



UNIVERSIDAD AUTONOMA DE
AGUASCALIENTES
CENTRO DE CIENCIAS BÁSICAS

Centro de Ciencias básicas

Departamento de Ciencias de la computación

TESIS

**Desarrollo de un clasificador para la detección de cáncer
cervicouterino mediante técnicas de aprendizaje
automático**

PRESENTA

José Guadalupe Domínguez García

Para Obtener el Grado de Maestro en Ciencias con opción a la
Computación, Matemáticas Aplicadas

TUTOR:

Dr. Julio Cesar Ponce Gallegos

INTEGRANTES DEL COMITÉ TUTORAL:

Dr. Ángel Eduardo Muñoz Zavala

Dr. Alejandro Padilla Díaz.

Aguascalientes, Ags, 23 de junio del 2025

CARTA DE VOTO APROBATORIO
INDIVIDUAL

JORGE MARTÍN ALFÉREZ CHÁVEZ
DECANO DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como TUTOR designado del estudiante JOSÉ GUADALUPE DOMÍNGUEZ GARCÍA con ID 214938 quien realizó la tesis titulada: DESARROLLO DE UN CLASIFICADOR PARA LA DETECCIÓN DE CÁNCER CERVICOUTERINO MEDIANTE TÉCNICAS DE APRENDIZAJE AUTOMÁTICO, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el VOTO APROBATORIO, para que él pueda proceder a imprimirla así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE
"Se Lumen Proferre"
Aguascalientes, Ags., a 23 de junio de 2025


Dr. Julio Cesar Ponce Gallegos
Tutor de tesis

c.c.p.- Interesado
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado.
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad.
Aprobado por: Depto. Control Escolar/ Depto. Apoyo al Posgrado.

Código: DO-SEE-FO-07
Actualización: 01
Emisión: 17/05/19

CARTA DE VOTO APROBATORIO
INDIVIDUAL

JORGE MARTÍN ALFÉREZ CHÁVEZ
DECANO DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

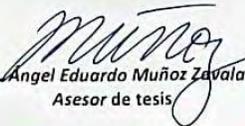
Por medio del presente como ASESOR designado del estudiante JOSÉ GUADALUPE DOMÍNGUEZ GARCÍA con ID 214938 quien realizó la tesis titulada: DESARROLLO DE UN CLASIFICADOR PARA LA DETECCIÓN DE CÁNCER CERVICOUTERINO MEDIANTE TÉCNICAS DE APRENDIZAJE AUTOMÁTICO, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el VOTO APROBATORIO, para que él pueda proceder a imprimirla así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE

"Se Lumen Proferre"

Aguascalientes, Ags., a 23 de junio de 2025


Dr. Angel Eduardo Muñoz Zavala
Asesor de tesis

c.c.p.- Interesado
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado.
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad.
Aprobado por: Depto. Control Escolar/ Depto. Apoyo al Posgrado.

Código: DO-SEE-FO-01
Actualización: 01
Emisión: 17/05/19

CARTA DE VOTO APROBATORIO
INDIVIDUAL

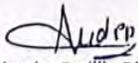
JORGE MARTÍN ALFÉREZ CHÁVEZ
DECANO DEL CENTRO DE CIENCIAS BÁSICAS

PRESENTE

Por medio del presente como ASESOR designado del estudiante JOSÉ GUADALUPE DOMÍNGUEZ GARCÍA con ID 214938 quien realizó la tesis titulada: **DESARROLLO DE UN CLASIFICADOR PARA LA DETECCIÓN DE CÁNCER CERVICOUTERINO MEDIANTE TÉCNICAS DE APRENDIZAJE AUTOMÁTICO**, un trabajo propio, innovador, relevante e inédito y con fundamento en el Artículo 175, Apartado II del Reglamento General de Docencia doy mi consentimiento de que la versión final del documento ha sido revisada y las correcciones se han incorporado apropiadamente, por lo que me permito emitir el **VOTO APROBATORIO**, para que él pueda proceder a imprimirla así como continuar con el procedimiento administrativo para la obtención del grado.

Pongo lo anterior a su digna consideración y sin otro particular por el momento, me permito enviarle un cordial saludo.

ATENTAMENTE
"Se Lumen Proferre"
Aguascalientes, Ags., a 23 de junio de 2025


Dr. Alejandro Padilla Díaz
Asesor de tesis

c.c.p.- Interesado
c.c.p.- Secretaría Técnica del Programa de Posgrado

Elaborado por: Depto. Apoyo al Posgrado
Revisado por: Depto. Control Escolar/Depto. Gestión de Calidad
Aprobado por: Depto. Control Escolar/ Depto. Apoyo al Posgrado

Código: DO-SEE-FC-07
Actualización: 03
Emisión: 17/05/18



DICTAMEN DE LIBERACIÓN ACADÉMICA PARA INICIAR LOS TRÁMITES DEL EXAMEN DE GRADO



Fecha de dictaminación dd/mm/aaaa: 21/07/2025

NOMBRE: José Guadalupe Domínguez García ID 214938

PROGRAMA: Maestría en ciencias con opción a la computación, matemáticas aplicadas LGAC (del posgrado): Ingeniería de software

TIPO DE TRABAJO: (X) Tesis () Trabajo Práctico

TÍTULO: Desarrollo de un clasificador para la detección de cáncer cervicouterino mediante técnicas de aprendizaje automático

IMPACTO SOCIAL (señalar el impacto logrado): Con el desarrollo del trabajo correspondiente, se aporta una investigación que ayuda al análisis mediante imágenes para detección de cáncer cervicouterino. Se deja una base de conocimiento que aporta a la reducción de tiempo en el diagnóstico y precisión de la enfermedad.

INDICAR SI NO N.A. (NO APLICA) SEGÚN CORRESPONDA:

Table with 2 columns: INDICAR (SI, NO, N.A.) and DESCRIPCIÓN. Rows include criteria for thesis review, graduation requirements, and publication status.

Con base a estos criterios, se autoriza se continúen con los trámites de titulación y programación del examen de grado: SI X No

FIRMAS

Elaboró: Dr. Jaime Muñoz Arteaga
NOMBRE Y FIRMA DEL SECRETARIO TÉCNICO: Dra. Marián Alfaro Gomez
Revisó: Dr. Alejandro Padilla Diaz
Autorizó: M. en C. Jorge Martín Alléruz Chávez

Nota: procede el trámite para el Depto. de Apoyo al Posgrado
En cumplimiento con el Art. 105C del Reglamento General de Docencia que a la letra señala entre las funciones del Consejo Académico: ... Cuidar la eficiencia terminal del programa de posgrado y el Art. 103F las funciones del Secretario Técnico, llevar el seguimiento de los alumnos.

Elaborado por: D. Apoyo al Posgrado
Revisado por: D. Control Escolar/D. Gestión de Calidad
Aprobado por: D. Control Escolar/ D. Apoyo al Posgrado.

Código: DD-341-F0-11
Actualización: 01
Emisión: 24/04/24

Agradecimientos

Al concluir esta etapa tan significativa de mi formación profesional, expreso mi más sincero agradecimiento a las instituciones, personas y seres queridos que hicieron posible la realización de este trabajo.

En primer lugar, agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT), hoy Secretaría de Ciencia, Humanidades, Tecnología e Innovación en México (SECIHTI), por el apoyo económico brindado a través de la beca que me permitió llevar a cabo mis estudios de posgrado. Asimismo, expreso mi reconocimiento a la Universidad Autónoma de Aguascalientes (UAA) por ofrecer el espacio académico y formativo donde se desarrolló este proyecto.

De manera especial, agradezco al comité de tesis por su acompañamiento y compromiso durante todo el proceso. Sus observaciones, sugerencias y constante guía fueron esenciales para enriquecer el presente trabajo y fortalecer mis habilidades como investigador.

A mis padres, gracias por su amor incondicional, su ejemplo de perseverancia y su apoyo inquebrantable. Esta meta es tanto mía como de ustedes. A mi novia, por su paciencia, comprensión y por ser un pilar de apoyo emocional durante los momentos más exigentes de esta etapa. A mi hermana y su familia, por brindarme siempre su cariño, apoyo constante y un segundo hogar donde refugiarme.

Finalmente, a todos los amigos, profesores y colegas que me acompañaron en este camino: gracias por cada palabra de aliento, cada intercambio de ideas y cada gesto que me impulsó a seguir adelante.

Índice

- Resumen 5
- Abstract 6
- 1. Introducción..... 7
- 1.1 Problemática por resolver 7
- 1.2 Cáncer cervicouterino 8
- 1.3 Aprendizaje máquina 10
- 2. Contenido del trabajo de investigación 12
- 3. Objetivos 14
- 4. Preguntas de investigación 14
- 5. Justificación del problema 15
- 6. Metodología de investigación..... 17
- 7. Marco Teórico 20
- 7.1 Aprendizaje Máquina..... 20
- 7.1.1 Características del aprendizaje máquina 20
- 7.1.2 Tipos de Aprendizaje Máquina..... 20
- 7.2 Clasificadores 21
- 7.3 Redes neuronales 22
- 7.3.1 Arquitectura de una red neuronal 22
- 7.3.1.1 Capa de entrada 23
- 7.3.1.2 Capa oculta 23
- 7.3.1.3 Capa de salida 23
- 7.3.2 Modelos de redes neuronales 24
- 7.3.2.1 Redes de capa simple (perceptrón) 24
- 7.3.2.2 Red neuronal Multicapa – Perceptrón multicapa 25
- 7.3.2.3 Redes neuronales recurrentes (RNN) 25
- 7.3.2.4 Redes neuronales convolucionales (CNN) 26
- 7.3.2.5 Capa de agrupamiento 27
- 7.3.2.6 Capa totalmente conectada 28
- 7.4 Tipos de redes neuronales convolucionales 28
- 7.4.1 LeNet 28
- 7.4.2 AlexNet 29

- 7.4.3 VGG 30
- 7.4.4 Google Net 31
- 7.4.5 ResNet 32
- 8. Trabajos relacionados a la investigación 33
- 9. Desarrollo..... 35
- 9.1 Descripción de los datos correspondientes a utilizar 37
- 9.2 Visualización de datos y exploración 42
- 9.3 Preprocesamiento de las imágenes 47
- 9.3.1 Redimensionamiento y procesamiento de imágenes en paralelo 47
- 9.3.2 Normalización y escalado de imágenes 48
- 9.4 Implementación de los modelos de redes neuronales convolucionales seleccionados 49
- 9.4.1 VGG 16 49
- 9.4.2 Google Inception 52
- 9.4.3 XceptionNet 54
- 10. Resultados y discusiones 56
- 11. Conclusiones y trabajo futuro 71
- 11.1 Trabajo futuro 73
- 12. Referencias bibliográficas 74

Índice de Tablas

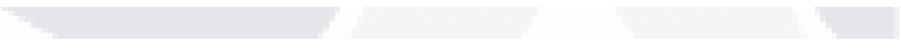
- Tabla 1: Trabajos relacionados al tema de investigación propuesto 35
- Tabla 2: Distribución de los datos correspondientes al conjunto de entrenamiento y validación 57
- Tabla 3: Resultados finales de la precisión de los modelos 60

Índice de Figuras

- Figura 1: Zona donde se desarrolla el cáncer cervicouterino 8
- Figura 2: Evolución de los métodos de diagnóstico del cáncer cervicouterino10
- Figura 3: Arquitectura de una red neuronal artificial11
- Figura 4: Etapas del desarrollo de la enfermedad15
- Figura 5: Metodología de la investigación aplicada 19
- Figura 6: Modelo de una neurona artificial (perceptrón) 24
- Figura 7: Ejemplo visual de una red neuronal multicapa 25

- Figura 8: Ejemplo visual de redes neuronales recurrentes 26
- Figura 9: Modelo de una red neuronal convolucional 26
- Figura 10: Red Neuronal Convolucional LENET 29
- Figura 11: Red Neuronal Convolucional ALEXNET 30
- Figura 12: Red Neuronal Convolucional VGG-16 30
- Figura 13: Red Neuronal Convolucional VGG-19 31
- Figura 14: Red Neuronal Convolucional GOOGLE NET 31
- Figura 15: Red Neuronal Convolucional RESNET 32
- Figura 16: Anatomía del cuello uterino 38
- Figura 17: Composición anatómica interna del cuello uterino 39
- Figura 18: Clasificación de daño en cuello uterino 40
- Figura 19: Cuello uterino sin lesiones aparentes 40
- Figura 20: Cuello uterino con probabilidad de cáncer 41
- Figura 21: Cuello uterino con signos cancerosos 41
- Figura 22: Visualización de imágenes categorizadas en el conjunto de entrenamiento ... 42
- Figura 23: Descripción básica de los datos (totales, clases, ubicación de los archivos y frecuencia de una clase) 43
- Figura 24: División de los datos y frecuencia de estos en todo el conjunto de datos 43
- Figura 25: Gráfico de barras para las tres clases correspondientes (precanceroso, canceroso y no canceroso) 44
- Figura 26: Gráfico de porcentajes para las tres clases de los datos presentados 45
- Figura 27: Visualización de imágenes después del preprocesamiento 48
- Figura 28: La función ReLU simplemente rectifica los datos negativos y los vuelve cero a la salida. Las entradas con valores positivos no sufren modificación alguna a la salida 50
- Figura 29: Explicación de la función de activación softmax 51
- Figura 30: Comparativa de resultados preliminares en la etapa de validación y entrenamiento del modelo VGG16 52
- Figura 31: Capas de convolución para el modelo de InceptionV3 53
- Figura 32: Comparativa de resultados preliminares en la etapa de validación y entrenamiento del modelo Inception V3 54
- Figura 33: Capas de convolución para el modelo de Xception 55

- Figura 34: Comparativa de resultados preliminares en la etapa de validación y entrenamiento del modelo Xception 56
- Figura 35: Comparativa de los resultados en el entrenamiento de los tres modelos correspondientes 58
- Figura 36: Comparativa de los resultados en la validación de los tres modelos correspondientes 59
- Figura 37: Comparativa de la clasificación de los modelos asociados a clase No.1..... 61
- Figura 38: Técnica de Grad-cam para la visualización de la zona o no afectada en base al tipo y al modelo utilizado (VGG16) 62
- Figura 39: Técnica de Grad-cam para la visualización de la zona o no afectada en base al tipo y al modelo utilizado (Xception) 63
- Figura 40: Técnica de Grad-cam para la visualización de la zona o no afectada en base al tipo y al modelo utilizado (Inception V3)..... 63
- Figura 41: Valores AUC para el modelo VGG16 64
- Figura 42: Matriz de confusión para el modelo VGG 16 65
- Figura 43: Valores AUC para el modelo Xception 65
- Figura 44: Matriz de confusión para el modelo Xception 66
- Figura 45: Valores AUC para el modelo Inception V3 67
- Figura 46: Matriz de confusión para el modelo Inception 68
- Figura 47: CURVA ROC por clase aplicando técnica de ensemble learning 69
- Figura 48: Matriz de confusión para los tres tipos y los modelos aplicados 70
- Figura 49: Datos resultantes de la aplicación de ensemble learning 71



Resumen

En este trabajo se explora la implementación de algoritmos de aprendizaje automático, en particular las redes neuronales convolucionales (*CNN*), para potenciar la detección temprana del cáncer cervicouterino a partir de imágenes colposcópicas. El estudio se centra en la problemática global del cáncer cervicouterino, subrayando su alta incidencia y mortalidad, especialmente en naciones en desarrollo y en América Latina.

La metodología adoptada es cualitativa, lo que implica la interpretación de patrones en imágenes médicas. Se utiliza un conjunto de datos que incluye imágenes del cuello uterino clasificadas como cancerosas, precancerosas y no cancerosas, extraídas del repositorio *Intel ODT*. El proceso abarca el preprocesamiento de las imágenes para corregir imperfecciones y normalizarlas, seguido de la implementación y el entrenamiento de varios modelos de *CNN* preentrenados, tales como *VGG16*, *Google Inception* y *XceptionNet*.

Se analizan los resultados de estos modelos en relación con su precisión, tiempo de entrenamiento y capacidad para generalizar a nuevos datos. Se aplican técnicas como *Early Stopping* y aumento de datos para optimizar el entrenamiento y evitar el sobreajuste. Además, se emplean mapas de activación *Grad-CAM* para visualizar las áreas de las imágenes que influyen en las decisiones de clasificación del modelo, lo que proporciona una mayor interpretación y confianza en los resultados.

Los hallazgos indican que los modelos de *CNN* son eficaces en la clasificación de imágenes colposcópicas, presentando variaciones en precisión y tiempo de entrenamiento entre los distintos modelos. El modelo *Inception V3* de *Google* mostró la mayor precisión en el conjunto de validación tras el aumento de datos, mientras que *XceptionNet* destacó por tener el menor tiempo de entrenamiento.

Abstract

This study investigates the use of machine learning algorithms, particularly convolutional neural networks (CNNs), to improve the early detection of cervical cancer through the analysis of colposcopic images. It highlights the worldwide issue of cervical cancer, noting the significant incidence and mortality rates, especially in underdeveloped regions and Latin America. The research employs a qualitative approach aimed at interpreting the patterns found in medical imagery. A collection of cervical images, classified as cancerous, precancerous, and non-cancerous, is utilized, sourced from the Intel ODT repository. The methodology includes preprocessing the images to rectify flaws and standardize the input data, followed by the application and training of multiple pre-trained CNN models, namely VGG16, Google Inception, and XceptionNet.

The models are evaluated based on their accuracy, training duration, and ability to generalize. Strategies like early stopping and data augmentation are incorporated to enhance training efficiency and mitigate overfitting. Furthermore, Grad-CAM activation maps are utilized to illustrate the specific regions of the images that impact the model's classification outcomes, thereby improving the interpretability and reliability of the findings.

The findings suggest that CNN models are proficient in categorizing colposcopic images, with variations noted in accuracy and training duration among the different models. Google's Inception V3 records the highest validation accuracy following data augmentation, while XceptionNet is distinguished by its reduced training time.

1. Introducción

1.1 Problemática por resolver

A nivel mundial, el cáncer cervicouterino (CCU) se clasifica como el cuarto tipo de cáncer más común y es una de las principales causas de mortalidad por esta enfermedad entre mujeres. Se estima que anualmente se diagnosticaron aproximadamente 604,127 casos y ocurrieron 341,831 muertes, de las cuales el 85% se registra en países en desarrollo (IMSS, 2022). La tasa de incidencia es mayor en estas regiones, variando desde el 42.7% en África Oriental hasta el 4.4% por cada 100,000 mujeres en Asia Occidental (Medio Oriente). Aunque la mortalidad ha mostrado una tendencia a la baja, atribuida a la disminución de la incidencia por mejoras en las condiciones sociales y el fortalecimiento de los sistemas de salud, el CCU sigue siendo un indicador de desigualdad, ya que las muertes se concentran en las áreas más vulnerables.

En América Latina, durante 2022, se diagnosticaron más de 78,000 casos de cáncer cervicouterino y se reportaron más de 40,000 muertes en la Región de las Américas (OPS, 2025). Las tasas de mortalidad en esta región son tres veces superiores a las de América del Norte, lo que pone de manifiesto las marcadas desigualdades en los sistemas de salud.

En México, el cáncer cervicouterino ocupó el segundo lugar en diagnósticos y en muertes entre mujeres en 2020, con un estimado de 9,439 nuevos casos y 4,335 muertes. La tasa de incidencia fue del 12.6% y la de mortalidad del 5.7% por cada 100,000 habitantes, aunque comparado con 2012, se observa una notable disminución en la incidencia.

En 2021, el Instituto Mexicano del Seguro Social (IMSS) reportó 1,155 nuevos casos y 1,059 defunciones, con una tasa de incidencia del 2.26% y una tasa de mortalidad del 5.23% por cada 100,000 habitantes (IMSS, 2022). Las entidades con las tasas más altas de mortalidad en el país fueron Morelos (18.6%), Chiapas (17.2%) y Veracruz (16.4%). Ese año, México registró un total de 269,332 muertes en mujeres, de las cuales 37,361 (13.8%) fueron atribuidas a tumores malignos.

Las variaciones en la incidencia de esta enfermedad están relacionadas con desigualdades en el acceso a servicios de vacunación, diagnóstico (cribado) y tratamiento, así como con factores de riesgo como la prevalencia del VIH y determinantes socioeconómicos, incluyendo la pobreza, el sexismo y los prejuicios de género (Ghoneim, A., et al., 2020). Las mujeres que viven con VIH tienen seis veces más probabilidad de desarrollar este tipo

de cáncer en comparación con la población general, y se estima que el 5% de los casos de CCU se debe a esta enfermedad. La influencia del VIH en el cáncer cervicouterino afecta de manera desproporcionada a mujeres jóvenes, lo que también repercute significativamente en los niños, ya que el 20% de los menores que pierden a su madre por cáncer lo hacen a causa del CCU.

1.2 Cáncer cervicouterino

El cáncer cervicouterino se considera la neoplasia maligna más común y mortal que afecta a las mujeres a nivel global. La detección y tratamiento de esta enfermedad son esenciales para prevenir complicaciones severas (Suphalakshmi, A., 2022). Este tipo de tumor se origina en las células del cuello uterino, que es la parte inferior del útero que conecta con la vagina (Figura 1). El principal desencadenante es la infección por el virus del papiloma humano (VPH), que se transmite a través del contacto sexual y afecta a 8 de cada 10 personas en algún momento de su vida. Sin embargo, solo una de cada mil mujeres infectadas desarrollará cáncer. Este proceso se inicia en el cuello uterino, específicamente en la zona que se une a la vagina. Antes de la aparición del cáncer, las células cervicales experimentan cambios, como displasia, lo que provoca la formación de células anormales en el tejido cervical.

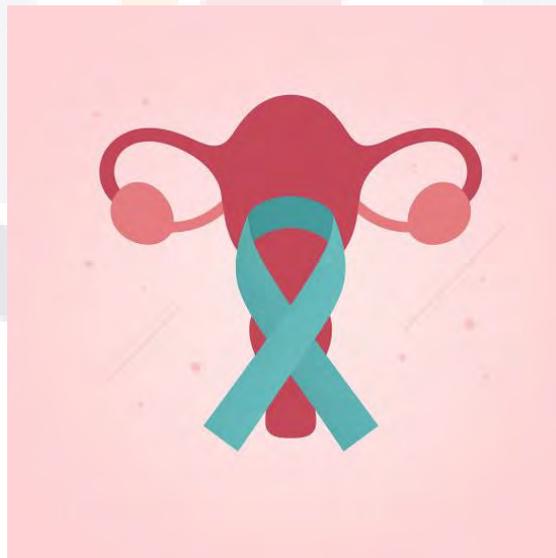


Figura 1: Zona donde se desarrolla el cáncer cervicouterino

En la actualidad, las opciones terapéuticas principales para el cáncer cervicouterino (CCU) incluyen cirugía, radioterapia y quimioterapia. Aunque la inmunoterapia ha mostrado un gran potencial en el tratamiento del CCU, las tasas de supervivencia siguen siendo

preocupantemente bajas. Un desafío significativo en el diagnóstico actual es la fase preclínica prolongada, que puede durar décadas. Por lo tanto, es fundamental identificar nuevos biomarcadores, técnicas o algoritmos que permitan un diagnóstico temprano y una predicción precisa, capaces de detectar lesiones precoces antes de que avancen a formas más avanzadas de cáncer.

Durante el seguimiento médico, se realiza comúnmente la prueba de Papanicolaou tras una histerectomía radical o un tratamiento de quimio radiación, con el fin de detectar cambios displásicos en etapas tempranas en la cúpula vaginal o el cuello uterino. No obstante, esta prueba no es recomendada en el primer año posterior a la radioterapia y la histerectomía, debido a las alteraciones celulares y tisulares causadas por estos tratamientos. Estudios previos han indicado un mayor riesgo de recaída en casos de CCU relacionados con infecciones persistentes por VPH de alto riesgo (Coudray, N., 2018).

La integración del genoma del VPH en el genoma del huésped es un paso crucial en la progresión del CCU, que va desde una célula normal a una lesión intraepitelial y, eventualmente, a un cáncer invasivo. Los dos principales métodos moleculares para analizar el ADN del VPH son el ensayo de captura híbrida y la reacción en cadena de la polimerasa (PCR). El ensayo de captura híbrida es una técnica de amplificación de señal que combina la captura de anticuerpos con la detección de señales quimio luminiscentes. Este sistema utiliza sondas de ARN que identifican 13 tipos de VPH de alto riesgo y 5 de bajo riesgo, detectando los híbridos ARN: ADN a través de anticuerpos conjugados con fosfatasa alcalina y midiendo la señal quimio luminiscente resultante. Por otro lado, los ensayos de PCR amplifican el ADN de muestras clínicas usando cebadores generales o específicos dirigidos a la región de la cápside L1. Este método se utiliza para la detección, genotipado y cuantificación de la carga viral del VPH, abarcando una amplia variedad de sus tipos.

Las pruebas de VPH identifican los tipos de alto riesgo de infección en el cuello uterino, mientras que los exámenes citológicos analizan las células extraídas del cuello uterino mediante microscopía, con el objetivo de detectar posibles lesiones precoces o cáncer. Aunque la biopsia guiada por colposcopia sigue siendo el método estándar para diagnosticar el CCU, se complementa con la estadificación basada en el examen clínico y

los resultados de pruebas de imagen (Liston, C., 2018). La Figura 2 ilustra los principales hitos en la evolución del diagnóstico del CCU.

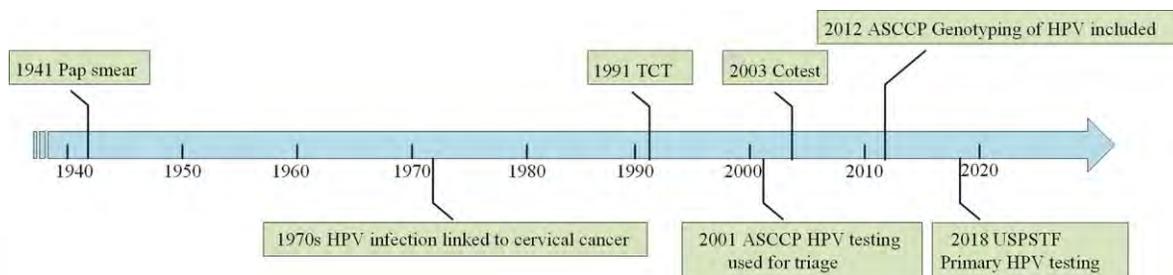


Figura 2: Evolución de los métodos de diagnóstico del cáncer cervicouterino.

En la actualidad, hay diversas técnicas que resultan muy útiles para llevar a cabo análisis a través de la imagenología. Entre estas se encuentran las colposcopías, que consisten en la utilización de un instrumento específico que permite ampliar de 5 a 40 veces el cuello uterino completamente expuesto. Esto facilita una evaluación visual en tiempo real del cuello uterino, centrándose especialmente en la zona de transformación, con el objetivo de identificar patrones malignos como lesiones intraepiteliales escamosas (LEI) y cáncer invasor. Además, se realiza una biopsia guiada por colposcopia en el área sospechosa para determinar la necesidad de un tratamiento adicional (Hou. X, et al., 2022).

1.3 Aprendizaje máquina

El aprendizaje automático es una rama de la inteligencia artificial que se enfoca en el desarrollo de algoritmos y modelos que permiten a las computadoras aprender y mejorar su desempeño en tareas específicas, sin necesidad de ser programadas de manera explícita para tal fin. Los sistemas de aprendizaje automático analizan datos, identifican patrones, realizan predicciones o toman decisiones basadas en dichos patrones.

En el contexto de utilizar esta tecnología para optimizar el diagnóstico o asistir en la obtención de resultados para este tipo de cáncer, existen diversas propuestas y algoritmos aplicados a este objetivo. Esta investigación se centrará en las redes neuronales, que constituyen un subconjunto del aprendizaje automático, incluyendo también los algoritmos de aprendizaje profundo. Su denominación y estructura se inspiran en el cerebro humano, reproduciendo la manera en que las neuronas biológicas transmiten información entre sí (M. Anousouya Devi., et al., 2016).

Las redes neuronales están formadas por capas de nodos, que incluyen una capa de entrada, una o más capas ocultas, y una capa de salida. Cada nodo, conocido como neurona artificial, está interconectado y tiene un peso y un umbral asignados. Si la salida de un nodo supera el umbral establecido, el nodo se activa y envía datos a la siguiente capa de la red. En caso contrario, no se transmiten datos a la siguiente capa.

Estas redes dependen de datos de entrenamiento para aprender y mejorar su precisión a lo largo del tiempo. No obstante, una vez que estos algoritmos han sido ajustados con eficacia, se convierten en herramientas sumamente efectivas en las ciencias de la computación y la inteligencia artificial, permitiéndonos clasificar y agrupar datos a gran velocidad (Figura 3).

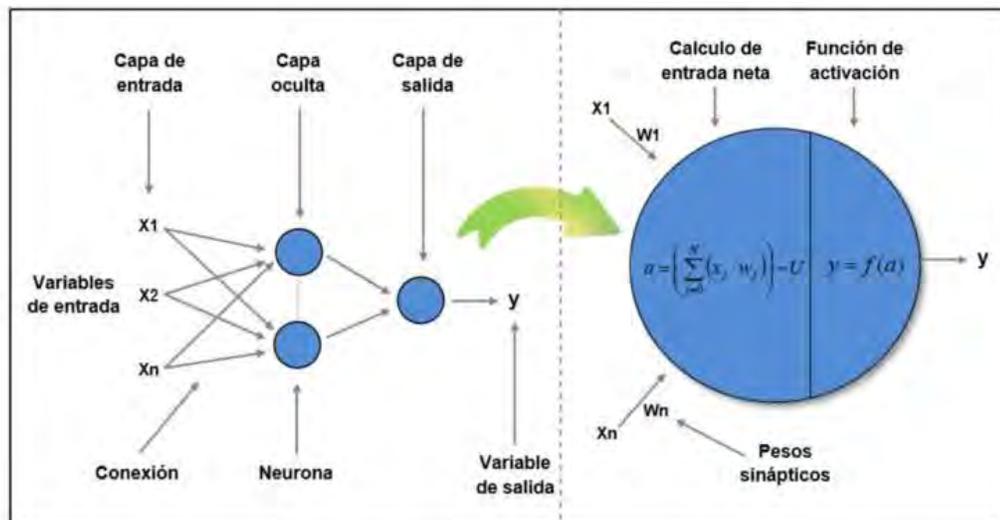


Figura 3: Arquitectura de una red neuronal artificial

Al abordar la detección del cáncer cervicouterino mediante imágenes de esa área, que es el enfocado del cuello uterino con el espéculo ampliado, es fundamental emplear redes neuronales, en particular aquellas basadas en aprendizaje profundo, gracias a su habilidad para interpretar y aprender características complejas a partir de datos visuales como las imágenes. En esta investigación, se implementarán redes neuronales convolucionales (profundizaremos más sobre este tema más adelante) con el fin de identificar patrones jerárquicos. A través de su estructura de capas, estas redes pueden reconocer bordes y texturas, transformándolos en formas o regiones, para finalmente identificar el patrón o el objeto correspondiente. Estas redes son especialmente efectivas en el análisis de imágenes

médicas, ya que facilitan la extracción de información relevante incluso en condiciones de ruido o variaciones en la calidad de las imágenes.

2. Contenido del trabajo de investigación

Se establece la siguiente estructura para abordar el problema del cáncer cervicouterino a través del aprendizaje automático:

1. Introducción

Se presenta el problema a resolver, enfatizando la magnitud del cáncer cervicouterino a nivel global, así como en las regiones de América Latina y en México. Se introduce el cáncer cervicouterino y el aprendizaje automático como métodos para su detección, ofreciendo un panorama general del contenido de la investigación.

2. Objetivos

Se definen el objetivo general del estudio, que consiste en implementar un algoritmo de aprendizaje automático eficaz para clasificar imágenes en la identificación de tumores asociados al cáncer cervicouterino, y los objetivos específicos que orientan la investigación.

3. Preguntas de Investigación

Se plantean preguntas que fundamentan la investigación, cuestionando aspectos relevantes y estableciendo un sólido punto de partida para el estudio.

4. Justificación del Problema

Se argumenta la relevancia de la detección temprana del cáncer cervicouterino y cómo la aplicación de algoritmos de aprendizaje automático puede ayudar a disminuir costos y mejorar la precisión en los diagnósticos. Se abordan factores de riesgo y se justifica la necesidad de nuevas técnicas diagnósticas.

5. Metodología de Investigación

Se describe el enfoque metodológico cualitativo adoptado, que se centra en la interpretación de patrones en imágenes. Se detallan las etapas de la investigación, desde la recopilación del conjunto de datos de imágenes hasta la presentación de los resultados.

6. Marco Teórico

Se revisan conceptos clave, tales como aprendizaje automático, clasificadores, y redes neuronales (incluyendo arquitecturas simples, multicapa y convolucionales), así como tipos específicos de redes neuronales convolucionales relevantes para la investigación (*LeNet*, *AlexNet*, *VGG*, *GoogleNet*, *ResNet*).

7. Trabajos Relacionados a la Investigación

Se presenta una revisión de estudios previos que han explorado la detección del cáncer cervicouterino mediante técnicas de inteligencia artificial y aprendizaje automático, resaltando sus enfoques, resultados y contribuciones.

8. Desarrollo

Se describe el proceso de implementación, que incluye la descripción de los datos utilizados, su visualización y exploración, el preprocesamiento de imágenes (redimensionamiento, procesamiento en paralelo, normalización y escalado), así como la implementación de los modelos (*VGG16*, *Google Inception*, *ExceptionNet*).

9. Resultados y Discusiones

Se presentan y comparan los resultados obtenidos de los diferentes modelos implementados, incluyendo métricas de precisión y tiempos de entrenamiento. Se muestra un ejemplo de predicción utilizando una imagen y se analizan los mapas de activación para comprender cómo los modelos toman decisiones.

10. Conclusiones y Trabajo Futuro

Se resumen las conclusiones del estudio, subrayando el cumplimiento de los objetivos y el potencial del sistema desarrollado. Se sugieren futuras líneas de investigación para mejorar y ampliar los resultados obtenidos.

11. Referencias Bibliográficas

Se enumeran todas las fuentes citadas en el documento, proporcionando una base sólida para la investigación.

3. Objetivos

General

- Desarrollar un algoritmo de aprendizaje automático orientado a la clasificación de imágenes médicas para la identificación de tumores relacionados con el cáncer cervicouterino.

Específicos

- Obtener la base de datos correspondiente con imágenes relacionadas con el cáncer cervicouterino y realizar la clasificación correspondiente para las mismas
- Realizar preprocesamiento para corrección de imperfecciones en imágenes colposcópicas (provenientes de la base de datos ya mencionada) con diferentes técnicas de procesamiento de imágenes
- Análisis y selección de los algoritmos de redes neuronales convolucionales
- Implementación de los algoritmos basados en redes neuronales convolucionales para la detección de los tumores para el cáncer cervicouterino mediante el procesamiento de imágenes
- Validación de la eficiencia y análisis comparativo entre los algoritmos implementados en comparación con los del estado del arte.

4. Preguntas de investigación

- ¿Cuál es la eficacia de un algoritmo de aprendizaje automático basado en redes neuronales convolucionales para detectar tumores relacionados con el cáncer cervicouterino en imágenes de colposcopia?
- ¿Qué conjuntos de datos de imágenes colposcópicas son apropiados para la formación y evaluación de modelos de clasificación vinculados al cáncer cervicouterino?
- ¿Cuáles son las técnicas de preprocesamiento más efectivas para mejorar la calidad de las imágenes colposcópicas empleadas en el entrenamiento de modelos de clasificación?
- ¿Qué algoritmos de redes neuronales convolucionales proporcionan el mejor desempeño en la clasificación de tumores en imágenes colposcópicas?
- ¿Cómo se comparan los algoritmos aplicados con otros modelos de vanguardia en términos de precisión, sensibilidad y especificidad?

5. Justificación del problema

La detección temprana del cáncer cervicouterino en las mujeres (Figura 4) puede salvarles la vida antes de un desarrollo crónico (Subarna, T., Sukumar, P., 2022). Es importante destacar que la detección temprana no solo mejora el pronóstico clínico, sino que también tiene un impacto significativo en el ámbito económico.



Figura 4: Etapas del desarrollo de la enfermedad

Con esto, se contribuye a la disminución de los costos relacionados con el tratamiento y análisis. En México, al evaluar las disparidades entre las posibles fuentes de subestimación derivadas de la revisión de expedientes y la sobreestimación procedente del panel de expertos, se establece que el costo promedio real del tratamiento en el primer año del padecimiento oscila entre 50 y 70 mil pesos, según datos proporcionados por entidades médicas autorizadas. La reducción de estos costos es crucial, dado que para los sectores vulnerables estas cifras son inalcanzables para financiar dicho proceso y tratamiento (IMSS, 2022).

Las causas, aunque se mencionan principalmente como atribuibles al virus del papiloma humano (Instituto Nacional del Cáncer, 2023), incluyen varios factores desencadenantes:

- **Consumo de tabaco:** Fumar incrementa el riesgo de desarrollar cáncer cervical. Las infecciones por VPH en fumadores tienden a persistir más tiempo y tienen

menos probabilidad de resolverse. El VPH es responsable de la mayoría de los tipos de cáncer cervical.

- **Número elevado de parejas sexuales:** Cuantas más parejas sexuales tenga una persona y mayor sea el número de parejas sexuales de su pareja, más alta será la probabilidad de contraer el virus.
- **Inicio temprano de la actividad sexual:** Comenzar a tener relaciones sexuales a una edad temprana aumenta el riesgo de infección por VPH.
- **Presencia de otras infecciones STI:** Tener otras infecciones transmitidas sexualmente (ITS) eleva el riesgo de contraer VPH, que puede conducir al cáncer cervical. Entre estas ITS se encuentran herpes, clamidia, gonorrea, sífilis y VIH/SIDA.
- **Exposición a medicamentos para prevenir abortos espontáneos:** Si la madre consumió dietilestilbestrol durante el embarazo, existe un riesgo incrementado para desarrollar cáncer cervical. Este medicamento fue utilizado en los años cincuenta para evitar abortos espontáneos y está asociado con un tipo específico llamado adenocarcinoma de células claras.

La detección precoz ha sido fundamental para reducir tanto la mortalidad como la incidencia del cáncer cervicouterino en las últimas cinco décadas. No obstante, debido al aumento en la carga laboral, las exploraciones visuales pueden resultar en diagnósticos erróneos e ineficiencias diagnósticas. El cáncer cervicouterino se identifica frecuentemente mediante técnicas avanzadas como imágenes médicas tridimensionales (3D), tomografía computarizada (TC), resonancia magnética nuclear (RMN) y tomografía por emisión de positrones (PET), así como colposcopias en dos dimensiones (Goker, H., 2024). La prueba Papanicolaou detecta alteraciones celulares en el cuello uterino que podrían ser indicativas del cáncer; implica un raspado ligero del tejido cervical seguido por su análisis en laboratorio. Las muestras celulares se recogen en un envase con líquido adicional (Luo, W., 2021).

El cáncer cervicouterino es fundamentalmente prevenible. La prevención primaria junto con un diagnóstico adecuado son las estrategias más efectivas para reducir tanto la carga asistencial como la mortalidad asociada a esta condición. Desde 2006 está disponible una vacuna contra el VPH destinada a prevenir este tipo canceroso. Los miembros del equipo interdisciplinario deben educar a las jóvenes pacientes—preferiblemente antes del inicio de

su vida sexual—y sus familias acerca de esta efectiva vacuna. Esta iniciativa aborda estrategias preventivas primarias, evaluaciones diagnósticas actuales, estadificación y opciones terapéuticas específicas para casos invasivos del cáncer cervicouterino (Fowler, 2023).

En esta propuesta para abordar el problema se presentan diversas soluciones destacando especialmente aquellas que implican aplicar diferentes enfoques terapéuticos y preventivos (Fowler, 2023), tales como:

- Identificar factores riesgosos, así como signos y síntomas asociados al cáncer cervicouterino incluyendo los tipos altos riesgosos del VPH.
- Implementar tratamientos basados en evidencia para lesiones precoces cervicales.
- Elegir adecuadamente modalidades terapéuticas para casos invasivos tomando en cuenta características individuales del paciente, estadificación y opciones disponibles que incluyan cirugía, radioterapia o quimioterapia.

6. Metodología de investigación

Para llevar a cabo el trabajo requerido, se adoptó una metodología fundamentada en un enfoque cualitativo. Esta elección se justificó por su capacidad para proporcionar una comprensión más profunda de procesos complejos, interacciones sociales y fenómenos culturales. El método permite la recolección de datos relacionados con experiencias vividas, emociones y comportamientos, así como los significados que las personas atribuyen a ellos. La selección de esta metodología estuvo motivada por la necesidad de considerar las percepciones, es decir, las características de las imágenes desde sus representaciones. Esto facilita la exploración de cómo y por qué ocurren ciertos eventos, los patrones que estos siguen, la interpretación de situaciones y la descripción de lo que se observa.

Se requirió un análisis enfocado en el reconocimiento de patrones en imágenes para la detección del cáncer cervicouterino. Por consiguiente, fue crucial realizar una adecuada interpretación del análisis y los resultados para emitir un diagnóstico preciso.

Al examinar el diagrama correspondiente (Figura 5), se identificaron los pasos que conformaron la estructura de la investigación para esta implementación, enumerando y describiendo cada una de las etapas metodológicas:

- **Obtención de un conjunto de datos de imágenes del cuello uterino**

Para este estudio, se obtuvo un conjunto de datos de imágenes del cuello uterino proveniente del repositorio *Intel ODT*. Este conjunto de datos contiene imágenes clasificadas en tres categorías principales: cancerosas, precancerosas y no cancerosas. Esta clasificación permitió entrenar modelos de aprendizaje profundo con un enfoque supervisado, facilitando la distinción entre los niveles de severidad en las lesiones cervicales. La diversidad y organización de este resulta fundamental para evaluar el desempeño del modelo en la detección y clasificación de anomalías en imágenes de colposcopia.

- **Identificar el tipo de lesiones indicativas de cáncer cervicouterino**

Se requirió realizar el desarrollo del marco teórico, de qué es y que causa dicho padecimiento. Se enfocó en cómo se identifican patrones malignos en el cuello uterino a partir de imágenes obtenidas mediante colposcopia. Por último, se exploró el uso del aprendizaje automático, en particular las redes neuronales convolucionales, para la detección de lesiones indicativas de cáncer en estas imágenes.

- **Realizar el preprocesamiento de las imágenes**

Se requirió realizar una corrección de imperfecciones y realizar mejoras a las imágenes como el acercamiento a la zona correspondiente el cuello uterino, el aumento de nitidez, limpieza de ruido, así como la redimensión de la imagen a 224 píxeles para el trabajo del modelo de manera óptima

- **Desarrollar el sistema clasificador basado en redes neuronales que identifique patrones malignos conocidos en colposcopias**

Se realizó la codificación del método, así como el entrenamiento de este con las imágenes ya preprocesadas que consistió en imágenes de tipo colposcopia y así se obtuvo el patrón maligno identificado en las mismas.

- **Realizar pruebas de implementación del algoritmo**

Se llevarán a cabo pruebas exhaustivas para evaluar el rendimiento del modelo durante su entrenamiento. Se validará la precisión, sensibilidad y especificidad del algoritmo en la detección de lesiones malignas en imágenes colposcópicas. Además, se realizarán pruebas de generalización con conjuntos de datos no vistos y técnicas como validación cruzada para evitar sobreajuste

- **Obtener resultados y comparar los mismos con los del estado del arte para validar su eficiencia**

Una vez que se implementó el modelo se procedió a obtener los resultados, arrojando la clasificación de las imágenes, además del porcentaje de precisión alcanzado. Se tomó en cuenta los trabajos relacionados debido a la validez y al sustento de la información de una buena o mala eficiencia de este

- **Informar y presentar resultados con base en precisión y clasificación de los resultados**

Mediante el trabajo escrito y una presentación, se comunicó los resultados obtenidos, comprobando así el trabajo realizado y cuáles fueron los objetivos alcanzados. También se analizó si la metodología del desarrollo de todo el proyecto alcanzó las metas propuestas y todos los pasos mencionados fueron atendidos

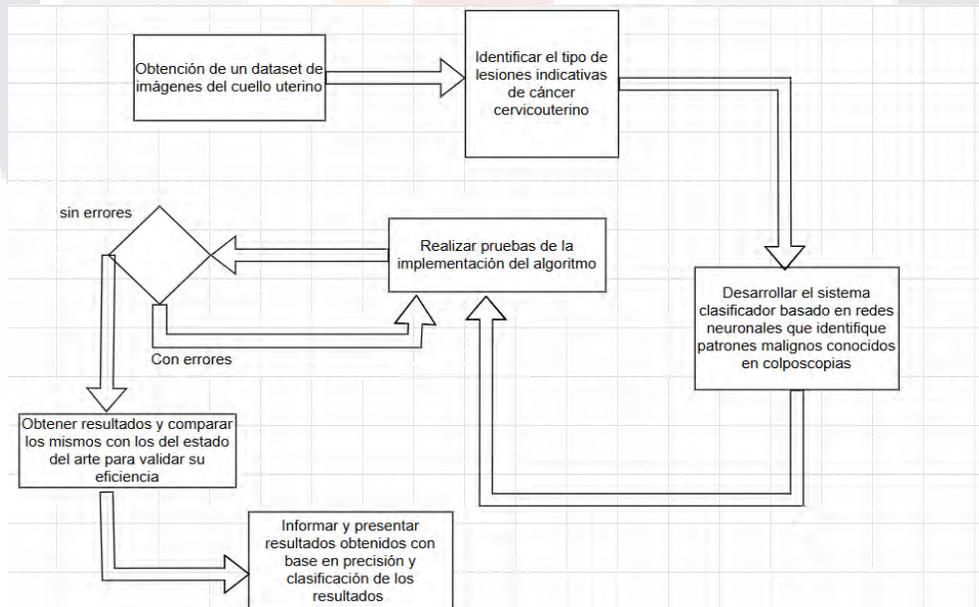


Figura 5: Metodología de la investigación aplicada

7. Marco Teórico

7.1 Aprendizaje Máquina

El aprendizaje automático, conocido como *Machine Learning* (ML), es una rama de la inteligencia artificial (IA) que se centra en desarrollar sistemas capaces de aprender a partir de datos y perfeccionar su rendimiento con el tiempo, sin necesidad de ser programados explícitamente para cada tarea. Este proceso se basa en algoritmos matemáticos y estadísticos que permiten analizar e interpretar datos, reconocer patrones y realizar predicciones o decisiones fundamentadas en la información recibida. A medida que los sistemas informáticos acumulan más "experiencias", tienen la capacidad de ajustarse y mejorar continuamente; por lo tanto, cuanto mayor sea la cantidad de datos introducidos, más precisos serán los resultados obtenidos. A continuación, se presentan algunas características clave junto con los diferentes tipos de aprendizaje (Corporativa Iberdrola, 2023):

7.1.1 Características del aprendizaje máquina

- Capacidad de aprendizaje: Habilidad para aumentar la eficacia en una tarea particular mediante la experiencia acumulada (datos).
- Generalización: Aptitud para procesar datos nuevos y no previamente conocidos tras haber adquirido conocimientos de un conjunto de datos de entrenamiento.
- Adaptabilidad: Capacidad para ajustarse a alteraciones en los datos o en el entorno que los rodea.
- Automatización de decisiones: Permite la toma de decisiones fundamentadas en datos sin necesidad de intervención humana.

7.1.2 Tipos de Aprendizaje Máquina

- Aprendizaje supervisado: Los algoritmos aprenden de un conjunto de datos etiquetados, intentando predecir la etiqueta para nuevos datos basándose en las relaciones aprendidas. Ejemplos incluyen regresión lineal y redes neuronales.
- Aprendizaje no supervisado: Trabaja con datos sin etiquetar, intentando encontrar estructuras ocultas o patrones en los datos. Ejemplos incluyen agrupamiento (*clustering*) y reducción de dimensionalidad.

- Aprendizaje por refuerzo: Un agente aprende a tomar decisiones optimizando alguna noción de recompensa a través de la interacción con un entorno. Es común en juegos y en problemas de optimización secuencial.
- Aprendizaje semi-supervisado y auto-supervisado: Métodos que utilizan tanto datos etiquetados como no etiquetados para mejorar el aprendizaje, aprovechando grandes cantidades de datos no etiquetados junto con pequeñas cantidades de datos etiquetados

7.2 Clasificadores

Es un algoritmo destinado a categorizar o clasificar una instancia de datos según sus atributos. Los clasificadores son algoritmos de aprendizaje supervisado que asignan una clase a una entrada, fundamentándose en un conjunto de datos previamente etiquetados (Sánchez, J., 2004). A continuación, se presentan algunos ejemplos de clasificadores y su funcionamiento:

- Árboles de Decisión: Este método utiliza una estructura arbórea en la cual cada nodo representa un atributo, cada enlace indica una decisión relacionada con ese atributo, y cada hoja simboliza un resultado final (una clase). Aunque los árboles de decisión son fáciles de interpretar y visualizar, pueden ser propensos al sobreajuste con respecto a los datos de entrenamiento.
- Regresión Logística: A pesar del término utilizado, este es un modelo de clasificación que estima la probabilidad de que una instancia pertenezca a una determinada clase. La regresión logística resulta eficaz para problemas de clasificación binaria y es comúnmente empleada en áreas como la medicina y las finanzas.
- Máquinas de Vectores de Soporte (SVM): Este enfoque busca el hiperplano óptimo que separa mejor las distintas clases en un espacio característico multidimensional. Las SVM son particularmente efectivas en espacios con alta dimensionalidad y cuando el número de dimensiones supera el número total de muestras.
- Redes Neuronales Artificiales (ANN): Inspiradas en el funcionamiento del cerebro humano, estas redes están formadas por neuronas organizadas en capas que procesan la información en diversos niveles de abstracción. Son altamente flexibles y potentes, capaces de detectar patrones complejos; se dará especial énfasis a este tema ya que es fundamental para la implementación.

- K-Vecinos más Cercanos (*K-NN*): Este algoritmo clasifica nuevas muestras basándose en las etiquetas correspondientes a las 'k' muestras más cercanas dentro del conjunto de entrenamiento. Es un método simple y efectivo que no requiere un modelo explícito, aunque puede volverse computacionalmente exigente conforme aumenta el tamaño del conjunto de datos.

7.3 Redes neuronales

Una red neuronal es un enfoque de inteligencia artificial que permite a las computadoras analizar datos de manera similar a cómo lo hace el cerebro humano. Este método se inscribe dentro del aprendizaje automático conocido como Aprendizaje Profundo, que emplea nodos o neuronas organizadas en capas interconectadas, emulando la estructura cerebral (AWS, 2023). Esto da lugar a un sistema que se adapta y aprende de sus errores, permitiendo una mejora continua. Así, las redes neuronales artificiales buscan abordar problemas complejos con mayor precisión, como la generación de resúmenes de documentos o el reconocimiento facial.

Las redes neuronales encuentran aplicación en diversos sectores, incluyendo:

- Diagnóstico médico mediante la clasificación de imágenes clínicas.
- Marketing dirigido a través del análisis de datos comportamentales y filtrado en redes sociales.
- Predicciones financieras basadas en el análisis de datos históricos de instrumentos financieros.
- Pronósticos sobre carga eléctrica y demanda energética.
- Procesos y control de calidad.
- Identificación de compuestos químicos.

7.3.1 Arquitectura de una red neuronal

Una red neuronal básica tiene diversas neuronas artificiales interconectadas en tres capas que se describen de manera detallada en los siguientes puntos (Acevedo-Lipes, A. M., 2021):

7.3.1.1 Capa de entrada

La información correspondiente entra en la red neuronal artificial desde la capa de entrada. Los nodos de entrada se encargan de procesar los datos, los analizan o los clasifican y los pasan a la siguiente capa.

7.3.1.2 Capa oculta

Las capas ocultas reciben su información de la capa de entrada o de otras capas ocultas. Las redes neuronales artificiales pueden contener múltiples capas ocultas. Cada una de estas capas examina la salida de la anterior, la procesa adicionalmente y la transmite a la siguiente capa.

7.3.1.3 Capa de salida

La capa de salida es responsable de generar el resultado final del procesamiento de datos realizado por una red neuronal artificial. Esta capa puede contener uno o varios nodos. Por ejemplo, en un escenario de clasificación binaria (sí/no), la capa de salida consiste en un único nodo que devuelve 1 o 0. En contraste, en problemas de clasificación multiclase, la capa de salida puede estar compuesta por múltiples nodos.

Las Redes Neuronales Artificiales (RNA) se inspiran en procesos biológicos; esto implica que están constituidas por elementos que funcionan de manera comparable a las neuronas en sus funciones más básicas y organizadas similarmente al cerebro (Acevedo-Lipes, A. M., 2021). Sin embargo, las similitudes son limitadas. Las principales características de las RNA incluyen:

- **Aprendizaje a partir de la experiencia:** Las RNA tienen la capacidad de modificar su comportamiento en respuesta a cambios en su entorno. Dado un conjunto específico de entradas (posiblemente junto con salidas deseadas), estas redes se ajustan para ofrecer respuestas consistentes. Se han desarrollado diversos algoritmos de entrenamiento, cada uno con sus propias ventajas y desventajas.
- **Generalización a nuevos ejemplos:** Una vez entrenada, la red puede responder con cierta estabilidad ante pequeñas variaciones en las entradas, lo que las vuelve efectivas para el reconocimiento de patrones.
- **Abstracción de información esencial:** Algunas RNA pueden extraer información relevante a partir de un conjunto dado de entradas. Por ejemplo, en tareas

relacionadas con el reconocimiento de patrones, una red puede ser entrenada usando secuencias distorsionadas que representan una letra. Tras un entrenamiento adecuado, esta red será capaz de producir resultados correctos incluso ante entradas distorsionadas, lo que indica que ha aprendido algo nuevo y previamente desconocido.

7.3.2 Modelos de redes neuronales

A continuación, se mencionan algunos de los modelos de redes neuronales más representativos. Estos modelos destacan por su arquitectura y rendimiento, y sirven como referencia para comprender las distintas estrategias aplicadas en el procesamiento de datos visuales.

7.3.2.1 Redes de capa simple (perceptrón)

A pesar de que una neurona individual puede llevar a cabo modelos básicos de funciones, su eficacia se maximiza cuando están dispuestas en redes. La forma más elemental de red consiste en un grupo de perceptrones que reciben un conjunto de entradas y generan la salida correspondiente. Cada perceptrón en la red produce una salida, calculada de la misma manera que para un único perceptrón, mediante la suma de las entradas multiplicadas por sus respectivos pesos. Al graficar una red neuronal, se introduce una "capa" inicial que no cuenta para los cálculos computacionales; su única función es distribuir las entradas entre los perceptrones (Sun, Y. et al., 2020). Esta capa será denominada capa 0. Por lo tanto, la representación gráfica de una red neuronal simple sería la siguiente (Figura 6):

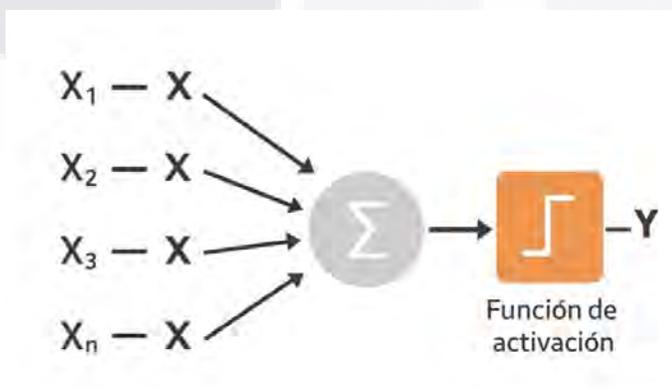


Figura 6: Modelo de una neurona artificial (perceptrón).

7.3.2.2 Red neuronal Multicapa – Perceptrón multicapa

La red neuronal multicapa se considera una extensión de la red neuronal monocapa. La principal distinción radica en que, a diferencia de la red monocapa, que consta únicamente de una capa de neuronas de entrada y otra de salida, la multicapa incluye varias capas intermedias, comúnmente denominadas capas ocultas, situadas entre la capa de entrada y la de salida (Figura 7).

La conectividad de esta red puede ser total o parcial, dependiendo del número de conexiones presentes. Una escasez en las conexiones puede llevar a que la red no logre resolver problemas debido a un número insuficiente de parámetros ajustables. Por otro lado, un exceso en las conexiones puede resultar en un sobreajuste durante el entrenamiento con los datos (Ludemir, T., 2005).

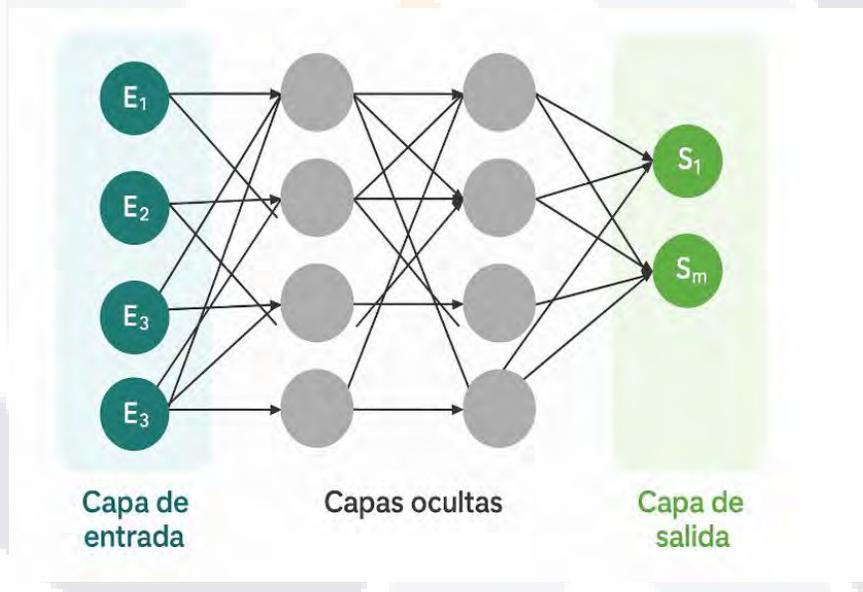


Figura 7: Ejemplo visual de una red neuronal multicapa

7.3.2.3 Redes neuronales recurrentes (RNN)

Las redes neuronales recurrentes (RNN) representan una categoría de modelos de aprendizaje profundo que están específicamente desarrollados para procesar datos en secuencia (Alhajeri, M., 2024). Estas se distinguen por su uso de ciclos de retroalimentación, los cuales permiten reenviar información hacia atrás con el fin de evaluar la precisión de sus predicciones y aprender a partir de sus errores, con el fin de llegar de mejor manera al resultado (Figura 8).

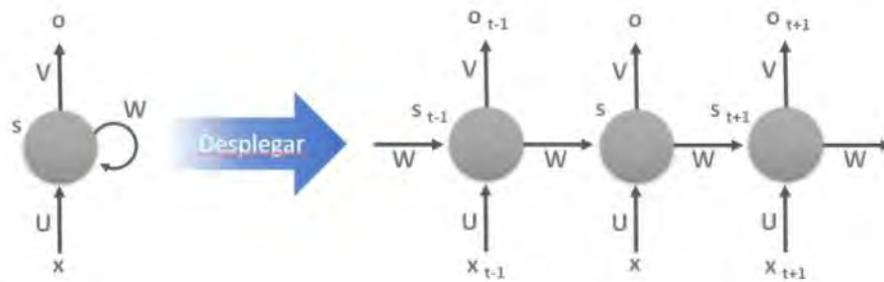


Figura 8: Ejemplo visual de redes neuronales recurrentes

7.3.2.4 Redes neuronales convolucionales (CNN)

Están formadas por numerosas capas interconectadas que permiten descomponer, examinar y clasificar imágenes y videos (Figura 9). Cada capa se enfoca en un aspecto específico de la imagen; por lo tanto, un mayor número de capas implica una complejidad analítica superior. Estas redes son cruciales para el avance de la visión artificial y requieren ser entrenadas con grandes volúmenes de imágenes. Además, se destaca una estructura general y el funcionamiento de cada capa involucrada en este proceso (Sun, Y. et al., 2020).

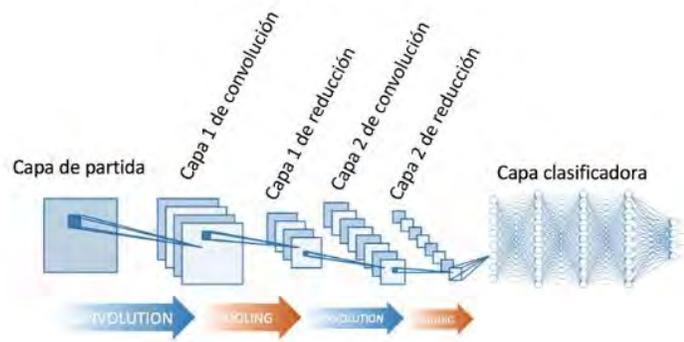


Figura 9: Modelo de una red neuronal convolucional

Las Redes Neuronales Convolucionales (CNN) se caracterizan por su rendimiento superior al procesar entradas que involucran imágenes, voz o audio. Estas redes están compuestas principalmente por tres tipos de capas:

- Capa convolucional
- Capa de agrupamiento
- Capa totalmente conectada (FC)

La capa convolucional es fundamental en una CNN, ya que es donde se lleva a cabo la mayor parte del procesamiento. Para su funcionamiento, se requieren varios componentes: datos de entrada, un filtro y un mapa de características. Si consideramos como entrada una imagen a color, esta se representará como una matriz tridimensional de píxeles, lo que

implica que tiene tres dimensiones (altura, ancho y profundidad), correspondientes al modelo RGB asociado con las imágenes. Un elemento clave en este proceso es el detector de características, también conocido como *kernel* o filtro, que se desplaza a través de los campos receptivos de la imagen para identificar la presencia de ciertas características. Este procedimiento recibe el nombre de convolución.

El detector de características consiste en una matriz bidimensional (2-D) de pesos que representa una sección específica de la imagen. Aunque su tamaño puede variar, comúnmente el filtro se define como una matriz 3 x 3; este tamaño también influye en las dimensiones del campo receptivo. Al aplicar el filtro a un área determinada de la imagen, se calcula un producto punto entre los píxeles de entrada y los valores del filtro. El resultado se introduce en una matriz de salida correspondiente. Posteriormente, el filtro avanza por pasos definidos y repite este proceso hasta haber cubierto toda la imagen. El resultado final obtenido tras realizar múltiples productos punto entre la entrada y el filtro es conocido como mapa de características, mapa de activación o característica convolucionada. Tras cada operación convolucional, las CNN aplican una transformación mediante la función *ReLU* (Unidad Lineal Rectificada) al mapa resultante para introducir no linealidades en el modelo (Zing, D., 2018).

Como se mencionó anteriormente, puede haber otra capa convolucional después de la inicial. En tales casos, la estructura general de la CNN tiende a tener un carácter jerárquico; esto ocurre porque las capas posteriores pueden analizar los píxeles dentro del contexto proporcionado por las capas previas. Por ejemplo, si el objetivo es determinar si una imagen incluye una bicicleta, se puede conceptualizar como compuesta por diversas partes: estructura principal, manubrio, llantas y pedales entre otros elementos. Cada uno de estos componentes individuales representa patrones básicos dentro del sistema neuronal y su combinación completa da lugar a un patrón más complejo; así se establece una jerarquía característica dentro de la red CNN.

7.3.2.5 Capa de agrupamiento

La agrupación de capas, conocida también como reducción de muestreo, disminuye la dimensionalidad y, por ende, el número de parámetros en la entrada. De manera análoga a la capa convolucional, esta operación aplica un filtro sobre toda la entrada; sin embargo, a diferencia del anterior, este filtro no cuenta con pesos. En su lugar, el *kernel* utiliza una

función de agregación sobre los valores presentes en el campo receptivo para llenar la matriz de salida. Existen dos tipos principales de agrupación:

- **Agrupación máxima:** mientras el filtro se desplaza a lo largo de la entrada, selecciona el píxel que posee el valor máximo para incluirlo en la matriz de salida. Este método es más comúnmente utilizado que la agrupación promedio.
- **Agrupación promedio:** al igual que en el caso anterior, el filtro se mueve a través de la entrada y calcula el valor promedio dentro del campo receptivo antes de enviarlo a la matriz de salida.

A pesar de que se pierde una cantidad considerable de información en esta capa, aporta varios beneficios a las redes neuronales convolucionales (CNN). Contribuye a reducir la complejidad del modelo, mejora su eficiencia y ayuda a mitigar el riesgo de sobreajuste.

7.3.2.6 Capa totalmente conectada

El término "capa totalmente conectada" es bastante descriptivo. Como se mencionó anteriormente, en las capas parcialmente conectadas, los valores de píxeles de la imagen de entrada no tienen una conexión directa con la capa de salida. En contraste, en la capa totalmente conectada, cada nodo en la capa de salida se conecta directamente a un nodo de la capa anterior (Goker, H., 2023).

Esta capa es responsable de realizar la clasificación basada en las características que han sido extraídas por las capas anteriores y sus diversos filtros. Mientras que las capas convolucionales y agrupadas comúnmente emplean funciones *ReLU*, las capas completamente conectadas (FC) generalmente utilizan una función de activación *softmax* para clasificar adecuadamente las entradas, generando probabilidades entre 0 y 1. La función *softmax* convierte el vector de salidas producido por la red en una distribución de probabilidad, lo que permite interpretar los resultados como la probabilidad de que una imagen pertenezca a una clase particular.

7.4 Tipos de redes neuronales convolucionales

7.4.1 LeNet

Desarrollada por Yann LeCun y su equipo, se logró implementar una red neuronal convolucional que puede reconocer caracteres mediante la aplicación de los principios de retro propagación y alimentación hacia adelante (Figura 10). Entre las características más notables se encuentran (Swapna, M. et al., 2020):

- Un elevado número de capas ocultas.

- Múltiples mapas de unidades replicadas en cada capa.
- Agrupación de salidas que pertenecen a unidades replicadas cercanas.

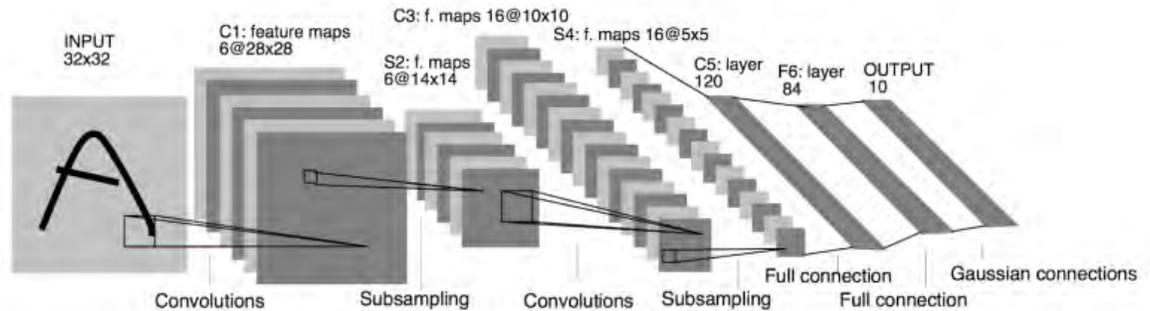


Figura 10: Red Neuronal Convolutiva *LENET*

7.4.2 AlexNet

Ha tenido un impacto significativo en el área del aprendizaje automático, especialmente en la implementación del aprendizaje profundo en la visión por computadora. Su arquitectura es bastante similar a la de *LeNet*, aunque se trata de una red más profunda, con un mayor número de filtros por cada capa y capas convolucionales apiladas. Los aspectos destacados de esta red incluyen:

- La utilización de la función de activación *ReLU* en lugar de *Tanh* para evitar añadir linealidad. Esto resulta en un aumento aproximado de seis veces en la velocidad del proceso manteniendo la misma precisión.
- La adopción de *dropout* como método para abordar el sobreajuste, en lugar de utilizar regularización. Sin embargo, esto provoca que el tiempo de entrenamiento se duplique cuando se aplica una tasa de *dropout* del 0.5.

La arquitectura presentada en la figura contiene ocho capas junto con sus respectivos pesos (Figura 11); las primeras cinco son capas convolucionales y las tres restantes son completamente conectadas. La función de activación *ReLU* se activa inmediatamente después de cada capa completamente conectada y cada capa convolucional (Swapna, M. et al., 2020). El *dropout* se implementa antes de la primera y segunda capa totalmente conectada. Esta red tiene un total de 62,3 millones de parámetros y requiere 1,1 millones de unidades computacionales para realizar una pasada hacia adelante.

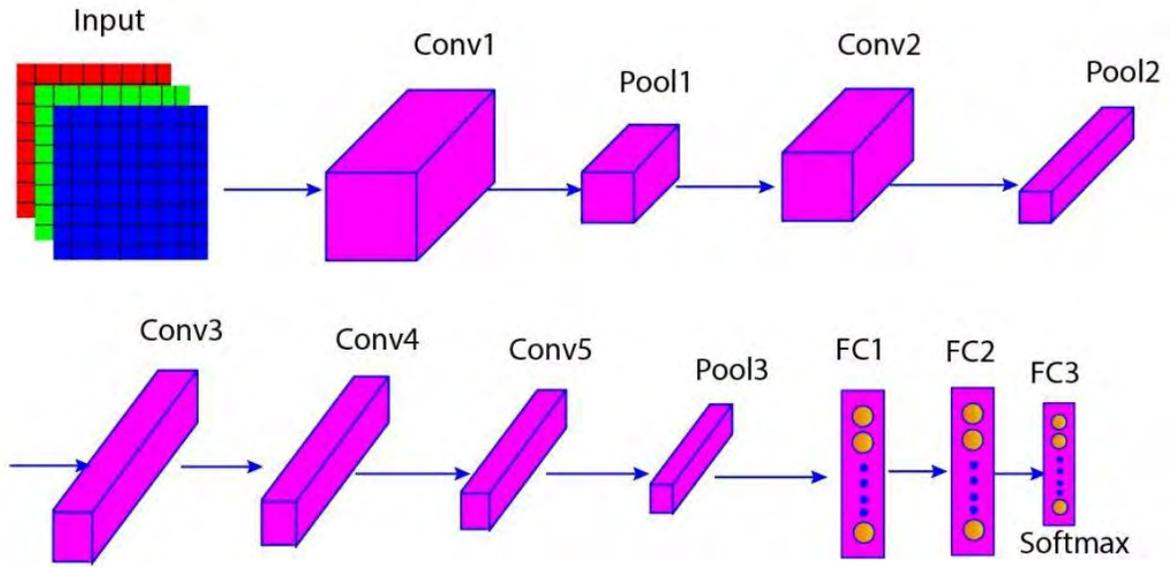


Figura 11: Red Neuronal Convolucional ALEXNET

7.4.3 VGG

Esta red fue creada por el Grupo de Geometría Visual de la Universidad de Oxford. Existen varias variantes bajo este nombre que se diferencian por la cantidad de capas que incluyen. Es una arquitectura similar a AlexNet, pero utiliza únicamente capas de convolución de 3x3 y cuenta con un gran número de filtros (Yang, F., 2019). Las variantes más utilizadas son:

- **VGG-16:** Consta de un total de 16 capas, incluyendo 13 capas de convolución, 2 completamente conectadas y una capa *softmax* al final para la clasificación (Figura 12).
- **VGG-19:** Compuesta por un total de 19 capas, donde hay 16 capas de convolución, 2 totalmente conectadas y una capa *softmax* final para clasificar (Figura 13).

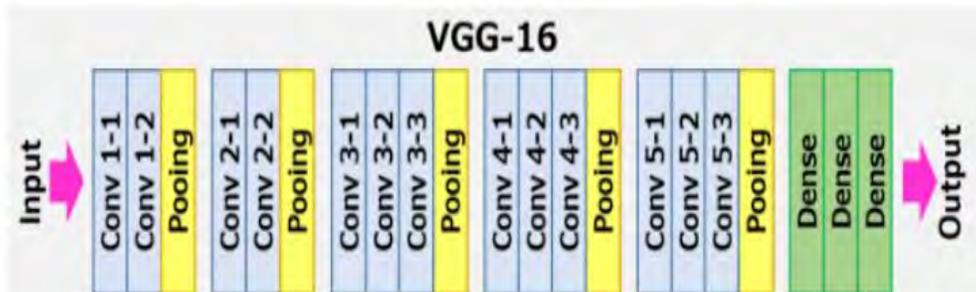


Figura 12: Red Neuronal Convolucional VGG-16

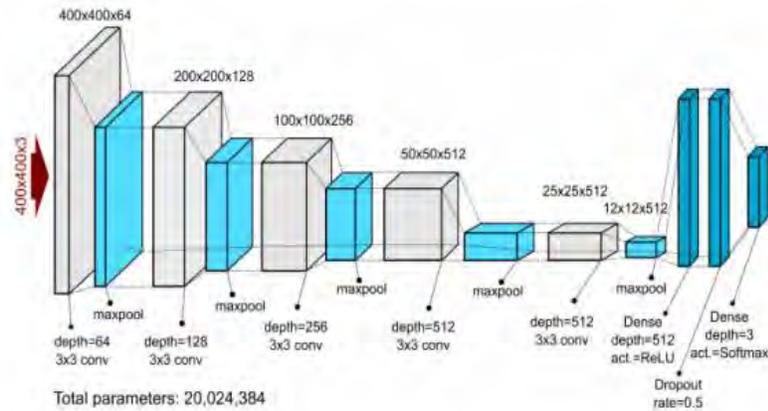


Figura 13: Red Neuronal Convolutional VGG-19

7.4.4 Google Net

Está formada por 22 capas y presenta aproximadamente 12 veces menos parámetros que, por ejemplo, *AlexNet*, lo que la convierte en una opción significativamente más rápida y precisa. *GoogleNet* incluye sus capas de inicio o *inception layers*. La idea fundamental detrás de estas capas es abarcar un área más amplia mientras se mantiene una buena resolución para obtener información detallada sobre las imágenes (Figura 14). Se busca que un conjunto de filtros de diferentes tamaños pueda gestionar mejor las escalas de diversos objetos, beneficiándose del hecho de que todos los filtros en la capa de inicio son más accesibles para el aprendizaje (Swapna, M. et al., 2020). Una manera efectiva de mejorar el rendimiento en el aprendizaje profundo es incrementar el número de capas y datos utilizados. Sin embargo, esto conlleva el riesgo de sobreajuste debido a la mayor cantidad de parámetros; por lo tanto, para prevenir dicho fenómeno desde el principio, se implementan diversas técnicas de cuello de botella.

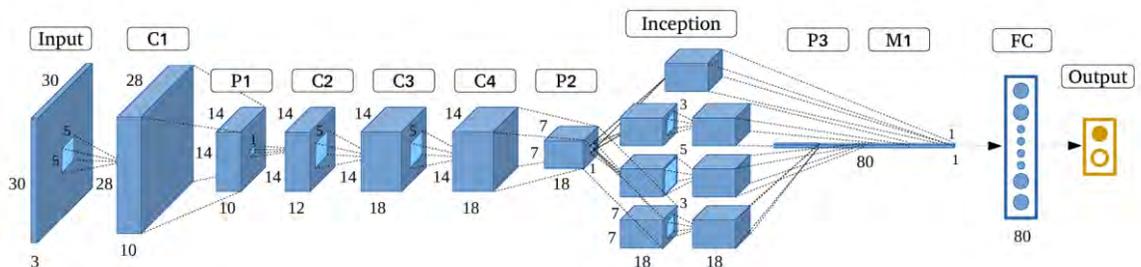


Figura 14: Red Neuronal Convolutional GOOGLE NET

7.4.5 ResNet

Las redes residuales tienen la capacidad de aprender funciones más complejas, lo que puede resultar en un rendimiento superior (ver Figura 15). No obstante, en algunas ocasiones, la adición de más capas puede perjudicar el rendimiento final. Este fenómeno se conoce como el problema de degradación, que hace referencia a que, aunque las técnicas avanzadas de inicialización de parámetros y la normalización por lotes facilitan la convergencia de redes más profundas, a veces estas pueden llegar a tasas de error más elevadas en comparación con redes menos profundas (Law, H. et al., 2018).

Para abordar esta situación, se sugiere implementar bloques residuales donde las capas intermedias dentro de un bloque aprenden una función residual en relación con la entrada del bloque. Esta función puede considerarse un proceso de refinamiento en el que se optimiza el mapa de características de entrada para mejorar su calidad. Esto contrasta con una red completa, donde se espera que cada capa adquiera nuevas características del mapa (Swapna, M. et al., 2020). En situaciones donde no es necesario realizar un refinamiento, las capas intermedias pueden gradualmente ajustar sus pesos hacia cero, haciendo que el bloque residual actúe como una función identidad.

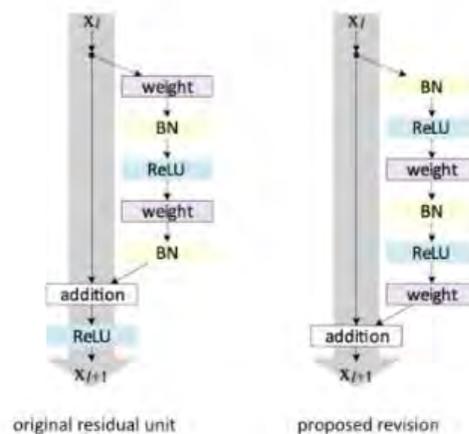


Figura 15: Red Neuronal Convolutacional *RESNET*

8. Trabajos relacionados a la investigación

En esta sección se describen los estudios previos y trabajos relacionados que han constituido el fundamento teórico y práctico de la investigación presentada en este estudio. Estas investigaciones anteriores no solo han guiado la dirección y enfoque del trabajo actual, sino que también han sido cruciales para el desarrollo de la metodología y la implementación de las soluciones tecnológicas sugeridas.

La mayoría de estos estudios se centran en la detección y análisis del cáncer cervicouterino a través del uso de clasificadores y técnicas de inteligencia artificial, con un énfasis particular en el aprendizaje profundo (*Deep Learning*) y el aprendizaje automático (*Machine Learning*). Estas metodologías han demostrado ser altamente efectivas en el reconocimiento de patrones en imágenes médicas, facilitando una detección más rápida y precisa de lesiones o anomalías en comparación con los métodos tradicionales (Yuo, H, 2021).

Como se mencionó anteriormente, la identificación del cáncer cervicouterino mediante el análisis de imágenes médicas representa un método eficiente, ampliamente validado y con un impacto significativo tanto en contextos clínicos como investigativos. Sin embargo, este trabajo propone una mejora respecto a enfoques anteriores: no solo se encarga de clasificar imágenes para detectar la presencia de cáncer, sino que también localiza visualmente las áreas que exhiben patrones malignos. Esto permite ofrecer una verificación visual del diagnóstico, lo cual constituye una ventaja significativa al validar los resultados obtenidos.

Para proporcionar un contexto más claro y facilitar comparaciones sobre estos avances, se ha elaborado la Tabla 1. Esta tabla resume estructuradamente los estudios más relevantes en el área, abarcando las técnicas utilizadas, los conjuntos de datos empleados, los resultados alcanzados y las principales contribuciones de cada investigación. Así, sirve como referencia para resaltar cómo este estudio se sitúa en relación con investigaciones previas, enfatizando tanto las similitudes como las innovaciones presentadas.

Título del recurso analizado	Idea principal
Cervical cancer classification using efficient net and fuzzy extreme machine learning (Suphalakshmi, et al., 2022)	Detección de patrones malignos en estudios de Papanicolau utilizando algoritmos de Machine Learning
Detection and classification of cervical cancer images using CEENET deep learning approach (Subarna & Sukumar ,2022).	Clasificador de imágenes positivas de cáncer cérvico uterino utilizando algoritmos de Deep Learning y CEENET
Dose Prediction for Cervical Cancer Brachytherapy Using 3-D Deep Convolutional Neural Network (Fahimian, et al., 2022)	Método de predicción de dosis para braquiterapia de cáncer de cuello de útero mediante una red neuronal convolucional profunda (CNN) para guiar la planificación del tratamiento
Classification of Cervical Cancer Using Artificial Neural Networks (Devi, M. A. et al., 2016).	Clasificador mediante redes neuronales para detección de cáncer cérvico uterino a través de colposcopias.
Cervical cancer classification using convolutional neural networks and extreme learning machines (Ghoneim, A et al, 2020)	Sistema basado en aprendizaje profundo para detectar y clasificar células cancerosas cervicales
Artificial Intelligence based Cervical Cancer Risk Prediction Using M1 Algorithms (H, N. C., Sai et al., 2022)	Sistema basado en algoritmo de redes neuronales convolucionales para detección de riesgo de cáncer cervicouterino mediante la clasificación de colposcopias
Deep Learning System for the Automatic Classification of Normal and Dysplastic Peripheral Blood Cells as a Support Tool for the Diagnosis (Acevedo Lipes, A. M., 2021)	Sistema basado en algoritmo de redes neuronales para clasificar automática de células sanguíneas periféricas normales y displásicas
Detection of cervical cancer from uterine cervix images using transfer learning architectures (Goker Hanife, 2024)	Detección de cáncer cervicouterino a través del algoritmos de RNC (Comparación de resultados entre Resnet, GoogleNet, VGG16 y AlexNet)
Predicting cervical cancer outcomes: statistics, images, and machine learning. Frontiers in Artificial Intelligence (Luo W., 2021)	Estudio de cáncer cervicouterino mediante técnicas estadísticas y a aprendizaje automático
Evolving Deep Convolutional Neural Networks for Image Classification (Sun, Y. et al., 2020)	Clasificador con redes neuronales convolucionales y redes neuronales recurrentes

Automatic Diagnosis Of Rheumatoid Arthritis From Hand Radiographs Using Convolutional Neural Networks (Betancourt-Hernández, M. et al., 2018)	Análisis de artritis reumatoide a través de imágenes. Identificación de zonas afectadas con redes neuronales convolucionales
Artificial Intelligence In Cervical Cancer Screening And Diagnosis (Hou, X. et al., 2022)	Uso de diferentes algoritmos de redes neuronales convolucionales para diagnóstico en imágenes de cáncer cervicouterino a través de células
Cervical cancer (Fowler, J. R., 2023)	Teoría de la inteligencia artificial aplicada al tratamiento de cáncer cervicouterino
Dense Convolutional Neural Network for Detection of Cancer from CT Images (Sreenivasu, S. V. N et al., 2022)	Redes neuronales convolucionales aplicadas a imágenes celulares de cuello uterino
Cancer Diagnosis Using Deep Learning (Munir, K. et al., 2019)	Uso de algoritmos de aprendizaje profundo para verificar zonas de lesiones cancerosas en cuello uterino a través de imágenes endocervicales
Advancement in Deep Learning Methods for Diagnosis and Prognosis of Cervical Cancer (Gupta, A. et al., 2022)	Uso de aprendizaje automático para diagnóstico de cáncer cervicouterino a través de datos históricos de pacientes.

Tabla 1: Trabajos relacionados al tema de investigación propuesto

9. Desarrollo

El análisis de imágenes médicas y los sistemas inteligentes son fundamentales en la identificación de células cancerosas. Con el avance de nuevas metodologías, estos procesos han demostrado ser más eficientes y requerir menos tiempo en comparación con técnicas tradicionales como la citología vaginal, colposcopia y cervicografía (Zing, D, 2018). Aunque estas nuevas técnicas no dependen de la experiencia humana, es importante destacar que no pueden reemplazar la evaluación subjetiva realizada por médicos especializados, aunque sí ofrecen un apoyo significativo.

Las tecnologías avanzadas de aprendizaje automático y las comunicaciones inalámbricas han facilitado la creación de un sistema integral para diagnóstico médico que opera en tiempo real con alta precisión y mínima intervención humana. Sin embargo, persisten

desafíos como la pérdida de datos durante la transmisión, el amplio ancho de banda requerido para enviar vídeos médicos y la necesidad de algoritmos robustos que manejen las variaciones en los datos (Hossain, M., 2020).

Recientemente, el aprendizaje profundo ha mejorado notablemente la exactitud en diversas aplicaciones. Su alta eficacia lo ha posicionado como una técnica líder en el ámbito del aprendizaje automático. Se han realizado estudios significativos sobre su aplicación en la detección de distintos tipos de células cancerosas. Este enfoque ha demostrado ser eficaz en áreas como la identificación de anomalías electroencefalográficas (EEG), clasificación ambiental, localización de nódulos pulmonares, así como en el diagnóstico del cáncer de mama y piel, análisis de imágenes médicas, reconocimiento emocional a través del audio y predicción de enfermedades. La creciente variedad de sensores ha añadido una nueva dimensión al procesamiento masivo de datos (*Big Data*), que se ha aplicado exitosamente en contextos médicos (Hossain, M., 2020).

Para implementar esto se utilizarán redes neuronales convolucionales mediante la biblioteca *Keras* en *Python* debido a sus capacidades para mejorar el tratamiento y reconocimiento de imágenes y objetos, además de poder identificar áreas afectadas si hay presencia cancerosa. También se aplicarán técnicas específicas para el procesamiento visual y métricas precisas a fin de evaluar el modelo desarrollado.

Los autores mencionados están enfocados en encontrar soluciones efectivas al cáncer cervicouterino utilizando técnicas avanzadas del aprendizaje automático, específicamente mediante aprendizaje profundo. Por ejemplo, Park y su equipo sostienen que estas metodologías permiten abordar problemas complejos utilizando conjuntos de datos variados e interconectados (Park, Y.R. et al., 2021). El cerebro humano es un sistema extremadamente complejo capaz de realizar billones de cálculos rápidamente; por ello se han creado redes neuronales que simulan su funcionamiento. Estas máquinas procesan información para reconocer patrones, clasificar elementos y ejecutar cálculos complejos sobre grandes volúmenes de información (Gupta, A., 2022).

De esta manera surgen las redes neuronales convolucionales gracias a sus capas densas que optimizan el procesamiento informático; su adaptabilidad ante variaciones en los datos les confiere ventajas sobre otros modelos tradicionales del aprendizaje automático. Modelos destacados incluyen *Inception (Google)*, *ResNet*, *Vgg16* y *Xception*. En un estudio

comparativo se evaluó el modelo *ResNet-50* junto con *Inception (Google)*, *Vgg16* y *Xception* contra modelos clásicos del aprendizaje automático como *Extreme Gradient Boosting (XGBoost)*, *Support Vector Machine (SVM)* y *Random Forest*; los hallazgos mostraron que estos modelos lograron detectar cáncer cervicouterino a partir imágenes cervicográficas con mayor precisión que los métodos convencionales mencionados (Jusman, Y., et al., 2014). Los resultados fueron alentadores ya que el algoritmo aplicado presentó una sensibilidad de especificidades aproximadas del 91.7% por imagen para estos modelos.

9.1 Descripción de los datos correspondientes a utilizar

Los datos que alimentarán el modelo se obtienen del repositorio de *Intel en Kaggle* (Intel & MobileODT, 2020), donde las imágenes están clasificadas en tres categorías relacionadas con el cuello uterino, tomadas mediante herramientas colposcópicas. Este recurso divide las imágenes en cancerosas, no cancerosas y precoces o sospechosas, teniendo en cuenta también el ángulo de captura y la proximidad al componente. Tras describir cómo está constituido el conjunto de datos y sus categorías (que se analizarán más a fondo más adelante), es fundamental entender la representación anatómica del cuello uterino, que se detalla a continuación:

La representación anatómica del sistema reproductor femenino suele observarse durante un examen ginecológico con la ayuda de un espéculo, que permite visualizar tanto los componentes internos como externos del cuello uterino (Figura 16):

- **Útero:** Es un órgano muscular con forma de pera invertida donde se desarrolla el feto durante el embarazo. Se localiza en la pelvis femenina entre la vejiga y el recto.
- **Cérvix:** Esta es la parte inferior estrecha del útero que se extiende hacia la vagina. Funciona como un canal entre ambos órganos y durante el parto se dilata para facilitar el paso del bebé desde el útero.
- **Vagina:** Se trata de un canal muscular que conecta el cérvix con el exterior del cuerpo. Actúa como canal de nacimiento y también permite la salida del flujo menstrual.
- **Os:** Este es el orificio cervical que comunica al cérvix con la vagina. Durante el ciclo menstrual, permite la salida del flujo menstrual; además, durante la ovulación, facilita

el paso de los espermatozoides hacia el útero. En el momento del parto, este orificio se dilata considerablemente para permitir que nazca el bebé (Munir, K., et al., 2019).

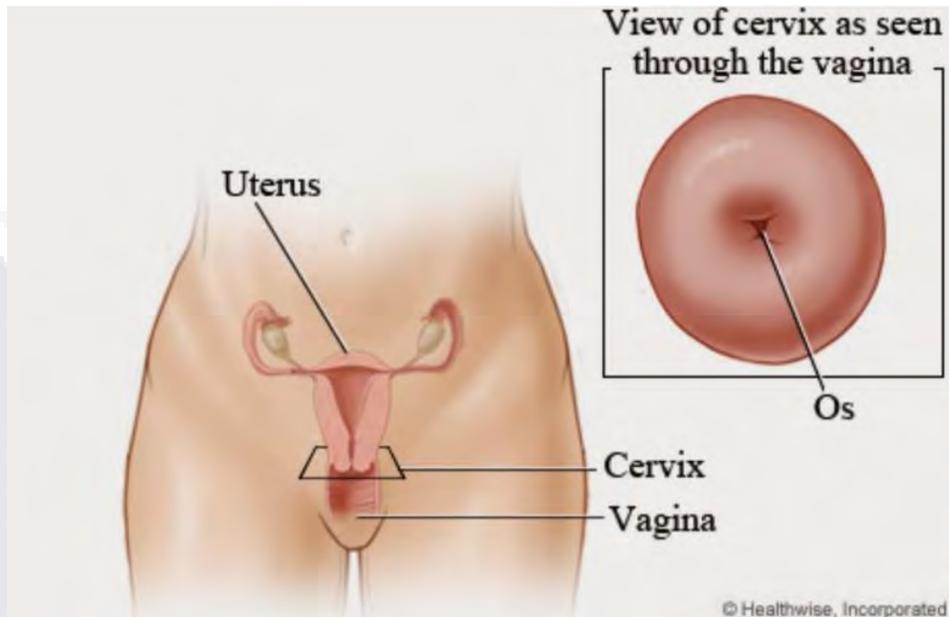


Figura 16: Anatomía del cuello uterino

Enfocándose en la región del cuello uterino, a continuación, se presenta una imagen colposcópica que resalta distintos tipos de células presentes en esta zona, así como la unión escamocolumnar. Se proporciona una descripción de los elementos destacados en la imagen:

- **Células escamosas:** Estas son células planas que constituyen la superficie externa del cérvix y la vagina. Representan el tipo más común de células en la parte externa del cérvix.
- **Células endocervicales o columnares:** Localizadas en el canal endocervical, que es la sección interna del cérvix que conecta con el útero. Estas células son más alargadas y estrechas en comparación con las escamosas.
- **Unión escamocolumnar (SCJ):** Este es el área donde las células escamosas del ectocérvix (parte externa) se encuentran con las células columnares del endocérvix (canal cervical). La unión escamocolumnar es un punto crucial para la detección de cáncer cervical, ya que es una región donde pueden surgir cambios precoces asociados al cáncer.

La imagen ofrece una vista detallada del cérvix mediante un colposcopio (Figura 17), un dispositivo utilizado para examinar minuciosamente el cuello uterino, la vagina y la vulva. Durante este procedimiento, es posible identificar áreas anormales y realizar biopsias si se considera necesario para evaluar posibles alteraciones precoces o malignas (Intel & MobileODT, 2020).

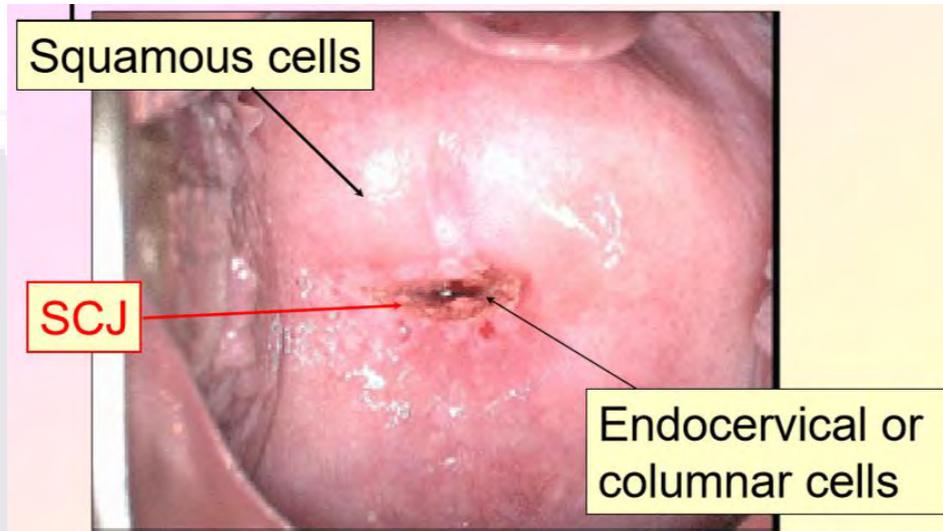


Figura 17: Composición anatómica interna del cuello uterino

El conjunto de datos que se utilizará para la implementación estará compuesto por imágenes colposcópicas del cuello uterino, las cuales estarán etiquetadas con diversos tipos de células y características anatómicas esenciales. Estas imágenes capturarán detalles sobre las células escamosas, las células endocervicales o columnares, así como la unión escamocolumnar (SCJ). La finalidad del conjunto de datos es ofrecer información visual precisa y bien anotada, que servirá para entrenar modelos de aprendizaje automático enfocados en la detección de anomalías cervicales y en la identificación de los diferentes tipos celulares presentes en el cérvix.

Los datos recopilados abarcan imágenes de colposcopia que se clasifican en tres categorías: Tipo 1 (normal), Tipo 2 (precoces) y Tipo 3 (imágenes no adecuadas para evaluación debido a problemas como mala iluminación o enfoque incorrecto). El objetivo es desarrollar modelos de aprendizaje automático capaces de clasificar estas imágenes de manera efectiva, lo que permitirá un diagnóstico más rápido y preciso (Figura 18). Este esfuerzo tiene como meta mejorar las herramientas disponibles para la detección y prevención del cáncer cervical.

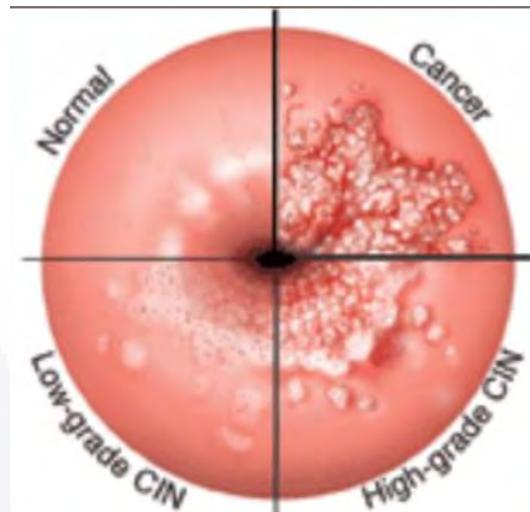


Figura 18: Clasificación de daño en cuello uterino

Tipo 3 o normal: Imágenes que no muestran signos de anomalías o enfermedades precancerosas o cancerosas. Además, son completamente ectocervicales, es decir, que la zona es visible sin ningún daño (Figura 19)



Figura 19: Cuello uterino sin lesiones aparentes

Tipo 2 o sospechosas: Imágenes que presentan signos de lesiones o cambios celulares que podrían evolucionar a cáncer. Además, tienen componentes endocervicales que son visibles y podrán contener componentes ectocervicales largos o pequeños (Figura 20). Pueden confundirse con otras patologías debido a los bultos anormales que se presentan o a alguna afección acompañada de sangrado.



Figura 20: Cuello uterino con probabilidad de cáncer

Tipo 1 o cancerosas: Imágenes que muestran signos que son altamente indicativos de cáncer, requiriendo intervención médica urgente (Figura 21). Se tienen componentes endocervicales que no son completamente visibles. Usualmente se presentan anomalías y en ocasiones bultos no normales dentro de la anatomía de este. Es importante evitar retrasos en la evaluación y tratamiento de este tipo de lesiones, ya que su progresión puede ser rápida y comprometer órganos vecinos.

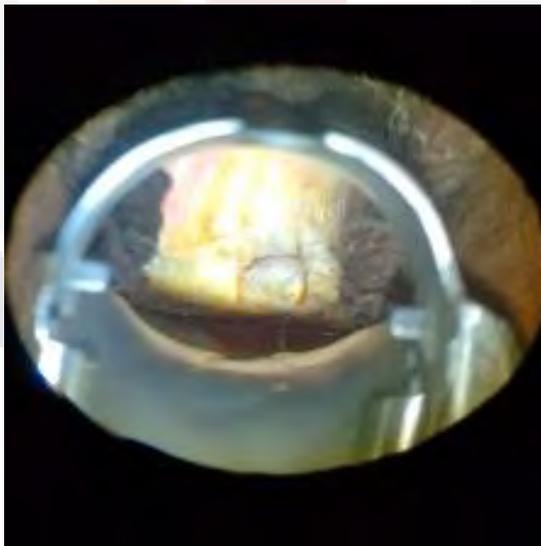


Figura 21: Cuello uterino con signos cancerosos

9.2 Visualización de datos y exploración

Se va a utilizar una carpeta de entrenamiento y una de pruebas con imágenes para el modelo, con tamaños de imagen de 224 x 224 píxeles. Para las imágenes de entrenamiento la división es entre las tres clases correspondientes que se mencionaron en la sección anterior (Figura 22), y para las imágenes de prueba están sin categorizar ya que son las imágenes que se espera que el modelo las clasifique de acuerdo con la patología o no presentada.



Figura 22: Visualización de imágenes categorizadas en el conjunto de entrenamiento

Es importante mencionar que se realiza la creación de una estructura de datos para poder visualizar la información de acuerdo con la clasificación que se considera con respecto a la carpeta de entrenamiento antes mencionada. Tras haber realizado esto, se procede a verificar la descripción de detalles básicos de los datos, como lo es cuántas imágenes se va a tomar para el entrenamiento, cuántas categorías existen en los mismos y de cuál de los tipos de imágenes que tienen las clasificaciones se presenta con mayor frecuencia (figura 23). Dicha estructura ayuda a poder ver la información de manera estructurada y que se puedan presentar en el mismo el nombre del archivo y su etiqueta o categoría a la que esta pertenece (Gupta, A., 2022).

	filepath	label
count	1481	1481
unique	1481	3
top	D:/respaldo/todo/tesis_implementacion/kaggle\t...	Type 2
freq	1	781

Figura 23: Descripción básica de los datos (totales, clases, ubicación de los archivos y frecuencia de una clase)

Siguiendo con la implementación de la sección correspondiente a la visualización de los datos se revisan imágenes duplicadas y archivos dañados, es decir, imágenes que contienen un ruido inminente, han sido cortadas o alteradas y que presentan un formato incompatible que no serían de buena ayuda para el entrenamiento de nuestros modelos incluso se le aplique un preprocesamiento. Como se pudo mostrar, las imágenes que se toman en cuenta cumplen con los requisitos mínimos para la implementación, categorizadas cada una de ellas y como se puede observar, mostrando el cuello uterino de una forma perceptible para el ojo humano y por lo tanto entendible para la red neuronal convolucional por implementar. Al ver la división de nuestros datos en las categorías mencionadas, se visualiza que se va a trabajar con estos tipos y esta frecuencia (Figura 24)

	label	Num_Values
0	Type 2	4346
1	Type 3	2426
2	Type 1	893

Figura 24: División de los datos y frecuencia de estos en todo el conjunto de datos

Como siguiente paso aquí en la sección de la exploración se realiza la distribución de clases para el conjunto de datos de entrenamiento, esto es muy útil para saber si las clases están

equilibradas y la frecuencia o el número de datos que cada una maneja (Munir, k., 2019). Esto se hace mediante la implementación de un gráfico de barras (Figura 25)

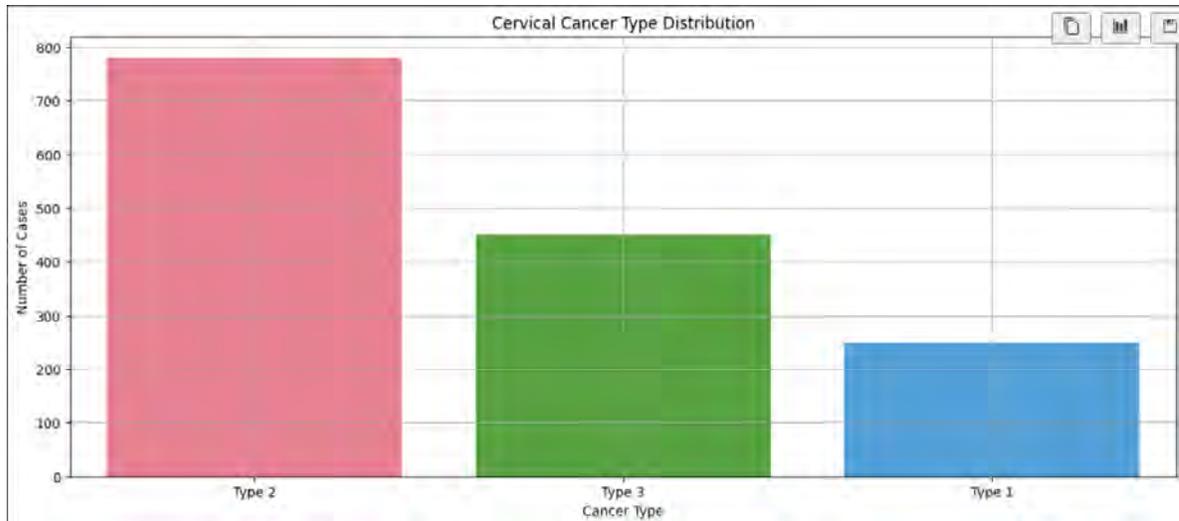


Figura 25: Gráfico de barras para las tres clases correspondientes (precanceroso, canceroso y no canceroso)

Como podemos observar la frecuencia que más sobresale es para el tipo 2 de la clasificación de las imágenes, con lo que es de consideración en el sesgo que existe entre una clase y otra. Las cantidades no son extremadamente altas, así que se puede dejar con esa distribución para el entrenamiento correspondiente con los modelos por implementar.

Aunado a esto, para clarificar la distribución de los datos, se realiza un gráfico de pastel, que puede mostrar el porcentaje de los tres tipos y de igual manera visualizar que el tipo 2 incluye la mayoría de los datos con un 52% del total de todos los datos de la muestra para el entrenamiento del modelo posterior (Figura 26). Como se puede observar en la siguiente figura, así se representan los demás porcentajes y la mayor frecuencia en el segundo tipo. En el caso del cáncer, la clasificación del riesgo puede ser un desafío, ya que en algunas mujeres el virus sobrevive durante años y, en última instancia, hace que las células cervicales se conviertan en células cancerosas.

Pie Plot of Type Distribution

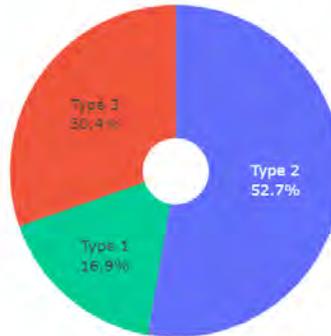


Figura 26: Gráfico de porcentajes para las tres clases de los datos presentados

Por otro lado, siguiendo con las implementaciones que este apartado estipula, también se realiza la exploración estadística de píxeles y color, ya que se verifica si es necesario normalizar los valores de los píxeles y la distribución de color en las imágenes. A continuación, se enumera el proceso matemático que se utiliza y que se explica de una manera detallada:

Cálculo de la media de los píxeles: Para cada imagen la convierte en un array de píxeles, donde cada pixel es un valor numérico (o un conjunto de tres valores para imágenes RGB). Luego la media de los valores se calcula de la siguiente manera (Figura 23)

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i$$

Donde:

- μ es la media
- x_i es el valor del i -ésimo píxel
- N es el número total de píxeles en la imagen

Cálculo de la desviación estándar: Para cada imagen, se calcula la desviación estándar de los valores de los píxeles. Esta mide cuán dispersos están los valores de los píxeles con respecto a la media (Figura 24)

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$$

Donde:

- σ es la desviación estándar
- x_i es el valor de i -ésimo pixel
- μ es la media de los valores de los píxeles
- N es el número total de píxeles

Finalmente, el código calcula el promedio de las medias y de las desviaciones estándar para todas las imágenes del conjunto de entrenamiento utilizando:

Media global de las medias de los píxeles:

$$\mu_{global} = \frac{1}{M} \sum_{j=1}^M \mu_j$$

donde M es el número de imágenes

Media global de las desviaciones estándar de píxeles:

$$\sigma_{global} = \frac{1}{M} \sum_{j=1}^M \sigma_j$$

donde M es el número de imágenes

Se obtiene un resultado de media para el valor de los píxeles de 108.76 unidades y para la desviación estándar de 62.39 unidades, lo que significa que tienden a ser un poco más oscuras, por lo tanto, para la desviación estándar implica que hay una amplia variedad de valores de píxeles en las imágenes. Si la desviación estándar fuera baja, significa que los valores de los píxeles están más agrupados alrededor de la media, y si fuera alta, significa que hay más contraste en las imágenes. Se requiere un proceso de normalización que ayudará a ajustar el modelo y que aprenda de mejor manera en la fase de entrenamiento posterior.

9.3 Preprocesamiento de las imágenes

Cuando se trabaja con datos de imágenes del cuello uterino, es común realizar una serie de pasos de preprocesamiento antes de construir modelos de aprendizaje automático. Estos pasos son fundamentales para corregir imperfecciones, verificar las dimensiones, y realizar otros procesos esenciales que aseguran la calidad de los datos. Este preprocesamiento es crucial porque las imágenes pueden presentar variaciones significativas que podrían afectar el rendimiento del modelo. Para realizar este trabajo de manera eficiente, se utilizan librerías especializadas en procesamiento de imágenes, como *OpenCV*, que es ampliamente reconocida por sus herramientas poderosas para manipulación y análisis de imágenes.

9.3.1 Redimensionamiento y procesamiento de imágenes en paralelo

Una parte importante del preprocesamiento es leer y ajustar las imágenes de manera eficiente. Aunque el tamaño de la base de datos de imágenes puede no ser masivo, procesar cada imagen de forma secuencial implicaría un costo computacional elevado y tiempo de espera considerable. Para abordar este problema, se utiliza el procesamiento en paralelo, que permite ejecutar múltiples tareas al mismo tiempo, optimizando el uso de los recursos del sistema. En este caso, se implementa un *ThreadPoolExecutor* que distribuye las tareas de preprocesamiento en diferentes hilos, acelerando así el tiempo total requerido para el procesamiento.

El redimensionamiento de imágenes es un paso específico en el que cada imagen se ajusta a un tamaño uniforme de 224 x 224 píxeles, necesario para que el modelo pueda procesarlas correctamente. Este tamaño es común en redes neuronales pre entrenadas como las arquitecturas de *Deep Learning* (por ejemplo, *VGG* o *ResNet*). Se utiliza interpolación cúbica durante el redimensionamiento, un método avanzado que calcula los nuevos valores de los píxeles basándose en una función cúbica, lo que permite que las imágenes redimensionadas mantengan detalles y suavidad.

Después del proceso de preprocesamiento, se visualizan ejemplos de las imágenes en las tres categorías definidas, seleccionadas de manera aleatoria. Se ajustan las imágenes para mostrarlas sin los ejes y con títulos que indican la categoría o clase correspondiente (Figura 27).

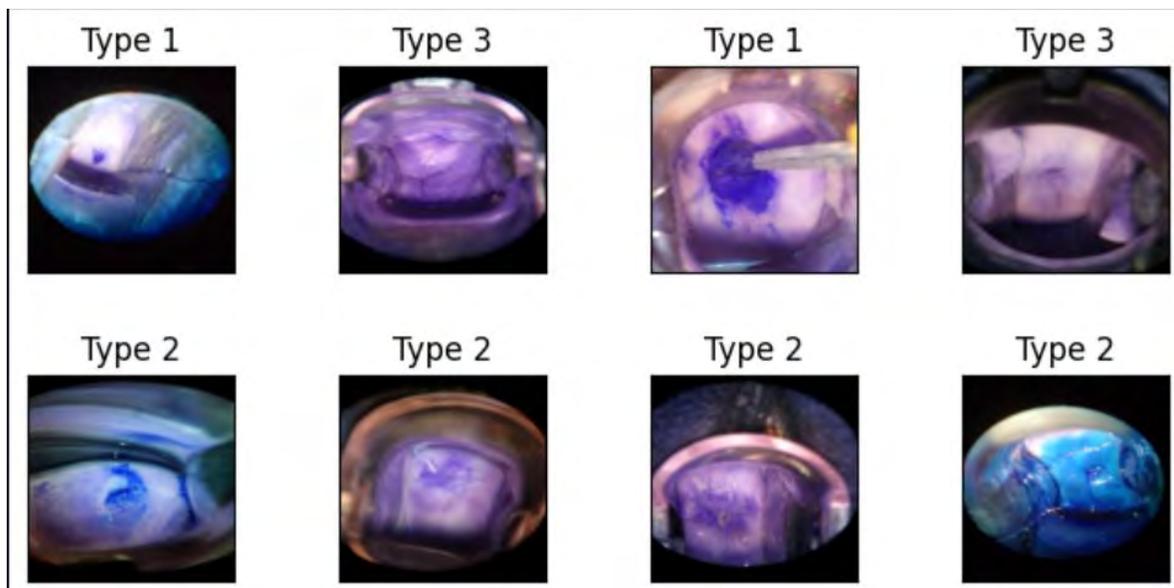


Figura 27: Visualización de imágenes después del preprocesamiento

9.3.2 Normalización y escalado de imágenes

La normalización y el escalado son pasos esenciales que se realizan antes de entrenar el modelo. Primero, se establece un tamaño de lote de 64, lo que significa que el modelo procesa 64 imágenes a la vez durante el entrenamiento, un tamaño que es eficiente para el procesamiento paralelo en *GPUs*. Luego, se define que hay tres clases posibles para la clasificación, asegurando que la red neuronal conozca el número total de categorías que debe distinguir.

El proceso de escalado consiste en normalizar los valores de los píxeles para que se encuentren en un rango de $[0, 1]$, en lugar de su rango original de $[0, 255]$. Matemáticamente, esto se expresa como:

$$\text{valor normalizado} = \frac{\text{valor original}}{255}$$

Posteriormente, las etiquetas de las imágenes, que inicialmente están en formato de texto (como "Type 1", "Type 2", etc.), se convierten en valores numéricos mediante un proceso de codificación de etiquetas. Por ejemplo, "Type 1" se convierte en 0, "Type 2" en 1, y así sucesivamente. Este paso es necesario para que las etiquetas puedan ser entendidas por el modelo.

Finalmente, se aplica el algoritmo de codificación *one-hot*, que transforma las etiquetas numéricas en vectores binarios. Por ejemplo, si hay tres clases, la etiqueta 0 se convierte en [1, 0, 0], la etiqueta 1 en [0, 1, 0], y la etiqueta 2 en [0, 0, 1]. Esta representación es especialmente útil para modelos de redes neuronales porque evita que el modelo interprete una relación de orden entre las clases, lo cual no existe. Las redes neuronales funcionan mejor con valores numéricos continuos y codificaciones como *one-hot*, ya que aseguran que cada clase se trate de manera independiente, sin suposiciones erróneas sobre su relación matemática.

9.4 Implementación de los modelos de redes neuronales convolucionales seleccionados

En este análisis se llevará a cabo una comparativa exhaustiva de diferentes modelos de redes neuronales convolucionales (CNN) como se pudo explicar al inicio de la sección de desarrollo, específicamente *XceptionNet*, *VGG 16*, *Inception Net* y *Resnet*. El objetivo es evaluar y contrastar sus resultados en términos de precisión y rendimiento en tareas de clasificación de imágenes para poder detectar anomalías que favorezcan la presencia de cáncer en colposcopias de cuello uterino.

La implementación de estos modelos se realizó en un entorno de cómputo configurado con un equipo Lenovo, equipado con una tarjeta gráfica *NVIDIA GeForce GTX 1650*, un procesador *Intel Core i5-10300H*, 16 GB de memoria RAM, y sistema operativo *Windows 10*. Para optimizar el uso de los recursos computacionales y mejorar el rendimiento en el entrenamiento y evaluación de los modelos, se implementó el uso de *ThreadPool*, una técnica de programación concurrente que permite distribuir de manera eficiente las tareas de procesamiento en múltiples hilos.

A medida que la inteligencia artificial y el aprendizaje profundo continúan avanzando, es fundamental entender cómo estos modelos se desempeñan bajo diversas condiciones y conjuntos de datos, lo que proporcionará información valiosa para la selección del modelo más adecuado en aplicaciones prácticas (Ramaham, M., et al., 2021).

9.4.1 VGG 16

Iniciando con este tipo de modelo de red neuronal convolucional, se carga la arquitectura de este con pesos pre entrenados en el conjunto de datos de *ImageNet*. Esto significa que

el modelo tendrá conocimiento previo de características útiles para la clasificación de las imágenes correspondientes.

Se hará una exclusión de la capa de salida, lo que significa que se omiten las capas completamente conectadas al final, dejando solo las capas convolucionales. También se congela la mayor parte del modelo, esto con el fin de evitar que los pesos ya mencionados sean modificados durante el entrenamiento. Gracias a esto, conserva las características que han sido aprendidas anteriormente.

Yendo hacia la descripción de la arquitectura implementada, se añade una capa de *flatten*, la cual convierte las características 2D de la salida de las capas convolucionales en un vector unidimensional para ser procesado por las capas densas. Ahora se añaden dos capas completamente conectadas con 512 neuronas cada una, con activación *Relu*. Esta función se utiliza cuando se sufre del problema de saturación en los datos, pero ofrece la ventaja de tener una salida simétrica lo cual facilita el proceso de entrenamiento (Figura 28).

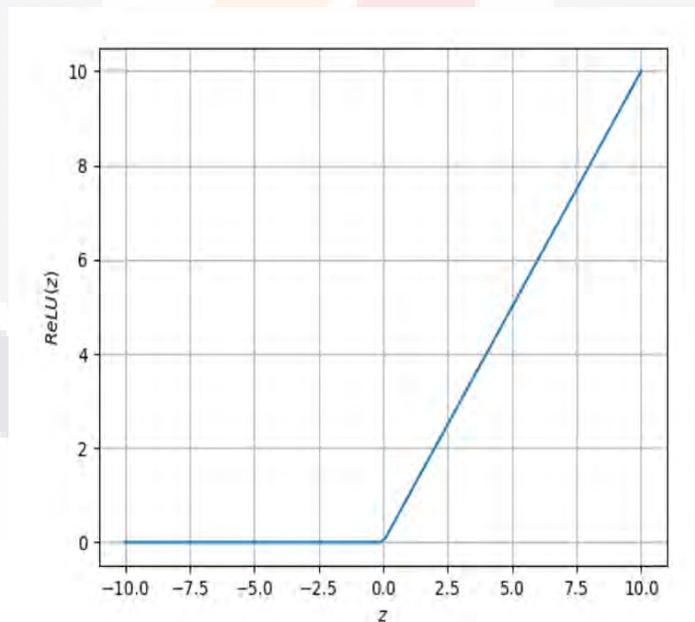


Figura 28: La función *ReLU* simplemente rectifica los datos negativos y los vuelve cero a la salida. Las entradas con valores positivos no sufren modificación alguna a la salida

Esta función generará una salida igual a cero cuando la entrada (z) sea negativa, y una salida igual a la entrada cuando dicha última sea positiva. Siguiendo en la descripción de la arquitectura mencionada, se tiene una capa de salida con tres neuronas y la función de

activación *softmax*, que realiza la clasificación en tres categorías. Esta última función de activación se utiliza en la capa de salida de una red neuronal para realizar la clasificación multiclase. Esta función se utiliza para convertir las salidas de la capa anterior en probabilidades que suman uno. Las probabilidades se utilizan para medir la confianza del modelo en la pertenencia a cada clase y se calcula mediante la siguiente fórmula (Figura 29).

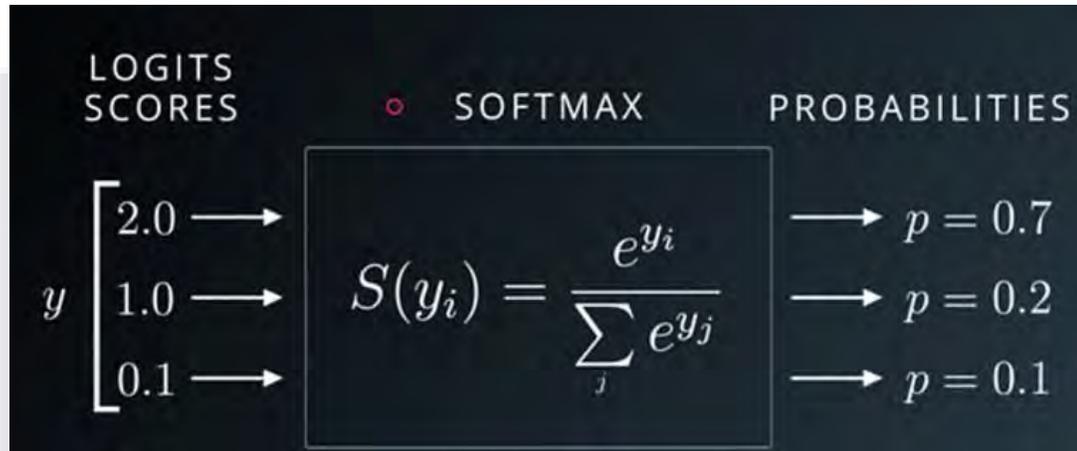


Figura 29: Explicación de la función de activación *softmax*

Por último, se incluyen capas de *dropout* después de cada capa densa para prevenir el sobreajuste, desactivando un 30% de las neuronas de forma aleatoria durante cada etapa del entrenamiento.

Ahora para la compilación del modelo se utiliza un optimizador Adam con una tasa de aprendizaje *learning rate*=1e-5 (ajusta con fórmula), el cual es un algoritmo de optimización que utiliza una estimación del momento y de la magnitud para actualizar los parámetros. Este adapta la tasa de aprendizaje de cada parámetro en función de estimación del momento, lo que conlleva a una mayor precisión de predicción. Para la función de pérdida, se utiliza *categorical_crossentropy*, la cual es adecuada para los problemas de clasificación con etiquetas codificadas con *onehot encoding* (proceso visto y aplicado en el preprocesamiento). Por último, se añade la métrica de evaluación, que medirá el porcentaje de las predicciones que ha hecho de manera correcta.

Para la etapa del entrenamiento del modelo una vez que se ha creado la estructura de la arquitectura, este comienza utilizando el conjunto de entrenamiento y el conjunto de validación, además de que se miden los tiempos de entrenamiento. Se implementa dentro

de todo este proceso la técnica de *Early Stopping* (parada anticipada) que detiene el entrenamiento si la pérdida de validación no mejora después de 5 épocas, lo que nos ayuda a prevenir el sobre entrenamiento (Figura 30).

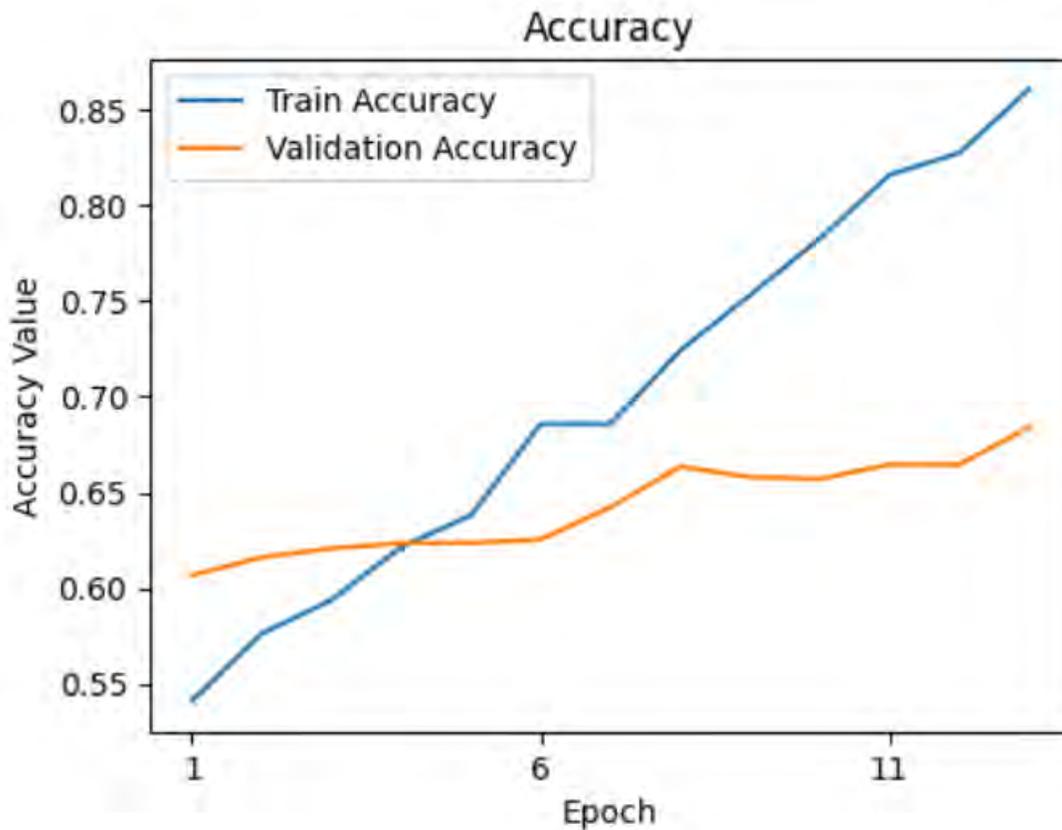


Figura 30: Comparativa de resultados preliminares en la etapa de validación y entrenamiento del modelo VGG16.

9.4.2 Google Inception

Hablando del modelo de *Inception*, perteneciente a Google que es un tipo de modelo de red neuronal convolucional, se carga la arquitectura del mismo con pesos pre entrenados de igual manera en el conjunto de datos de *ImageNet*. Esto significa que el modelo tendrá conocimiento previo de características útiles para la clasificación de las imágenes correspondientes y en este modelo es más importante por que su arquitectura es la más robusta en comparación con los modelos tratados en la investigación.

Se hará una exclusión de la capa de salida, lo que significa que se omiten las capas completamente conectadas al final, dejando solo las capas convolucionales que constituyen

al modelo. También se congela la mayor parte de este, esto con el fin de evitar que los pesos ya mencionados sean modificados durante el entrenamiento. Gracias a esto, conserva las características que han sido aprendidas anteriormente y hace que los resultados sean más precisos conforme la arquitectura vaya siendo más sofisticada, como es este el caso (Figura 31). Se aplican convoluciones con diferentes tamaños de filtro, sobre todo de 1×1 y de 3×3 , que a su vez se incluyen operaciones de *polling* para reducir dimensionalidad y evitar sobreajuste, esto conforma las capas iniciales (extracción básica de características).

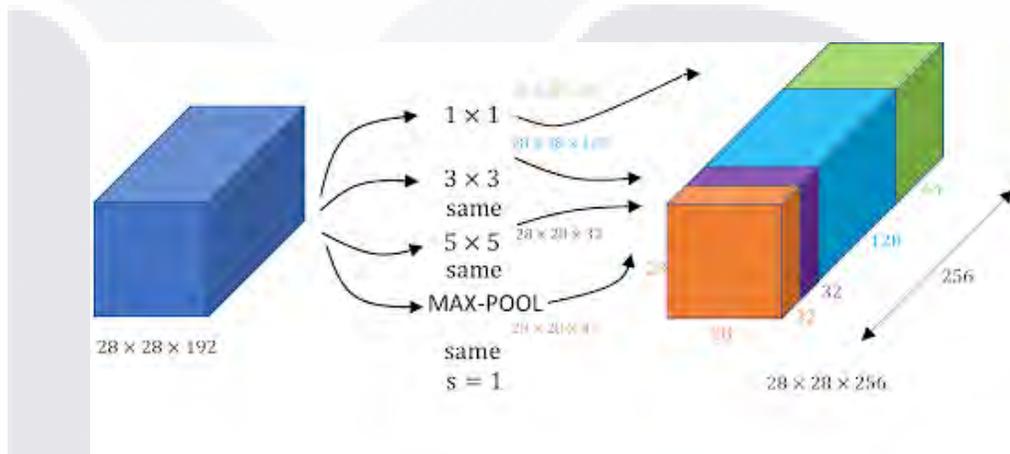


Figura 31: Capas de convolución para el modelo de *InceptionV3*

Los bloques *Inception* integran las salidas en una única representación, que posteriormente se envía a las capas intermedias o auxiliares. Estas capas contribuyen a mejorar la propagación del gradiente y facilitan el entrenamiento del modelo. Para la compilación, se emplea el optimizador Adam con una tasa de aprendizaje de 0.0001 y un algoritmo de optimización que utiliza estimaciones de momento y magnitud para ajustar los parámetros. Este enfoque adapta dinámicamente la tasa de aprendizaje para cada parámetro, lo que mejora la precisión en las predicciones, incluso en arquitecturas complejas como esta, que destacan por su robustez. La función de pérdida utilizada es *categorical_crossentropy*, adecuada para problemas de clasificación con etiquetas codificadas mediante *one-hot encoding*, proceso que se realiza durante el preprocesamiento.

Finalmente, se incluye una métrica de evaluación que calcula el porcentaje de predicciones correctas realizadas por el modelo. Por último, se mide el tiempo en el que tarda el entrenamiento de este, aunado ya con el proceso propio del entrenamiento, que contiene las imágenes preprocesadas y escaladas, las etiquetas correspondientes a las imágenes

de entrenamiento a través de un vector binario a través de 20 épocas. Al último, se indican las métricas como precisión, pérdida y comportamiento del modelo a lo largo de todas las iteraciones (Figura 32).

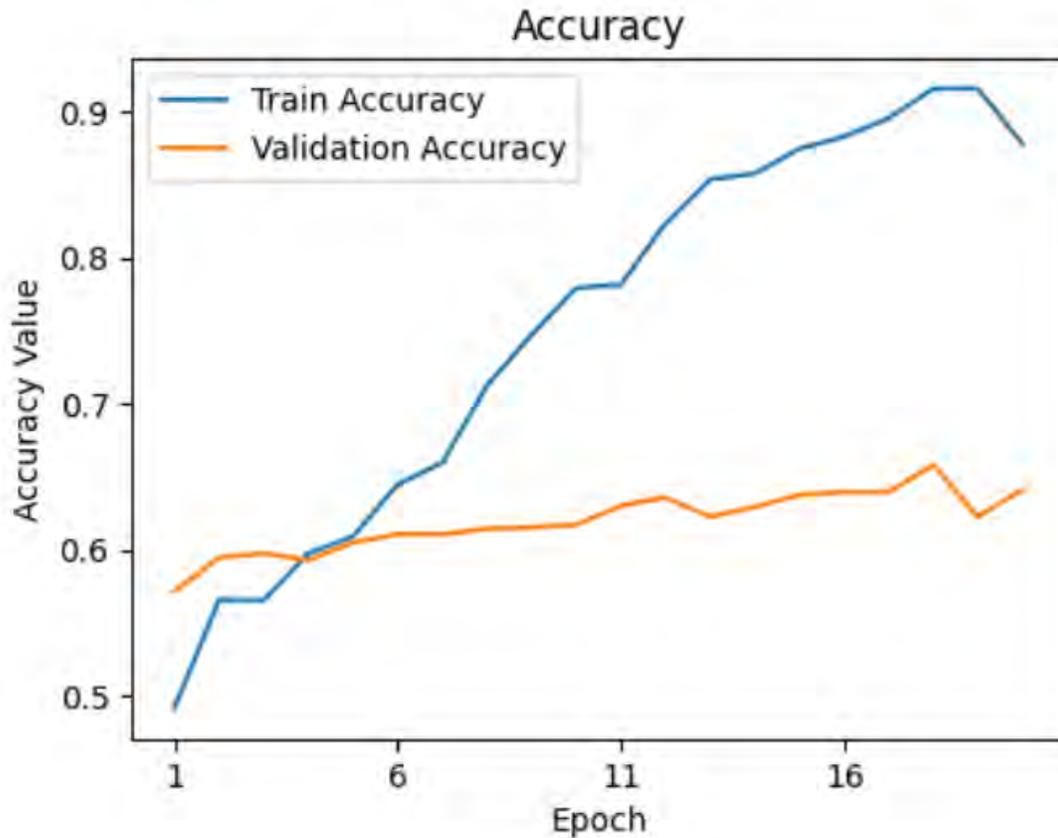


Figura 32: Comparativa de resultados preliminares en la etapa de validación y entrenamiento del modelo *Inception V3*.

9.4.3 XceptionNet

Xception se inspira en la arquitectura *Inception* de Google, que utiliza bloques con múltiples tipos de filtros convolucionales en paralelo. A diferencia de *Inception*, *Xception* simplifica el diseño al reemplazar los bloques *Inception* con convoluciones separables en profundidad, logrando un modelo más eficiente y potente (Figura 33).

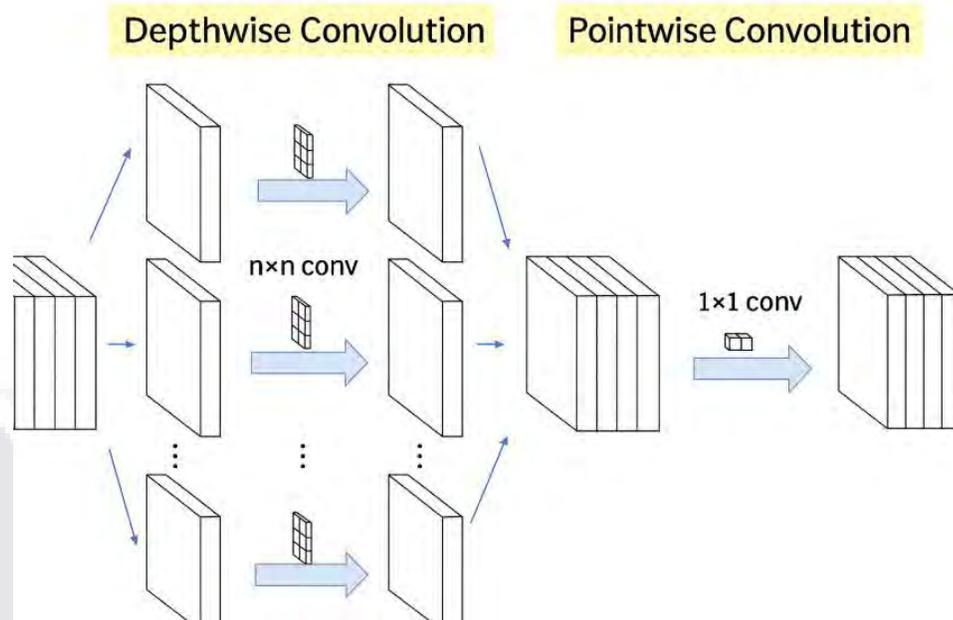


Figura 33: Capas de convolución para el modelo de *Xception*

Se carga el modelo correspondiente, el cual excluye la capa densa final (clasificación) del modelo *Xception* original. Esto permite usar las capas convolucionales pre entrenadas para extraer características. Aunado a esto se cargan pesos pre entrenados. Después de este proceso se define el tamaño de las imágenes de entrada, que son de $224 \times 224 \times 3$. Al igual que los modelos anteriores, se congelan todas las capas del modelo pre entrenado, ya que no se ajustarán durante el entrenamiento, lo que reduce el riesgo de sobreajuste y por ende impacta de manera positiva para el rendimiento del modelo.

Continuando con la implementación correspondiente, se utilizan capas convolucionales del modelo para extraer características de las imágenes. Se compone de capas densas de 512 neuronas, con una función de activación *RELU*. De igual manera se aplica la técnica de *dropout*, que reduce el riesgo de sobreajuste desactivando aleatoriamente el 30% de las neuronas durante el entrenamiento. Se aplica de igual manera una capa densa con tres neuronas, que es una por clase y la activación *softmax* para devolver las probabilidades de la clasificación mencionada. Para la compilación del modelo, se ajustan las capas densas utilizando el optimizador *Adam* con una tasa de aprendizaje de 0.00001.

Cabe mencionar que también se tuvo que agregar una técnica de *Early Stopping*, la cual consta de una verificación que revisa si el rendimiento en el conjunto de validación no mejora después de 5 épocas y en este caso ajusta sólo las capas personalizadas utilizando

imágenes y etiquetas en formatos escalado y codificado en *one-hot*. Al último se indican las métricas como precisión, pérdida y comportamiento del modelo a lo largo de todas las iteraciones pertenecientes al entrenamiento (Figura 34).

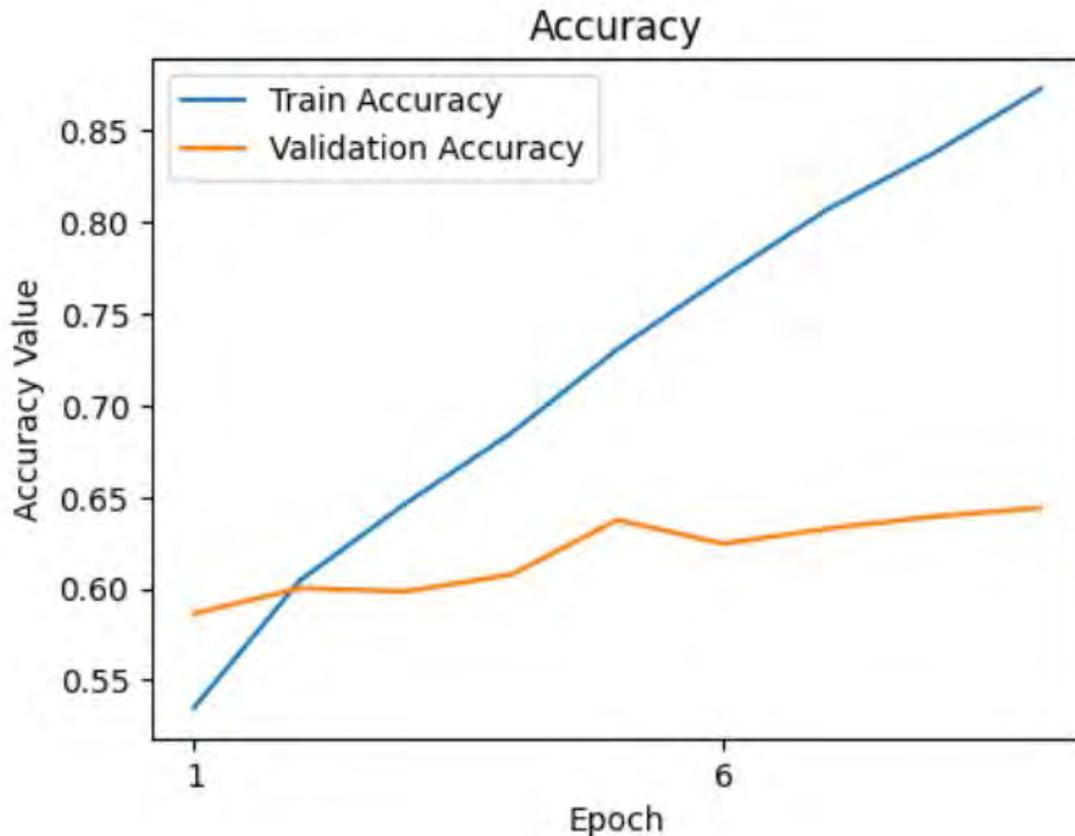


Figura 34: Comparativa de resultados preliminares en la etapa de validación y entrenamiento del modelo *Xception*.

10. Resultados y discusiones

Una vez que se ha mencionado cómo se han realizado las implementaciones, se arrojaron los resultados correspondientes derivado del entrenamiento de cada modelo, así como de sus métricas correspondientes. Cabe destacar que estos fueron los resultados iniciales y que se demostró la mejora de la precisión en los resultados para el conjunto de datos de validación del algoritmo, además de la aplicación de técnicas para conjuntar las habilidades de cada modelo en la clasificación de cada clase.

Durante el desarrollo del proyecto se utilizaron tres modelos de redes neuronales convolucionales preentrenadas: *Xception*, *InceptionV3* y *VGG16*, con el objetivo de evaluar

su capacidad para clasificar imágenes médicas de manera precisa. Para el entrenamiento, los datos fueron divididos en un 80% para el conjunto de entrenamiento y un 20% para el conjunto de pruebas (Tabla 2). Aunque inicialmente se definió un máximo de 20 épocas, no se utilizó un número fijo, ya que se implementó una estrategia de *early stopping*, permitiendo detener el entrenamiento de forma anticipada cuando no se observaban mejoras en el desempeño sobre el conjunto de validación. Esta técnica resultó útil para evitar el sobreajuste y mejorar la capacidad de generalización de los modelos.

	Porcentaje	Total de imágenes
Total	100%	8215
Entrenamiento	80%	6572
Pruebas	20%	1643

Tabla 2: Distribución de los datos correspondientes al conjunto de entrenamiento y validación

La validación del rendimiento se realizó mediante el análisis del comportamiento de cada modelo sobre el conjunto de pruebas, prestando especial atención a la manera en que clasificaban las distintas clases presentes en los datos. Las métricas de evaluación seleccionadas fueron la exactitud (*accuracy*) y la curva *ROC*, debido a su relevancia para evaluar tanto el rendimiento global como el específico por clase. Además, se emplearon matrices de confusión como herramienta complementaria para observar con mayor detalle los aciertos y errores cometidos por cada modelo en la clasificación. Estas matrices permitieron identificar patrones de confusión frecuentes y evaluar de forma más crítica el comportamiento de las arquitecturas analizadas, especialmente en relación con el desempeño por clase.

Para comenzar, se visualizó a través de un gráfico de líneas la comparación de los tres modelos correspondientes (Figura 35). Se tuvo la intersección del porcentaje a lo largo de las épocas en las que el modelo necesitó para obtener su mejor precisión en el conjunto de pruebas, el cual indicó el performance que tuvo para aprender los patrones reconocidos en las imágenes, además de los bordes, contornos y otras características importantes para la predicción. El modelo de *VGG16* y *Xception* tuvieron casi la misma mejora a lo largo del proceso como se pudo observar, mientras el modelo de *Inception V3* tuvo una mejor representación de la precisión / épocas en el modelo (Figura 35).

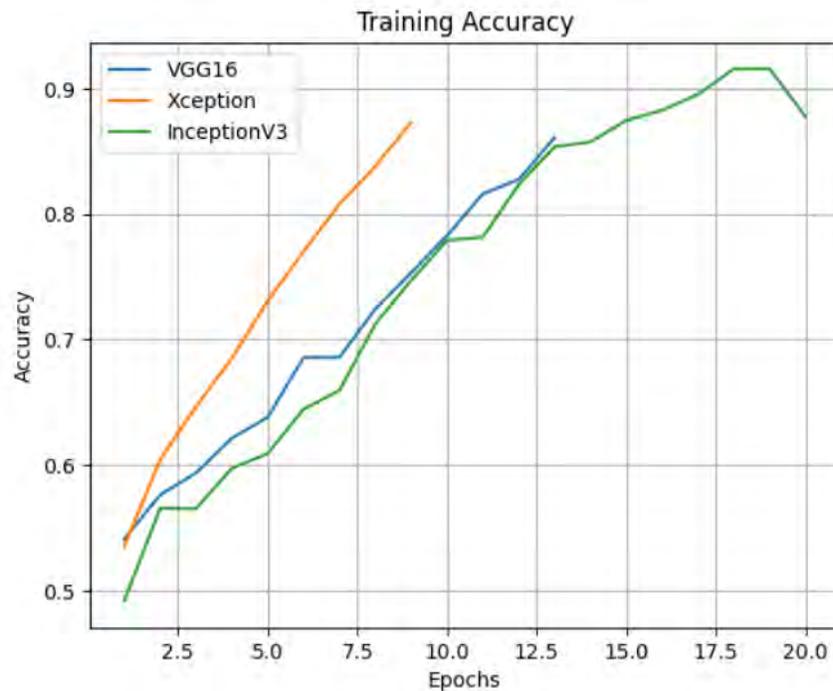


Figura 35: Comparativa de los resultados en el entrenamiento de los tres modelos correspondientes

Aunado a estos resultados, también se presentaron los mismos para el conjunto de validación, los cuales fueron de suma importancia conocer ya que estos dieron la pauta de que tan preciso fue el modelo para poder aprender de nuevas características relacionadas a patrones malignos que pudieran ser fácilmente visibles en las imágenes. Se tiene debajo la representación de los tres modelos de acuerdo con la precisión y las épocas. Como se pudo mostrar, la precisión de los tres modelos estuvo entre el 60 y 70% de precisión a lo largo de no más de 20 épocas, esto debido a que cuando el modelo dejó de mejorar en precisión, existe una técnica para detener el entrenamiento y arrojar los resultados en donde deja de aprender el modelo significativamente (Figura 36).

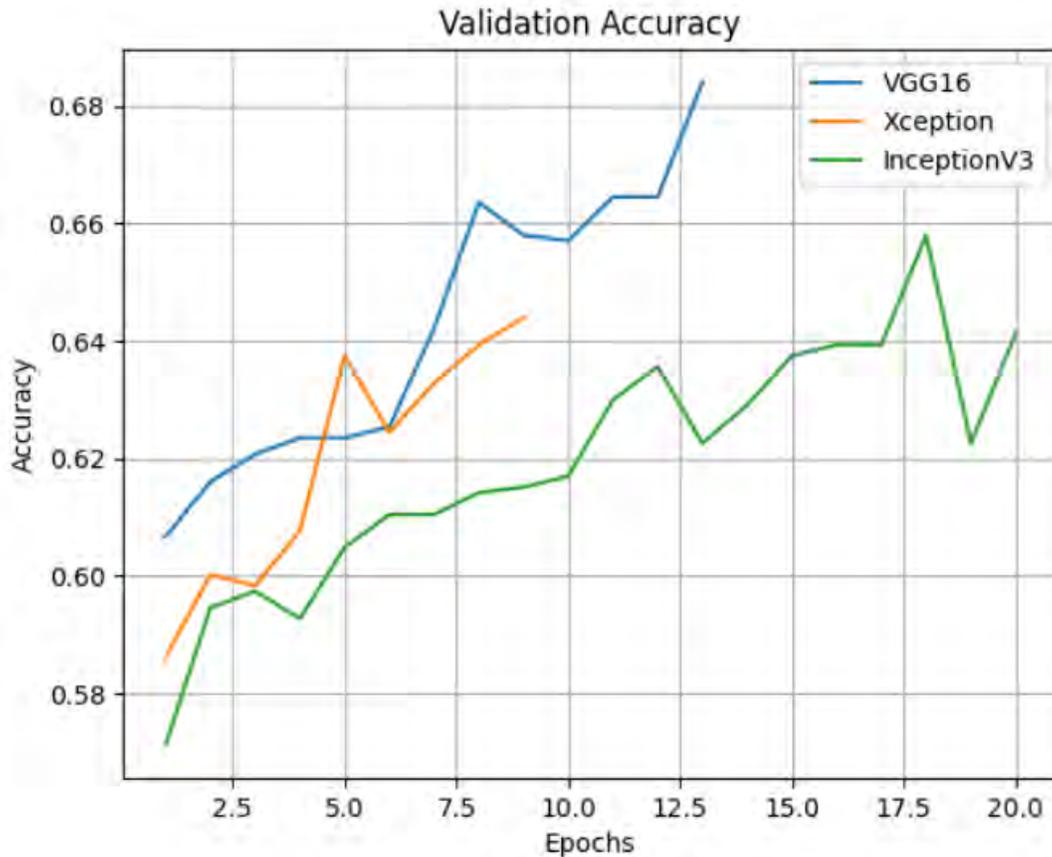


Figura 36: Comparativa de los resultados en la validación de los tres modelos correspondientes

Se evaluaron tres modelos: *VGG16*, *Inception V3* y *Xception*, considerando tres métricas principales:

- Precisión en el conjunto de entrenamiento: Indica qué tan bien la red aprendió los datos de entrenamiento.
- Precisión en el conjunto de validación: Midió el desempeño del modelo en datos no vistos durante el entrenamiento, lo que ayudó a evaluar su capacidad de generalización.
- Tiempo de entrenamiento: Representó el tiempo total requerido para entrenar cada modelo.

Como se pudo observar, para el algoritmo de *VGG16* la precisión en el conjunto de entrenamiento fue de 86.70%, que este parámetro indicó cómo es que aprendió a reconocer los patrones en las imágenes durante las épocas correspondientes en las que fue entrenado el modelo, mientras que la precisión en el conjunto de validación fue del 69%, que

representó la capacidad del modelo para aprender nuevos patrones. Se presentó el mismo caso para el modelo de *Inception V3* de *Google*, con un 91.54% de precisión en entrenamiento y un 64% en validación respectivamente. Para el último modelo la precisión fue de 87.28% y un 63% respectivamente en los conjuntos correspondientes. Por último, de acuerdo con los tiempos de entrenamiento que tomó cada modelo, se pudo notar que el que menos tiempo empleó en realizar dicho proceso fue el de *VGG16*, ya que su arquitectura es más ligera, por lo que en los modelos siguientes el tiempo fue más prolongado. Debido a que el modelo de *Inception V3* tiene una arquitectura robusta de capas conectadas, como consecuencia impactó en este parámetro.

Dentro de las dos mejoras que se realizaron aunado a la mejora del porcentaje de precisión, se utilizó una técnica de aumento de datos correspondiente, el cual consta de una serie de pasos para tomar el conjunto de datos y hacer variaciones, de las cuales el modelo pueda seguir aprendiendo diferentes patrones ante nuevos datos de entrada. Una vez que se implementó el aumento de datos, se realizó el entrenamiento nuevamente durante 20 épocas. A continuación, en la siguiente tabla (Tabla 3) se muestran las mejoras realizadas en la precisión del conjunto mencionado.

Red neuronal	Precisión en el conjunto de entrenamiento	Precisión en el conjunto de validación	Tiempo de entrenamiento
Vgg16	86.70%	78%	10036 segundos
Inception V3(Google Net)	91.54%	85%	3271.79 segundos
Xception (Google Net)	87.28 %	82%	2778 segundos

Tabla 3: Resultados finales de la precisión de los modelos

Para finalizar, como se puede visualizar en la Figura 37, se tuvo la misma imagen que se le ha introducido al sistema, la cual fue una imagen etiquetada y que cumplió las características para su análisis correspondiente. Después de que se realizó el análisis de predicción correspondiente a las tres categorías que se mencionan en secciones anteriores, se obtuvieron los resultados de cómo el proceso mismo clasificó a la imagen, dando el porcentaje más alto para el tipo 1 (lesiones visiblemente cancerosas). Para explicar de una mejor manera los porcentajes, la máquina analizó de acuerdo con la probabilidad de que esté en uno o en otro tipo, por lo que manejó sus resultados reconociendo las lesiones correspondientes o no, además de que tomó el modelo correspondiente y para cada uno se registró su porcentaje de clasificación aunado a la etiqueta de la lesión correspondiente.

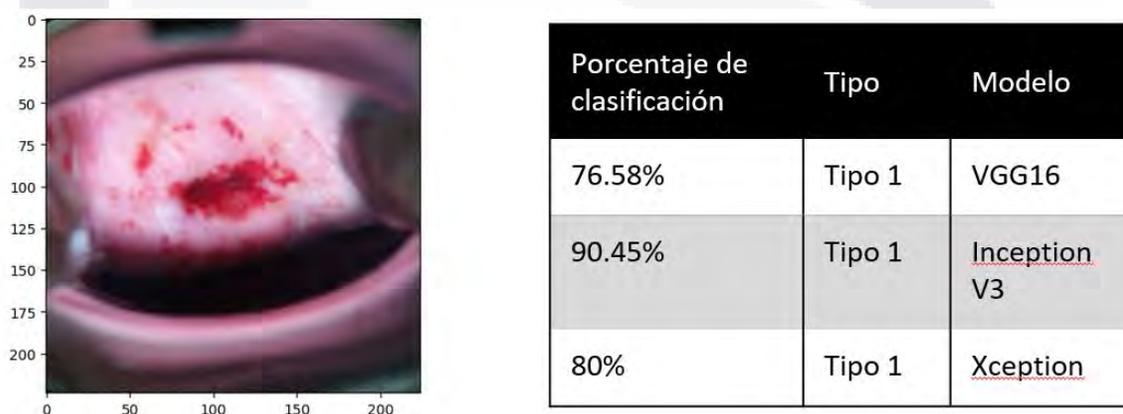


Figura 37: Comparativa de la clasificación de los modelos asociados a clase No.1

Para terminar de comprender mejor el rendimiento del modelo, se analizó los mapas de activación que indican las áreas específicas en las que el modelo se enfocó al tomar decisiones de clasificación para cada imagen (Figura 38). Para calcular y visualizar estas activaciones, se utilizó una técnica llamada *Gradient-weighted Class Activation Mapping (Grad-CAM)*. Estos mapas de activación emplean mapas de calor (*heatmaps*), donde las regiones asociadas a la clase predicha se representan mediante una escala de colores tipo *jet colormap* (Darwish M., 2023). En esta escala, las áreas con mayor activación se muestran en rojo intenso, mientras que las de menor activación aparecen en azul profundo.

La técnica *Grad-CAM*, basada en *Vision Transformer (ViT)* con *Shifted Patch Tokenization (SPT)*, se utilizó para obtener localizaciones en imágenes de colposcopia cervical del conjunto de prueba. En las visualizaciones presentadas, la primera columna muestra las imágenes originales, mientras que la segunda fila exhibe los mapas de activación superpuestos sobre las imágenes. Estos mapas resaltan las regiones que más

contribuyeron a la decisión de clasificación del modelo. En un mapa de calor de *Grad-CAM*, los colores varían desde rojo (que indica mayor importancia) hasta azul (que indica menor importancia), y la intensidad del color representa el grado de relevancia de cada región en la imagen. Además, se mostró el tipo, qué porcentaje obtuvo para cada uno y qué zona es donde detectó la presencia de anomalías.

Además, la visualización de mapas de activación permitió detectar posibles errores o sesgos en el modelo (Figura 38, 39 y 40). Si las zonas resaltadas por *Grad-CAM* no corresponden con áreas clínicamente relevantes, podría indicar la necesidad de realizar ajustes en el entrenamiento del modelo, ya sea refinando la arquitectura de la red neuronal, utilizando un conjunto de datos más diverso o aplicando técnicas de preprocesamiento de datos para mejorar la calidad de la información ingresada.

En el desarrollo de modelos de inteligencia artificial para el ámbito médico, la interpretabilidad no solo mejora la confianza de los profesionales de la salud en estas herramientas, sino que también facilita su aceptación y uso en la práctica clínica. La integración de *Grad-CAM* y otros métodos de explicabilidad contribuye significativamente a la transparencia del modelo, permitiendo una auditoría más rigurosa y asegurando que sus decisiones estén fundamentadas en características relevantes desde el punto de vista médico.

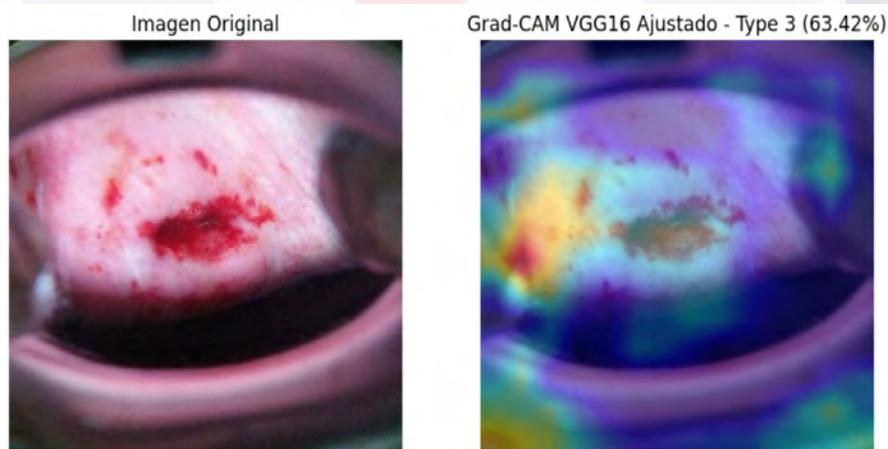


Figura 38: Técnica de *Grad-cam* para la visualización de la zona o no afectada en base al tipo y al modelo utilizado (VGG16)



Figura 39: Técnica de *Grad-cam* para la visualización de la zona o no afectada en base al tipo y al modelo utilizado (*Xception*)

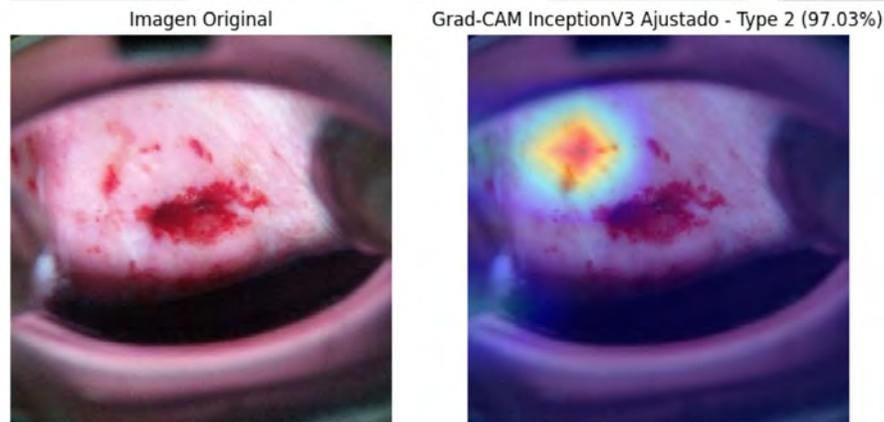


Figura 40: Técnica de *Grad-cam* para la visualización de la zona o no afectada en base al tipo y al modelo utilizado (*Inception V3*)

Para proporcionar un análisis más profundo del rendimiento del modelo, se calculó métricas de evaluación adicionales, incluyendo sensibilidad (recuerdo), precisión, puntuación F1 y el área bajo la curva característica operativa del receptor (*AUROC*). Estas métricas se calcularon por separado para cada clase (Tipo 1, Tipo 2, Tipo 3) con el fin de comprender mejor el comportamiento del modelo más allá de la precisión general.

El modelo *VGG16* alcanzó valores *AUC* de 0.92, 0.85 y 0.87 para el tipo 1, el tipo 2 y el tipo 3, respectivamente (Figura 41).

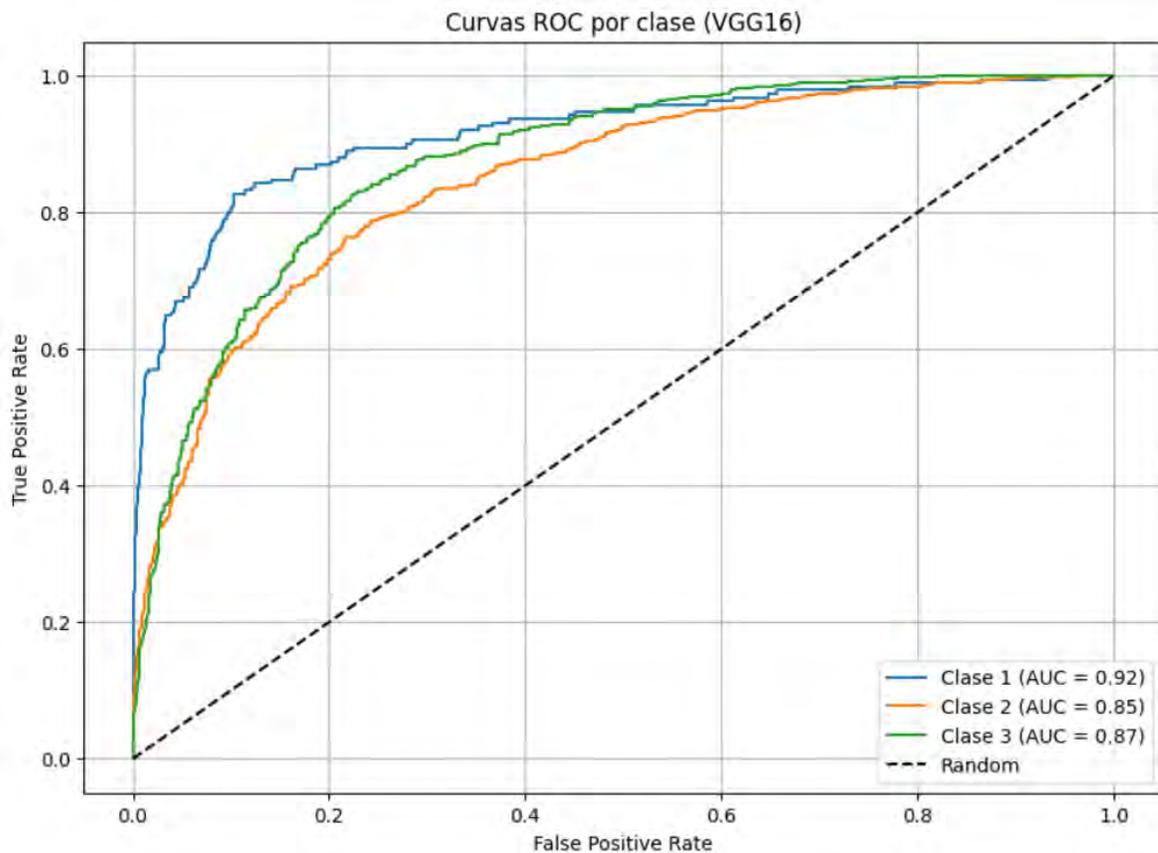


Figura 41: Valores AUC para el modelo VGG16

A pesar de sus puntuaciones AUC relativamente altas, la recuperación para el tipo 1 fue notablemente baja (39 %), lo que indica una dificultad para identificar correctamente la clase. La precisión global de la prueba fue del 70 % y la puntuación F1 media macro alcanzó el 0.68%. La matriz de confusión (Figura 42) muestra que la mayoría de las clasificaciones erróneas se produjeron entre el tipo 1 y el tipo 2.

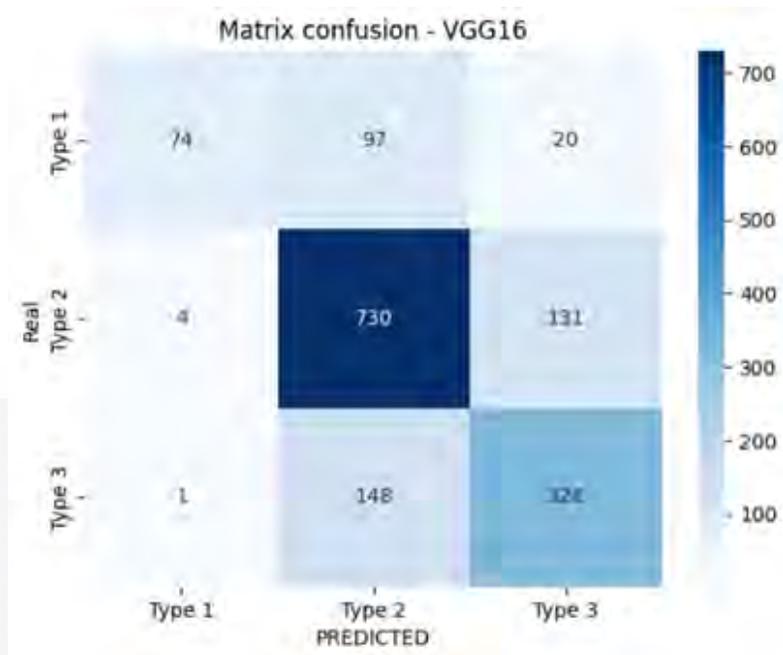


Figura 42: Matriz de confusión para el modelo VGG 16

El modelo *Xception* obtuvo valores *AUC* de 0.89, 0.82 y 0.85 para las tres clases (Figura 43).

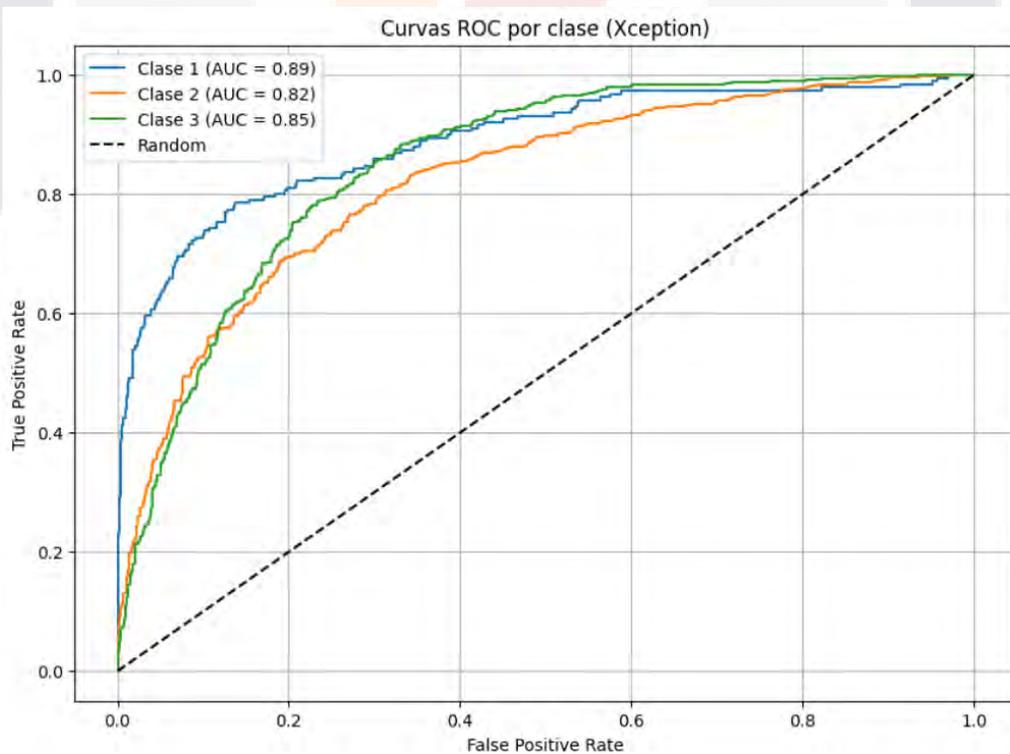


Figura 43: Valores AUC para el modelo *Xception*

Aunque logró una alta recuperación (88%) para el tipo 2, su recuperación para el tipo 1 fue significativamente menor (35%), lo que afectó negativamente al rendimiento general. La precisión de la prueba fue del 62.30% y la puntuación F1 macro media fue de 0,63. La matriz de confusión (Figura 44) destaca una considerable clasificación errónea entre el tipo 1 y el tipo 2. Específicamente, se observa que una gran cantidad de instancias reales del tipo 1 fueron clasificadas como tipo 2 (113 casos), lo que sugiere que el modelo tiende a sobre predecir la clase tipo 2, posiblemente debido a un sesgo hacia esta clase en los patrones aprendidos. Esta tendencia también se evidencia en el hecho de que las predicciones tipo 2 fueron las más frecuentes, incluso para datos que realmente pertenecían a las clases tipo 1 y tipo 3. Este desbalance en la capacidad de discriminación del modelo afecta su utilidad práctica, ya que compromete la detección precisa de clases menos representadas o más difíciles de distinguir. En resumen, el modelo *Xception* muestra una fuerte inclinación hacia la clase tipo 2, lo que, aunque mejora su rendimiento en esa clase, perjudica la equidad del modelo en el contexto multicategoría.

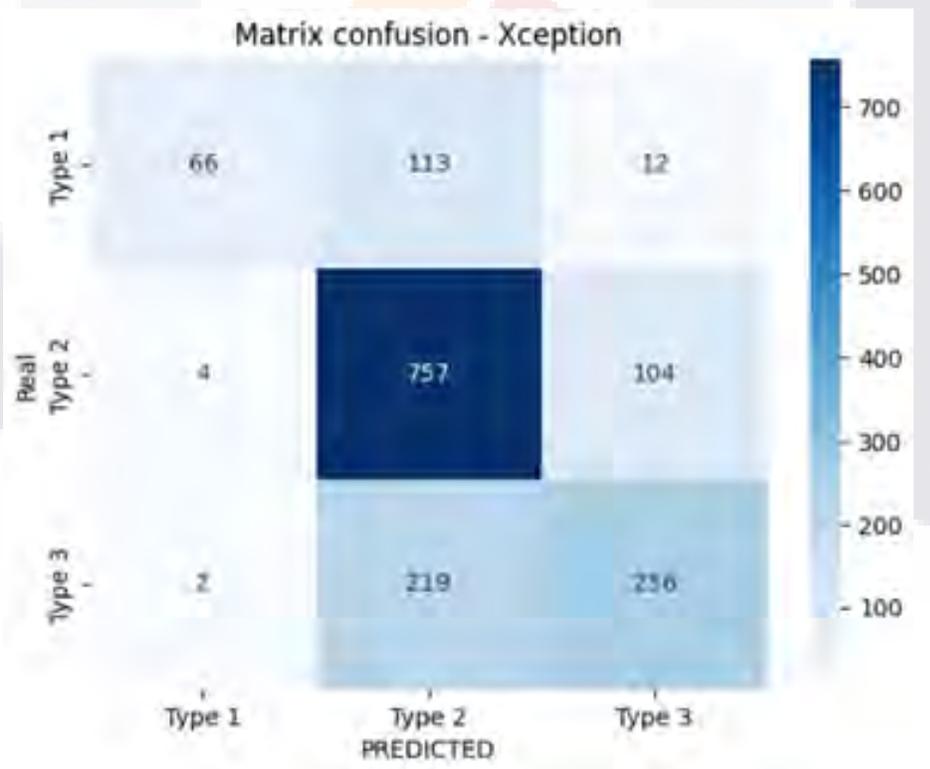


Figura 44: Matriz de confusión para el modelo *Xception*

InceptionV3 ofreció el mejor rendimiento, con valores *AUC* de 0,94, 0,91 y 0,92 para el tipo 1, el tipo 2 y el tipo 3 (Fig. 45).

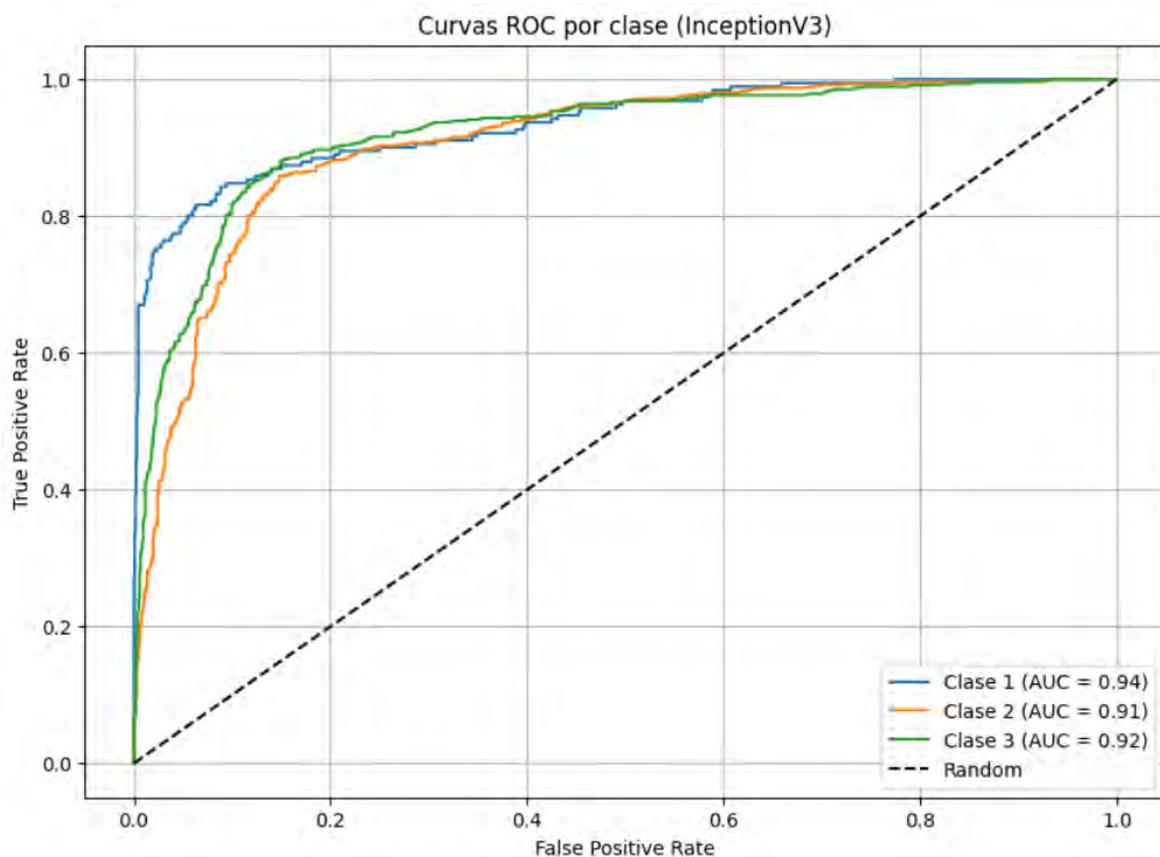


Figura 45: Valores AUC para el modelo *Inception V3*

Se logró una recuperación del 70% para el tipo 1, del 87% para el tipo 2 y del 83% para el tipo 3. La precisión global alcanzó el 84%, con una puntuación F1 macro media de 0.82, lo que indica un rendimiento sólido y equilibrado en todas las categorías. La matriz de confusión (Figura 46) respalda aún más la capacidad del modelo para minimizar las clasificaciones erróneas en comparación con otras arquitecturas. Se observa una distribución más equilibrada de los errores, con menores confusiones entre clases, especialmente entre los tipos 1 y 2, donde el modelo redujo notablemente los errores cometidos por *Xception*. Además, la clase tipo 3 también fue identificada con mayor precisión, lo que se refleja en sus 398 aciertos frente a solo 75 clasificaciones incorrectas como tipo 2. A diferencia de *Xception*, *InceptionV3* no muestra una sobrecarga clara hacia ninguna categoría, lo que sugiere una mejor generalización del modelo y un aprendizaje más representativo de los patrones entre clases. Esta estabilidad en el desempeño lo

convierte en una opción más confiable para escenarios donde la correcta identificación de todas las clases es crítica.

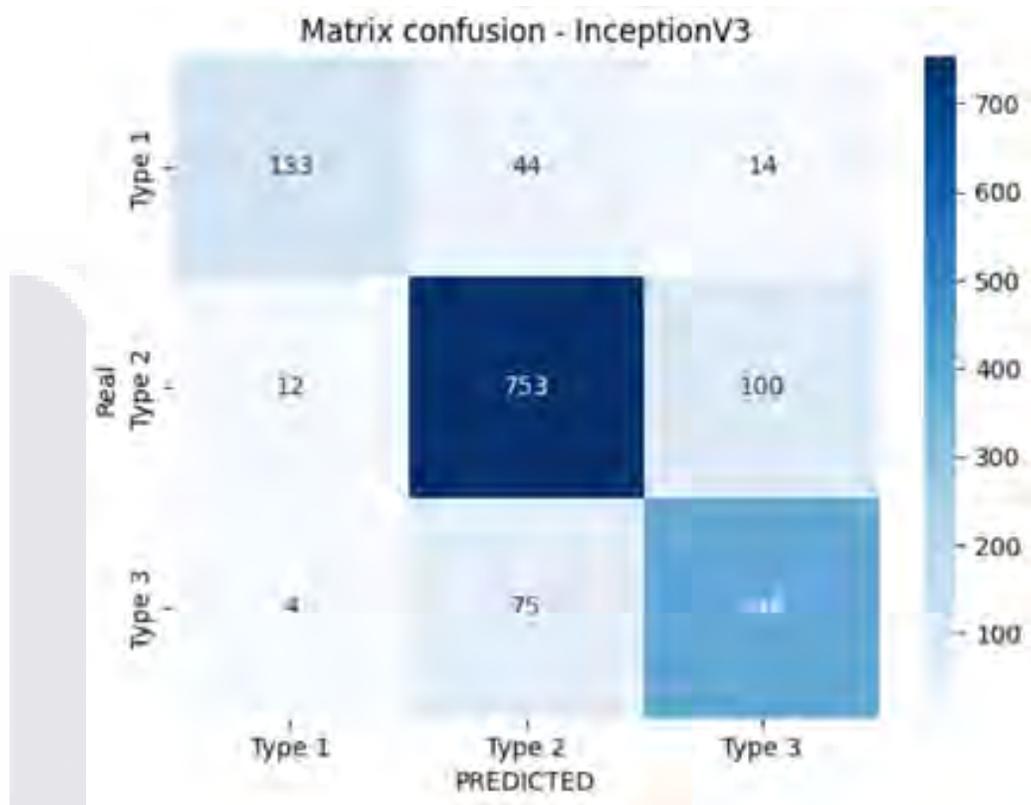


Figura 46: Matriz de confusión para el modelo *Inception*

Tras evaluar el desempeño individual de los modelos de clasificación implementados, se observó que cada uno presentaba fortalezas particulares en la predicción de ciertas clases. Sin embargo, también se evidenciaron limitaciones específicas en otros casos, lo que motivó la búsqueda de una estrategia que pudiera aprovechar lo mejor de cada modelo para mejorar la robustez y generalización del sistema predictivo.

En este contexto, se optó por aplicar *Ensemble Learning* con votación suave (*soft voting*), una técnica que combina las salidas probabilísticas de múltiples clasificadores para generar una predicción final más estable y precisa. A diferencia de la votación dura (*hard voting*), que considera únicamente la clase predicha por mayoría, la votación suave tiene en cuenta las probabilidades asignadas por cada modelo a las distintas clases, ponderando la decisión final de manera más informada.

Este enfoque permitió obtener una mejora significativa en el rendimiento global del sistema. En particular, las curvas *ROC* (*Receiver Operating Characteristic*) obtenidas para el modelo ensamblado muestran un alto poder de discriminación para todas las clases, con valores de *AUC* superiores a 0.90, lo que indica una excelente capacidad para distinguir entre las diferentes categorías diagnósticas (Figura 47). Estos resultados confirman que la integración de múltiples modelos mediante *Ensemble Learning* contribuye a una predicción más robusta y confiable, especialmente en escenarios clínicos donde la precisión diagnóstica es fundamental.

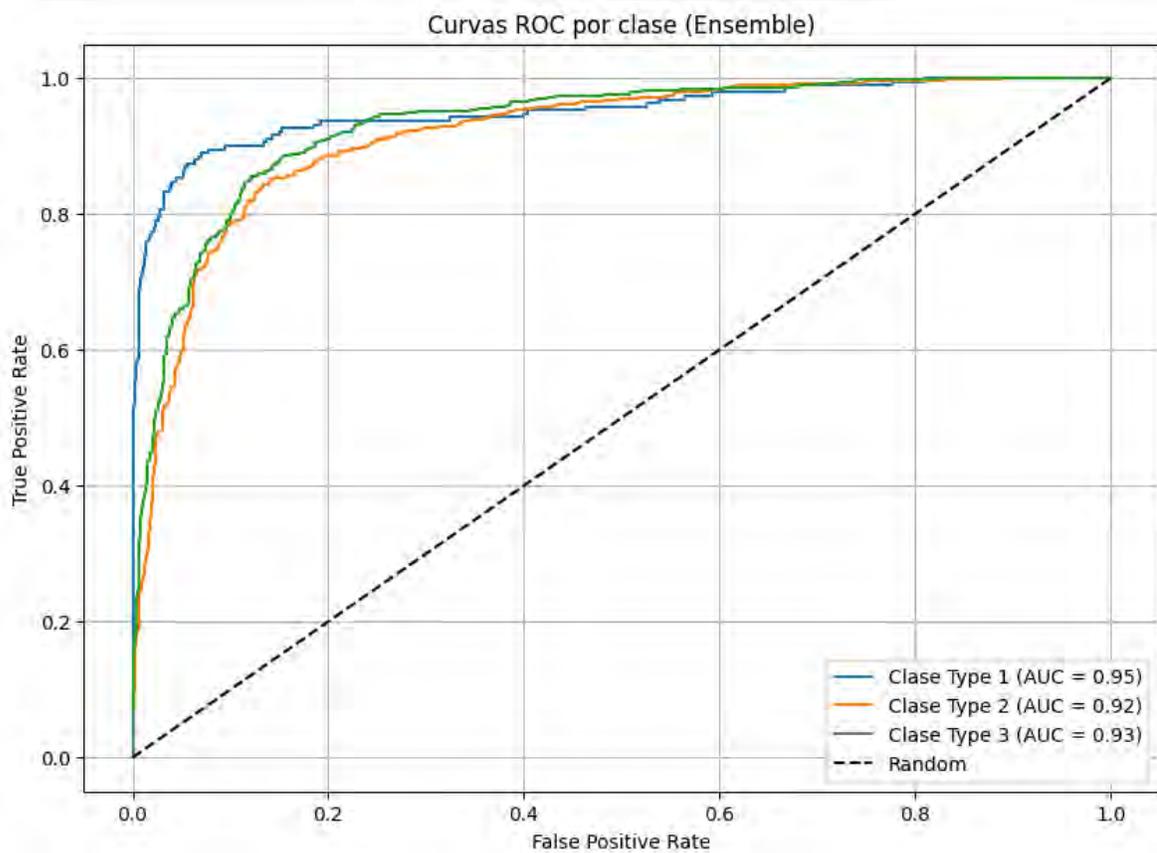


Figura 47: CURVA ROC por clase aplicando técnica de *ensemble learning*

La matriz de confusión obtenida para el modelo de ensamblado (*ensemble*) demuestra un desempeño sólido en la clasificación de las tres clases evaluadas, siendo *Type 2* la categoría con mayor precisión (774 clasificaciones correctas). Sin embargo, se observa una confusión significativa entre *Type 1* y *Type 2*, con 72 muestras de *Type 1* clasificadas erróneamente como *Type 2*, lo que sugiere similitudes en sus características que podrían estar dificultando su diferenciación por el modelo. Asimismo, se evidencia una confusión

moderada entre *Type 3* y *Type 2*, con 96 instancias clasificadas incorrectamente. A pesar de estas limitaciones, el modelo logra un buen equilibrio general, con un rendimiento que mejora la precisión global respecto a los modelos individuales *Xception* e *InceptionV3*, en especial al combinar sus fortalezas y compensar sus debilidades. Se destaca que el ensemble evita una sobre predicción marcada hacia una clase específica, como ocurría con *Xception*, y mantiene un nivel de recuperación aceptable para las tres clases. Esta capacidad de balancear los errores y mejorar la generalización valida el enfoque de ensamblado como una estrategia eficaz para incrementar la robustez y confiabilidad del sistema de clasificación en tareas multiclase (Figura 48).

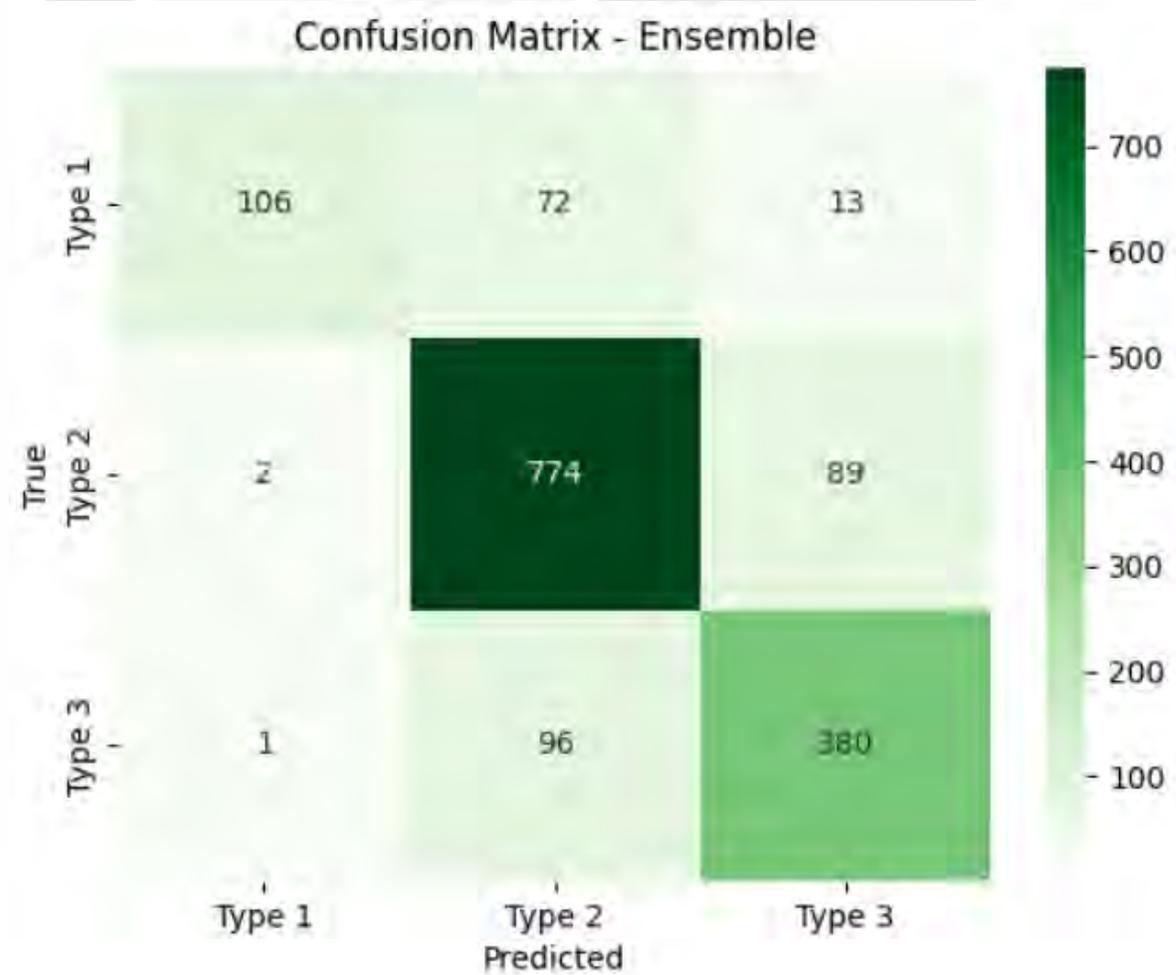


Figura:48: Matriz de confusión para los tres tipos y los modelos aplicados

El modelo de clasificación multiclase evaluado presenta un buen rendimiento general, con un *F1-score* equilibrado entre las clases, lo que evidencia una buena capacidad de

generalización. A pesar del desbalance en el soporte de clases (por ejemplo, *Type 2* tiene significativamente más ejemplos que *Type 1*), los promedios macro y ponderado del *F1-score* se mantienen cercanos a 0.85.

Esto indica un balance adecuado entre *precision* y *recall*, lo cual es fundamental en problemas multiclase con clases desiguales. Además, la matriz de confusión muestra que los errores de clasificación más comunes ocurren entre clases similares (*Type 2* y *Type 3*), pero el modelo aún logra una precisión del 90% en términos de *accuracy* global (Figura 49).

Métrica	Precision	Recall	F1-score	Soporte
Type 1	0.94	0.85	0.81	191
Type 2	0.90	0.87	0.85	865
Type 3	0.88	0.82	0.80	477
Accuracy			0.90	1533
Macro avg	0.86	0.84	0.92	1533
Weighted avg	0.84	0.82	0.82	1533

Figura 49: Datos resultantes de la aplicación de *ensemble learning*

11. Conclusiones y trabajo futuro

A lo largo de este trabajo de investigación, se logró cumplir satisfactoriamente con los objetivos planteados, estableciendo un sistema robusto de clasificación automática de imágenes colposcópicas mediante redes neuronales convolucionales. Este sistema demostró ser una herramienta eficaz para la identificación de lesiones asociadas al cáncer cervicouterino, una problemática de gran impacto en la salud pública.

En primer lugar, se logró recopilar y organizar una base de datos adecuada utilizando el conjunto *Intel ODT for Cervix Images*, que contiene 8,215 imágenes colposcópicas clasificadas en tres categorías clínicas (normales, sospechosas y cancerosas), abarcando tanto vistas endocervicales como ectocervicales. Esto permitió construir un conjunto de datos representativo y diverso, adecuado para el entrenamiento y evaluación de modelos de clasificación automática.

El proceso de preprocesamiento, que incluyó corrección de imperfecciones, normalización, reducción de ruido y aumento de datos, fue clave para mejorar la calidad de la información utilizada por los modelos. Esto permitió minimizar el sesgo y aumentar la generalización del sistema. Fue importante porque permitió estandarizar las condiciones visuales de las imágenes, eliminando variaciones no relacionadas con la patología (como iluminación desigual, manchas u otros artefactos), lo cual facilitó que el modelo se enfocara en las características relevantes de las lesiones. Además, el aumento de datos ayudó a mitigar el desbalance de clases y mejorar la robustez del modelo. En comparación con el análisis de imágenes tradicionales sin procesar, el preprocesamiento contribuyó a reducir el riesgo de sobreajuste, aumentó la precisión del modelo y permitió una extracción de características más eficaz, mejorando significativamente el rendimiento del sistema de detección automática.

El desarrollo e implementación del modelo se llevó a cabo utilizando arquitecturas de redes neuronales ampliamente validadas como *VGG16*, *InceptionV3* y *Xception*. Para ello, se utilizó el lenguaje de programación *Python*, junto con bibliotecas especializadas en *deep learning* y procesamiento de datos. Entre las más relevantes se encuentra *Keras*, utilizada para construir y entrenar las redes neuronales de forma eficiente y modular. También se emplearon *OpenCV* para el procesamiento y la manipulación de imágenes, *NumPy* y *Pandas* para el manejo de estructuras de datos y análisis numérico, y *Matplotlib* y *Plotly* para la visualización de resultados y métricas de evaluación del modelo. Estas herramientas facilitaron el desarrollo de un flujo de trabajo robusto y reproducible. Se demostró que estos modelos pueden alcanzar niveles de precisión superiores a los métodos tradicionales.

Desde el punto de vista de la validación, se emplearon métricas como la precisión, obteniendo resultados prometedores que indican el potencial del sistema para su aplicación en entornos reales. Se evidenció también una buena capacidad de generalización, incluso

en imágenes con variaciones de iluminación, calidad y posición, lo cual es crucial para un posible uso clínico.

Al comparar las tres arquitecturas, *InceptionV3* superó claramente a *VGG16* y *Xception* en todos los parámetros principales (*AUROC*, sensibilidad, precisión, puntuación F1). Aunque *VGG16* y *Xception* lograron precisiones de entrenamiento comparables, su rendimiento en el conjunto de validación y, especialmente, en la sensibilidad para las lesiones de tipo 1 fue notablemente inferior.

InceptionV3, por el contrario, mostró una mayor capacidad para identificar correctamente las diferentes clases, incluyendo las lesiones cancerosas, sin comprometer la precisión en las clases benignas. Su arquitectura, que incorpora múltiples convoluciones de distintos tamaños dentro de los mismos módulos, permite una mejor extracción de características relevantes a distintas escalas, lo cual es crucial en imágenes médicas donde los patrones pueden ser sutiles y variados.

Además, *InceptionV3* logró mantener un equilibrio óptimo entre el tiempo de entrenamiento y el rendimiento final. A pesar de ser una red más profunda y compleja que *VGG16*, su eficiencia computacional fue adecuada y permitió entrenar el modelo sin tiempos excesivos ni necesidad de recursos desproporcionados. Este aspecto es particularmente relevante en contextos de implementación clínica o en regiones con infraestructura tecnológica limitada.

11.1 Trabajo futuro

Este trabajo sienta una base sólida para continuar con el desarrollo de herramientas inteligentes aplicadas al diagnóstico médico. Algunas líneas de investigación que podrían enriquecer y extender los resultados obtenidos incluyen:

- Detección en tiempo real mediante la optimización de los modelos y su implementación en dispositivos móviles o portátiles, utilizando *frameworks* como *TensorFlow Lite* o *ONNX*.
- Aumento y diversificación del conjunto de datos, incorporando imágenes provenientes de distintas fuentes, regiones geográficas y dispositivos colposcópicos, para mejorar la robustez y aplicabilidad del modelo.

- Integración de datos clínicos (edad, antecedentes, resultados de citologías o pruebas de VPH) junto con las imágenes, mediante arquitecturas multimodales, para aumentar la precisión del diagnóstico.
- Desarrollo de modelos explicativos (*Explainable AI*) que puedan justificar sus decisiones a los profesionales médicos, favoreciendo la confianza y adopción en entornos clínicos.
- Implementación en contextos reales como clínicas rurales o unidades móviles de salud, evaluando su impacto en la detección temprana y en la reducción de tiempos de diagnóstico.

12. Referencias bibliográficas

- Subarna, T. G., & Sukumar, P. (2022). Detection and classification of cervical cancer images using CEENET deep learning approach. *Journal of Intelligent & Fuzzy Systems*, 43(3), 3695-3695–3707. <https://doi.org/10.3233/JIFS-220173>
- Suphalakshmi, A., Ahilan, A., Jeyam, A., & Subramanian, M. (2022). Cervical cancer classification using efficient net and fuzzy extreme learning machine. *Journal of Intelligent & Fuzzy Systems*, 43(5), 6333-6333–6342. <https://doi.org/10.3233/JIFS-220296>
- Fowler, J. R. (2023). *Cervical cancer*. StatPearls - NCBI Bookshelf. <https://www.ncbi.nlm.nih.gov/books/NBK431093/>
- Hou, X., Shen, G., Zhou, L., Li, Y., Wang, T., & Ma, X. (2022). Artificial Intelligence in Cervical Cancer Screening and Diagnosis. *Frontiers In Oncology*, 12. <https://doi.org/10.3389/fonc.2022.851367>
- Devi, M. A., Ravi, S., Vaishnavi, J., & Punitha, S. (2016). Classification of Cervical Cancer Using Artificial Neural Networks. *Procedia Computer Science*, 89, 465–472. <https://doi.org/10.1016/j.procs.2016.06.105>

- Corporativa, Iberdrola. (2023). *Descubre los principales beneficios del Machine Learning*. Iberdrola. <https://www.iberdrola.com/innovacion/machine-learning-aprendizaje-automatico>
- *¿Qué es una red neuronal? - Explicación de las redes neuronales artificiales - AWS*. (s. f.). Amazon Web Services, Inc. <https://aws.amazon.com/es/what-is/neural-network/#:~:text=Una%20red%20neuronal%20es%20un,lo%20hace%20el%20cerebro%20humano.>
- Ghoneim, A., Muhammad, G., & Hossain, M. S. (2020). Cervical cancer classification using convolutional neural networks and extreme learning machines. *Future Generation Computer Systems*, 102, 643–649. <https://doi.org/https://doi.org/10.1016/j.future.2019.09.015>
- Darwish M, Altabel MZ, Abiyev RH. Enhancing Cervical Pre-Cancerous Classification Using Advanced Vision Transformer. *Diagnostics (Basel)*. 2023 Sep 8;13(18):2884.
¿Qué son las redes neuronales convolucionales? | IBM. (s. f.). <https://www.ibm.com/mx-es/topics/convolutional-neural-networks>
- H, N. C., Sai, P. P., Madhuri, G., Reddy, K. S., & Reddy, D. V. B. (2022). Artificial Intelligence based Cervical Cancer Risk Prediction Using M1 Algorithms. 2022 International Conference on Emerging Smart Computing and Informatics (ESCI), 1–6. <https://doi.org/10.1109/ESCI53509.2022.9758241>
- Acevedo Lipes, A. M. (2021). *Deep Learning System for the Automatic Classification of Normal and Dysplastic Peripheral Blood Cells as a Support Tool for the Diagnosis*. <http://hdl.handle.net/10803/671387>
- GÖKER, H. (2024). DETECTION OF CERVICAL CANCER FROM UTERINE CERVIX IMAGES USING TRANSFER LEARNING ARCHITECTURES. *Eskişehir Technical*

University Journal of Science & Technology A - Applied Sciences & Engineering,
25(2), 222-222–239. <https://doi.org/10.18038/estubtda.1384489>

- Luo W. Predicting cervical cancer outcomes: statistics, images, and machine learning. *Frontiers in Artificial Intelligence* 2021; 4: 627369, 1-5.
- Sun, Y., Xue, B., Zhang, M., & Yen, G. G. (2020). Evolving Deep Convolutional Neural Networks for Image Classification. *IEEE Transactions on Evolutionary Computation*, 24(2), 394–407. <https://doi.org/10.1109/TEVC.2019.2916183>
- *Causas, riesgos y prevención del cáncer de cuello uterino*. (Instituto Nacional del Cancer, 2023). Cancer.gov. <https://www.cancer.gov/espanol/tipos/cuello-uterino/causas-factores-riesgos-prevencion>
- BETANCOURT-HERNÁNDEZ, M., VIERA-LÓPEZ, G., & SERRANO-MUÑOZ, A. (2018). AUTOMATIC DIAGNOSIS OF RHEUMATOID ARTHRITIS FROM HAND RADIOGRAPHS USING CONVOLUTIONAL NEURAL NETWORKS. *Revista Cubana de Física*, 35(1), 39-39–43.
- Sreenivasu, S. V. N., Gomathi, S., Kumar, M. J., Prathap, L., Madduri, A., Almutairi, K. M. A., Alonazi, W. B., Kali, D., & Jayadhas, S. A. (2022). Dense Convolutional Neural Network for Detection of Cancer from CT Images. *BioMed research international*, 2022, 1293548. <https://doi.org/10.1155/2022/1293548> (Retraction published Biomed Res Int. 2023 Dec 29; 2023:9871637. doi: 10.1155/2023/9871637)
- Munir, K., Elahi, H., Ayub, A., Frezza, F., & Rizzi, A. (2019). Cancer Diagnosis Using Deep Learning: A Bibliographic Review. *Cancers*, 11(9), 1235. <https://doi.org/10.3390/cancers11091235>
- Ma, M., Kidd, E., Fahimian, B. P., Han, B., Niedermayr, T. R., Hristov, D., Xing, L., & Yang, Y. (2022). Dose Prediction for Cervical Cancer Brachytherapy Using 3-D Deep Convolutional Neural Network. *IEEE Transactions on Radiation and Plasma Medical*

Sciences, Radiation and Plasma Medical Sciences, IEEE Transactions on, IEEE Trans. Radiat. Plasma Med. Sci., 6(2), 214-214–221.

<https://doi.org/10.1109/TRPMS.2021.3098507>

- Gupta A, Parveen A, Kumar A, Yadav P. Advancement in Deep Learning Methods for Diagnosis and Prognosis of Cervical Cancer. Curr Genomics. 2022 Aug 11;23(4):234-245. doi: 10.2174/1389202923666220511155939. PMID: 36777879; PMCID: PMC9875539.
- Park YR, Kim YJ, Ju W., Nam K., Kim S., Kim KG Comparación de aprendizaje automático y profundo para la clasificación del cáncer de cuello uterino en función de imágenes de cervicografía. Sci. Rep. 2021;11(1):16143. doi: 10.1038/s41598-021-95748-3
- Jusman Y., Ng SC, Abu Osman NA Sistemas de detección inteligente del cáncer de cuello uterino. Sci. World J., 2014;2014:810368. doi: 10.1155/2014/810368
- Rahaman M, Li C., Yao Y., Kulwa F., Wu X., Li X., Wang Q. DeepCervix: un marco basado en aprendizaje profundo para la clasificación de células cervicales mediante técnicas híbridas de fusión de características profundas
- *Cáncer cervicouterino*. (OPS, 2025). OPS/OMS | Organización Panamericana de la Salud. <https://www.paho.org/es/temas/cancer-cervicouterino>
- Intel & MobileODT Cervical Cancer Screening. 2020. [(último acceso 5 de marzo del 2025)]. Available online: <https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening/data>
- Liston, C. J. L. C. (2018). New machine-learning technologies for computer-aided diagnosis. Nature Medicine, 24(9), 1306–1307. <https://doi.org/10.1038/s41591-018-0178-4>

- Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A. L., Razavian, N., & Tsirigos, A. (2018). Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature Medicine*, 24(10), 1559–1567. <https://doi.org/10.1038/s41591-018-0177-5>
- Zing, D. S. W., Han, N. K. C., Lee, Z. Y. K., & Tan, K. L. (2018). AI for medical imaging goes deep. *Nature Medicine*, 24(5), 539–540. <https://doi.org/10.1038/s41591-018-0029-3>
- Sánchez, J. S., & Pla, F. (2004). Modelos y algoritmos de clasificación supervisada. *Revista Iberoamericana de Inteligencia Artificial*, 8(24), 33–43.
- T.B Ludermir “Hybrid Optimization Algorithm for the Definition of MLP Neural Network Architectures and Weights” Proceedings of the Fifth International Conference on Hybrid Intelligent Systems (HIS’05) 0-7695- 2457-5/05 20.00 2005 IEEE.
- Alhajeri, M.S.; Ren, Y.M.; Ou, F.; Abdullah, F.; Christofides, P.D. Model predictive control of nonlinear processes using transfer learning-based recurrent neural networks. *Chem. Eng. Res. Des.* 2024, 205, 1–12.
- Shahinzadeh, H.; Mahmoudi, A.; Asilian, A.; Sadrarhami, H.; Hemmati, M.; Saberi, Y. Deep Learning: A Overview of Theory and Architectures. In Proceedings of the 2024 20th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP), Babol, Iran, 21–22 February 2024; pp. 1–11.
- Swapna, M.; Sharma, D.Y.K.; Prasad, D.B. CNN Architectures: Alex Net, Le Net, VGG, Google Net, Res Net. *Int. J. Recent Technol. Eng.* **2020**, 8, 953–959
- Yang, F.; Fan, H.; Chu, P.; Blasch, E.; Ling, H. Clustered object detection in aerial images. In *Proceedings of the IEEE International Conference on Computer Vision 2019*; IEEE: Piscataway, NJ, USA, 2019; pp. 8311–8320.

- Law, H.; Deng, J. Cornernet: Detecting objects as paired keypoints. In Proceedings of the European Conference on Computer Vision (ECCV) 2018, Munich, Germany, 8–14 September 2018; pp. 734–750.

