con 256 neuronas y con función de activación ReLU, y como complemento, en la capa 13, se aplica nuevamente la normalización, pero esta vez a la capa densa. En la capa 14, se aplica otro Dropout del 15%. El proceso continúa en las capas 15 y 16, con otra capa densa de 128 neuronas y función de activación ReLU y normalización, respectivamente. En la capa 17, se agrega una capa densa con 64 neuronas y ReLU, en la capa 18 se normaliza, y en la capa 19 se aplica otro Dropout del 15%. La capa 20 es una capa densa de 32 neuronas y ReLU, y la capa 21 se refiere a la

última capa de normalización del modelo. La última capa es la capa de salida con 7 neuronas (dado el problema de clasificación multiclase del conjunto de datos con los 7 tipos de lesiones cutáneas). Para entrenar el modelo, se utiliza un optimizador Adam que combina técnicas RMSprop y Momentum, con una tasa de aprendizaje de 0.0001, que siendo baja, es útil para garantizar una convergencia suave y evitar oscilaciones alrededor de mínimos locales. El tamaño del lote (batch size) para cada actualización de peso durante el entrenamiento es de 128, y el número de épocas utilizadas es 50. Para introducir variabilidad en cada iteración, se activa la función de mezcla (shuffle).

En complemento, en la tabla 4 se muestra un resumen detallado del modelo desarrollado, donde se muestran las dimensiones de salida capa a capa, y, además, se presenta el número total de parámetros de la red, siendo 504103 parámetros, de los cuales 502983 son entrenables y 1120 no entrenables, referentes a los parámetros de normalización por lotes. La red utiliza técnicas como MaxPooling, Dropout y BatchNormalization para mejorar la generalización y prevenir el sobreajuste durante el entrenamiento.

Tabla 4. Resumen capa a capa del modelo desarrollado.

Capa (tipo)	Forma de salida	Número de parámetros
Conv2d (Conv2D)	(None, 28, 28, 16)	448
max_pooling2d (MaxPooling2D)	(None, 14, 14, 16)	0
batch_normalization (BatchNormalization)	(None, 14, 14, 16)	64
conv2d_1 (Conv2D)	(None, 12, 12, 32)	4640
conv2d_2 (Conv2D)	(None, 10, 10, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 5, 5, 64)	0
batch_normalization_1 (BatchNormalization)	(None, 5, 5, 64)	256
conv2d_3 (Conv2D)	(None, 3, 3, 128)	73856
conv2d_3 (Conv2D)	(None, 1, 1, 256)	295168
flatten (Flatten)	(None, 256)	0
dropout (Dropout)	(None, 256)	0
dense (Dense)	(None, 256)	65792
batch_normalization_2 (BatchNormalization)	(None, 256)	1024
dropout_1 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 128)	32896
batch_normalization_3 (BatchNormalization)	(None, 128)	512
dense_2 (Dense)	(None, 64)	8256
batch_normalization_4 (BatchNormalization)	(None, 64)	256

dropout_2 (Dropout)	(None, 64)	0
dense_3 (Dense)	(None, 32)	2080
batch_normalization_5 (BatchNormalization)	(None, 32)	128
dense_4 (Dense)	(None, 7)	231
Cantidad total de parámetros		504103 (1.92 MB)
Cantidad de parámetros entrenables		502983 (1.92 MB)
Cantidad de parámetros no entrenables		(4.38 KB)

Conociendo la estructura de la red neuronal convolucional posteriormente se entrenó el modelo. A lo largo de las 50 épocas de entrenamiento, se observó una evolución positiva en la capacidad del modelo para clasificar las imágenes tanto en el conjunto de entrenamiento como en el conjunto de prueba tal y como se muestra en la tabla 5. Inicialmente, la función de pérdida en el conjunto de entrenamiento fue elevada, registrando 1.8453, lo que indicaba una falta de ajuste inicial del modelo. Sin embargo, a medida que avanzaban las épocas, la pérdida en el conjunto de entrenamiento disminuyó gradualmente, llegando a 0.0272 al final del proceso. Este descenso constante refleja la capacidad del modelo para aprender patrones complejos en los datos de entrenamiento. De manera similar, en el conjunto de prueba, la pérdida también disminuyó de manera consistente, sugiriendo una generalización efectiva a datos no vistos. La evolución de la exactitud sigue una tendencia similar, mejorando continuamente en ambos conjuntos a medida que avanzaban las épocas. Este comportamiento indica que el modelo logró adaptarse a la complejidad de los datos de manera progresiva, demostrando así su capacidad para aprender y generalizar de manera efectiva. Asimismo, se destaca una tendencia a mantener relativamente estables los valores tanto del tiempo de entrenamiento como el tiempo por paso (ms/paso). El tiempo de entrenamiento oscila alrededor de los 47 segundos por época, sin mostrar una clara disminución o aumento significativo. De manera similar, los ms/paso permanecen en torno a los 200 ms, con fluctuaciones mínimas. Aunque no hay una mejora significativa en la eficiencia del tiempo de entrenamiento o en los ms/paso, la estabilidad en estos valores sugiere que el modelo mantiene una consistencia en su rendimiento a lo largo del proceso de entrenamiento. Es importante destacar que, a pesar de no haber una disminución marcada en estos valores, su estabilidad puede indicar una convergencia efectiva del modelo.

Tabla 5. Evolución del proceso de entrenamiento de la CNN propuesta.

Época	Pérdida en entrenamiento	Exactitud entrenamiento	Pérdida en validación	Exactitud validación	Tiempo de entrenamiento	ms/paso
1	1.8453	0.3084	1.6457	0.3542	00:00:52	203

2	1.4013	0.4733	1.1778	0.5818	00:00:49	209
3	1.1849	0.5683	0.9827	0.6510	00:00:49	205
4	1.0145	0.6308	0.8435	0.7029	00:00:48	204
5	0.8809	0.6858	0.7173	0.7483	00:00:48	202
6	0.7637	0.7305	0.6230	0.7840	00:00:47	199
7	0.6717	0.7646	0.5480	0.8132	00:00:48	203
8	0.5928	0.7915	0.4879	0.8305	00:00:47	196
9	0.5181	0.8188	0.4257	0.8509	00:00:48	203
10	0.4555	0.8416	0.3810	0.8641	00:00:47	199
11	0.4027	0.8587	0.3540	0.8772	00:00:50	210
12	0.3557	0.8772	0.3052	0.8928	00:00:47	197
13	0.3165	0.8901	0.2880	0.9014	00:00:48	204
14	0.2803	0.9045	0.2509	0.9122	00:00:47	198
15	0.2508	0.9146	0.2464	0.9165	00:00:47	199
16	0.2287	0.9237	0.2267	0.9195	00:00:47	199
17	0.1959	0.9342	0.2175	0.9253	00:00:50	210
18	0.1780	0.9401	0.1959	0.9333	00:00:47	197
19	0.1588	0.9479	0.1952	0.9361	00:00:48	203
20	0.1429	0.9525	0.1853	0.9389	00:00:47	198
21	0.1309	0.9563	0.1686	0.9451	00:00:47	197
22	0.1210	0.9602	0.1686	0.9455	00:00:50	204
23	0.1080	0.9660	0.1733	0.9434	00:00:47	197
24	0.0977	0.9678	0.1629	0.9506	00:00:50	210
25	0.0893	0.9702	0.1787	0.9452	00:00:47	198
26	0.0832	0.9726	0.1574	0.9535	00:00:48	201
27	0.0799	0.9734	0.1473	0.9533	00:00:48	202
28	0.0701	0.9766	0.1586	0.9489	00:00:47	199
29	0.0696	0.9776	0.1658	0.9527	00:00:49	205
30	0.0631	0.9790	0.1470	0.9547	00:00:47	200
31	0.0635	0.9787	0.1520	0.9567	00:00:48	204
32	0.0546	0.9823	0.1639	0.9525	00:00:47	200
33	0.0535	0.9823	0.1569	0.9572	00:00:48	203
34	0.0492	0.9850	0.1384	0.9612	00:00:47	200
35	0.0490	0.9837	0.1371	0.9632	00:00:57	239
36	0.0418	0.9867	0.1539	0.9609	00:00:47	198
37	0.0430	0.9864	0.1563	0.9585	00:00:50	211
38	0.0426	0.9869	0.1507	0.9617	00:00:47	199
39	0.0375	0.9877	0.1467	0.9618	00:00:48	203
40	0.0406	0.9865	0.1507	0.9603	00:00:47	197
41	0.0369	0.9875	0.1510	0.9630	00:00:47	200
42	0.0400	0.9873	0.1462	0.9601	00:00:48	204
43	0.0365	0.9874	0.1339	0.9626	00:00:47	199

44	0.0330	0.9891	0.1401	0.9641	00:00:47	198
45	0.0322	0.9895	0.1453	0.9658	00:00:50	210
46	0.0307	0.9903	0.1607	0.9592	00:00:47	197
47	0.0341	0.9894	0.1699	0.9600	00:00:48	205
48	0.0284	0.9903	0.1499	0.9630	00:00:47	199
49	0.0263	0.9914	0.1335	0.9679	00:00:50	210
50	0.0272	0.9909	0.1488	0.9630	00:00:47	198

Asimismo, en la figura 7 se muestran las gráficas que describen la evolución de la exactitud y la función de pérdida durante el entrenamiento de la red neuronal convolucional propuesta para la clasificación del cáncer de piel.

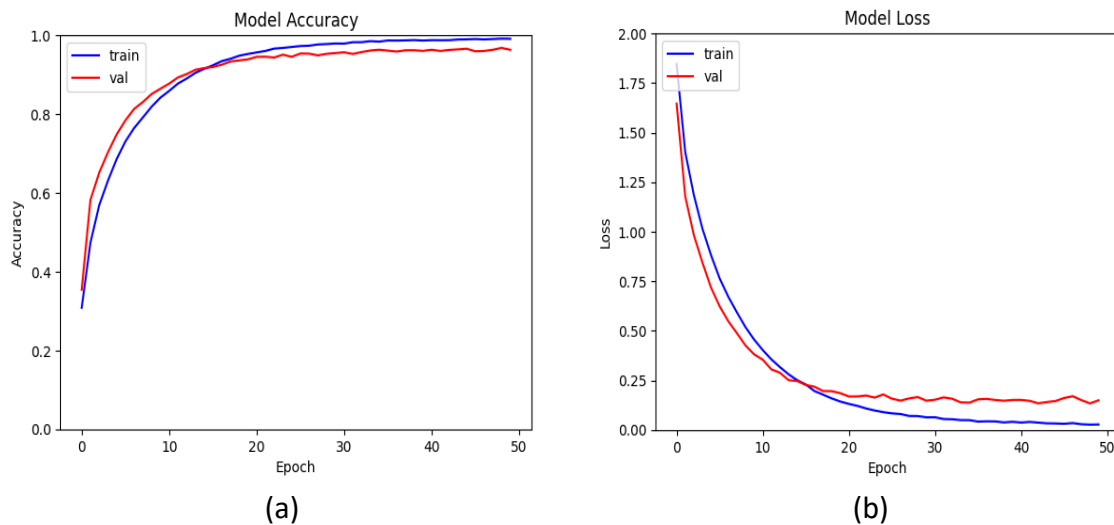


Figura 7. Evolución del entrenamiento de la CNN. (a) Datos de entrenamiento, (b) Datos de validación.

En complemento a lo anterior, se presenta la matriz de confusión que resume el desempeño del modelo computacional desarrollado para la clasificación de los 7 tipos de lesiones dermatológicas. Para la clase 0 ('akiec'), se obtuvo una tasa de aciertos de 99.50 %, mientras que para la clase 1 ('bcc') la tasa de aciertos en la clasificación fue de 99.33 %. En adición, la clase 2 ('bkl') presentó un acierto en la clasificación de 96.39 %. La clase 3 ('df') obtuvo un valor de 98.10 % en la clasificación, mientras que la clase 4 ('nv'), presentó una tasa de acierto un poco mas baja en la clasificación con un valor de 85.26 %. En cuanto a la clase 5 ('vasc') la tasa de acierto en la clasificación se elevó a 99.71 %. Finalmente, la clase 6 ('mel') presentó una tasa de acierto en la clasificación de 93.92 %. Estos resultados subrayan la complejidad de la tarea de clasificación en múltiples clases dermatológicas y la capacidad del modelo desarrollado para clasificar los 7 tipos de lesiones cutáneas que se presentan en el conjunto de datos HAM10000. En la figura 8 se presenta la matriz de confusión en porcentajes.

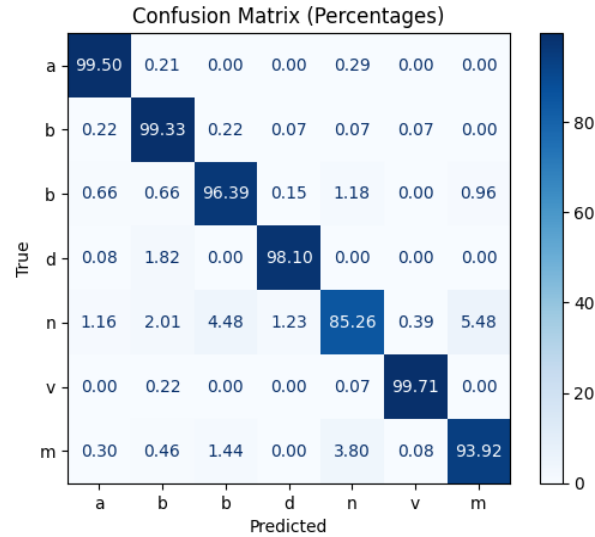


Figura 8. Matriz de confusión porcentual del modelo computacional desarrollado.

Adicionalmente, en la tabla 6 se muestra el desempeño del modelo convolucional desarrollado para la clasificación de lesiones cutáneas.

Tabla 6. Desempeño del modelo de clasificación propuesto.

Métricas de desempeño	Clases						
	AKIEC	BCC	BKL	DF	NV	VASC	MEL
Recall	99.5 %	99.33 %	93.69 %	98.1 %	85.26 %	99.71 %	93.92 %
Precision	97.76 %	94.94 %	94.23 %	98.55 %	93.88 %	99.49 %	93.64 %
F1-Score	98.62 %	97.09 %	95.30 %	98.33 %	89.37 %	99.6 %	93.78 %
Especificidad	99.6 %	99.12 %	99 %	99.76 %	99.11 %	99.91 %	98.96 %
Exactitud	96.11 %						
Tiempo de entrenamiento	46 minutos						
CPU utilizada	53 %						
RAM utilizada	31 %						

Para la clase AKIEC, el modelo demostró un alto rendimiento con un recall del 99.5%, precisión del 97.76% y F1-Score del 98.62%. Similarmente, las clases BCC, BKL, y DF también presentaron sólidos resultados, con altos valores de recall, precisión y F1-Score. La clase NV mostró un recall ligeramente más bajo (85.26%), pero con una precisión del 93.88%. La clase VASC demostró un rendimiento excepcional, con un recall del 99.71% y precisión del 99.49%. La clase MEL mostró resultados sólidos con un recall del 93.92%, precisión del 93.64% y F1-Score del 93.78%. En general, el modelo exhibió una especificidad elevada en todas las clases, indicando su capacidad para identificar correctamente las instancias negativas. La exactitud general del modelo fue del 96.11%, resaltando su eficacia en la clasificación de múltiples clases dermatológicas. Además, para evaluar la capacidad del

modelo para distinguir entre clases se calcula la Curva de Características Operativa del Receptor (ROC) y el Área bajo la Curva (AUC). La ROC visualiza la relación entre la tasa de verdaderos positivos y la tasa de falsos positivos y el AUC cuantifica la habilidad del modelo para discriminar clases. La curva de ROC con los valores de AUC se muestra en la figura 9.

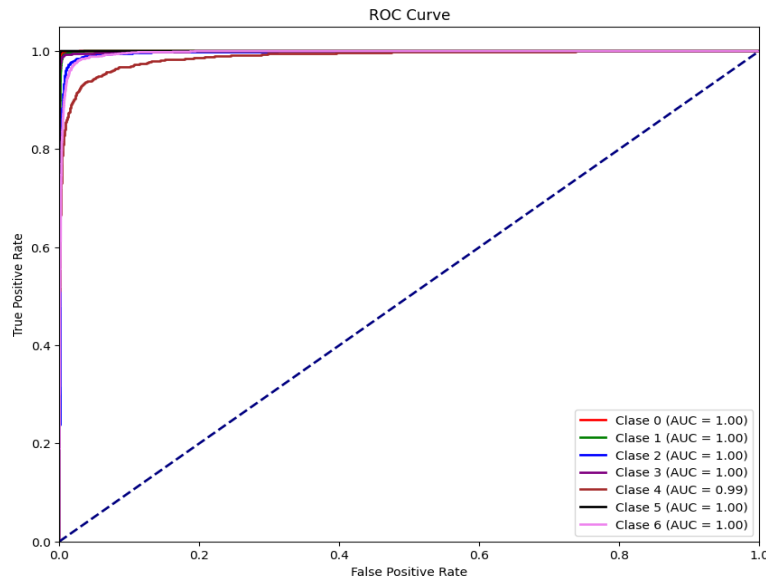


Figura 9. Curva de ROC con valores de AUC para cada clase de lesiones cutáneas.

Se observa un buen rendimiento del modelo de clasificación en siete clases dermatológicas, con valores de Área bajo la Curva (AUC) altos. Todas las clases, a excepción de la Clase 4 ('nv') que obtuvo un AUC de 0.99, demostraron una AUC perfecta de 1. Este rendimiento sugiere una capacidad casi ideal del modelo para discriminar entre las clases, destacando su eficacia en la tarea de clasificación dermatológica. La alta consistencia en los valores de AUC refuerza la robustez del modelo en la identificación precisa de cada categoría.

Para consolidar la prueba de la red neuronal convolucional propuesta para la clasificación de cáncer de piel se probaron los modelos de aprendizaje automático Random Forest y XGBoost, con hiperparámetros encontrados mediante barrido de variables. Además, se comparó el desempeño de la clasificación respecto a artículos científicos publicados en revistas científicas indexadas en categoría Q1 y Q2 por Scopus donde se implementan técnicas basadas en aprendizaje profundo.

Respecto a la técnica de Random Forest, se genera un conjunto de datos de entrenamiento aleatorio a partir del conjunto de datos original, y, además, se entrena un árbol de decisión en el conjunto de datos de entrenamiento aleatorio generando un conjunto de árboles de decisión. Para realizar una predicción, se consultan todos los árboles de decisión del bosque. La predicción final es el resultado de la votación de los árboles. Para este caso, el número de árboles se probó entre 50, 100 y 150. Además, el parámetro de profundidad máxima que controla la profundidad máxima de cada árbol del bosque se probó con 10, 20 y 30, así como sin límite de profundidad. En caso de emplearse sin límite de profundidad, los árboles se

expanden hasta que cada hoja contenga menos muestras requeridas. El parámetro número mínimo de muestras para dividir un nodo interno se probó entre 2, 5 y 10, considerando que entre mayor sea este valor menos divisiones se realizan, ayudando a prevenir el sobreajuste. Además, el parámetro de número mínimo de muestras requeridas en una hoja fue probado con valores de 1, 2 y 4. Con la búsqueda para la optimización, se definieron los siguientes hiperparámetros con sus valores: Número de árboles = 100, profundidad máxima: 30, número mínimo de muestras en una hoja: 1, número mínimo de muestras para dividir un nodo interno: 2. Las cifras de desempeño al probar las métricas se muestran en la tabla 7. Destacan altos valores de recall, precisión y F1-Score para diversas clases como AKIEC, BCC, DF, y VASC. Sin embargo, es importante señalar que los resultados anteriores de la CNN propuesta superaron consistentemente estos valores. Aunque el modelo basado en Random Forest muestra un rendimiento aceptable, con una exactitud general del 90.08%, los resultados previos de la CNN propuesta, con una exactitud del 96.11%, aún destacan como superiores. En adición, el tiempo de entrenamiento con búsqueda de hiperparámetros fue de 225 minutos, exigiendo la CPU en el 87 % de su capacidad, y el 49 % de la RAM disponible.

Tabla 7. Desempeño de la técnica de Random Forest en la clasificación.

Métricas de desempeño	Clases						
	AKIEC	BCC	BKL	DF	NV	VASC	MEL
Recall	98.35 %	94.30 %	78.61 %	99.64 %	72.77 %	99.22 %	87.55 %
Precision	92.45 %	87.37 %	82.92 %	96.95 %	89.34 %	98.37 %	82.72 %
F1-Score	95.31 %	90.70 %	80.71 %	98.28 %	80.21 %	98.80 %	85.06 %
Especificidad	92.45 %	87.37 %	82.92 %	96.95 %	89.34 %	98.37 %	82.72 %
Exactitud	90.08 %						
Tiempo de entrenamiento	225 minutos						
CPU utilizada	87 %						
RAM utilizada	49 %						

Igualmente, en la figura 10 se muestra tanto la matriz de confusión porcentual como la curva de ROC con valores de AUC al entrenar el modelo basado en Random forest. En términos generales, la exactitud por clases superó el 95 % en 3 clases, el 90 % en 1 clase, el 80 % en 1 clase y las 2 clases restantes alcanzaron valores de 78.61 % y 72.77 %. Respecto al comportamiento de la curva de ROC y el área bajo la curva, en 3 clases se alcanzó el valor de 100 % en AUC, en una clase 99 %, en una clase 98 % y en 2 clases el 97 %. Si bien son resultados prometedores, siguen siendo inferiores a los obtenidos con la CNN propuesta.

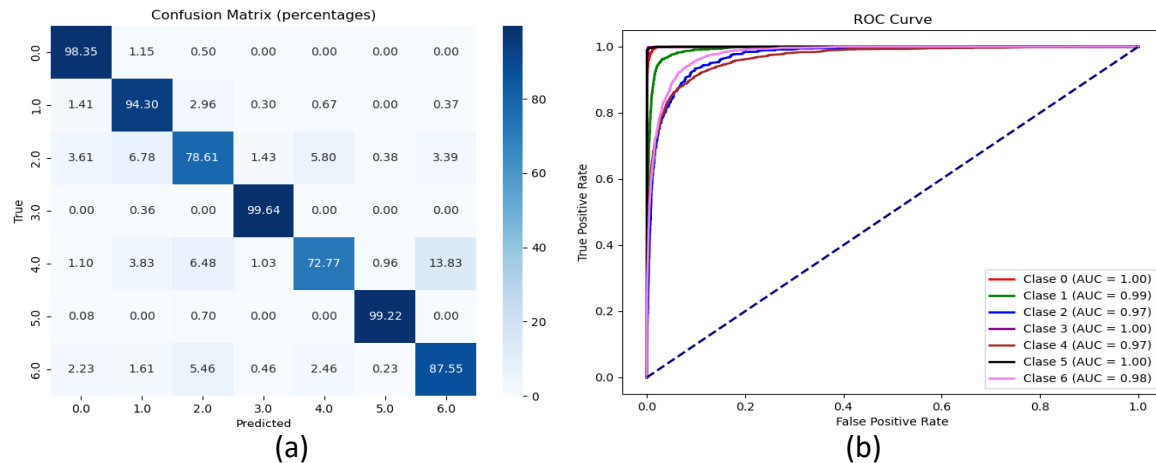


Figura 10. Desempeño general del modelo basado en Random forest.

El entrenamiento se replicó, pero esta vez probando XGBoost, un algoritmo de aprendizaje automático supervisado que es una implementación eficiente y escalable de la técnica de boosting, que combina múltiples modelos más débiles, generalmente árboles de decisión, para crear un modelo más fuerte y preciso. Para este caso, hiperparámetros como número de árboles a construir fue probado con valores de 50, 100 y 150, la profundidad máxima de cada árbol fue probada con valores de 3, 5 y 7. La tasa de aprendizaje que controla la contribución de cada árbol al modelo fue probada con valores de 0.01, 0.1 y 0.2, considerando que valores más bajos hacen que el modelo sea más robusto, pero puede requerir de mayor número de árboles. La mínima suma de pesos de instancias en un nodo hoja se probó con valores de 1, 3 y 5. La proporción de instancias de entrenamiento a utilizar en cada paso se probó con valores de 0.8 y 1.0. En complemento, la proporción de instancias de entrenamiento a utilizar en cada paso de construcción del árbol se probó con esos mismos valores. La mínima reducción de pérdida requerida para realizar una partición inicial en un nodo dada por gamma se probó con 0, 0.1 y 0.2, mismos valores empleados para la regularización L1 y L2. Con la búsqueda para la optimización, se definieron los siguientes hiperparámetros con sus valores: Tasa de aprendizaje: 0.001, fracción de características para el entrenamiento de cada árbol: 1, profundidad máxima: 7, número de árboles a construir: 150, muestras: 1, y regularización L1 y L2: 0.1. En la tabla 8 se muestra el desempeño del modelo basado en XGBoost. El modelo exhibe un rendimiento general positivo, según las métricas presentadas.

Tabla 8. Desempeño de la técnica de XGBoost en la clasificación.

Métricas de desempeño	Clases						
	AKIEC	BCC	BKL	DF	NV	VASC	MEL
Recall	99.86 %	98.67 %	91.94 %	100 %	80.43 %	100 %	93.24 %
Precision	97.75 %	95.01 %	90.18 %	99.35 %	93.98 %	99.46 %	88.35 %
F1-Score	98.79 %	96.80 %	91.05 %	99.67 %	86.68 %	99.73 %	90.73 %
Especificidad	97.75 %	95.01 %	90.18 %	99.35 %	93.98 %	99.46 %	88.35 %

Exactitud	94.88 %
Tiempo de entrenamiento	960 minutos
CPU utilizada	96 %
RAM utilizada	80 %

El recall, que mide la capacidad del modelo para identificar positivos, es notablemente alto para la mayoría de las clases, destacando un 100% para la clase "DF". Sin embargo, la clase "NV" muestra un recall del 80.43%, indicando que el modelo podría pasar por alto algunos casos positivos de esta categoría. En términos de precisión, que evalúa la precisión de las predicciones positivas, el modelo se desempeña bien, con altos valores en todas las clases. La clase "VASC" destaca con una precisión del 99.46%, sugiriendo que las predicciones positivas para esta clase son altamente confiables. El "F1-Score", que equilibra precisión y recall, refleja un rendimiento sólido en general. No obstante, la clase "NV" muestra un F1-Score relativamente bajo del 86.68%, indicando que podría haber un desequilibrio entre la capacidad del modelo para identificar positivos y la precisión de esas predicciones. La especificidad, que mide la capacidad del modelo para identificar correctamente los casos negativos, es alta y coincide con las métricas de precisión para todas las clases. En cuanto a la exactitud, el modelo logra un 94.88%, indicando que la proporción de predicciones correctas, tanto positivas como negativas, es alta. Si bien el desempeño del modelo en términos de software es bueno, cabe resaltar que demora aproximadamente 20 veces más que el modelo de CNN propuesto, y eleva aproximadamente al doble los niveles de CPU y RAM utilizados.

Asimismo, en la figura 11 se muestra la matriz de confusión porcentual y la curva de ROC con valores de AUC para el modelo basado en XGBoost.

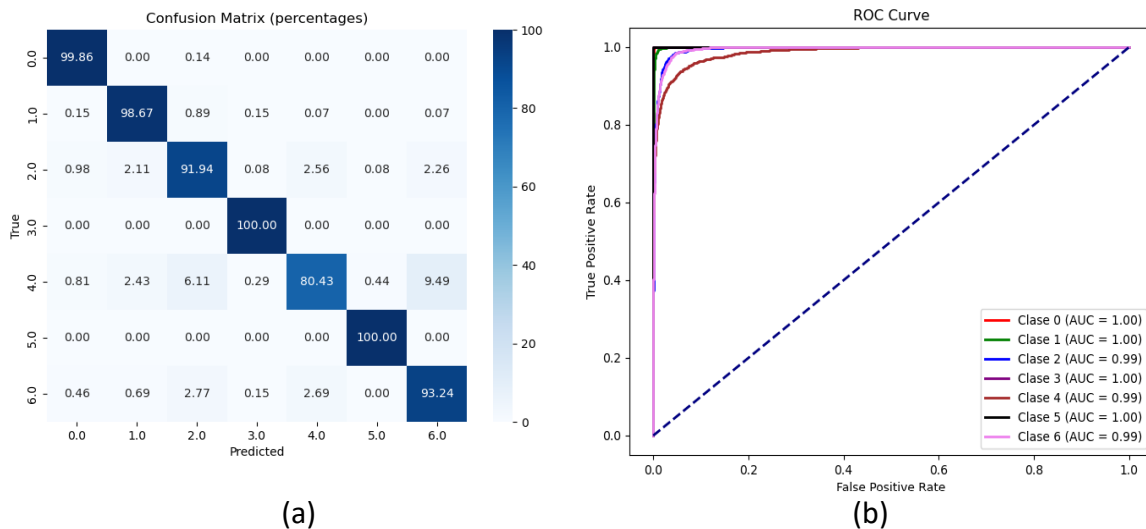


Figura 11. Desempeño general del modelo basado en Random forest.

La exactitud por clases superó el 95 % en 4 clases, el 90 % en 2 clases, y el 80 % en 1 clase. Respecto al comportamiento de la curva de ROC y el área bajo la curva, en 4 clases se alcanzó el valor de 100 % en AUC, y en las tres restantes el valor del AUC alcanzó el 99 %.

Para continuar con la prueba de la CNN propuesta para la clasificación del cáncer de piel, se compara ahora con trabajos del estado del arte donde se hayan involucrado estructuras basadas en redes neuronales convolucionales y se comparó el desempeño de las métricas, tal y como se muestra en la tabla 9. La red propuesta de 22 capas destaca en comparación con varias técnicas, mostrando un rendimiento superior en términos de recall (0.9603), precisión (0.9607), F1 Score (0.9601), especificidad (0.9935) y exactitud (0.9611). Estas métricas sugieren una capacidad robusta para identificar correctamente tanto los casos positivos como los negativos. Esta red supera a varias otras técnicas, incluida la EW-FCM + ShuffleNet, que, aunque muestra una especificidad razonable (0.7071) y exactitud (0.848), tiene un rendimiento inferior en términos de recall y precisión. Al observar otras técnicas, la Xception modificada (SBXception) exhibe una alta recall (0.9543) y precisión modesta (0.8534), indicando una buena capacidad para identificar positivos, aunque con menos precisión. La CNN modificada (DSCC_Net) también presenta un rendimiento competitivo, con un alto recall (0.9376) y precisión (0.9428). La estructura convolucional desarrollada mostró mejor desempeño que trabajos desarrollados y publicados en revista científicas de alto impacto categorizadas en Q1 y Q2 de Scopus, mostrando así que el enfoque propuesto es eficiente y fiable en su implementación.

Esta mejora en el rendimiento se obtuvo gracias a la estrategia de prescindir de modelos pre-entrenados y optar por una construcción gradual de la red, lo que permitió una adaptación más efectiva de la arquitectura a las complejidades específicas de la clasificación del cáncer de piel. Al diseñar la red desde cero, se tuvo la libertad de ajustar cada capa y parámetro para maximizar la capacidad de la red para aprender y generalizar patrones relevantes en las imágenes de lesiones cutáneas. Un aspecto destacado de este enfoque fue la estrategia de adaptación de descriptores de imagen a una resolución más baja de 28x28, en contraste con las imágenes originales de 600x600. Esta reducción de resolución mitigó efectivamente el impacto del ruido y las áreas con vello corporal, mejorando la capacidad de la red para identificar características clave asociadas con el cáncer de piel. Además, la aplicación de técnicas de inducción de ruido controlado proporcionó un equilibrio más uniforme entre las clases, sin comprometer la integridad del conjunto de datos.

Tabla 9. Comparación de rendimiento entre el modelo propuesto y trabajos del estado del arte publicados en revistas Q1 y Q2 de Scopus.

Trabajo	Scopus	Técnica	Recall	Precision	F1 Score	Especificidad	Exactitud
CNN		Red propia de 22 capas	0.9603	0.9607	0.9601	0.9935	0.9611
[31]	Q2	EW-FCM + ShuffleNet	-	0.7515	0,7261	0.7071	0.848
[32]	Q1	Red propia de 24 capas	-	0.8701	0.8628	0.857	0.865
[33]	Q2	Xception modificada: SBXception	0.9543	0.8534	0.8896	-	0.9697
[34]	Q2	DeepLabv3+ Fine KNN	-	0.9185	0.9214	0.9185	0.9201
[35]	Q2	Red propia de 24 capas	-	-	-	-	0.80
[36]	Q1	Mask RCNN+DenseNet201+ LS-SVM	-	0.8886	0.8860	0.8854	0.8850
[37]	Q1	ResNet-50+ Naïve Bayes classifier	-	0.8628	0.8614	0.86	0.858
[38]	Q1	Red propia de 15 capas	0.85	0.88	0.86	-	0.94
[39]	Q2	MESbS+ C-SVM	0.8241	0.8297	-	-	0.828
[40]	Q2	NASNET Large + WOA + ELM	-	0.931	-	-	0.934
[41]	Q2	CNN modificada: DSCC_Net	0.9376	0.9428	0.9376	-	0.9417
[42]	Q2	Transfer Learning: Xception	0.8957	0.8876	0.8902	-	0.8966
[43]	Q1	CNN optimizada: X-AI	-	-	-	-	0.8124
[44]	Q1	DRANet	0.897	0.905	0.897	-	0.897
[45]	Q1	G-AI + GAN + ESRGAN + VGG19 + SVM	0.93	0.93	0.93	-	0.93
[46]	Q1	Modelo convolucional propio	-	-	-	0.943	0.9749
[47]	Q1	CUSS-Net no supervisada	-	-	-	0.959	0.898
[48]	Q1	DL Ensemble (VGG, CapsNet, ResNet)	-	0.94	0.92	0.870	0.935
[49]	Q1	EfficientNetB4-CLF	0.8061	0.86	-	-	0.8761
[50]	Q1	Transfer Learning: DenseNet169	-	-	0.9170	-	0.9120

7. HERRAMIENTA DE DIAGNÓSTICO ASISTIDO EN DISPOSITIVO DE BAJO CONSUMO

Para la implementación y uso del modelo de clasificación en un contexto y ambiente real, se diseñó una herramienta de diagnóstico asistido basado en interfaz gráfica centrada en el usuario. Para el diseño de la herramienta se utilizó QtDesigner, una herramienta de diseño gráfico que forma parte del framework Qt. De esta forma, se facilitó la disposición y diseño de elementos gráficos para el diagnóstico asistido, y, además, permitió que después del diseño se conecte el código Python para generar funcionalidad total a la interfaz. Esta herramienta es compatible en multiplataforma, en sistemas operativos como Windows, Linux y macOS.

La primera etapa de la herramienta de diagnóstico asistido se centra en el diseño de una interfaz de inicio de sesión, dado que para el diagnóstico y clasificación de cáncer de piel se maneja información sensible y confidencial sobre pacientes. Además, la identificación de usuarios mediante inicio de sesión permite a futuro llevar registro de las acciones realizadas por cada usuario. Asimismo, se reduce significativamente el riesgo de accesos no autorizados para velar por la integridad y confidencialidad de la información. En la figura 12 se muestra la interfaz de usuario desarrollada para el inicio de sesión, en donde se aprecia la presentación de la herramienta y la instrucción de ingresar credenciales, los espacios para el ingreso de usuario y de contraseña, y botones para acceder y recordar las credenciales de ingreso al sistema.



Figura 12. Interfaz de inicio de sesión al sistema.

Los dos botones presentes en la interfaz de inicio de sesión representan los dos escenarios posibles en este punto, considerando la necesidad de resguardar la información del proceso y de los pacientes. En caso de que el usuario de la herramienta (médico especialista) ingrese correctamente las credenciales accederá a la interfaz de diagnóstico médico, y en caso de que las credenciales sean incorrectas no podrá acceder a la interfaz. Si el usuario olvidó la contraseña, mediante el botón de recordar recibirá un mensaje de texto al teléfono celular registrado con las credenciales para acceder al sistema. Para la generación automática de

mensajes de texto se utilizó un servicio de Twilio configurado, con identificador de cuenta, token y un número virtual generado por Twilio, tal y como se muestra en la sección (a) de la figura 13. En adición, en la sección (B) se muestra el mensaje de texto generado al usuario de la herramienta (como prueba se muestra un usuario creado con el nombre CarlosVicenteNinoRondon y la contraseña 12345)

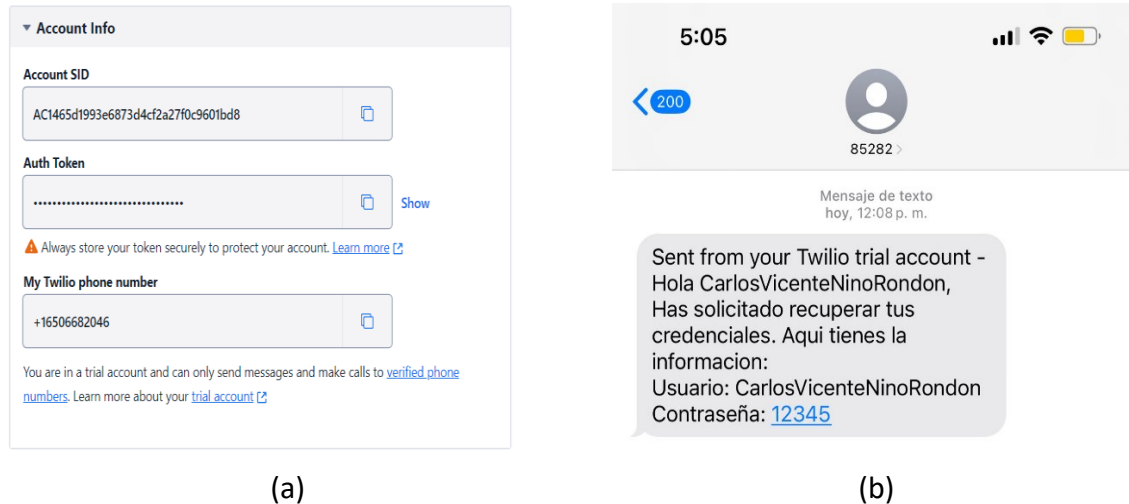


Figura 13. Recuperación de credenciales para ingresar al sistema.

Al ingresar correctamente con las credenciales de inicio de sesión se accede directamente a la interfaz de diagnóstico médico. Para el desarrollo de la interfaz fue necesario el uso de los paquetes de librerías mostrados en la tabla 10, donde se relaciona el nombre de la librería, la versión empleada, el tamaño de la misma y la funcionalidad de estas en la interfaz de diagnóstico de cáncer de piel.

Tabla 10. Librerías utilizadas para el desarrollo de la interfaz gráfica de usuario.

Librería	Versión	Tamaño	Funcionalidad
PyQt5	PyQt5 5.15.7	50.1 MB	Proporciona clases y funciones para crear interfaces gráficas de usuario (GUI), manejar eventos, realizar operaciones de impresión, etc.
	Qt 5.15.2		
TensorFlow	2.15.0- dev20230920	300.9 MB	Carga modelos de aprendizaje profundo previamente entrenados.
OpenCV	4.8.0	38.6 MB	Procesamiento de imágenes de lesiones cutáneas
Pillow	9.4.0	16.4 MB	Manipulación de imágenes. Ofrece soporte para abrir, manipular y guardar varios formatos de imagen

Numpy	1.24.3	15.8 MB	Manipulación de arreglos y matrices multidimensionales. Proporciona funciones para realizar operaciones matemáticas
Pandas	1.5.3	12.3 MB	Análisis y manipulación de datos. Facilita la manipulación y análisis de datos tabulares.

Asimismo, en la tabla 11 se muestra un pseudocódigo que representa las principales funcionalidades de la herramienta de diagnóstico asistido.

Tabla 11. Pseudocódigo con funcionalidades básicas de la interfaz de usuario desarrollada.

Función	
1.	Importar bibliotecas necesarias
2.	Definir la clase de la interfaz de usuario
3.	Definir el método para inicializar la interfaz de usuario
4.	Definir elementos de la interfaz como botones, y etiquetas
5.	Definir el método para capturar y guardar un PDF de la interfaz
6.	Capturar la interfaz como una imagen y guardarla como PDF
7.	Definir el método para cargar una imagen en la interfaz
8.	Abrir un cuadro de diálogo para seleccionar una imagen y mostrarla en la interfaz
9.	Definir el método para mostrar información de los desarrolladores
10.	Mostrar un cuadro de diálogo con información sobre los desarrolladores
11.	Definir el método para confirmar la salida de la aplicación
12.	Mostrar un cuadro de diálogo de confirmación y cerrar la aplicación si es afirmativo
13.	Definir el método para mostrar un mensaje de uso
14.	Mostrar un cuadro de diálogo con instrucciones de uso
15.	Definir el método para realizar el diagnóstico de una lesión
16.	Cargar un modelo previamente entrenado y realizar el diagnóstico
17.	Definir el método para guardar información del paciente en un archivo CSV
18.	Obtener información del paciente y guardarla en un archivo CSV
19.	Ejecutar el programa principal
20.	Mostrar la interfaz
21.	Iniciar el bucle de eventos de la aplicación

En primer lugar, se importan las bibliotecas necesarias para la interfaz gráfica y el manejo de datos. Luego, se define una clase que representa la interfaz de usuario, con métodos para inicializarla y manejar diversas funcionalidades. Los elementos de la interfaz, como botones y etiquetas, son definidos y configurados. Existen funciones para capturar y guardar la interfaz como un archivo PDF, cargar imágenes en la interfaz, mostrar información sobre los desarrolladores, confirmar la salida de la aplicación, mostrar mensajes de uso y realizar diagnósticos de lesiones utilizando un modelo previamente entrenado. Además, hay un método para guardar la información del paciente en un archivo CSV. En el programa principal, se instancia la aplicación, se crea la interfaz y se muestra al usuario. La ejecución del programa entra en un bucle de eventos que permite interactuar con la interfaz hasta que el usuario decide cerrar la aplicación.

En la figura 14 se muestra la interfaz gráfica diseñada para la etapa de diagnóstico de la lesión cutánea mediante la red neuronal convolucional propuesta.

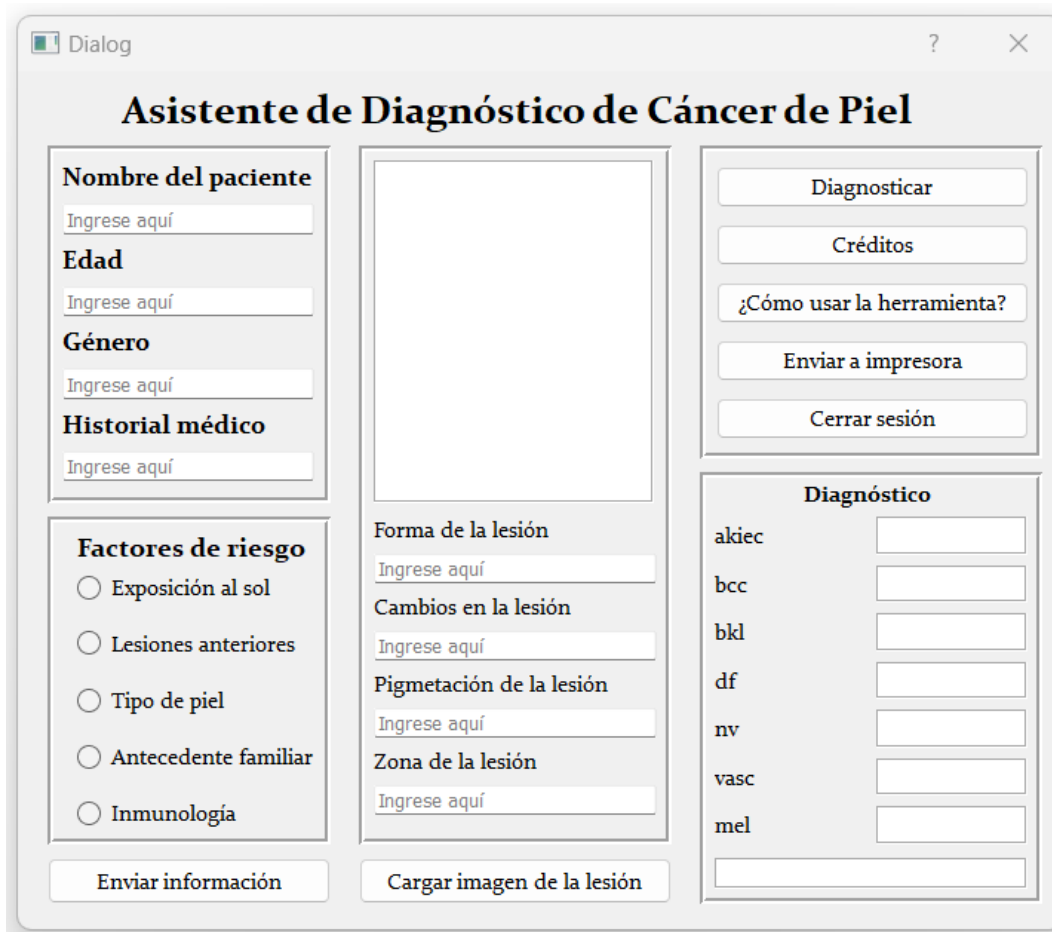


Figura 14. Interfaz gráfica desarrollada para el diagnóstico de la lesión cutánea.

En la interfaz gráfica se utilizó un total de 29 etiquetas distribuidas en distintas secciones. Estas etiquetas se emplean para mostrar información como el título de la aplicación,

imágenes, resultados de diagnóstico, y campos de entrada para información del paciente y la lesión. En cuanto a los botones de tipo Push, se emplean un total de 7, con funcionalidades específicas tales como guardar la información del paciente en un archivo CSV, cargar la imagen de la lesión, cerrar la sesión de la aplicación con una confirmación, capturar y guardar la interfaz como un archivo PDF, mostrar mensajes de ayuda sobre el uso de la herramienta, visualizar los créditos de los desarrolladores y realizar el diagnóstico de la lesión. Adicionalmente, se utilizan cinco botones a modo de lista de chequeo asociados a los factores de riesgo de las lesiones de cáncer de piel.

En adición, en la tabla 12 se relacionan los casos de uso de la interfaz gráfica de usuario. La tabla de casos de uso describe las interacciones clave en la interfaz gráfica de la aplicación. Comienza con el "Iniciar Aplicación", donde el usuario da comienzo al uso de la aplicación. A continuación, el "Ingresar Información del Paciente" permite al usuario introducir datos relevantes del paciente, seguido por la opción de "Seleccionar Factores de Riesgo". Posteriormente, el usuario puede "Cargar Imagen de la Lesión" para que el sistema realice el "Diagnóstico de la Lesión". Una vez completado, el usuario puede "Visualizar Resultado del Diagnóstico". Los datos se almacenan en un archivo CSV mediante el caso de uso "Almacenar Información en Archivo CSV" y se puede generar un resumen en formato PDF con "Generar Archivo PDF de la Interfaz Actual". Finalmente, el usuario puede "Cerrar Sesión" para concluir la interacción con la aplicación.

Tabla 12. Casos de uso de la interfaz gráfica desarrollada para el diagnóstico.

Caso de uso	Actor principal	Descripción del caso de uso
Iniciar aplicación	Usuario	El usuario inicia la aplicación desde el dispositivo.
Ingresar información del paciente	Usuario	El usuario introduce datos del paciente.
Seleccionar factores de riesgo	Usuario	El usuario elige factores de riesgo.
Cargar imagen de la lesión	Usuario	El usuario carga una imagen de la lesión.
Realizar diagnóstico de la lesión	Sistema	El sistema analiza la imagen y datos para realizar un diagnóstico.
Visualizar resultado del diagnóstico	Usuario	Presenta al usuario el resultado del diagnóstico.
Almacenar información en archivo CSV	Sistema	Los datos del paciente y el diagnóstico se almacenan en un archivo CSV.
Generar archivo PDF de la interfaz	Sistema	Genera un archivo PDF que resume la interfaz y sus resultados.
Cerrar sesión	Usuario	El usuario cierra la sesión y finaliza la aplicación.

Asimismo, en la figura 15 se muestra la interfaz gráfica funcional, donde se observa la imagen cargada, las probabilidades de diagnóstico y la información de caracterización del paciente y de la lesión cutánea ingresadas al sistema.

Dialog
?
×

Asistente de Diagnóstico de Cáncer de Piel

Nombre del paciente
Carlos Vicente Niño Rondón

Edad
25

Género
Masculino

Historial médico
No aplica

Factores de riesgo

Exposición al sol

Lesiones anteriores

Tipo de piel

Antecedente familiar

Inmunología

Enviar información



Forma de la lesión
Circular

Cambios en la lesión
No aplica

Pigmentación de la lesión
Rosado

Zona de la lesión
Antebrazo

Cargar imagen de la lesión

Diagnosticar

Créditos

¿Cómo usar la herramienta?

Enviar a impresora

Cerrar sesión

Diagnóstico

akiec	0.0000
bcc	0.0000
bkl	0.0000
df	0.0000
nv	0.0080
vasc	0.9920
mel	0.0000

Lesión predicha: vasc-pyogenic gra

Figura 15. Interfaz gráfica desarrollada con funcionalidad.

8. DESEMPEÑO DEL SISTEMA DE CLASIFICACIÓN EN DISPOSITIVO DE BAJO CONSUMO

Para la validación del sistema de clasificación de cáncer de piel se emplea un dispositivo de bajo consumo. La elección del dispositivo se hace mediante la técnica de matriz de priorización donde se consideran parámetros de selección y se les asigna un peso específico, de modo que se le asigna el valor de 10 si el parámetro mucho más importante que el comparado, 5 si es más importante, 1 si tienen la misma importancia, 1/5 si es menos importante y 1/10 si es mucho menos importante. Los valores obtenidos se suman obteniendo el peso específico del parámetro en cuestión, tal y como se muestra en la tabla 13, en donde le asigna la denotación “A” al parámetro de velocidad de procesamiento, “B” al parámetro de disponibilidad de interfaces de entrada y salida, “C” al consumo de energía, “D” al parámetro de capacidad de disipación de calor, y “E” al parámetro de precio y disponibilidad en el mercado. Se observa que el consumo de energía (Parámetro C) y la velocidad de procesamiento (Parámetro A) son los más relevantes en la selección del dispositivo embebido para ejecutar el sistema de clasificación de cáncer de piel. Estos parámetros, con factores normalizados de 0.3508 cada uno, indican su importancia preeminente en comparación con otros criterios evaluados. Por otro lado, la disponibilidad de interfaces de entrada/salida (Parámetro B) y la capacidad de disipación de calor (Parámetro D) muestran factores normalizados de 0.1403, lo que sugiere su importancia relativa un poco inferior en la decisión final. El parámetro de precio y disponibilidad en el mercado (Parámetro E), es el menos significativo en la elección con un factor de 0.0178.

Tabla 13. Peso de los parámetros para la selección del dispositivo de bajo consumo.

	A	B	C	D	E	Suma	Factor
A	X	5	1	5	5	16	0,3508
B	1/5	X	1/5	1	5	6.4	0,1403
C	1	5	X	5	5	16	0,3508
D	1/5	1	1/5	X	5	6.4	0,1403
E	1/5	1/5	1/5	1/5	X	0.8	0,0178
Total						45,6	1

Asimismo, en la tabla 14 se comparan las posibles herramientas de hardware respecto al parámetro “A”, referente a la velocidad de procesamiento. Las herramientas a comparar son las placas Raspberry Pi 4B, BeagleBone Black, NVIDIA Jetson Nano, y Odroid M1s. La NVIDIA Jetson Nano destaca como la opción más favorable al comparar la velocidad de procesamiento, con un factor de normalización de 0.5446, seguida por la placa Raspberry Pi 4B con 0.3311, indicando un rendimiento y características superiores en comparación con los otros dispositivos comparados. Las placas BeagleBone Black y Odroid M1s muestran factores de normalización más bajos, lo que indica que son menos adecuados respecto a la velocidad de procesamiento.

Tabla 14. Comparación de las herramientas de hardware respecto a la velocidad de procesamiento.

	Raspberry Pi 4B	BeagleBone Black	NVIDIA Jetson Nano	Odroid M1s	Suma	Factor
Raspberry Pi 4B	X	5	1/5	10	15.2	0.3311
BeagleBone Black	1/5	X	1/10	5	5.3	0.1154
NVIDIA Jetson Nano	5	10	X	10	25	0,5446
Odroid M1s	1/10	1/5	1/10	X	0.4	0,0089
Total					45.9	1

En adición, en la tabla 15 se comparan las herramientas de hardware respecto al parámetro “B”, la disponibilidad de interfaces de entrada y salida. La Raspberry Pi 4B obtiene la puntuación más alta con un factor de normalización de 0.6079, lo que la posiciona como la mejor opción en términos de disponibilidad de E/S. La NVIDIA Jetson Nano y el Odroid M1s comparten el segundo lugar con un factor de normalización de 0.1885 cada uno, mientras que el BeagleBone Black queda rezagado con un factor de normalización de 0.0151, lo que indica una disponibilidad relativamente limitada de interfaces de E/S en comparación con las otras opciones.

Tabla 15. Comparación de las herramientas de hardware respecto a la disponibilidad de interfaces de entrada y salida.

	Raspberry Pi 4B	BeagleBone Black	NVIDIA Jetson Nano	Odroid M1s	Suma	Factor
Raspberry Pi 4B	X	10	5	5	20	0,6079
BeagleBone Black	1/10	X	1/5	1/5	0.5	0.0151
NVIDIA Jetson Nano	1/5	5	X	1	6.2	0.1885
Odroid M1s	1/5	5	1	X	6.2	0.1885
Total					32.9	

De igual forma, en la tabla 16 se comparan las herramientas de hardware respecto al parámetro “C”, el consumo de energía. Tanto la Raspberry Pi 4B y como la BeagleBone Black obtienen la misma puntuación más alta en términos de consumo de energía, con un factor de normalización de 0.4435 cada una. Esto sugiere que ambas placas son igualmente

eficientes en términos de consumo de energía. Por otro lado, tanto la NVIDIA Jetson Nano como el Odroid M1s tienen un factor de normalización de 0.0565, lo que indica que aproximadamente son siete veces menos eficientes en cuanto al consumo energía que la Raspberry Pi 4B y la BeagleBone Black.

Tabla 16. Comparación de las herramientas de hardware respecto al consumo de energía.

	Raspberry Pi 4B	BeagleBone Black	NVIDIA Jetson Nano	Odroid M1s	Suma	Factor
Raspberry Pi 4B	X	1	5	5	11	0.4435
BeagleBone Black	1	X	5	5	11	0.4435
NVIDIA Jetson Nano	1/5	1/5	X	1	1.4	0.0565
Odroid M1s	1/5	1/5	1	X	1.4	0.0565
Total					24.8	1

En complemento, en la tabla 17 se muestra la comparación respecto al parámetro “D”, la de capacidad de disipación de calor. La NVIDIA Jetson Nano sobresale con un factor de normalización de 0.6613, indicando su capacidad superior para gestionar el calor generado durante el funcionamiento. La Raspberry Pi 4B sugiere una capacidad moderada para disipar calor con un factor de normalización de 0.2698. Finalmente, las placas BeagleBone Black y Odroid M1s son las de menor rendimiento en cuanto a la disipación de calor con factores normalizados muy inferiores respecto a las demás placas.

Tabla 17. Comparación de las herramientas de hardware respecto a la capacidad de disipación de calor

	Raspberry Pi 4B	BeagleBone Black	NVIDIA Jetson Nano	Odroid M1s	Suma	Factor
Raspberry Pi 4B	X	5	1/5	5	10.2	0.2698
BeagleBone Black	1/5	X	1/10	1	1.3	0.0343
NVIDIA Jetson Nano	5	10	X	10	25	0.6613
Odroid M1s	1/5	1	1/10	X	1.3	0.0343
Total					37.8	1

En cuanto al parámetro “E”, el precio y la disponibilidad en el mercado, en la tabla 18 se muestra dicha comparación. La placa Raspberry Pi 4B destaca con un factor de normalización de 0.6613, indicando una notable mejora en cuanto al precio y disponibilidad. Las placas BeagleBone Black y Odroid M1s comparten el segundo lugar con un factor de normalización de 0.1884, mostrando ser un poco más costosas y menos comunes en el mercado. Por su parte, la placa NVIDIA Jetson Nano presenta un precio muy elevado respecto a sus similares, lo que la ubica en la última posición con un factor normalizado de 0.0149.

Tabla 18. Comparación de las herramientas de hardware respecto al precio y la disponibilidad en el mercado.

	Raspberry Pi 4B	BeagleBone Black	NVIDIA Jetson Nano	Odroid M1s	Suma	Factor
Raspberry Pi 4B	X	5	10	5	20	0.6079
BeagleBone Black	1/5	X	5	1	6.2	0.1884
NVIDIA Jetson Nano	1/10	1/5	X	1/5	0.5	0.0149
Odroid M1s	1/5	1	5	X	6.2	0.1884
Total					32.9	1

Finalmente, en la tabla 19 se muestra la matriz de selección de la herramienta de hardware de bajo consumo para ejecutar el sistema de clasificación de cáncer de piel. Para calcular el valor total para la elección se multiplica el valor del factor para cada herramienta en cada uno de los parámetros, por el valor del peso de cada parámetro, obtenido en la tabla 14.

Tabla 19. Selección de herramienta de hardware de bajo consumo.

	A	B	C	D	E	Suma
Raspberry Pi 4B	0.3311 x 0,3508	0,6079 x 0,1403	0.4435 x 0,3508	0.2698 x 0,1403	0.6079 x 0,0178	0.4054
BeagleBone Black	0.1154 x 0,3508	0.0151 x 0,1403	0.4435 x 0,3508	0.0343 x 0,1403	0.1884 x 0,0178	0.2062
NVIDIA Jetson Nano	0,5446 x 0,3508	0.1885 x 0,1403	0.0565 x 0,3508	0.6613 x 0,1403	0.0149 x 0,0178	0.3329
Odroid M1s	0,0089 x 0,3508	0.1885 x 0,1403	0.0565 x 0,3508	0.0343 x 0,1403	0.1884 x 0,0178	0.0555

Como se observa, la placa Raspberry Pi 4B se muestra como la opción más favorable con un puntaje total de 0.4054, lo que indica su destacado desempeño en relación con la velocidad

de procesamiento, disponibilidad de interfaces, capacidad de disipación de calor, precio y disponibilidad en el mercado. Le sigue la NVIDIA Jetson Nano con un puntaje de 0.3329, demostrando su solidez en varias áreas clave. La BeagleBone Black obtiene un puntaje de 0.2062, mientras que el Odroid M1s registra un puntaje mucho más bajo de 0.0555. Estos resultados sugieren que la Raspberry Pi 4B es la opción más equilibrada y versátil entre las cuatro placas.

Una vez definida la placa, se procede a realizar la validación del desempeño del sistema de diagnóstico se hace mediante pruebas de ejecución. El dispositivo cuenta con un procesador Broadcom BCM2711 de cuatro núcleos ARM Cortex-A72 a 1.5 GHz y opciones de memoria RAM de 4 GB LPDDR4 SDRAM. Ofrece conectividad versátil con dos puertos micro HDMI capaces de 4K a 60Hz, cuatro puertos USB (dos USB 3.0 y dos USB 2.0), un puerto Gigabit Ethernet y soporte para Wi-Fi 802.11ac de doble banda y Bluetooth 5.0. Además, incluye un conector GPIO de 40 pines para la conexión de periféricos y componentes electrónicos. La arquitectura del sistema de clasificación y su implementación en el sistema embebido se muestra en la figura 16, en donde se observa que, para la funcionalidad, a la placa Raspberry Pi 4B se conecta un teclado y un mouse por USB, la conexión a internet se hace vía Wi-fi y la visualización se hace por conexión de un monitor al puerto micro-HDMI. Para acceder se ingresan inicialmente sus credenciales, donde en el caso 1 se expone la situación con credenciales incorrectas, donde estas se recuperan mediante un mensaje de texto gestionado por Twilio, y el caso 2, donde al acceder con credenciales correctas, se carga la imagen de la lesión, se ingresa la información del paciente, así como los factores de riesgo, y se diagnostica la lesión según el modelo .h5 desarrollado.

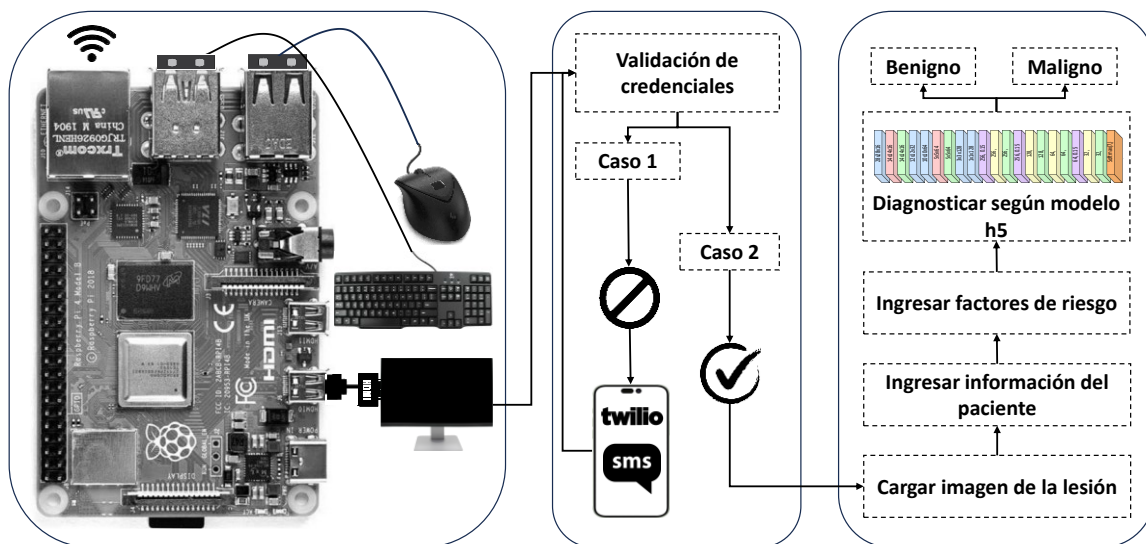


Figura 16. Implementación en sistema embebido.

En la figura 17 se observa el consolidado de tiempos de ejecución del sistema de clasificación sobre la placa Raspberry Pi 3B+, donde se enfatiza en el tiempo empleado en el cargue del modelo desarrollado, el almacenamiento de la información y el diagnóstico. El

tiempo de carga del modelo convolucional presenta una distribución que oscila entre 4 y 9 segundos, con una media de 6.35 y una mediana de 6.5 segundos. Esto indica que, en promedio, la carga del modelo tiende a tomar alrededor de 6 o 7 segundos con un valor de desviación estándar de aproximadamente 1.867. El rango intercuartílico sugiere que el 50% central de los datos está contenido en un rango de 2.75 segundos. En cuanto al almacenamiento de información, los datos varían entre 2 y 5 segundos, con una media de 3.6 y una mediana de 4 segundos. Esto indica que, en general, la cantidad de información tarda en almacenarse alrededor de 3 o 4 segundos, con una desviación estándar de aproximadamente 1.187, lo que sugiere una dispersión moderada de los datos. El rango intercuartílico es de 2 segundos. El tiempo de diagnóstico muestra una distribución similar a la del almacenamiento de información, con valores que van de 2 a 5 segundos. La media y la mediana son iguales, con un valor de 3.6 segundos. La desviación estándar de aproximadamente 1.059 indica una dispersión relativamente baja de los datos en comparación con los otros aspectos. El rango intercuartílico es de 2 segundos, al igual que en el almacenamiento de información. Finalmente, el tiempo total empleado en cada prueba varía de 9 a 18 segundos, con una media de 13.6 segundos y una mediana de 13 segundos. La desviación estándar de aproximadamente 2.732 indica una variabilidad moderada en los tiempos totales. El rango intercuartílico es de 3.75 segundos, lo que sugiere que la mitad central de los tiempos totales se encuentra dentro de este rango.

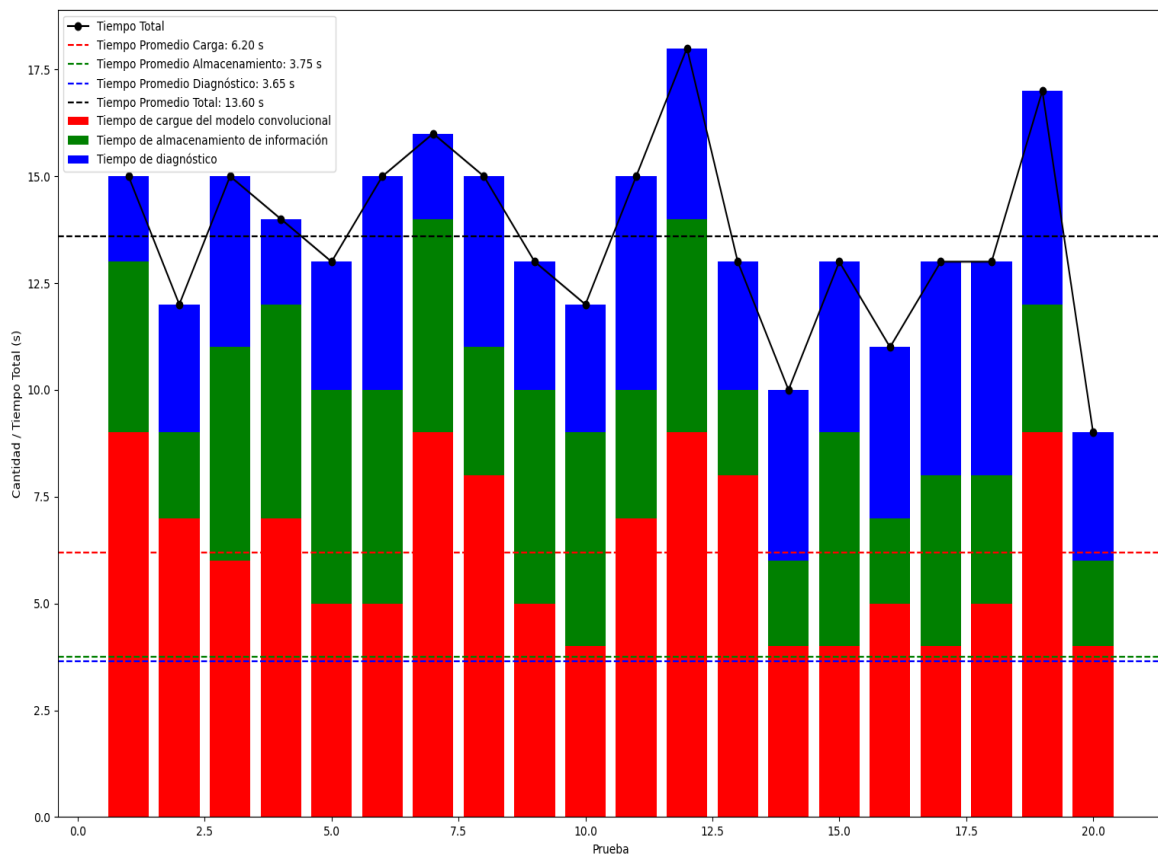


Figura 17. Desempeño del sistema de clasificación respecto al tiempo de respuesta.

En adición, en la figura 18 se muestra el desempeño de la herramienta de clasificación desarrollado en el sistema embebido considerando el nivel de CPU y el nivel de RAM empleado según las funcionalidades de la interfaz gráfica de usuario. el porcentaje de CPU, se observa que los valores oscilan entre 22% y 42% en las diferentes pruebas realizadas. El promedio de uso de CPU es aproximadamente del 32.65%, con una desviación estándar de alrededor del 5.74%. Esto indica que, en promedio, el sistema tiende a utilizar alrededor de un tercio de su capacidad de procesamiento, pero con cierta variabilidad en los niveles de carga observados durante las pruebas. Además, la distribución de los datos parece ser relativamente simétrica, con una ligera tendencia hacia valores más altos de uso de CPU. Asimismo, en cuanto al uso de RAM, los valores de uso oscilan entre 34% y 56%. El promedio de uso de RAM es aproximadamente del 46.70%, con una desviación estándar de alrededor del 6.01%. Esto sugiere que, en promedio, cerca de la mitad de la memoria RAM disponible se utiliza durante las pruebas, nuevamente con cierta variabilidad en los niveles de uso observados. Al igual que con la CPU, la distribución de los datos parece ser relativamente simétrica, aunque con una tendencia hacia valores más altos de uso de RAM. Es importante monitorear estos niveles de uso de recursos para garantizar un rendimiento óptimo del sistema y evitar posibles cuellos de botella o problemas de rendimiento debido a la falta de recursos disponibles.

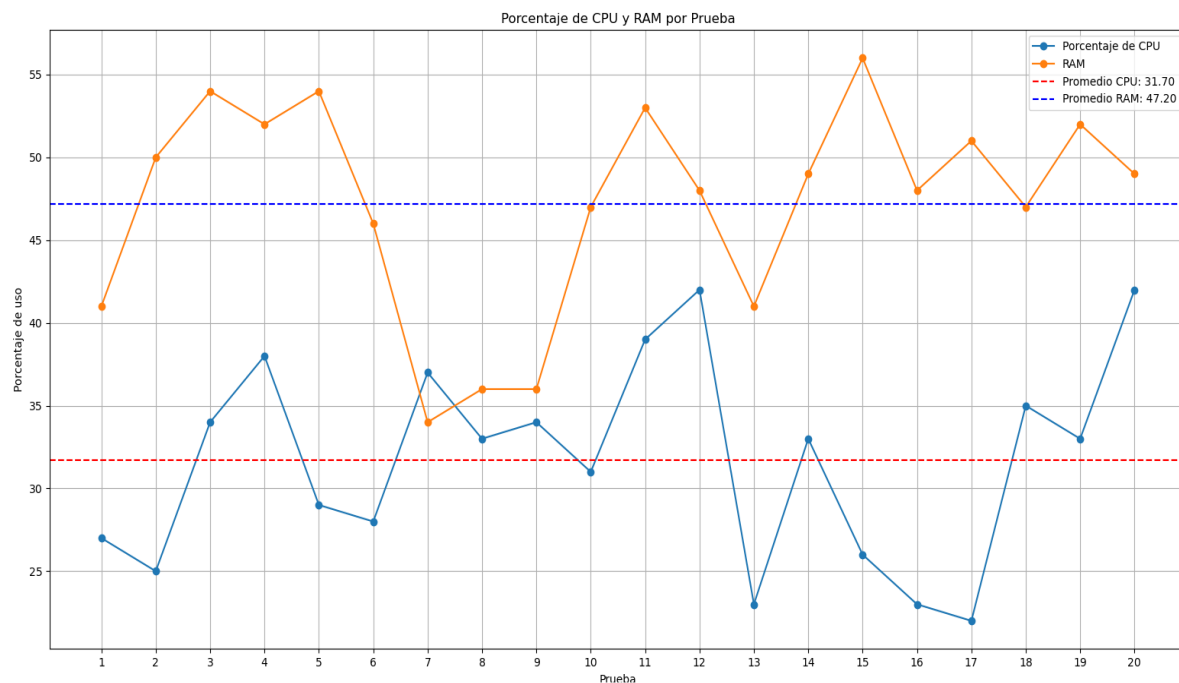


Figura 18. Desempeño del sistema de clasificación en cuanto a niveles de CPU y RAM requerido.

El análisis de desempeño del sistema de clasificación respecto a la temperatura del procesador del sistema embebido al ejecutar la herramienta se muestra en la figura 19. En tonalidad verde se muestra la temperatura inicial del procesador con valores fluctuantes

entre 0° y 12° con valor promedio de 6.4°. En estas condiciones solo se considera el encendido del dispositivo sin funcionalidad de la interfaz. En adición, el cargue del modelo computacional eleva el nivel de temperatura, obteniendo valores mínimos de 6.3° y valores máximos de 27°, donde en este instante, el nivel medio de temperatura en el sistema embebido aumentó a 24°, representado en color azul. En un tercer instante, se pondera el almacenamiento de información como siguiente tarea del sistema de clasificación, en donde ilustrado en tonalidad violeta, donde el nivel de temperatura alcanzó en algunas pruebas los 40°, mientras que, en la etapa de diagnóstico se alcanzaron temperaturas de hasta 55°.

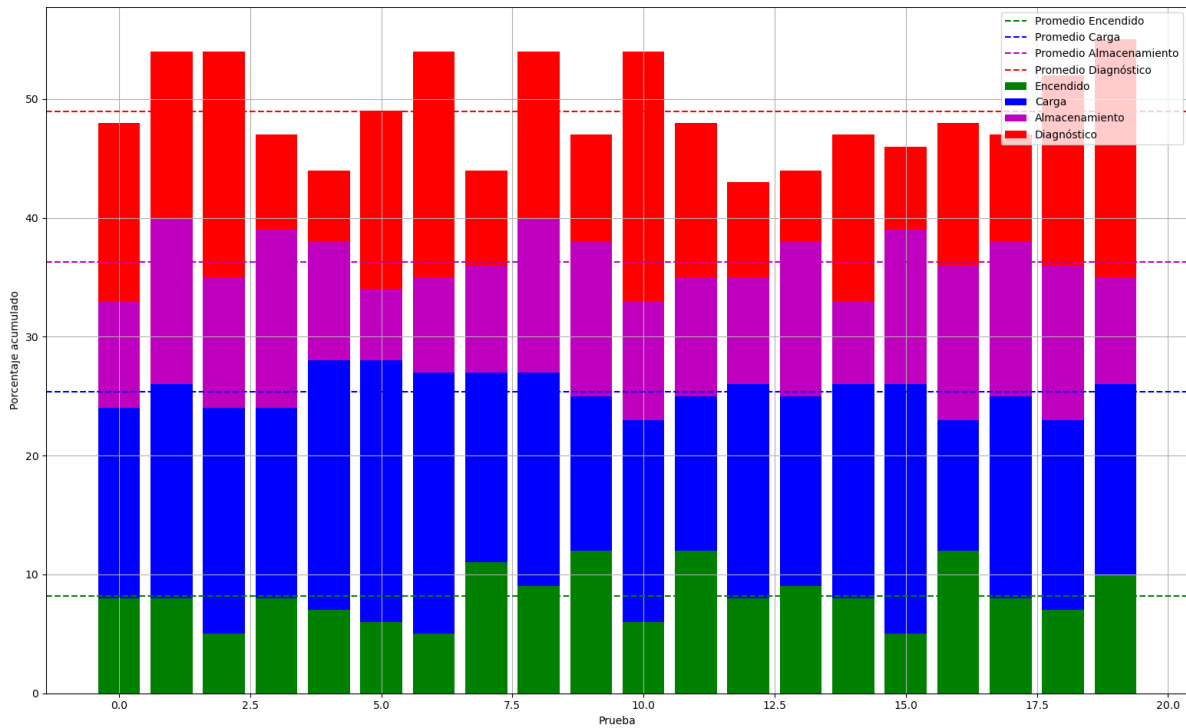


Figura 19. Desempeño del sistema de clasificación en cuanto a la temperatura del procesador del sistema embebido.

9. CONCLUSIONES

El análisis realizado al conjunto de datos HAM10000 revela su valor para la clasificación del cáncer de piel, junto con una significativa disparidad entre las clases benignas y malignas. Para abordar esta disparidad y mejorar la capacidad del modelo para generalizar, se opta por el sobremuestreo mediante la inducción de ruido a los descriptores de las imágenes, buscando introducir pequeñas fluctuaciones que no comprometan la información central de las clases originales. Este proceso de aumento de datos se valida mediante la comparación estructural entre los descriptores de las imágenes originales y aumentadas, utilizando análisis de componentes principales (PCA) y curvas de estimación de densidad de núcleo (KDE). Los resultados muestran que el aumento de datos mantiene la estructura subyacente de la distribución, sin distorsionar significativamente las características fundamentales. Además, la adaptación del conjunto de datos a un formato más manejable, como HMNIST, y el redimensionamiento de las imágenes facilitan la exploración eficiente de las características y la implementación en dispositivos de bajo consumo, sin comprometer la calidad del diagnóstico dermatológico.

Se aplicó una estrategia de sobremuestreo por inducción de ruido para equilibrar el conjunto de datos posterior a la división en conjuntos de entrenamiento y prueba. Esta estrategia es crucial para garantizar que el modelo no esté sesgado hacia clases dominantes y que pueda generalizar bien en datos no vistos. Asimismo, se desarrolló una arquitectura de red neuronal convolucional específicamente adaptada para la tarea de clasificación de imágenes de lesiones cutáneas. Esta arquitectura incluye capas convolucionales para la extracción de características espaciales y capas densas para la clasificación final, con técnicas como normalización por lotes y Dropout para mejorar la generalización del modelo. El modelo se entrenó durante 50 épocas y se observó una mejora gradual tanto en la pérdida como en la precisión en los conjuntos de entrenamiento y prueba. Esto indica que el modelo aprendió eficazmente los patrones complejos en los datos y generalizó bien en datos no vistos. El modelo logró una alta precisión tanto en el conjunto de entrenamiento como en el de prueba, lo que sugiere que es capaz de clasificar con precisión diferentes tipos de lesiones cutáneas. Además, se mantuvo una estabilidad en el tiempo de entrenamiento y los ms/paso a lo largo de las épocas, lo que indica una convergencia efectiva del modelo. A nivel de recall y exactitud, el modelo desarrollado supera a técnicas como EW-FCM + ShuffleNet en aproximadamente un 27.7% en ambas métricas, lo que destaca su capacidad para identificar correctamente los verdaderos positivos y minimizar los falsos positivos. Además, en términos de Precision, supera a varios modelos, como Xception modificada: SBXception, y Mask RCNN+DenseNet201+ LS-SVM, en aproximadamente un 12.3%, y 8.1% respectivamente, lo que sugiere su capacidad para proporcionar resultados precisos. Estos resultados respaldan la eficacia del modelo de 22 capas construido para la clasificación del cáncer de piel, mostrando un rendimiento competitivo en términos de estas métricas clave en comparación con una variedad de modelos alternativos.

El modelo desarrollado muestra mejoras significativas en comparación con Random Forest en términos de Recall (media de 1.17%), Precision (media de 5.75%), F1-Score (media de 3.48%) y Exactitud (media 6.69%), así como una eficiencia notable en tiempo de entrenamiento, utilizando aproximadamente 5 veces menos tiempo (46 minutos frente a 225 minutos). Además, el modelo requiere menos recursos computacionales, utilizando un 34% menos de CPU (53% frente a 87%) y un 37% menos de RAM (31% frente a 49%). Aunque respecto a XGBoost las mejoras son menos pronunciadas a nivel software, sigue siendo una opción competitiva, destacando especialmente por su eficiencia en recursos computacionales en comparación con XGBoost.

La implementación del modelo de clasificación de cáncer de piel se logra mediante la creación de una herramienta de diagnóstico asistido con una interfaz gráfica centrada en el usuario. Esta herramienta, diseñada utilizando QtDesigner y Python, ofrece una solución práctica y eficiente para el diagnóstico de lesiones cutáneas, manteniendo la confidencialidad de la información del paciente y permitiendo un registro detallado de las acciones realizadas por los usuarios. La interfaz de inicio de sesión asegura que solo los usuarios autorizados, como médicos especialistas, puedan acceder al sistema, mitigando así el riesgo de accesos no autorizados y protegiendo la integridad de los datos sensibles. Además, la compatibilidad multiplataforma de la herramienta garantiza su accesibilidad en diversos entornos operativos, lo que amplía su utilidad y aplicabilidad. El diseño de la interfaz gráfica se lleva a cabo con un enfoque cuidadoso en la usabilidad y la funcionalidad. Se incluyen elementos como botones, etiquetas y campos de entrada para facilitar la interacción del usuario y proporcionar una experiencia intuitiva. La utilización de librerías como PyQt5, TensorFlow y OpenCV permite integrar funcionalidades avanzadas, como la carga de modelos de aprendizaje profundo previamente entrenados y el procesamiento de imágenes de lesiones cutáneas. Además, se emplean técnicas como la manipulación de imágenes con Pillow y el análisis de datos con Pandas para optimizar el flujo de trabajo y mejorar la precisión del diagnóstico. El desarrollo de la interfaz se complementa con un conjunto de funcionalidades clave, como la captura y guardado de la interfaz como un archivo PDF, la carga de imágenes de lesiones, la visualización de resultados de diagnóstico y la gestión de información del paciente.

A través de la aplicación de técnicas de selección y priorización, se identificó la placa Raspberry Pi 4B como la opción más favorable para ejecutar el sistema, considerando parámetros como velocidad de procesamiento, disponibilidad de interfaces de entrada/salida, consumo de energía, capacidad de disipación de calor, precio y disponibilidad en el mercado. La evaluación detallada de cada parámetro y la comparación entre diferentes herramientas de hardware arroja resultados claros sobre las fortalezas y debilidades de cada opción. La Raspberry Pi 4B destaca en varios aspectos, incluida la velocidad de procesamiento, la disponibilidad de interfaces y el precio asequible, lo que la convierte en la opción más equilibrada y versátil para el sistema de clasificación. Además, el análisis del desempeño del sistema en la Raspberry Pi 4B reveló tiempos de ejecución

promedio, donde el tiempo total empleado en cada prueba varió de 9 a 18 segundos, con una media de 13.6 segundos. En términos de uso de recursos, el promedio de uso de CPU fue de aproximadamente 32.65%, mientras que el promedio de uso de RAM fue de alrededor del 46.70%. Por último, el monitoreo de la temperatura del procesador durante la ejecución del sistema mostró un aumento significativo en la temperatura, alcanzando valores de hasta 55°C durante la etapa de diagnóstico. Estos valores proporcionan una base sólida para el desarrollo y la implementación exitosa de soluciones de diagnóstico asistido que pueden tener un impacto significativo en el campo médico.

10. TRABAJOS FUTUROS

Como trabajos futuros se proponen las siguientes líneas de investigación y/o desarrollo:

- Uso de técnicas de compresión de modelos de aprendizaje automático específicamente diseñadas para sistemas embebidos como la Raspberry Pi 4B, incluyendo la cuantización de parámetros, o el uso de arquitecturas de modelos eficientes para reducir el tamaño del modelo y la carga computacional necesaria para la inferencia.
- Implementación de técnicas de aceleración de hardware en la Raspberry Pi 4B. Esto implicaría explorar la integración de unidades de procesamiento gráfico (GPU) o de aprendizaje automático especializadas para mejorar la velocidad y eficiencia del sistema de clasificación, requiriendo adaptaciones específicas del software para aprovechar al máximo el hardware disponible.
- Desarrollo de algoritmos de procesamiento de imágenes optimizados para la detección y clasificación de lesiones cutáneas en imágenes dermatoscópicas. Estos algoritmos podrían incluir técnicas de segmentación de imágenes y extracción de características específicas de lesiones.
- Uso de técnicas de transferencia de conocimiento, como el aprendizaje por transferencia o el ajuste fino de modelos pre-entrenados, para adaptar modelos de clasificación de imágenes desarrollados en otros dominios al diagnóstico de cáncer de piel en la Raspberry Pi 4B.
- Desarrollar un sistema de clasificación de cáncer de piel en tiempo real que pueda analizar y diagnosticar lesiones dermatoscópicas mientras se capturan las imágenes. Esto requeriría la optimización de algoritmos y la implementación de técnicas de procesamiento de imágenes en tiempo real en la Raspberry Pi 4B, con posibles aplicaciones prácticas en la detección temprana y el monitoreo continuo de lesiones cutáneas.

11. REFERENCIAS BIBLIOGRÁFICAS

- [1] S. Sajadimajd *et al.*, “Advances on Natural Polyphenols as Anticancer Agents for Skin Cancer,” *Pharmacol Res*, vol. 151, p. 104584, 2020, doi: 10.1016/j.phrs.2019.104584.
- [2] A. Khamparia, P. K. Singh, P. Rani, D. Samanta, A. Khanna, and B. Bhushan, “An internet of health things-driven deep learning framework for detection and classification of skin cancer using transfer learning,” *Transactions on Emerging Telecommunications Technologies*, vol. 32, no. 7, Jul. 2021, doi: 10.1002/ett.3963.
- [3] D. Bisla, A. Choromanska, R. S. Berman, J. A. Stein, and D. Polsky, “Towards automated melanoma detection with deep learning: Data purification and augmentation,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, IEEE Computer Society, Jun. 2019, pp. 2720–2728. Doi: 10.1109/CVPRW.2019.00330.
- [4] N. Razmjoooy *et al.*, “Computer-aided Diagnosis of Skin Cancer: A Review,” *Current Medical Imaging Formerly Current Medical Imaging Reviews*, vol. 16, no. 7, pp. 781–793, Jan. 2020, doi: 10.2174/1573405616666200129095242.
- [5] Z. Xu, F. R. Sheykhahmad, N. Ghadimi, and N. Razmjoooy, “Computer-aided diagnosis of skin cancer based on soft computing techniques,” *Open Medicine (Poland)*, vol. 15, no. 1, pp. 860–871, 2020, doi: 10.1515/med-2020-0131.
- [6] Ş. Öztürk and U. Özkaya, “Skin Lesion Segmentation with Improved Convolutional Neural Network,” *J Digit Imaging*, vol. 33, no. 4, pp. 958–970, May 2020, doi: 10.1007/S10278-020-00343-Z.
- [7] P. Tschandl *et al.*, “Expert-Level Diagnosis of Nonpigmented Skin Cancer by Combined Convolutional Neural Networks,” *JAMA Dermatol*, vol. 155, no. 1, pp. 58–65, 2019, doi: 10.1001/jamadermatol.2018.4378.
- [8] M. Arshad *et al.*, “A Computer-Aided Diagnosis System Using Deep Learning for Multiclass Skin Lesion Classification,” *Comput Intell Neurosci*, vol. 2021, 2021, doi: 10.1155/2021/9619079.
- [9] S. Jain, V. Jagtap, and N. Pise, “Computer aided melanoma skin cancer detection using image processing,” *Procedia Comput Sci*, vol. 48, no. C, pp. 735–740, 2015, doi: 10.1016/j.procs.2015.04.209.
- [10] A. S. Jahn *et al.*, “Over-Detection of Melanoma-Suspect Lesions by a CE-Certified Smartphone App: Performance in Comparison to Dermatologists, 2D and 3D Convolutional Neural Networks in a Prospective Data Set of 1204 Pigmented Skin Lesions Involving Patients’ Perception,” *Cancers (Basel)*, vol. 14, no. 15, Aug. 2022, doi: 10.3390/cancers14153829.

- [11] P. Bibiloni, M. González Hidalgo, and S. Massanet, "Skin Hair Removal in Dermoscopic Images Using Soft Color Morphology," in *Conference on Artificial Intelligence in Medicine in Europe*, A. ten Teije, C. Popow, J. H. Holmes, and L. Sacchi, Eds., in *Lecture Notes in Computer Science*, vol. 10259. Cham: Springer International Publishing, 2017. Doi: 10.1007/978-3-319-59758-4.
- [12] J. Ramya, H. C. Vijaylakshmi, and H. Mirza Saifuddin, "Segmentation of skin lesion images using discrete wavelet transform," *Biomed Signal Process Control*, vol. 69, Aug. 2021, doi: 10.1016/j.bspc.2021.102839.
- [13] G. Carleo *et al.*, "Machine learning and the physical sciences," *Rev Mod Phys*, vol. 91, no. 4, p. 45002, 2019, doi: 10.1103/RevModPhys.91.045002.
- [14] A. J. Larrazabal, N. Nieto, V. Peterson, D. H. Milone, and E. Ferrante, "Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis," *Proc Natl Acad Sci U S A*, vol. 117, no. 23, pp. 12592–12594, Jun. 2020, doi: 10.1073/pnas.1919012117.
- [15] L. Bi, J. Kim, E. Ahn, A. Kumar, M. Fulham, and D. Feng, "Dermoscopic Image Segmentation via Multistage Fully Convolutional Networks," *IEEE Trans Biomed Eng*, vol. 64, no. 9, pp. 2065–2074, Sep. 2017, doi: 10.1109/TBME.2017.2712771.
- [16] F. Mattioli Della Rocca *et al.*, "A 128 × 128 SPAD motion-triggered time-of-flight image sensor with in-pixel histogram and column-parallel vision processor," *IEEE J Solid-State Circuits*, vol. 55, no. 7, pp. 1762–1775, 2020, doi: 10.1109/JSSC.2020.2993722.
- [17] D. Ramírez-Jiménez and J. D. Quintero-Ospina, "Clasificación de patologías presentes en la columna vertebral mediante técnicas de máquinas de aprendizaje," *Ingeniería Solidaria*, vol. 15, no. 27, pp. 1–23, 2019, doi: 10.16925/2357-6014.2019.01.05.
- [18] M. N. Bajwa *et al.*, "Computer-aided diagnosis of skin diseases using deep neural networks," *Applied Sciences (Switzerland)*, vol. 10, no. 7, pp. 1–13, 2020, doi: 10.3390/app10072488.
- [19] D. P. Kaucha, P. W. C. Prasad, A. Alsadoon, A. Elchouemi, and S. Sreedharan, "Early detection of lung cancer using SVM classifier in biomedical image processing," *IEEE International Conference on Power, Control, Signals and Instrumentation Engineering, ICPCSI 2017*, pp. 3143–3148, Jun. 2018, doi: 10.1109/ICPCSI.2017.8392305.
- [20] N. Hameed, A. Shabut, and M. A. Hossain, "A Computer-Aided diagnosis system for classifying prominent skin lesions using machine learning," in *2018 10th Computer Science and Electronic Engineering Conference, CEEC 2018 – Proceedings*, IEEE, 2019, pp. 186–191. Doi: 10.1109/CEEC.2018.8674183.

- [21] J. Jaworek-Korjakowska, "Computer-aided diagnosis of micro-malignant melanoma lesions applying support vector machines," *Biomed Res Int*, vol. 2016, no. 1, pp. 1–9, 2016, doi: 10.1155/2016/4381972.
- [22] M. Alenezi, M. Akour, A. Hussien, and M. Z. Al-Saad, "Test suite effectiveness: An indicator for open source software quality," in *2016 2nd International Conference on Open Source Software Computing, OSSCOM, 2017*. Doi: 10.1109/OSSCOM.2016.7863677.
- [23] T. Lu, Y. Guan, Y. Zhang, S. Qu, and Z. Xiong, "Robust and efficient face recognition via low-rank supported extreme learning machine," *Multimed Tools Appl*, vol. 77, no. 9, pp. 11219–11240, 2018, doi: 10.1007/s11042-017-5475-2.
- [24] S. Wang and M. Hamian, "Skin Cancer Detection Based on Extreme Learning Machine and a Developed Version of Thermal Exchange Optimization," *Comput Intell Neurosci*, vol. 2021, 2021, doi: 10.1155/2021/9528664.
- [25] E. Shoen, "DermlA: Machine Learning to Improve Skin Cancer Screening," *Journal of Digital Imaging*, vol. 34, no. 6. Springer Science and Business Media Deutschland GmbH, pp. 1430–1434, Dec. 01, 2021. Doi: 10.1007/s10278-020-00395-1.
- [26] W. Gouda, N. U. Sama, G. Al-Waakid, M. Humayun, and N. Z. Jhanjhi, "Detection of Skin Cancer Based on Skin Lesion Images Using Deep Learning," *Healthcare (Switzerland)*, vol. 10, no. 7, Jul. 2022, doi: 10.3390/healthcare10071183.
- [27] M. Nawaz *et al.*, "Skin cancer detection from dermoscopic images using deep learning and fuzzy k-means clustering," *Microsc Res Tech*, vol. 85, no. 1, pp. 339–351, Jan. 2022, doi: 10.1002/jemt.23908.
- [28] M. Fraiwan and E. Faouri, "On the Automatic Detection and Classification of Skin Cancer Using Deep Transfer Learning," *Sensors*, vol. 22, no. 13, Jul. 2022, doi: 10.3390/s22134963.
- [29] T. Guergueb and M. A. Akhloufi, "Melanoma Skin Cancer Detection Using Recent Deep Learning Models," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Institute of Electrical and Electronics Engineers Inc., 2021*, pp. 3074–3077. Doi: 10.1109/EMBC46164.2021.9631047.
- [30] B. Shetty, R. Fernandes, A. P. Rodrigues, R. Chengoden, S. Bhattacharya, and K. Lakshmana, "Skin lesion classification of dermoscopic images using machine learning and convolutional neural network," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-22644-9.

- [31] 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," *Applied Sciences (Switzerland)*, vol. 12, no. 5, Mar. 2022, doi: 10.3390/app12052677.
- [32] M. A. Khan, Y. D. Zhang, M. Sharif, and T. Akram, "Pixels to Classes: Intelligent Learning Framework for Multiclass Skin Lesion Localization and Classification," *Computers and Electrical Engineering*, vol. 90, Mar. 2021, doi: 10.1016/j.compeleceng.2020.106956.
- [33] A. Mehmood, Y. Gulzar, Q. M. Ilyas, A. Jabbari, M. Ahmad, and S. Iqbal, "SBXception: A Shallower and Broader Xception Architecture for Efficient Classification of Skin Lesions," *Cancers (Basel)*, vol. 15, no. 14, Jul. 2023, doi: 10.3390/cancers15143604.
- [34] M. Zafar, J. Amin, M. Sharif, M. A. Anjum, G. A. Mallah, and S. Kadry, "DeepLabv3+-Based Segmentation and Best Features Selection Using Slime Mould Algorithm for Multi-Class Skin Lesion Classification," *Mathematics*, vol. 11, no. 2, Jan. 2023, doi: 10.3390/math11020364.
- [35] M. Saarela and L. Geogieva, "Robustness, Stability, and Fidelity of Explanations for a Deep Skin Cancer Classification Model," *Applied Sciences (Switzerland)*, vol. 12, no. 19, Oct. 2022, doi: 10.3390/app12199545.
- [36] M. A. Khan, T. Akram, Y. D. Zhang, and M. Sharif, "Attributes based skin lesion detection and recognition: A mask RCNN and transfer learning-based deep learning framework," *Pattern Recognit Lett*, vol. 143, pp. 58–66, Mar. 2021, doi: 10.1016/j.patrec.2020.12.015.
- [37] F. Afza, M. Sharif, M. Mittal, M. A. Khan, and D. Jude Hemanth, "A hierarchical three-step superpixels and deep learning framework for skin lesion classification," *Methods*, vol. 202, pp. 88–102, Jun. 2022, doi: 10.1016/j.ymeth.2021.02.013.
- [38] B. Shetty, R. Fernandes, A. P. Rodrigues, R. Chengoden, S. Bhattacharya, and K. Lakshmana, "Skin lesion classification of dermoscopic images using machine learning and convolutional neural network," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-22644-9.
- [39] A. Bibi *et al.*, "Skin lesion segmentation and classification using conventional and deep learning based framework," *Computers, Materials and Continua*, vol. 71, no. 2, pp. 2477–2495, 2022, doi: 10.32604/cmc.2022.018917.
- [40] F. Afza, M. Sharif, M. A. Khan, U. Tariq, H. S. Yong, and J. Cha, "Multiclass Skin Lesion Classification Using Hybrid Deep Features Selection and Extreme Learning Machine," *Sensors*, vol. 22, no. 3, Feb. 2022, doi: 10.3390/s22030799.
- [41] M. Tahir, A. Naeem, H. Malik, J. Tanveer, R. A. Naqvi, and S. W. Lee, "DSCC_Net: Multi-Classification Deep Learning Models for Diagnosing of Skin Cancer Using Dermoscopic Images," *Cancers (Basel)*, vol. 15, no. 7, Apr. 2023, doi: 10.3390/cancers15072179.

- [42] S. Jain, U. Singhania, B. Tripathy, E. A. Nasr, M. K. Aboudaif, and A. K. Kamrani, “Deep learning-based transfer learning for classification of skin cancer,” *Sensors*, vol. 21, no. 23, Dec. 2021, doi: 10.3390/s21238142.
- [43] K. Mridha, M. M. Uddin, J. Shin, S. Khadka, and M. F. Mridha, “An Interpretable Skin Cancer Classification Using Optimized Convolutional Neural Network for a Smart Healthcare System,” *IEEE Access*, vol. 11, pp. 41003–41018, 2023, doi: 10.1109/ACCESS.2023.3269694.
- [44] S. Jiang, H. Li, and Z. Jin, “A Visually Interpretable Deep Learning Framework for Histopathological Image-Based Skin Cancer Diagnosis,” *IEEE J Biomed Health Inform*, vol. 25, no. 5, pp. 1483–1494, May 2021, doi: 10.1109/JBHI.2021.3052044.
- [45] M. Saeed, A. Naseer, H. Masood, S. ur Rehman, and V. Gruhn, “The Power of Generative AI to Augment for Enhanced Skin Cancer Classification: A Deep Learning Approach,” *IEEE Access*, pp. 1–1, Nov. 2023, doi: 10.1109/access.2023.3332628.
- [46] M. A. Albahar, “Skin Lesion Classification Using Convolutional Neural Network with Novel Regularizer,” *IEEE Access*, vol. 7, pp. 38306–38313, 2019, doi: 10.1109/ACCESS.2019.2906241.
- [47] X. Zhou *et al.*, “CUSS-Net: A Cascaded Unsupervised-Based Strategy and Supervised Network for Biomedical Image Diagnosis and Segmentation,” *IEEE J Biomed Health Inform*, vol. 27, no. 5, pp. 2444–2455, May 2023, doi: 10.1109/JBHI.2023.3238726.
- [48] A. Imran, A. Nasir, M. Bilal, G. Sun, A. Alzahrani, and A. Almuhaimeed, “Skin Cancer Detection Using Combined Decision of Deep Learners,” *IEEE Access*, vol. 10, pp. 118198–118212, 2022, doi: 10.1109/ACCESS.2022.3220329.
- [49] T. C. Pham, A. Doucet, C. M. Luong, C. T. Tran, and V. D. Hoang, “Improving Skin-Disease Classification Based on Customized Loss Function Combined with Balanced Mini-Batch Logic and Real-Time Image Augmentation,” *IEEE Access*, vol. 8, pp. 150725–150737, 2020, doi: 10.1109/ACCESS.2020.3016653.
- [50] H. L. Gururaj, N. Manju, A. Nagarjun, V. N. Manjunath Aradhya, and F. Flammini, “DeepSkin: A Deep Learning Approach for Skin Cancer Classification,” *IEEE Access*, vol. 11, pp. 50205–50214, 2023, doi: 10.1109/ACCESS.2023.3274848.

ANEXOS

Anexo I. Manual de usuario de la herramienta de diagnóstico asistido por computador

Manual de Usuario: Asistente de Diagnóstico de Cáncer de Piel

Este software está diseñado para ayudar en la detección temprana de posibles lesiones cutáneas que podrían ser indicativas de cáncer de piel. A continuación, encontrará una guía detallada sobre cómo utilizar las diferentes funciones de esta aplicación.

Carlos Vicente Niño Rondón ¹

Sergio Alexander Castro Casadiego ¹

Manuel Guillermo Forero Vargas ²

¹ *Universidad Francisco de Paula Santander, Cúcuta, Colombia*

² *Universidad de Ibagué, Ibagué, Colombia*

COLOMBIA

2024

Inicio de sesión

La Interfaz de Inicio de Sesión es la primera parte de la aplicación diseñada para la detección de cáncer de piel. Esta interfaz permite a los usuarios acceder al sistema mediante la autenticación de sus credenciales.

- **Requisitos del sistema**

1. Dispositivo con acceso a Internet.
2. Instalación de Python y PyQt5.
3. Imágenes de lesiones cutáneas para el proceso de diagnóstico posterior.

- **Procedimiento de inicio de sesión**

1. Ejecute la aplicación desde la terminal utilizando el comando Python LOGIN.py. La interfaz de inicio de sesión se mostrará en la pantalla.
2. Ingrese su nombre de usuario y contraseña en los campos correspondientes. Haga clic en el botón "Ingresar" para verificar sus credenciales.
3. Automáticamente la aplicación verificará las credenciales ingresadas. Si las credenciales son válidas, se accederá al siguiente paso. Si las credenciales son incorrectas, se mostrará un mensaje de error.
4. Si desea recibir sus credenciales vía SMS, haga clic en el botón "Recordar". Se enviará un SMS con su nombre de usuario y contraseña asociados.

- **Consideraciones y uso**

Para obtener asistencia técnica o reportar problemas, comuníquese con el equipo de soporte técnico proporcionado en

la interfaz. La aplicación garantiza la seguridad y privacidad de tus credenciales. Para esto, no comparta esta información con terceros. No olvide cerrar sesión cuando haya terminado de usar la aplicación para proteger sus datos personales y los datos de los pacientes.

Asistente de Diagnóstico

Esta sección le servirá de guía para comprender las diversas funcionalidades de la aplicación para el diagnóstico preciso de lesiones cutáneas. Aprenderá cómo registrar los datos del paciente, cargar imágenes de las lesiones, seleccionar factores de riesgo, realizar diagnósticos y más. Además, encontrará instrucciones sobre cómo imprimir resultados y guardar información en un archivo CSV.

- **Registro de Datos del Paciente**

- 1. Información del Paciente:** Complete los campos solicitados con la información del paciente, incluyendo su nombre, edad, género y cualquier historial médico relevante.
- 2. Datos Clínicos:** Proporcione cualquier detalle adicional sobre la lesión o condiciones relacionadas que puedan ser útiles para el diagnóstico.

- **Cargue de Imagen de la Lesión**

- 1. Selección de la imagen:** Haga clic en el botón "Cargar imagen de la lesión" y seleccione la imagen deseada desde su dispositivo. Asegúrese de que la imagen esté clara y bien enfocada para una evaluación precisa.
- 2. Selección de factores de riesgo:** Marque las casillas correspondientes a los factores de riesgo presentes en el paciente.

Estos pueden incluir exposición al sol, antecedentes familiares de cáncer de piel, entre otros.

- **Diagnóstico**

Una vez cargada la imagen y seleccionados los factores de riesgo, haga clic en el botón "Diagnosticar" para que el software procese la información y genere un diagnóstico preliminar.

- **Impresión de Resultados**

Seleccione la opción "Enviar a impresora" para obtener un informe detallado de los resultados obtenidos, que incluya el diagnóstico preliminar, recomendaciones y cualquier otra información relevante.

- **Guardar Información en un Archivo CSV**

Haga clic en el botón "Enviar información" para guardar todos los datos del paciente y el diagnóstico realizado en un archivo CSV, que podrá utilizar para futuras referencias o análisis.

- **Funciones Adicionales**

1. **Créditos:** Explore la sección de "Créditos" para conocer más sobre los desarrolladores del software.
2. **Instrucciones de Uso:** Si necesita ayuda adicional sobre cómo utilizar la herramienta, consulte la sección de "Instrucciones de Uso" para obtener orientación detallada.
3. **Cerrar Sesión:** Para salir de la aplicación de manera segura, utilice la opción "Cerrar sesión" y asegúrese de que sus datos estén protegidos.