



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ingeniería en Electricidad y Computación

Segmentación de Lesiones en Imágenes de Colposcopia usando Redes
Neuronales Artificiales

PROYECTO DE TITULACIÓN

Previo la obtención del Título de:

Magister en Ingeniería Biomédica

Presentado por:

Angel Alexis Valladarez Lojan

GUAYAQUIL - ECUADOR

Año: 2025

DEDICATORIA

El presente proyecto lo dedico a mi ser superior, por permitirme alcanzar una de mis metas, por ser mi guía y orientarme en la toma de decisiones en el transcurso de mi existencia. A mi madre por su apoyo incondicional; y a todas aquellas personas que confiaron en mí desde el principio.

AGRADECIMIENTOS

Mi eterno agradecimiento a mi ser superior, gracias a ello he alcanzado culminar los estudios de posgrado, superando toda dificultad durante el transcurso del trabajo de titulación. A mi familia, a mi madre, hermanas, sobrino y mascota, por brindarme el apoyo incondicional a lo largo de todo el trayecto académico.

A la ESPOL, por contar con docentes de excelencia que compartieron su conocimiento, experiencia profesional y guía durante mi formación.

DECLARACIÓN EXPRESA

Yo Angel Alexis Valladarez Lojan acuerdo y reconozco que: La titularidad de los derechos patrimoniales de autor (derechos de autor) del proyecto de graduación corresponderá al autor o autores, sin perjuicio de lo cual la ESPOL recibe en este acto una licencia gratuita de plazo indefinido para el uso no comercial y comercial de la obra con facultad de sublicenciar, incluyendo la autorización para su divulgación, así como para la creación y uso de obras derivadas. En el caso de usos comerciales se respetará el porcentaje de participación en beneficios que corresponda a favor del autor o autores. El o los estudiantes deberán procurar en cualquier caso de cesión de sus derechos patrimoniales incluir una cláusula en la cesión que proteja la vigencia de la licencia aquí concedida a la ESPOL.

La titularidad total y exclusiva sobre los derechos patrimoniales de patente de invención, modelo de utilidad, diseño industrial, secreto industrial, secreto empresarial, derechos patrimoniales de autor sobre software o información no divulgada que corresponda o pueda corresponder respecto de cualquier investigación, desarrollo tecnológico o invención realizada por mí durante el desarrollo del proyecto de graduación, pertenecerán de forma total, exclusiva e indivisible a la ESPOL, sin perjuicio del porcentaje que me corresponda de los beneficios económicos que la ESPOL reciba por la explotación de mi innovación, de ser el caso.

En los casos donde la Oficina de Transferencia de Resultados de Investigación (OTRI) de la ESPOL comunique al autor que existe una innovación potencialmente patentable sobre los

resultados del proyecto de graduación, no se realizará publicación o divulgación alguna, sin la autorización expresa y previa de la ESPOL.

Guayaquil, 25 de junio del 2025.

Angel Alexis Valladarez Lojan

EVALUADORES

Ph.D. Edwin Valarezo Añazco

PROFESOR TUTOR

Ph.D. Francis Loayza Paredes

PROFESOR EVALUADOR

RESUMEN

El cáncer de cuello uterino representa el cuarto lugar de mortalidad para mujeres de acuerdo a al Observatorio Global de Cáncer y GLOBOCAN 2022 a nivel mundial. La detección temprana mediante técnicas como la colposcopia, es crucial para reducir la mortalidad. El proyecto tiene como objetivo desarrollar un algoritmo de segmentación automática de lesiones en imágenes colposcópicas utilizando redes neuronales artificiales basadas en U-Net. Este enfoque radica en la necesidad de automatizar el análisis de imágenes médicas para segmentar lesiones de cuello uterino y reducir los errores causados por la interpretación subjetiva del médico. El proyecto utilizó una base de datos de 200 imágenes de colposcopia de la plataforma Kaggle, en las cuales se segmentó manualmente las regiones de interés (ROI) con colaboración de personal médico capacitado para generar máscaras binarias necesarias para entrenar el modelo. Se emplearon técnicas de aumento de datos y el rendimiento se evaluó con las métricas del coeficiente de Dice e Intersección sobre Unión (IoU). Los resultados mostraron que el modelo alcanzó un coeficiente de Dice de 84% en el conjunto de validación y una IoU de 75%. El modelo U-Net demostró ser efectivo para segmentar lesiones en imágenes de colposcopia, con un desempeño adecuado en la validación. El modelo podría ser usado en una herramienta de asistencia para ayudar al médico en la segmentación de lesiones cáncer de cuello uterino.

Palabras Clave: Inteligencia Artificial, Colposcopia, Cérvix, U-Net, ROI.

ABSTRACT

Cervical cancer is the fourth leading cause of mortality in women, according to the Global Cancer Observatory and GLOBOCAN 2022 worldwide. Early detection through techniques such as colposcopy is crucial to reducing mortality. The project aims to develop an algorithm for automatic lesion segmentation in colposcopic images using U-Net-based artificial neural networks. This approach stems from the need to automate medical image analysis to segment cervical lesions and reduce errors caused by physicians' subjective interpretation. The project used a database of 200 colposcopy images from the Kaggle platform, in which regions of interest (ROIs) were manually segmented with the collaboration of trained medical personnel to generate binary masks necessary to train the model. Data augmentation techniques were employed, and performance was evaluated using the Dice coefficient and Intersection over Union (IoU) metrics. The results showed that the model achieved a Dice coefficient of 84% on the validation set and the IoU achieved 75%. The U-Net model proved effective in segmenting lesions in colposcopy images, with adequate validation performance. The model could be used in an assistance tool to assist physicians in segmenting cervical cancer lesions.

Keywords: Artificial Intelligence, Colposcopy, Cervix, U-Net, ROI.

ÍNDICE GENERAL

EVALUADORES.....	6
RESUMEN	I
<i>ABSTRACT</i>	II
ÍNDICE GENERAL	III
ABREVIATURAS.....	VI
ÍNDICE DE FIGURAS	VII
ÍNDICE DE TABLAS.....	VIII
INTRODUCCIÓN.....	IX
CAPÍTULO 1	13
1. Planteamiento de la problemática	13
1.1 Identificación del problema	13
1.2 Justificación	15
1.3 Objetivos	16
1.3.1 Objetivo General	16
1.3.2 Objetivos Específicos	16
1.4 Resultados esperados	17
1.5 Propuesta	19
1.6 Marco Teórico.....	20
1.6.1 Cáncer de cuello uterino	20
1.6.2 Anatomía	21
1.6.3 Precánceres	23
1.6.4 Antecedentes	24
1.6.5 Técnicas de imagenología	25
1.6.6 Segmentación de imagen médica	27

1.6.7	Técnicas de segmentación.....	29
1.6.8	Asistencia en la segmentación	29
CAPÍTULO 2		31
2.	ESTADO DEL ARTE.....	31
2.1	Herramientas de asistencia en segmentación	31
2.1.1	3D Slicer	31
2.1.2	ITK-SNAP	32
2.1.3	ImageJ - Fiji	32
2.2	Algoritmos de segmentación: Aprendizaje Automático (ML)	33
2.2.1	K-means	34
2.2.2	Random Forests (RF)	35
2.2.3	Máquinas de Soporte Vectorial (SVM)	36
2.3	Algoritmos de segmentación: Aprendizaje Profundo (DL).....	37
2.3.1	Mask R-CNN	37
2.3.2	DeepLab V3	38
2.3.3	Redes Neuronales Generativas (GANs).....	39
CAPÍTULO 3		41
3.	METODOLOGÍA	41
3.1	Base de datos	42
3.1.1	Extracción manual del ROI.....	43
3.1.2	Mascara binaria	43
3.1.3	Aumento de datos	44
3.2	U-Net.....	44
3.2.1	Red neuronal artificial	45
3.3	Marco Experimental.....	47
3.3.1	Validación	47

3.3.2	Entorno de trabajo.....	47
3.3.3	Métricas de evaluación	48
CAPÍTULO 4		51
4.	RESULTADOS	51
4.1	Resultados de la segmentación	52
4.1.1	Coeficiente de Sorensen-Dice.....	52
4.1.2	Intersección Sobre Unión o Índice de Jaccard	53
4.2	Análisis Cuantitativo	54
4.3	Discusión de los resultados	55
Conclusiones Y Recomendaciones		56
Conclusiones.....		56
Recomendaciones.....		57
BIBLIOGRAFÍA		58

ABREVIATURAS

IA	Inteligencia Artificial
ANN	Red Neuronal Artificial
CNN	Red Neuronal Convolucional
GCO	Observatorio Global de Cáncer
GLOBOCAN	Estadísticas Globales sobre el Cáncer
GPU	Unidad de Procesamiento Gráfico
ROI	Región de Interés
DB	Base de Datos
IoU	Intersección Sobre Unión
VPH	Virus del Papiloma Humano
OMS	Organización Mundial de la Salud
SCJ	Unión Escamosa Columnar
ML	Aprendizaje Automático
SVM	Máquinas de Soporte Vectorial
RF	Random Forest
AW	Región Acetoblanca
CE	Región Cuello Uterino
DL	Aprendizaje Profundo
GANs	Redes Neuronales Generativas
VIH	Virus de la Inmunodeficiencia Humana
VPH	Virus del Papiloma Humano
CIN	Neoplasia Intraepitelial Cervical
SIL	Lesión Intraepitelial Escamosa
ITS	Infección de Transmisión Sexual
Pap	Prueba de Papanicolau
HSIL	Lesiones Intraepiteliales Escamosas Alto Grado
TC	Tomografía Computarizada
RM	Resonancia Magnética
PET	Tomografía por Emisión de Positrones
CAD	Diagnóstico Asistido por Computador

ÍNDICE DE FIGURAS

Figura I. 1 a) Incidencia, b) Mortalidad de cánceres a nivel global [1]	X
Figura I. 2 a) Incidencia, b) Mortalidad de cánceres en América Latina y el Caribe [1]	XI
Figura I. 3 a) Incidencia, b) Mortalidad de cánceres en Ecuador [1]	XII
Figura 1. 1 Colposcopia [4]	13
Figura 1. 2 Propuesta	20
Figura 1. 3 Características anatómicas del cuello uterino [11]	22
Figura 1. 4 Anatomía del cérvix visto desde el canal vaginal [14]	22
Figura 1. 5 Tipos de CIN [14]	24
Figura 1. 6 Procedimiento de colposcopia [17]	26
Figura 1. 7 Segmentación: a) imagen del cérvix, b) máscara binaria	29
Figura 3. 1 Metodología: segmentación de lesiones en imagen de colposcopia	42
Figura 3. 2 Software ImageJ con distribución Fiji [32]	43
Figura 3. 3 a) Selección de lesión en cérvix, b) Generación de máscara binaria (lesión)	44
Figura 3. 4 Rotación imágenes	44
Figura 3. 5 Arquitectura U-Net modificada	46
Figura 3. 6 Tipos de predicciones en segmentación	49
Figura 3. 7 Métricas de evaluación para segmentación: IoU y Dice	50
Figura 4. 1 Gráficas: a) Exactitud (<i>Accuracy</i>), b) Coeficiente Dice, c) Función de pérdida (<i>Loss</i>)	52
Figura 4. 2 Segmentación: (a) Imágenes originales, (b - c) Imágenes Ground-truth, (d - e) Imágenes U-Net	53
Figura 4. 3 Segmentación: (a) Imágenes originales, (b - c) Imágenes Ground truth, (d - e) Imágenes U-Net	54

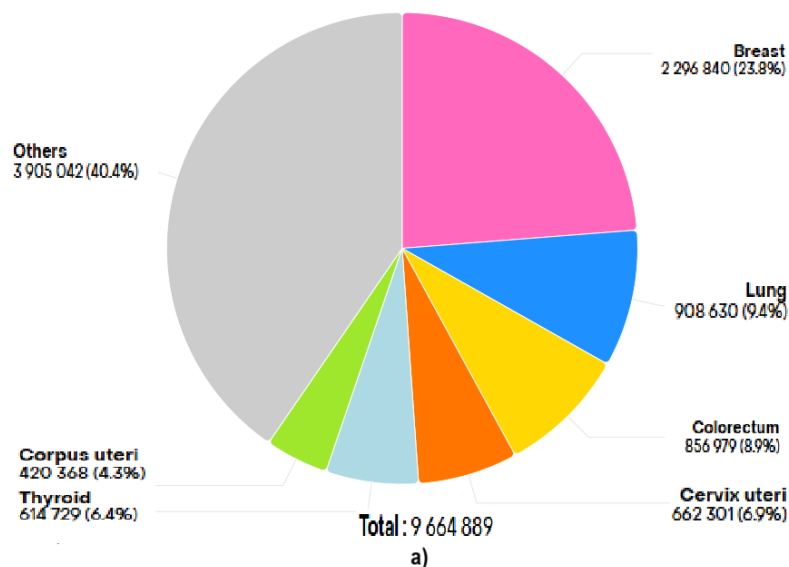
ÍNDICE DE TABLAS

Tabla 4. 1 Valor promedio de evaluación.....	55
--	----

INTRODUCCIÓN

El cáncer de cuello uterino (cérvix) representa un desafío significativo para los sistemas de salud a nivel global. Según las estadísticas de la plataforma web del Observatorio Global de Cáncer (GCO, por sus siglas en inglés) y datos de GLOBOCAN 2022 la incidencia de los diferentes cánceres a nivel global representa 9 664 889 casos diagnosticados [1].

El cáncer de cuello uterino a nivel global se encuentra en cuarto lugar con 662 301(6.9%) de casos nuevos [1]. En la Figura I.1 se tiene los diferentes cánceres a nivel global a) la incidencia representa en mama 2 296 840 (23.8%), pulmón 908 630 (9.4%), colon 856 979 (8.9%), cuello uterino 662 301(6.9%), tiroides 614 729 (6.4%), cuerpo del útero 420 368 (4.3%), y otros cánceres 3 905 042 (40.4%). En b) La mortalidad a nivel global es de 4 313 548 casos y la mortalidad del cáncer de cuello uterino se ubica en cuarto puesto con una representación de 348 874 (8.1%) de casos.



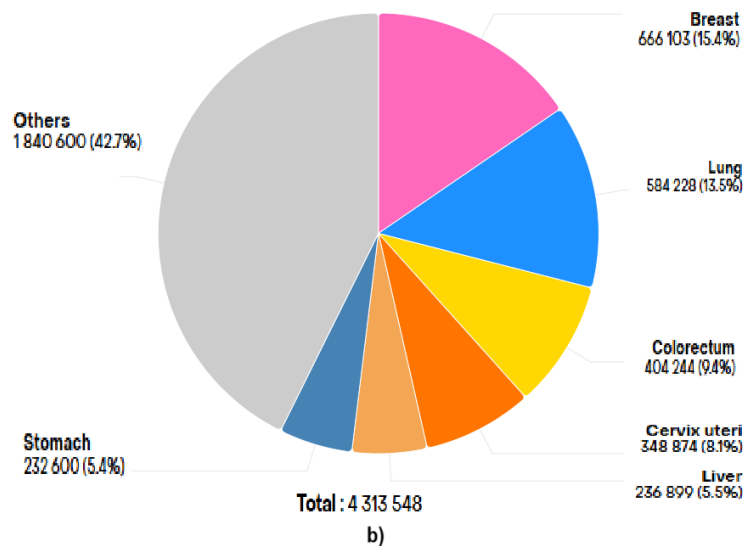
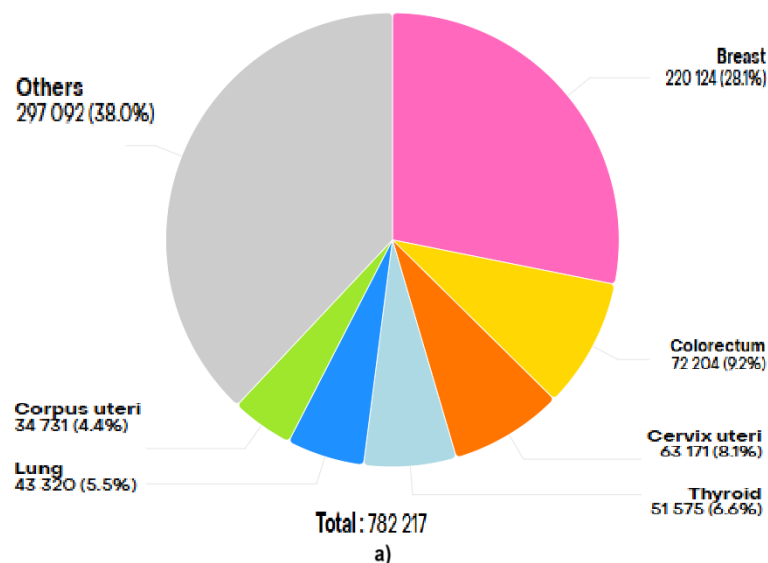


Figura I. 1 a) Incidencia, b) Mortalidad de cánceres a nivel global [1]

En la Figura I.2 representa América Latina y el Caribe a) incidencia de cáncer de cuello uterino es de 63 171 (8.1%) casos, ocupando el tercer lugar del total de 782 217 casos diagnosticados. b) La mortalidad con un total de 365 838 casos y en cáncer de cuello uterino representa el 33 514 (9.2%), ubicándose en cuarto lugar.



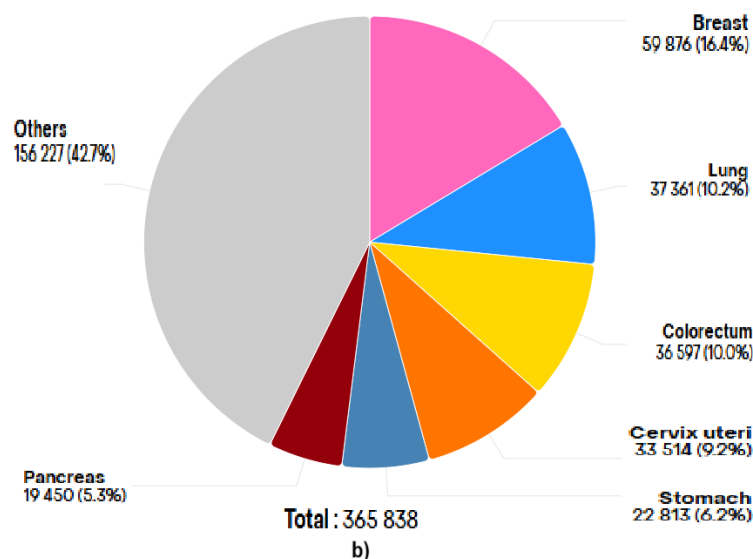
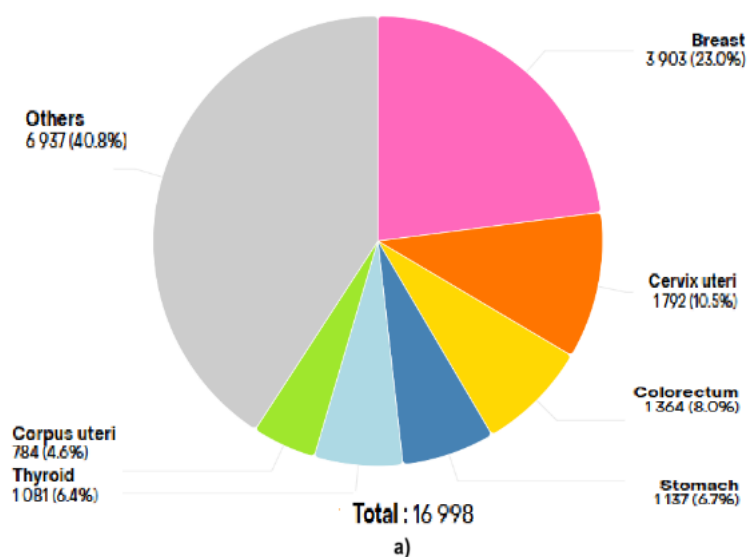


Figura I. 2 a) Incidencia, b) Mortalidad de cánceres en América Latina y el Caribe [1]

En la Figura I.3 se tiene a Ecuador: a) incidencia de cánceres de 16 998 casos diagnosticados. En cuanto, al cáncer de cuello uterino es de 1 792 (10.5%) y representa el segundo cáncer con más casos nuevos por debajo del cáncer de mama que es el principal cáncer con más casos de 1 154 (13.6%), entre otros cánceres. En b) la mortalidad es de 8 500 casos y el cáncer de cuello uterino representa un 939 (11.0%) es el segundo cáncer con casos de mortalidad por debajo del cáncer de mama en el país.



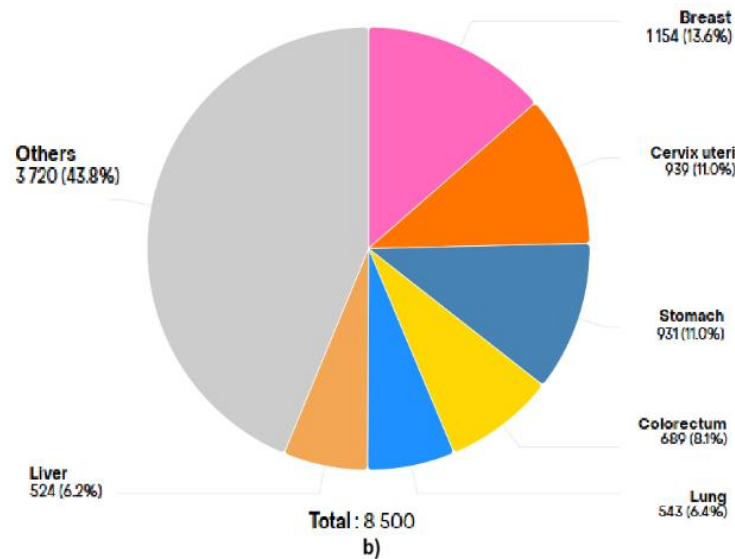


Figura I. 3 a) Incidencia, b) Mortalidad de cánceres en Ecuador [1]

El diagnóstico temprano de este tipo de cáncer es clave para reducir la mortalidad, y la segmentación de imágenes médicas juega un papel crucial en este proceso. En particular, la colposcopia, una técnica de imagen utilizada para examinar el cuello uterino, permite identificar anomalías precancerosas que podrían ser indicativas de cáncer. Sin embargo, la segmentación precisa de las lesiones en estas imágenes es un desafío técnico debido a la variabilidad en su apariencia, presencia de ruido y heterogeneidad en las imágenes.

Las Redes Neuronales Artificiales (ANN, por sus siglas en inglés) han demostrado ser herramientas eficaces para abordar problemas de segmentación en imágenes médicas. Estas redes, que emulan el funcionamiento del cerebro humano, son capaces de aprender patrones complejos y detectar detalles sutiles en las imágenes, lo que las hace ideales para identificar y segmentar lesiones en colposcopia. La investigación se centra en el potencial de la Inteligencia Artificial (IA) para ayudar a identificar lesiones en cáncer de cuello uterino, contribuyendo como herramienta de asistencia al médico. Por lo que, al desarrollar algoritmos avanzados basados en IA no solo podría acelerar el proceso de análisis, sino también aliviar la carga sobre los profesionales de la salud, mejorando así los resultados clínicos.

CAPÍTULO 1

1. PLANTEAMIENTO DE LA PROBLEMÁTICA

1.1 Identificación del problema

La colposcopia es un procedimiento que ayuda a identificar cambios anormales en el cuello uterino (cérvix), que podrían estar relacionadas con el cáncer de útero. El diagnóstico es realizado por el ginecólogo y consiste en examinar el cervix, la vagina y la vulva con un instrumento llamado colposcopio. Durante el examen, se aplican soluciones de ácido acético y/o Lugol en el cervix para mejorar la visibilidad de posibles lesiones. El ácido acético hace que las células con cambios anormales se blanqueen, mientras que el Lugol tiñe las células normales en color marrón oscuro y a las células anormales no las tiñe. Estas sustancias ayudan al ginecólogo identificar áreas sospechosas que podrían ser lesiones precancerosas o anormales (ver Figura 1.1). Este examen es comúnmente realizado en mujeres mayores a 35 años, debido a su mayor predisposición a desarrollar este tipo de patología [2][3].

Observación del cervix



Figura 1. 1 Colposcopia [4]

A pesar de la disponibilidad de métodos de diagnóstico como la citología cervical y la colposcopia, la interpretación de las imágenes obtenidas puede ser compleja y subjetiva, dependiendo en gran medida de la experiencia del médico especialista. La variabilidad en las características de las lesiones, la presencia de artefactos en las imágenes y las limitaciones en la resolución de las cámaras colposcópicas son

factores que dificultan un diagnóstico preciso de las lesiones. Este análisis manual implica una intensa carga de trabajo para el especialista, quien podría incurrir en errores involuntarios de diagnóstico por agotamiento. El inconveniente radica en la identificación precisa de las lesiones de la patología durante los cambios en el epitelio blanco del ácido acético en el cérvix [5].

La identificación de lesiones precancerosas en imágenes de colposcopia es un proceso manual realizado por el ginecólogo, quien interviene directamente para identificar, seleccionar y anotar la ubicación de lesiones sospechosas. Este proceso, conocido como segmentación manual, permite la creación de máscaras binarias, esenciales para la generación de bases de datos etiquetadas, que sirven para entrenar algoritmos de IA. Sin embargo, la intervención manual puede ser laboriosa, subjetiva y propensa a errores, lo que resalta la necesidad de automatizar este proceso.

Además, se presentan cuellos de botella en los centros de salud, lo que provoca largos tiempos de espera y afecta tanto la calidad como la prontitud del diagnóstico para el paciente. En zonas urbanas, la sobrecarga de trabajo de los profesionales de salud, debido a la cantidad de pacientes, genera un cuello de botella que ralentiza el proceso de detección y diagnóstico temprano del cáncer de cérvix, a pesar de contar con recursos. Este problema está directamente vinculado con la insuficiencia en la optimización de los flujos de trabajo, lo que perjudica la eficacia del sistema de salud. Por otro lado, en zonas rurales, el acceso a servicios ginecológicos es aún más limitado, lo que crea un cuello de botella aún más crítico, ya que no cuenta con médicos especialistas, si no con médicos generales y la lejanía de los centros de salud dificultan la prevención y tratamiento de enfermedades. La falta de infraestructura adecuada y el aislamiento geográfico refuerzan estas limitaciones, haciendo que la detección y tratamiento temprano se conviertan en desafíos aún más complejos.

1.2 Justificación

Actualmente, las Redes Neuronales Convolucionales (CNN, por sus siglas en inglés), se están utilizando en el análisis de imágenes médicas debido a su capacidad de extracción automática de características para identificar patrones complejos en diferentes tipos de cánceres. En particular, la arquitectura U-Net ha demostrado ser efectiva en la identificación y localización de lesiones en la segmentación automática en imágenes biomédicas [6].

El desarrollo de algoritmos de IA, han permitido la creación de herramientas de asistencia de diagnóstico, con el fin ayudar a los profesionales de salud a optimizar la carga laboral, dado el volumen de imágenes que deben analizar, y reducir los errores derivados del cansancio y agotamiento. Gracias a esta tecnología, se están logrando resultados comparables con el diagnóstico médico tradicional en cuanto a precisión de detección y clasificación [7].

Por lo tanto, la implementación de la IA representa una solución para abordar estas problemáticas, contribuyendo a la identificación temprana de lesiones en cérvix. Esto permitirá un flujo de trabajo más eficiente, reducirá los tiempos de espera y mejora la calidad del servicio. Asimismo, disminuirá los errores asociados al cansancio y agotamiento, ayudando al médico general y al ginecólogo como una segunda opinión en la identificación de patologías, así disminuyendo la subjetividad del diagnóstico visual.

Al presente, no se dispone de base de datos públicas en Ecuador de imágenes médicas para realizar investigaciones. Por lo que, se propone utilizar una base de datos pública internacional, que permita llevar la investigación y, a futuro, aplicar este modelo para realizar predicciones con imágenes de colposcopia de Ecuador. Cabe indicar que esta base de datos, no contiene las máscaras binarias que corresponde a las lesiones de cérvix. Por lo que, se crearán las máscaras binarias con la capacitación del médico especialista. Esta base no solo apoyaría la investigación en la identificación de lesiones de cérvix, sino que facilitarían el entrenamiento de modelos de IA.

1.3 Objetivos

1.3.1 Objetivo General

Desarrollar un algoritmo para la segmentación automática de lesiones de cuello uterino a partir de imágenes de colposcopia usando redes neuronales artificiales.

1.3.2 Objetivos Específicos

1. Seleccionar una base de datos con imágenes de colposcopia que esté debidamente avalada por un comité de ética de seres humanos.

El objetivo implica la búsqueda y selección de un conjunto de imágenes de colposcopia que contenga muestras representativas de lesiones de cuello uterino, tales como lesiones precancerosas. La base de datos seleccionada sea pública o privada debe cumplir con todos los estándares éticos y legales establecidos para el uso de datos médicos. Es crucial que las imágenes sean anonimizadas y que se haya obtenido el consentimiento informado de los pacientes para su uso en investigación.

2. Colaborar con ginecólogos y expertos en diagnóstico de cáncer de cuello uterino para la construcción de las máscaras binarias.

La colaboración con el médico especialista en diagnóstico es esencial para seleccionar las lesiones del cérvix y construir las imágenes de las máscaras binarias, que indican la localización de dichas lesiones en las imágenes de colposcopia. El especialista, con su experiencia, indicará las características a tener en cuenta al identificar las lesiones en el cérvix. Las máscaras binarias serán utilizadas como el ground-truth para entrenar el modelo, de modo que la red neuronal aprenda a identificar y segmentar las lesiones en nuevas imágenes.

3. Entrenar redes neuronales artificiales para la segmentación de lesiones de cuello uterino en imágenes de colposcopia.

Este objetivo se enfoca en el desarrollo y entrenamiento del modelo de red neuronal profunda U-Net para la segmentación automática de lesiones en las imágenes de colposcopia. La segmentación es el proceso de dividir una imagen en diferentes regiones, en este caso, para identificar áreas sospechosas que contienen lesiones en el cuello uterino. Se utiliza aprendizaje supervisado, en el que la red neuronal aprende a reconocer patrones en las imágenes a partir de las máscaras binarias previamente construidas. El entrenamiento requiere de imágenes de colposcopia con sus respectivas máscaras binarias, el ajuste de los hiperparámetros y la validación del modelo de segmentación.

4. Implementar métricas para la evaluación del rendimiento del modelo de segmentación en lesiones de cuello uterino.

Una vez entrenado el modelo, es necesario evaluar el rendimiento utilizando métricas de segmentación que permitan medir la precisión de las predicciones realizadas por la red neuronal. Las métricas para evaluar modelos de segmentación incluyen la Intersección Sobre Unión (IoU, por sus siglas en inglés) y el coeficiente de Dice, es la más utilizada para evaluar la segmentación en imágenes. Estas métricas son fundamentales para determinar la capacidad del modelo para segmentar lesiones en imágenes de colposcopia.

1.4 Resultados esperados

- Generar las máscaras binarias de la lesión de cáncer de cuello uterino a partir de la interpretación de médicos especialistas.

A través de un proceso de segmentación manual realizado por el médico especialista en imágenes de colposcopia, se generarán las máscaras binarias que indiquen las regiones afectadas por lesiones precancerosas en el cuello uterino. Estas máscaras binarias, que son representaciones en blanco y negro, sirven como ground-truth para entrenar el modelo de red neuronal, ya que

contienen la información sobre la ubicación y la extensión de las lesiones. Este proceso es crucial para proporcionar una base de datos para entrenar el modelo de segmentación.

- Obtener una base de datos de imágenes de colposcopia que incluya la máscara binaria para la segmentación automática.

Una vez construidas las máscaras binarias, se recopilarán para crear una base de datos de imágenes de colposcopia que incluya tanto las imágenes originales como las máscaras binarias correspondientes a las lesiones del cérvix. Esta base de datos servirá para alimentar la red neuronal durante el entrenamiento y validación del modelo. La cantidad de imágenes disponibles serán fundamental para que el modelo segmente las áreas de las lesiones en las imágenes no vistas durante el entrenamiento. Las imágenes deben ser variadas, representando distintas lesiones, condiciones de iluminación, ángulos de visión y resolución.

- Automatizar la segmentación de lesiones en imágenes de colposcopia usando redes neuronales artificiales.

Esto se logrará entrenando un modelo basado en redes neuronales artificiales, específicamente la arquitectura para segmentación U-Net. Esta CNN es adecuada para tareas de segmentación biomédica debido a su capacidad para trabajar con imágenes de alta resolución y el diseño permite una segmentación precisa incluso con pocos datos. El modelo entrenado será capaz de identificar y segmentar de manera automática las áreas afectadas por lesiones de cuello uterino en nuevas imágenes de colposcopia, lo que ahorrará tiempo al médico.

- Establecer métricas de evaluación para medir el desempeño del modelo U-Net en imágenes de colposcopia.

Para evaluar la eficacia del modelo de segmentación, se deben definir métricas de rendimiento. Las métricas para este tipo de tarea son: La intersección sobre unión, mide la superposición entre la máscara predicha por el modelo y la máscara real, y el coeficiente Dice, que también mide la similitud entre las áreas

segmentadas por el modelo y las áreas reales de la lesión. Esta métrica es ideal en imágenes biomédicas, ya que se enfoca en la calidad de la segmentación.

1.5 Propuesta

En la Figura 1.2 se presenta la propuesta para el desarrollo del proyecto de investigación:

- **Base de datos**
Seleccionar una base de datos pública y/o esté validada por un comité de ética en investigación con seres humanos. Para el proyecto de titulación se cuenta con una base de datos pública internacional de imágenes de colposcopia. Cabe indicar que esta no contiene las máscaras binarias que corresponden a las lesiones de cérvix.
- **Colaboración con el médico**
Mediante la capacitación y colaboración del médico, se procederán a identificar las áreas sospechosas que podrían ser lesiones precancerosas o anormales.
- **Generación de máscara binaria**
En las imágenes de colposcopia se procederá con la selección de la región de interés (ROI, por sus siglas en inglés) y la segmentación de lesiones en el cérvix. Para la construcción de la máscara binaria se empleará una herramienta informática de tratamiento digital de imagen orientada a las ciencias de la salud.
- **DB combinada (Base de datos)**
El conjunto de datos consta de imágenes de colposcopia e imágenes de las máscaras binarias que corresponde a las lesiones en el cérvix.
- **U-Net: entrenamiento y evaluación**
El entrenamiento y validación del algoritmo se realizarán de manera local, utilizando un equipo de cómputo con Unidad de Procesamiento Gráfico (GPU). Asimismo, se emplearán métricas de evaluación para medir la eficiencia del modelo de segmentación empleado, tales como Intersección sobre unión (IoU) o Índice de Jaccard y el Coeficiente de Sorensen-Dice.

Desarrollo del proyecto de investigación

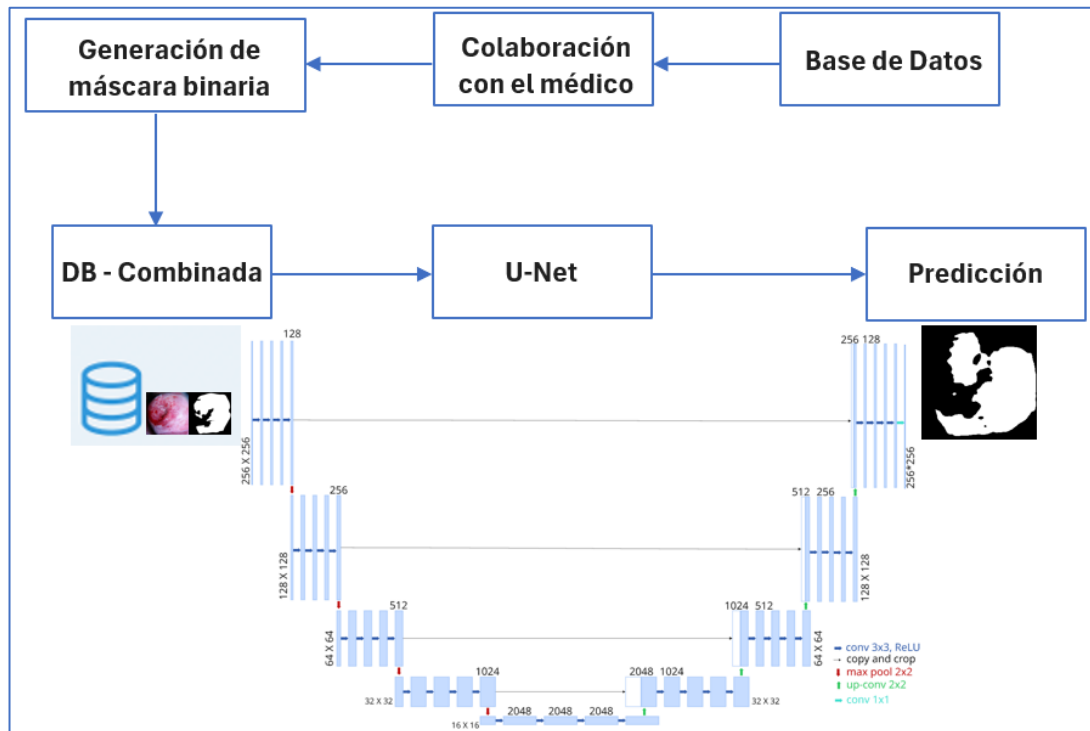


Figura 1. 2 Propuesta

1.6 Marco Teórico

1.6.1 Cáncer de cuello uterino

El cáncer de cuello uterino, también conocido como cáncer cervical o de cérvix, se caracteriza por la aparición de células malignas en los tejidos que forman el cuello del útero. Antes de desarrollarse en cáncer, las células experimentan alteraciones (displasia) que conducen a la formación de células anormales. Con el tiempo, las células anormales se transforman en cancerosas, se multiplican y se expanden hacia zonas más profundas del cuello uterino y las áreas cercanas. Este tipo de cáncer suele ser causado por el Virus del Papiloma Humano (VPH), una infección de transmisión sexual común que afecta a la zona genital. Según la Organización Mundial de la Salud (OMS), el 95% de los casos de cáncer de cuello uterino son causados por la infección VPH. Es importante señalar que las células anormales pueden tardar entre 15 y 20 años en transformarse en cáncer. Sin embargo, en mujeres con un sistema inmunitario debilitado, como en el caso de

quienes no reciben tratamiento para el Virus de la Inmunodeficiencia Humana (VIH), este proceso ocurre más rápidamente, en un período de 5 a 10 años. Algunos factores para la progresión es el grado de ecogenicidad del tipo de VPH, el estado del sistema inmunológico, la presencia de infecciones de transmisión sexual, la cantidad de partos, el embarazo a una edad temprana, el uso de anticonceptivos hormonales y el tabaquismo. Aunque la edad promedio de diagnóstico es alrededor de los 50 años, el cáncer se detecta con mayor frecuencia en mujeres entre 35 a 44 años [8], [9], [10].

1.6.2 Anatomía

El cuello uterino es la porción inferior y angosta del útero que enlaza la cavidad uterina con la vagina. Esta estructura fibromuscular posee una longitud de 3 a 4 cm, un diámetro de 2.5 cm y forma cilíndrica o cónica. La morfología de este órgano experimenta variaciones en tamaño y forma influenciadas por la edad, paridad y el ciclo menstrual. En el ámbito reproductivo, cumple una doble función esencial: sirve como vía para el transporte de espermatozoides hacia el útero durante la fertilización y actúa como un mecanismo de defensa que previene la entrada de agentes infecciosos al útero. Además, durante el parto la dilatación del cuello uterino es fundamental para facilitar el nacimiento del bebé [11], [12], [13]. En la Figura 1.3 se muestra las características anatómicas del cuello uterino:

- Orificio interno: abertura que une el cérvix con la región superior del útero.
- Endocérvis: abertura del cuello uterino que lleva hasta el útero y cubierto de epitelio cilíndrico/columnar de color rojizo.
- Ectocérvis o exocérvis: sección externa del cérvix, accesible a la observación directa durante un examen ginecológico. Histológicamente, presenta un recubrimiento de epitelio escamoso estratificado de color rosa.
- Orificio externo: conexión entre el cérvix y la vagina.

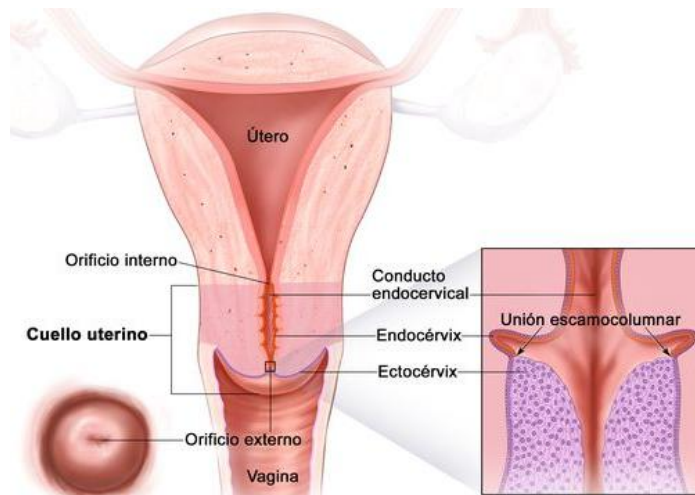


Figura 1. 3 Características anatómicas del cuello uterino [11]

En la Figura 1.4 se presenta una vista anatómica del cérvix desde la perspectiva del canal vaginal. Aquí es donde ocurren las lesiones precancerosas. En términos generales, es la zona en la que el endocérnix se encuentra con el ectocérnix se denomina unión escamoso-columnar (SCJ, por sus siglas en inglés). Esta área está conformada por células glandulares en forma de columna que producen moco del endocérnix, así como por células escamosas planas y delgadas del ectocérnix. Esta unión escamosa-columnar es la zona de transformación, ya que en ella se encuentran tanto la unión escamosa-columnar antigua como la nueva. El epitelio escamoso metaplásico hace referencia al cambio natural del epitelio cilíndrico del endocérnix hacia un epitelio escamoso, que se forma a partir de las células de reserva subcilíndricas. La mayoría de los cánceres cervicales se desarrollan a partir de las células ubicadas en la zona de transformación [14], [15].

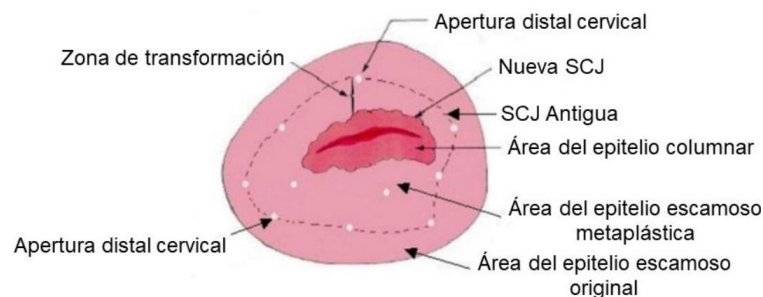


Figura 1. 4 Anatomía del cérvix visto desde el canal vaginal [14]

1.6.3 Precánceres

Las células de la zona de transformación no adquieren características cancerosas de manera repentina, sino que, previamente, las células normales del cuello uterino experimentan una serie de transformaciones anormales de manera gradual, se conocen como cambios precancerosos. Estos cambios se son designados con distintos términos, tales como Neoplasia Intraepitelial Cervical (CIN), Lesión Intraepitelial Escamosa (SIL) y displasia [14], [16].

En la Figura 1.5, se puede identificar los grados que tiene la CIN. A partir del grado II de la lesión es importante que se someta a la paciente a una biopsia, dado que, en estas etapas, es muy probable la presencia de cáncer de cérvix y no solo la lesión en tejido escamoso. Según la anormalidad del cuello uterino el CIN se clasifica en tres grados de acuerdo con la gravedad de las células [14], [16]:

1. CIN I: se denomina SIL de bajo grado o displasia leve y se clasifica como la forma menos severa, dado que la presencia de tejido anormal es mínima. Generalmente se resuelve por sí sola sin tratamiento. La lesión se encuentra solamente en el tejido del cérvix, afectando la zona de transformación.
2. CIN II: Lesión de grado intermedio, con cambios celulares más pronunciados, que puede progresar a un cáncer si no se trata.
3. CIN III: Lesión de alto grado, donde los cambios celulares son graves y tiene una alta probabilidad de progresar a cáncer cervical invasivo si no se trata.

El CIN II/CIN III, conocidos también como lesiones intraepiteliales escamosas de alto grado o displasias moderada/grave, presentan una mayor alteración del tejido. Estas lesiones representan la forma más avanzada de precáncer y conllevan un alto riesgo de progresar a cáncer. El diagnóstico se dificulta dependiendo del tipo de cérvix de la paciente, dado que esto genera que la zona de transformación presente una diferente ubicación y dependiendo del tipo, la CIN puede generar lesiones ocultas [14], [16].

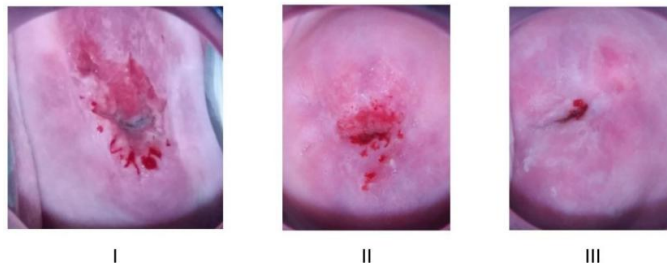


Figura 1. 5 Tipos de CIN [14]

1.6.4 Antecedentes

A continuación, se describen los antecedentes y factores vinculados al cáncer de cérvix:

El Virus del Papiloma Humano (VPH): La infección persistente por determinados tipos del Virus del Papiloma Humano (VPH) constituye la causa principal del cáncer de cérvix. Es una Infección de Transmisión Sexual (ITS) común que afecta las células del cuello uterino. Las infecciones por VPH, en su mayoría, no son permanentes y el sistema inmune se encarga de combatirlas, pero en algunos casos, el virus persiste y puede provocar cambios celulares que, con el tiempo, pueden convertirse en cáncer.

Historia clínica y factores de riesgo: Existen ciertos factores que incrementan la posibilidad de desarrollar el cáncer de cérvix:

- Infección por VPH: a infección persistente por VPH es la causa más importante.
- Edad: El diagnóstico de cáncer de cérvix es común en mujeres de entre 35 y 44 años.
- Actividad sexual temprana y múltiples parejas: las mujeres que inician su vida sexual a una edad temprana o que han tenido varias parejas sexuales presentan un mayor riesgo, ya que están más expuestas al VPH.
- Tabaquismo: El consumo de tabaco incrementa la probabilidad de padecer cáncer de cérvix, ya que debilita el sistema inmunológico y puede permitir que las células infectadas por el VPH se conviertan en cancerosas.
- Sistema inmunológico débil: Las mujeres con sistemas inmunológicos comprometidos, como las que viven con VIH/SIDA y las que toman

medicamentos inmunosupresores presentan un riesgo incrementado de desarrollar cáncer cervical.

Detección y diagnóstico: El Pap Test o prueba de Papanicolaou y la prueba del VPH son las herramientas clave para el diagnóstico temprano del cáncer de cérvix. La prueba de Papanicolaou analiza las células del cuello uterino para identificar posibles irregularidades, mientras que la prueba del VPH sirve para detectar los tipos de virus que podrían provocar alteraciones celulares. Si se detectan células anormales, es posible realizar pruebas adicionales (como una biopsia) utilizada para evaluar la presencia de cáncer o lesiones precancerosas.

Precancer y etapas tempranas: El cáncer cervical suele progresar de forma paulatina a partir de una condición precursora conocida como neoplasia cervical intraepitelial, que es un aumento inusual y descontrolado de células en el cuello uterino. Esta condición no siempre progresa a cáncer, y a menudo, si se detecta a tiempo, se puede tratar con éxito antes de que se convierta en cáncer CIN I, CIN II, CIN III.

1.6.5 Técnicas de imagenología

Métodos utilizados para capturar imágenes del interior del cuerpo humano con fines diagnósticos o investigación médica. Estas técnicas permiten a los profesionales de la salud visualizar estructuras internas sin necesidad de procedimientos invasivos. Las imágenes obtenidas se emplean para detectar enfermedades, identificar lesiones, planificar tratamientos y monitorear la evolución de una afección.

1.6.5.1 Colposcopia

Examen médico empleado para observar detalladamente el cuello uterino, la vagina y la vulva con el fin de detectar signos de enfermedades o alteraciones, como infecciones o cambios celulares que puedan indicar un riesgo de cáncer. Este procedimiento generalmente lo realiza un ginecólogo o un médico especialista y se emplea con frecuencia cuando los resultados del examen de Papanicolau (Pap) son anormales o cuando la paciente presenta síntomas como

sangrados anormales o lesiones visibles [14], [17]. Tiene la finalidad de identificar anomalías tales como lesiones intraepiteliales escamosas de alto grado (HSIL, por sus siglas en inglés), CIN entre otras.

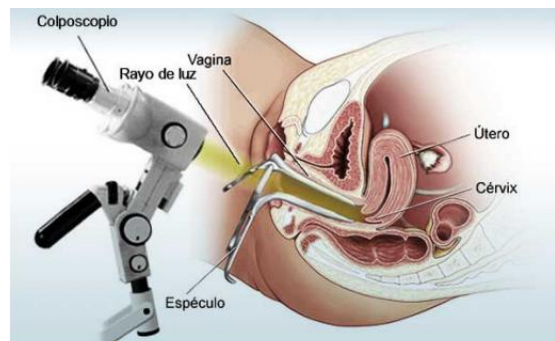


Figura 1. 6 Procedimiento de colposcopia [17]

En la Figura 1.6 se tiene el procedimiento de la colposcopia, el cual consiste en:

1. Se coloca en una mesa ginecológica, adoptando la posición típica de un examen pélvico (posición en litotomía). Luego, se inserta un espéculo en la vagina para separar las paredes vaginales y permitir una visualización clara del cuello uterino [14].
2. El colposcopio es un microscopio de luz brillante que se ubica a unos 30-40 cm de la vulva, sin hacer contacto directo con la paciente. El médico examina el cuello uterino a través del ocular del colposcopio para detectar posibles anomalías visibles, como alteraciones en el color, la textura o la vascularización del cuello uterino [14].
3. Para mejorar la visibilidad de las posibles lesiones, se aplican soluciones sobre el cuello uterino, estas son:
 - Ácido acético: Se aplica sobre el cuello uterino y se espera unos minutos. La sustancia hace que las áreas anormales se pongan blancas (signo de lesiones precoces o displasia).
 - Yodo (solución de Lugol): Tiñe las células normales de color marrón. Las células anormales no absorben el yodo y permanecen pálidas o no se tiñen, lo que ayuda a identificar áreas de preocupación.
4. Retiro del espéculo.

1.6.5.2 Biopsia

Técnica médica que consiste en obtener una pequeña porción de tejido del cuello uterino (cérvix) con el propósito de realizar un análisis de laboratorio. Este método es fundamental para verificar la presencia de células anormales en la muestra. La realización de la biopsia es crucial para establecer un diagnóstico certero y, a partir de los hallazgos, orientar el tratamiento más adecuado. Se realiza cuando se han detectado células anormales en estudios previos, como una citología (Papanicolaou) o una colposcopia. Su objetivo es confirmar si dichas células anormales son cancerosas, lo cual es crucial para identificar el tipo y el alcance de las lesiones, y así definir el pronóstico y las opciones terapéuticas para la paciente [18].

1.6.5.3 Estudios por imágenes

Se utilizan para evaluar la extensión del cáncer, es decir, si ha invadido otros órganos o estructuras cercanas. Las técnicas más comunes incluyen:

Ultrasonido transvaginal: Método más usados para examinar el cervix y órganos adyacentes, como el útero, los ovarios y la vejiga. Puede detectar anomalías, masas o modificaciones en el tamaño y la morfología del cuello uterino [19].

Tomografía Computarizada (TC): Es una técnica de imagen que utiliza rayos X para generar representaciones detalladas del interior del cuerpo. Para determinar la extensión del cáncer cervical (si se ha diseminado a otras zonas como los ganglios linfáticos, los huesos o los pulmones) y es útil en estadios más avanzados para evaluar metástasis [20].

Resonancia Magnética (RM): Obtiene imágenes precisas de los tejidos blandos y se utiliza para evaluar la invasión del cáncer en el útero, la vagina y los órganos cercanos [21].

1.6.6 Segmentación de imagen médica

Tiene un rol importante en el ámbito de la salud, ya que permite extraer información valiosa para el diagnóstico de forma semiautomática o automática, y facilitar el seguimiento de diversas patologías [22]. El propósito de la

segmentación es la separación y detección de la ROI de la imagen, esto dependerá del problema a resolver, por ejemplo, el caso de lesiones en cérvix se obtendría información como la textura, intensidad de color, forma, bordes, etc. [23].

La segmentación es un procedimiento mediante el cual se divide una imagen en distintas regiones u objetos de interés. En este proceso, a cada píxel se le asigna una etiqueta o categoría que indica la región a la que pertenece. Los píxeles que comparten la misma etiqueta presentan características similares entre sí, que pueden ser el nivel gris, brillo, color, textura, contraste, etc. [24].

Las técnicas de segmentación en imágenes se utilizan para localizar órganos, tejidos, tumores o lesiones asociadas a diversas patologías. Al identificar una lesión, se obtiene información relevante para el análisis de la región segmentada. Estas técnicas se aplican a imágenes obtenidas mediante distintas tecnologías, como colposcopia, mamografía, ultrasonido, resonancia magnética, Tomografía por Emisión de Positrones (PET, por sus siglas en inglés), las cuales suelen requerir preprocesamiento o mejoras de calidad de la imagen médica. Al trabajar con imágenes médicas, se enfrentan diversos desafíos, como la presencia de artefactos, baja resolución, volúmenes parciales, movimientos y ruido en la imagen [24].

En la Figura 1.7, se muestra la segmentación de lesión de cuello uterino (cérvix), donde a) es el cérvix y b) la máscara binaria que corresponde la lesión (color blanco) y no lesión (color negro). Por ende, el resultado de llevar la segmentación en la imagen médica es una nueva imagen donde se identifica la lesión de la patología conocida como máscara. Es decir, en cada píxel tiene establecido un color según la categoría o etiqueta asignada. Generalmente en segmentación de imágenes médicas se establece 2 categorías; lesión de color blanco y no lesión color negro. Para obtener una segmentación automática se utiliza técnicas de aprendizaje profundo de inteligencia artificial mediante del uso de redes neuronales artificiales.

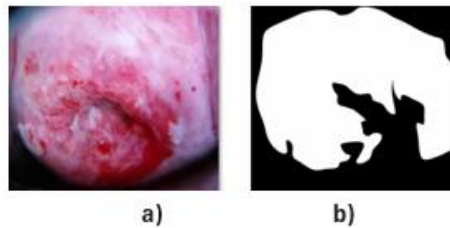


Figura 1. 7 Segmentación: a) imagen del cérvix, b) máscara binaria

1.6.7 Técnicas de segmentación

- Umbralización: establece umbrales de intensidad y clasifica los píxeles según el nivel de umbral al que corresponden [25]. Es decir, se analiza la imagen y se determina la región a segmentar utilizando el umbral óptimo, el cual se define a partir del histograma de la imagen. El histograma representa gráficamente cómo se distribuyen las diferentes intensidades de píxeles en la imagen [26].
- Detección de bordes: consisten en detectar las posiciones que se hallan en los bordes de las imágenes. El borde es la frontera entre dos regiones de la imagen que poseen distintas intensidades de color o niveles de gris [27]. Estos cambios marcan la separación entre diferentes regiones de la imagen, y permiten delimitar los contornos de los elementos presentes. A partir de estos bordes, es posible definir regiones cerradas que representan los objetos segmentados dentro de la imagen [24].
- Agrupamiento: reconocer y agrupar muestras de píxeles de la imagen, estos grupos de píxeles se los conoce como clústeres. El agrupamiento involucra dividir el área de la imagen en regiones de características similares [28]. Esta separación de grupos de píxeles pertenece a un espacio característico derivado de la imagen. Este espacio característico, es un conjunto de características que corresponde a funciones, bien sea de la imagen o provistas por el usuario [26].

1.6.8 Asistencia en la segmentación

La asistencia en la segmentación en imágenes médicas, tiene un rol importante en los Sistemas de Diagnóstico Asistido por Computador (CAD,

por sus siglas en inglés). El sistema CAD, es una herramienta auxiliar que facilita a los radiólogos a tomar decisiones de diagnóstico ante anomalías presentes en la imagen médica. Se utiliza, en el diagnóstico de cáncer de mama, cáncer de pulmón, enfermedad de las arterias, defectos cardíacos, detección patología del cerebro, etc. En general, se limita a marcar áreas posibles, donde puede existir una lesión o anomalías. Siendo el responsable el radiólogo de delimitar el área de la lesión para que pueda ser caracterizada y diagnosticada [29]. El sistema CAD consta de varias etapas:

- Preprocesamiento: reducción de artefactos de fondo, eliminación de ruido y mejorar la calidad de imagen.
- Segmentación: localizar y buscar regiones sospechosas que poseen anomalía como tejido, lesiones, masas, etc.
- Extracción de características: la región segmentada es analizada en busca de características sospechosas como textura, intensidad, tamaño, ubicación, borde, etc.
- Clasificación: se clasifican las regiones sospechosas utilizando las características de la etapa anterior.

CAPÍTULO 2

2. ESTADO DEL ARTE

2.1 Herramientas de asistencia en segmentación

La segmentación de imágenes médicas es una tarea importante en el procesamiento de imágenes biomédicas, ya que permite la identificación y delimitación precisa de estructuras anatómicas y regiones de interés. En las últimas décadas, se han desarrollado diversas herramientas de software que permiten asistir este proceso, facilitando tanto el trabajo clínico como la investigación.

2.1.1 3D Slicer

Una plataforma de software libre y de código abierto diseñada para el análisis y visualización de imágenes médicas. Permite realizar segmentación, registro, reconstrucción 3D, análisis cuantitativo y visualización interactiva. Su arquitectura modular y su compatibilidad con múltiples formatos de imagen médica (como DICOM, NIfTI, entre otros) lo hacen una herramienta versátil tanto en investigación como en entornos clínicos [30].

Una de las principales fortalezas de 3D Slicer es la posibilidad de integrar algoritmos avanzados de segmentación basados en inteligencia artificial, aprendizaje profundo o métodos tradicionales como crecimientos de regiones, umbralización y métodos basados en contornos activos. Además, permite la personalización a través de scripts en Python y la integración de módulos externos [30].

Características:

- Soporte multiplataforma (Windows, macOS, Linux).
- Interfaz gráfica intuitiva y altamente configurable.
- Repositorio extensivo de módulos desarrollados por la comunidad.
- Amplio soporte para imágenes volumétricas y reconstrucción 3D.

2.1.2 ITK-SNAP

Aplicación de software gratuita de código abierto y multiplataforma que se utiliza para segmentar estructuras en imágenes biomédicas 3D y 4D. Fue desarrollada originalmente en la Universidad de Carolina del Norte. Esta herramienta proporciona segmentación semiautomática mediante métodos de contorno activo, así como delineación manual y navegación por imágenes [31].

ITK-SNAP se centra específicamente en la segmentación, lo hace más ligero y fácil de utilizar para tareas concretas. Es particular en estudios donde se requiere una intervención manual mínima pero precisa, como en la segmentación de tumores, ventrículos cerebrales y lesiones focales [31].

Características:

- Acceso a algoritmos avanzados a través del servicio de segmentación distribuida.
- Visualización en 3D y en cortes ortogonales (axial, coronal, sagital).
- Una interfaz gráfica de usuario moderna basada en Qt6.
- Compatibilidad con imágenes en color, multicanal y 3D+tiempo.
- Soporte para visualización simultánea y vinculada, y segmentación de múltiples imágenes.
- Capacidades de registro de imágenes manuales y automáticas.

2.1.3 ImageJ - Fiji

ImageJ es un software de procesamiento de imágenes ampliamente utilizado en la ciencia y la investigación. Originalmente desarrollado por Wayne Rasband en el National Institutes of Health (NIH, por sus siglas en inglés) en los Estados Unidos, ImageJ se ha convertido en una herramienta fundamental para la visualización, el análisis y la manipulación de imágenes científicas. Es una aplicación de código abierto, lo que permite a los usuarios modificar y personalizar el software según sus necesidades. ImageJ soporta

una amplia gama de formatos de imágenes y proporciona una extensa colección de herramientas y funciones para el procesamiento de imágenes, como la medición de áreas, la filtración, la segmentación y la creación de imágenes en 3D. Además, su capacidad para integrar macros y plugins le otorga una gran flexibilidad y extensibilidad, lo que lo convierte en una opción popular entre los científicos de diversas disciplinas, como la biología, la medicina y la física [32].

Fiji una distribución de ImageJ que incluye una serie de complementos (plugins) y bibliotecas adicionales para facilitar el análisis de imágenes científicas de manera más eficiente. Aunque Fiji se basa en ImageJ, su principal ventaja radica en la integración de una gran cantidad de herramientas preinstaladas, lo que elimina la necesidad de que los usuarios descarguen e instalen complementos por separado. Esto convierte a Fiji en una opción más accesible para quienes necesitan un entorno listo para usar sin complicaciones. Además, incluye mejoras en la interfaz de usuario y proporciona soporte optimizado para trabajos de análisis en imágenes volumétricas y series de imágenes. La compatibilidad con sistemas operativos como Windows, macOS, Linux y su enfoque en la comunidad de código abierto le otorgan una gran flexibilidad para ser utilizado en entornos de investigación colaborativos [33].

2.2 Algoritmos de segmentación: Aprendizaje Automático (ML)

Los algoritmos de segmentación basados en Aprendizaje Automático permiten dividir conjuntos de datos en grupos con características similares de forma automática y eficiente. Entre los métodos más utilizados se encuentran K-means, Random Forest y Máquinas de Soporte Vectorial (SVM). Estos algoritmos pueden aplicarse tanto en tareas supervisadas como no supervisadas, dependiendo del tipo y disponibilidad de datos. Su uso se ha extendido en áreas como la visión por computadora y la medicina. Gracias a la capacidad de adaptación y aprendizaje, el ML ha superado a muchos enfoques tradicionales en precisión y escalabilidad.

2.2.1 K-means

El artículo de Bai et al. [34], se utiliza el algoritmo k-means para segmentar la imagen colposcópica transformada al espacio de color HSV, específicamente para extraer la componente V, que es la que contiene la mayor información clínica sobre la región cervical. En este proceso, los píxeles de la imagen se dividen en K centros de agrupamiento, y luego se calcula la distancia de cada muestra a los centros de los grupos, asignando cada muestra al centro de grupo más cercano. Este enfoque permite separar las distintas áreas de la imagen, como la región cervical, la zona extra-cervical y la región del endoscopio, facilitando así la segmentación precisa de la región de interés [34].

Para llevar a cabo esta segmentación, se siguen varios pasos. Primero, se seleccionan aleatoriamente K centros de agrupamiento en el espacio HSV. Luego, cada punto de la imagen se asigna al centro de agrupamiento más cercano, lo que se calcula mediante una fórmula de distancia. Posteriormente, los centros de agrupamiento se actualizan, y el proceso de asignación se repite iterativamente hasta que los centros ya no cambian, indicando que se ha alcanzado la convergencia. Con este procedimiento, se obtienen tres regiones principales: la zona cervical, la región externa del cuello uterino y la zona del endoscopio [34].

Una vez realizada la segmentación utilizando K-means, se aplica un filtro de área para suavizar los bordes de la región cervical. Este filtro elimina las pequeñas regiones conectadas que no corresponden a la zona cervical, asegurando que solo se mantenga la parte relevante de la imagen para el análisis. El filtro también ayuda a preservar las características detalladas de la imagen, mientras suprime el ruido, lo que mejora la precisión de la segmentación y facilita la evaluación clínica. Este enfoque ha demostrado ser efectivo para identificar y segmentar de manera precisa la región cervical, incluso en imágenes con detalles complejo [34].

2.2.2 Random Forests (RF)

El artículo de Yao & Li [35], este estudio se propone un método mejorado para el reconocimiento de cambios citopáticos cervicales, utilizando el algoritmo Random Forest (RF, por sus siglas en inglés) enfocado en el reconocimiento de características locales de las lesiones. Para optimizar el tiempo de ejecución del algoritmo, se aplicó una simplificación en la selección de características, contribuyó la precisión de identificación. Además, se empleó una selección equilibrada de árboles de decisión para fortalecer la robustez del modelo. La efectividad del método fue validada mediante simulaciones que mostraron mejoras significativas [35].

El algoritmo Random Forest se construye a partir de múltiples árboles de decisión generados mediante el método de muestreo Bootstrap. A partir del conjunto de entrenamiento, se extraen aleatoriamente múltiples subconjuntos de datos, y para cada uno de ellos se selecciona un subconjunto de características de manera aleatoria. En cada nodo de un árbol, se elige la mejor característica para dividir, utilizando criterios como la ganancia de información. La clasificación final se realiza mediante votación mayoritaria entre todos los árboles, lo que permite una toma de decisión más precisa y reduce la posibilidad de sobreajuste [35].

Para comprobar la eficacia del modelo, se utilizaron 280 imágenes TCT de células cervicales obtenidas de un hospital, divididas entre células normales y células con lesiones. El algoritmo fue implementado en un entorno Matlab, alcanzando una sensibilidad del 93.5% en la identificación de lesiones. Los resultados indican que el modelo RF no solo es factible y práctico, sino que también mejora la precisión en el reconocimiento de enfermedades cervicales, superando incluso a métodos tradicionales [35].

2.2.3 Máquinas de Soporte Vectorial (SVM)

El artículo de Artan & Huang [36], Máquinas de Soporte Vectorial (SVM, por sus siglas en inglés) es un método de clasificación utilizado con frecuencia en aplicaciones de procesamiento de imágenes y reconocimiento de patrones. El objetivo de SVM es construir un hiperplano de decisión que separe las diferentes clases con el margen máximo, es decir, la distancia máxima entre el hiperplano de decisión y el punto más cercano a este. Este hiperplano de decisión se obtiene mediante la resolución de un problema de programación cuadrática convexa. Para aplicaciones de clasificación, se utiliza la formulación C-SVM original, donde el parámetro C controla el sobreajuste. Sin embargo, en ciertas aplicaciones, la penalización del error es desigual entre las clases, lo que da lugar a la propuesta de un SVM ponderado, el 2C-SVM, donde se asignan diferentes costos a los errores de clasificación de una clase más importante que la otra [36].

En aplicaciones de clasificación donde el costo de etiquetar erróneamente una muestra de una clase de región acetoblanca (AW, por sus siglas en inglés) es mayor que el de cometer un error con la clase región de cuello uterino (CE, por sus siglas en inglés), se propone el uso de un SVM sensible al costo. Este enfoque se aplica en la segmentación de imágenes del cérvix para clasificar las regiones de tejido como AW o CE. Para optimizar el rendimiento de clasificación, se utiliza un sistema de clasificadores múltiples basado en SVM sensible al costo. Este método ha demostrado buenos resultados en un conjunto reducido de imágenes, donde cada imagen aporta cientos de muestras de píxeles AW y CE. Durante la fase de entrenamiento, se determinan los mejores parámetros para cada clasificador y se aplican durante la prueba a las imágenes nuevas [36].

El uso de un SVM sensible al costo ha mostrado ser ventajoso en la segmentación de lesiones en imágenes del cérvix. Al emplear este enfoque, se realiza una búsqueda de parámetros en una malla regular para minimizar la tasa de falsos positivos y maximizar la tasa de verdaderos positivos. Cada

clasificador 2v-SVM proporciona una segmentación de la imagen, y estas segmentaciones se combinan utilizando el algoritmo STAPLE maneja información de verosimilitud iterativa sobre píxeles individuales para estimar la segmentación verdadera oculta de las áreas afectadas (AW) de una imagen de prueba dada. Los resultados obtenidos con esta técnica son prometedores, y el sistema está siendo optimizado mediante la selección dinámica de clasificadores, lo que permite elegir los clasificadores más adecuados para cada imagen de prueba [36].

2.3 Algoritmos de segmentación: Aprendizaje Profundo (DL)

El Aprendizaje Profundo (Deep Learning) ha revolucionado los algoritmos de segmentación al permitir el análisis de datos complejos con un alto nivel de precisión. Modelos como las redes neuronales convolucionales, U-Net y DeepLab son ampliamente utilizados para tareas de segmentación, especialmente en el procesamiento de imágenes. Estos algoritmos pueden aprender representaciones jerárquicas de los datos, lo que mejora significativamente la detección de patrones sutiles y detallados. Su aplicación ha demostrado resultados sobresalientes en campos como la medicina. A pesar de requerir grandes volúmenes de datos y poder computacional, el DL se posiciona como una de las técnicas más efectivas en segmentación avanzada.

2.3.1 Mask R-CNN

El artículo de Saranya & Sujatha [37], propone un sistema de detección de cáncer de cuello uterino mediante la aplicación del modelo Mask R-CNN, una arquitectura de aprendizaje profundo en la detección de objetos y segmentación por instancia. Se adapta para analizar imágenes de colposcopia, permitiendo no solo identificar la presencia de lesiones cervicales, sino también delinear sus contornos. El modelo, preentrenado con el conjunto de datos COCO, es ajustado con imágenes etiquetadas de colposcopia para aprender las características de las lesiones cervicales. Esta adaptación permite al sistema detectar las lesiones precancerosas

como cancerosas, facilitando el análisis detallado y automatizado que puede complementar el diagnóstico clínico tradicional [37].

La arquitectura Mask R-CNN se basa en una red convolucional profunda que extrae características relevantes de las imágenes. Utiliza una red de propuestas de regiones para identificar áreas que podrían contener lesiones, las cuales son posteriormente refinadas mediante capas adicionales que predicen tanto la clase (lesión o no) como una máscara binaria que segmenta la región afectada. Este enfoque permite manejar imágenes con diferentes calidades, resoluciones y variabilidad en las condiciones de iluminación o estructura del tejido, es común en imágenes clínicas reales. Al generar una máscara de segmentación para cada región de interés, se proporciona una representación visual clara de las zonas sospechosas, lo que resulta de gran utilidad para el personal médico [37].

El modelo se optimiza mediante Adam que ajusta las tasas de aprendizaje para mejorar la convergencia y estabilidad durante la fase de entrenamiento. La combinación entre la arquitectura Mask R-CNN y la optimización de Adam permite el desarrollo del sistema robusto, capaz de distinguir entre células sanas y cancerosas. El modelo se evalúa con métricas como precisión, sensibilidad, especificidad, y se integra en una plataforma que permite a los profesionales de la salud cargar imágenes y recibir análisis automáticos en tiempo real. De este modo, ofrecen una herramienta eficiente para apoyar el diagnóstico temprano del cáncer de cérvix [37].

2.3.2 DeepLab V3

En Chun Li et al. [38], evaluó un estudio de la eficacia del modelo DeepLab V3 para la segmentación automática de tumores cervicales a partir de imágenes de resonancia magnética. La red empleó Xception como arquitectura base e implementó un proceso de normalización de las intensidades de imagen, técnicas de aumento de datos para fortalecer el

modelo. En total, fue entrenado con más de 10 mil imágenes derivadas de 256 pacientes con cáncer uterino [38].

Este modelo fue aplicado a un conjunto de datos específico del cérvix compuesto por 144 pacientes. Se desarrollaron tres enfoques para evaluar su rendimiento: un modelo exclusivo del útero entrenado desde cero, un modelo de transferencia de aprendizaje que reutilizaba el modelo cervical preentrenado, y un modelo agregado que combinaba datos uterinos y cervicales para evaluar su capacidad de generalización. Para el modelo de transferencia de aprendizaje, se exploraron diferentes niveles de congelamiento y ajuste de capas, evaluaron cómo la reutilización de características extraídas previamente influía en la precisión del modelo al segmentar lesiones cervicales [38].

Los resultados de la segmentación se evaluaron mediante el coeficiente de similitud Dice, teniendo una alta concordancia entre las segmentaciones automáticas y las manuales realizadas por especialistas. Asimismo, para validar la utilidad clínica de los modelos, extrajeron características radiómicas de los ROIs predichos y compararon con las obtenidas manualmente, teniendo consistencia entre ambos enfoques. Esto respalda la efectividad de DeepLab V3 como una herramienta precisa y confiable para la segmentación automática de lesiones del cérvix [38].

2.3.3 Redes Neuronales Generativas (GANs)

El artículo de Huang et al. [39], la segmentación de células cervicales en imágenes de smear es un paso crucial para la detección temprana del cáncer de cérvix. Esta tarea presenta grandes desafíos debido al bajo contraste, la variabilidad morfológica de las células y la superposición entre ellas. Para abordar este problema propusieron un enfoque basado en redes generativas adversariales denominado Cell-GAN, que segmenta células individuales incluso en imágenes densamente pobladas. La red aprende la distribución de probabilidad de la morfología celular mediante la comparación entre

imágenes generadas y anotadas manualmente, permitiendo identificar la estructura completa de una célula y descartar información no relevante, como células superpuestas o ruido [39].

El modelo Cell-GAN se basa en la arquitectura DCGAN y está diseñada como un autoencoder, donde el codificador utiliza módulos Inception y el decodificador emplea capas de deconvolución más simples. El sistema recibe como entrada una imagen con múltiples células y un factor guía el núcleo, que actúa como referencia para localizar y segmentar únicamente la célula de interés. A través de una fusión aditiva de características, enfatizan las regiones relevantes del núcleo, evitando la generación de imágenes con múltiples células. Durante el posprocesamiento, aplican un umbral binario alto y eliminan contornos irrelevantes para refinar los resultados, asegurando que la imagen final contenga una única célula segmentada [39].

Los resultados muestran que logran una segmentación precisa tanto en imágenes de células individuales como en aquellas con superposición. Alcanzaron valores de 94.3% en el coeficiente Dice y 7.9% de falsos negativos en imágenes individuales, y 89.9% Dice y 6.4% falsos negativos en imágenes con células superpuestas. Las métricas demuestran la eficacia del modelo para adaptarse a distintas condiciones morfológicas y posicionando al modelo como una alternativa para la segmentación automática en contextos clínicos [39].

CAPÍTULO 3

3. METODOLOGÍA

La metodología se presenta en la Figura 3.1, esta consista en:

- Base de datos (DB): se utilizó un conjunto de imágenes de colposcopia de la plataforma web Kaggle y se tomó 200 imágenes para realizar la investigación. Esta DB no tiene las máscaras binarias que corresponde a las lesiones del cérvix.
- Preprocesamiento: se seleccionó el cérvix como la ROI en las imágenes de colposcopia.
- Segmentación manual: se procedió con la selección, delimitación de las lesiones del cérvix en las imágenes con colaboración y capacitación del médico especialista. Posterior, se generó la máscara binaria donde estará la lesión del cérvix.

Se realizó un estandarizado de [256x256] píxeles en las imágenes de colposcopia y máscaras binarias, debido a que hay variedad en tamaño en las imágenes.

- Aumento de datos: para aumentar el conjunto de datos se utilizó 4 rotaciones para cada imagen de colposcopia y máscara binaria.
- Segmentación automática: se utilizó U-Net para segmentar las imágenes de colposcopia del conjunto de imágenes de validación.
- Predicción: con el algoritmo entrando se procedió con la predicción de la máscara binaria donde estará la lesión del cérvix y se evaluó el rendimiento del modelo con métricas para segmentación: Intersección sobre Unión (IoU) y el coeficiente de Dice.

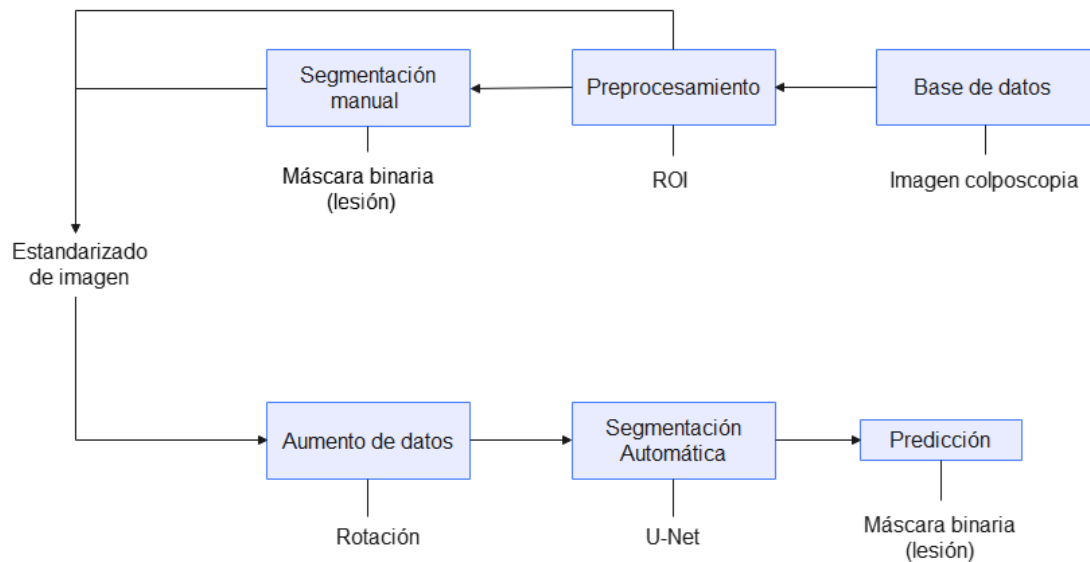


Figura 3. 1 Metodología: segmentación de lesiones en imagen de colposcopia

3.1 Base de datos

Se utilizó una Base de Datos (DB) internacional de la plataforma web Kaggle con el nombre Intel & MobileODT Cervical Cancer Screening. Esta DB contiene archivos con imágenes de colposcopia: conjunto para entrenamiento de 1482 imágenes; conjunto para test de 512 imágenes. Adicional, hay archivos de imágenes con lo siguiente: conjunto de test_stg2 de 3506 y un conjunto de imágenes adicionales de 6734 imágenes, sobre este conjunto menciona que son imágenes de pacientes duplicados, mismas que pueden parecer similares, son tomadas en la misma sesión y otras son casos que no se han seleccionado debido a la calidad de la imagen [40].

Para llevar la investigación se seleccionó el archivo de imágenes de colposcopia de entrenamiento tienen formato JPG (en ingles Joint Photographic Experts Group) con dimensiones aproximada [4000x3000] píxeles. El tamaño de las imágenes varía, algunas imágenes son borrosas con ruido que no se observan las patologías. Cabe indicar que la DB no contiene las máscaras binarias que corresponde a la identificación de la lesión en el cérvix. Por ende, el conjunto de imágenes de colposcopia que se utilizó son 200 y se generó 200 máscaras binarias que corresponden a la lesión del cérvix que se utilizó para el entrenamiento y validación del modelo.

3.1.1 Extracción manual del ROI

Para la extracción manual de la región de interés (ROI, por sus siglas en inglés) se utilizó ImageJ con distribución Fiji (ver Figura 3.2). Un software informático de tratamiento digital de imagen orientada a las ciencias de la salud, desarrollado en los Institutos Nacionales de Salud de Estados Unidos. Esta herramienta cuenta con opciones para seleccionar, editar, procesar, analizar imágenes, etc. [32]. Con fiji se procedió a seleccionar y extraer la ROI para este caso el cérvix en las 200 imágenes de colposcopia.

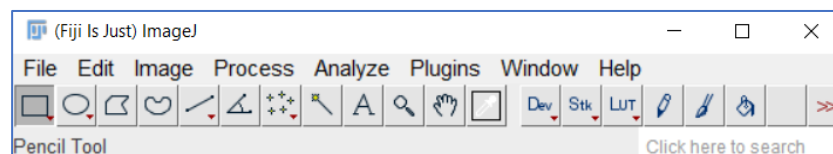


Figura 3. 2 Software ImageJ con distribución Fiji [32]

3.1.2 Mascara binaria

Se utilizó la herramienta de Fiji para generar la máscara binaria. A partir de la ROI del cérvix, se escogió una opción para delimitar la segmentación de la lesión en el cérvix para esto se empleó la selección de polígono ubicado en la parte izquierda de la herramienta.

El criterio que se consideró con la colaboración y capacitación del médico especialista fue fundamental para la segmentación manual e identificación de las lesiones en el cuello uterino en imagen de colposcopia (ver Figura 3.3 literal a) es observar la zona del ectocérnix la presencia blanquecina manto superpuesto y tipo coliflor en la zona debido a la aplicación de la sustancia ácido acético que permite observar las posibles lesiones, ya que esta zona es de color rosa antes de aplicar la sustancia.

Una vez identificada la lesión, se utilizó la opción de create mask para generar la máscara binaria correspondiente. Estas servirán como *ground-truth* para entrenar y validar el modelo de segmentación. En la Figura 3.3, se

muestra lo mencionado y este proceso se realizó para cada una de las respectivas imágenes de colposcopia del presente trabajo.

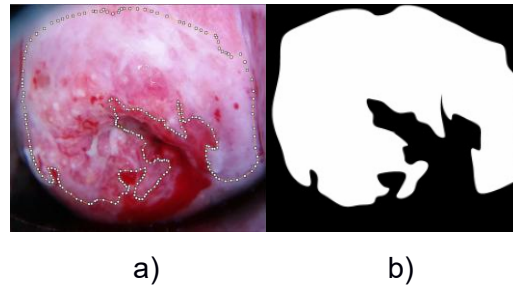


Figura 3. 3 a) Selección de lesión en cérvix, b) Generación de máscara binaria (lesión)

3.1.3 Aumento de datos

Antes de realizar el aumento de datos se procedió a redimensionar al tamaño [256x256] píxeles tanto para las ROI del cérvix con las macaras binaria de la lesión. Con el objetivo que el conjunto de imágenes sea de un mismo tamaño. Posterior, para realizar el aumento de datos se utilizó rotación en 4 giros para cada una de ROI y mascara binaria, como se muestra en la Figura 3.4. Por ende, el conjunto final de datos es de 800 imágenes.

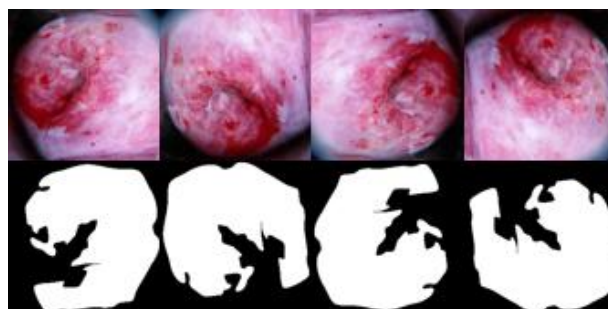


Figura 3. 4 Rotación imágenes

3.2 U-Net

Una arquitectura de red completamente convolucional, reconocida por su estructura simétrica en forma de U, y compuesta por 23 capas convolucionales. Fue propuesta en 2015 por Olaf Ronneberger, Philipp Fischer y Thomas Brox con el objetivo de realizar segmentación de imágenes biomédicas. Esta red tiene la capacidad de ser entrenada de manera eficiente incluso con un número reducido de imágenes [41].

La arquitectura no tiene ningún tipo de limitación de uso exclusivo a la medicina, también, es aplicable a diferentes tipos de problemas de segmentación.

La red U-Net se basa en aplicar convoluciones a la imagen de entrada para comprimir la información y extraer características, lo que se conoce como la ruta de contracción. Posteriormente, se genera la nueva imagen utilizando las características aprendidas en esta fase, en un proceso denominado ruta expansiva. Durante este proceso, se emplean deconvoluciones, que son operaciones inversas a las convoluciones, y que usan los filtros previamente aprendidos para restaurar la imagen comprimida [42].

3.2.1 Red neuronal artificial

Es un modelo computacional que toma como referencia el funcionamiento del cerebro humano, orientado al reconocimiento de patrones y al aprendizaje a partir de datos. Está formado por una estructura de neuronas en distintas capas: una capa inicial de entrada, una o varias capas intermedias denominadas ocultas, y una capa final de salida. Cada neurona recibe señales de entrada, las transforma mediante una función matemática y luego envía el resultado a las neuronas de la siguiente capa [43]. Este tipo de red es utilizado en tareas de segmentación, clasificación, entre otros.

En la Figura 3.5, se presenta la arquitectura U-Net modificada con 41 capas convolucionales que consiste en una ruta de contracción (parte codificador) y una ruta expansiva (parte decodificador). El recuadro azul corresponde a un mapa de características, con la cantidad de canales indicada en la parte superior del mismo. El tamaño(x,y) se especifican en la esquina inferior izquierda (caja azul). Los recuadros blancos representan mapas de características duplicados, mientras que las flechas señalan las distintas operaciones realizadas [41], [42].

La ruta de contracción, se basa en la aplicación repetida de cuatro convoluciones con filtros de tamaño 3x3, seguidas por la activación ReLU

(unidad lineal rectificada), y una capa de agrupamiento mediante max-pooling 2x2. Esta última se encarga de reducir a la mitad la resolución espacial del mapa de características. Con cada reducción en la resolución, es decir, al descender un nivel en la red se duplica la cantidad de canales en las capas convolucionales, comenzando con 128 y aumentando progresivamente a 256, 512, hasta alcanzar los 2048. Esta etapa tiene como finalidad principal capturar el contexto global de la imagen para facilitar el proceso de segmentación [41], [42].

La ruta de expansión, consiste en realizar una deconvolución o convolución transpuesta (up-convolution), para la deconvolución se hace uso de un upsampling seguido de una convolución de tamaño 2x2; el upsampling duplica las dimensiones del mapa de características, mientras que la convolución reduce a la mitad la cantidad de canales. Luego, se concatena este resultado con el mapa de características recortado correspondiente de la ruta de contracción, y a continuación se aplican cuatro capas convolucionales de 3x3, cada una seguida por la función de activación ReLU. Finalmente, en la capa de salida se emplea una convolución 1x1 para transformar cada vector de características de 64 elementos en el número deseado de clases [41], [42].

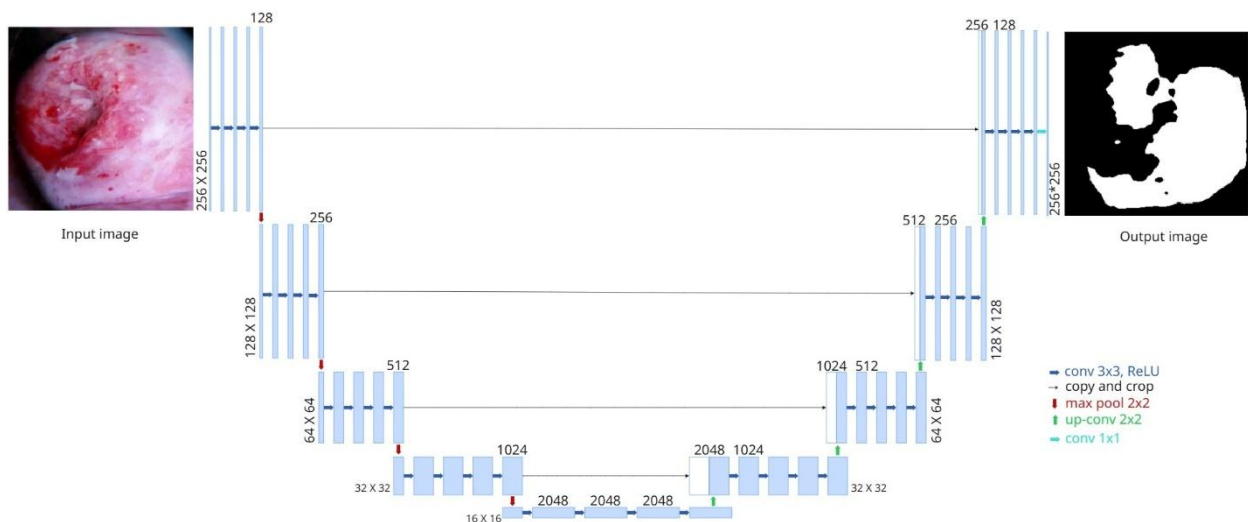


Figura 3. 5 Arquitectura U-Net modificada

3.3 Marco Experimental

3.3.1 Validación

El conjunto de datos de 800 ROIs del cérvix y 800 máscaras binarias se dividió en dos partes, uno de entrenamiento y otro de validación. El conjunto de entrenamiento como su nombre indica, se usa para entrenar el modelo y el conjunto de validación para validar el modelo. Para nuestro proyecto se dividió el 80% en entrenamiento equivalente a 640 imágenes ROIs con sus máscaras binarias, y 20% en validación que corresponden a 160 imágenes ROIs con sus máscaras binarias.

3.3.2 Entorno de trabajo

Para el desarrollo del Proyecto de Titulación se empleó las siguientes herramientas:

Anaconda: Interfaz gráfica de usuario de escritorio que facilita el lanzamiento de aplicaciones como Jupyter Notebook, Spyder, Rstudio. Se utilizó para crear un entorno de desarrollo controlado. En el cual se gestionó los paquetes como NumPy, Pandas, Matplotlib, entre otros para el procesamiento de imágenes. Además, la creación del entorno virtual permitió trabajar con versiones específicas de TensorFlow y Keras, lo que garantizó la compatibilidad entre las herramientas y librerías utilizadas [44].

Spyder: Se utilizó como Entorno de Desarrollo Integrado (IDE) disponible a través de Anaconda, debido a su compatibilidad con Python y las bibliotecas científicas orientadas al aprendizaje automático, como TensorFlow, Keras, OpenCV. Este entorno fue empleado para el desarrollo interactivo de código, permitiendo la ejecución y visualización en tiempo real de los resultados durante el proceso de desarrollo. Esto resultó útil en el contexto en la experimentación y ajuste de hiperparámetros durante la fase de desarrollo del proyecto [45].

Python: Se utilizó por su versatilidad y amplia disponibilidad de librerías especializadas en procesamiento de imágenes en inteligencia artificial. Su sintaxis clara y su capacidad de manejar grandes volúmenes de datos permitió un desarrollo rápido, facilitando la experimentación y ajuste de los modelos [46].

Keras: Se utilizó para la implementación y construcción del modelo U-Net, facilitando la creación y ajuste de la arquitectura del modelo sin complicaciones. Esta librería permitió experimentar con diferentes capas y parámetros, optimizando el modelo para la segmentación de las imágenes [47].

Tensorflow: Se utilizó en el proyecto debido a la necesidad de manejar cálculos intensivos y grandes volúmenes de datos. Las imágenes de colposcopia al ser de tamaño [256x256] píxeles, requieren un procesamiento considerable de cálculos de cómputo. Tensorflow garantizó que el modelo se ejecutara de forma optimizada para aprovechar la aceleración por GPU, para el entrenamiento del modelo. debido a la capacidad de la GPU para manejar eficientemente la gran cantidad de operaciones matemáticas involucradas en el entrenamiento del modelo [48].

3.3.3 Métricas de evaluación

Para comprobar la eficiencia del modelo de red, se procede a evaluar los resultados obtenidos mediante métricas de evaluación para la segmentación de imágenes. En la figura 3.6, se observa los 4 tipos de predicciones, el positivo como presencia de lesión (píxel activo) y el negativo como no lesión o fondo [49].

Verdaderos Positivos (TP): aquellos píxeles clasificados correctamente por el modelo, que corresponden con la región de interés (lesión), también conocidos como sensibilidad o recuperación (recall).

Falsos Positivos (FP): aquellos píxeles que no deben ser clasificados por el modelo, pero el modelo clasifica como parte de la región de interés (lesión).

Verdaderos Negativos (TN): aquellos píxeles no clasificados por el modelo, ya que, efectivamente no corresponden como parte de la región de interés (lesión), también conocidos como especificidad.

Falsos Negativos (FN): aquellos píxeles que deben ser clasificados por el modelo, pero el modelo no ha clasificado como parte de la región de interés (lesión).

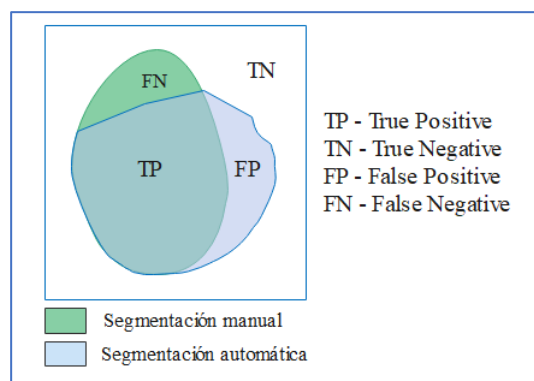


Figura 3. 6 Tipos de predicciones en segmentación

A continuación, se define las métricas para evaluar el modelo de segmentación en imágenes:

Intersección sobre Unión (IoU): También conocido como índice de jaccard, mide el grado de similitud entre dos conjuntos de muestras, A y B. Se considera A como el valor predicho y B como el valor real o conocido como ground-truth. Esta métrica también se conoce como Índice de Jaccard; medida de solapamiento entre los píxeles de dos imágenes.

En la figura 3.7, se representa un ejemplo de IoU. El cual, representa B el cuadro verde de la segmentación manual y A el cuadro rojo de la segmentación de la imagen predicha por el modelo [50], [51]. Se toman valores entre 0 y 1, considerándose 1 como igualdad entre ambos conjuntos,

valores mayores a 0.5 se considera buen resultado y valores menores a 0.5 como malos resultados, la fórmula:

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{TP}{TP + FN + FP}$$

Coeficiente de Sorensen-Dice: Es una medida que se utiliza comúnmente en el campo del procesamiento de imágenes. En la segmentación de imágenes es la más utilizada para evaluar los modelos de segmentación de inteligencia artificial. La métrica evalúa la similitud entre dos conjuntos, donde A, cuadro rojo representa la segmentación manual o ground-truth y B, cuadro verde es la segmentación predicha por el modelo, la fórmula:

$$DSC = \frac{2|A \cap B|}{|A| + |B|} = \frac{2 * TP}{2 * TP + FN + FP}$$

Se toman valores que varía entre 0 (sin similitud no hay coincidencia con el ground-truth) y 1 (similitud total coincide totalmente con el ground-truth) que es poco realista que suceda en segmentación de imágenes médicas. Entonces, para evaluar si una detección es mala, buena y excelente se utiliza un valor mínimo de Dice que es conocido como umbral, el valor es de 0.6, es decir que si es mayor a ese valor se considera una segmentación con buenos resultados, pero si es menor a ese valor se considera un mal resultado de segmentación para el modelo entrenado. En la figura 3.7 se representa.

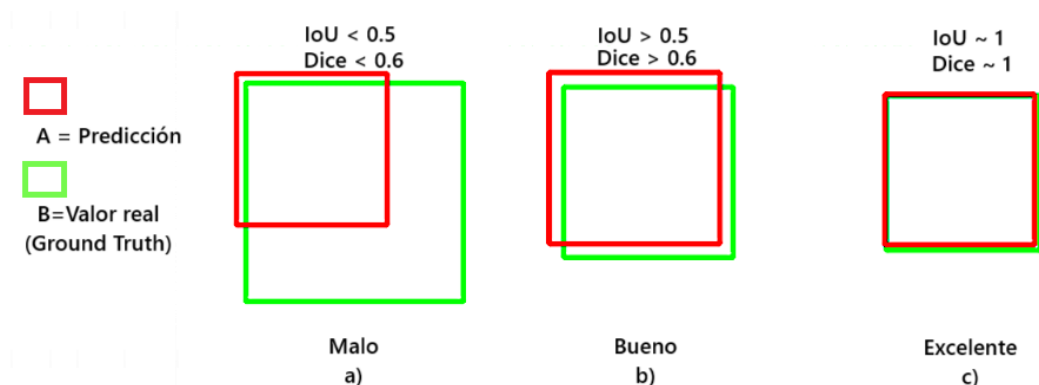


Figura 3. 7 Métricas de evaluación para segmentación: IoU y Dice

CAPÍTULO 4

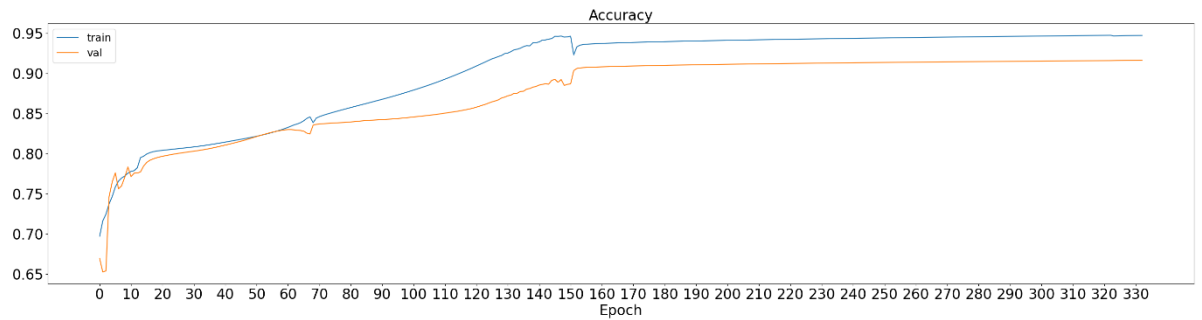
4. RESULTADOS

Se presentan los resultados del entrenamiento de la arquitectura U-Net para la segmentación de lesiones en el cérvix a partir de imágenes de colposcopia. El conjunto de datos consta de 800 ROIs del cérvix y 800 máscaras binarias. Este conjunto se dividió en dos partes: el 80% (640 imágenes de ROIs con sus máscaras binarias) para el entrenamiento, mientras que el 20% restante (160 imágenes de ROIs con sus máscaras binarias) se utilizó para validación.

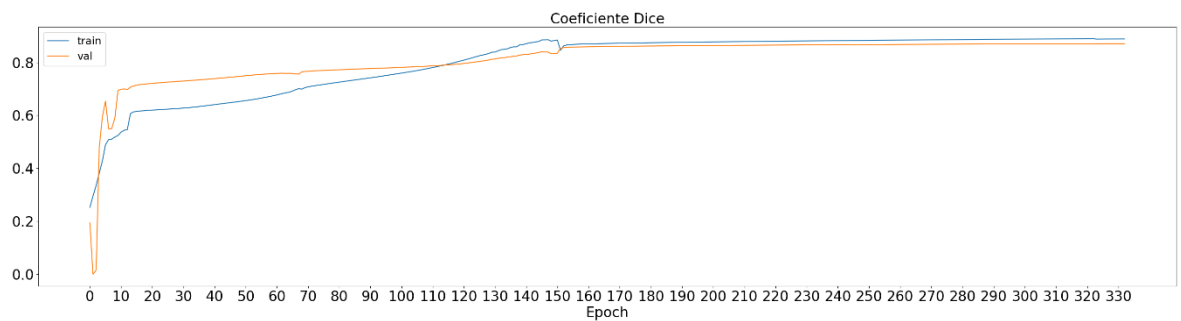
En la Figura 4.1, se presentan las gráficas de los resultados obtenidos durante el entrenamiento y validación de la arquitectura U-Net. En a) se representa la exactitud (*accuracy*), mostrando la evolución del entrenamiento del modelo U-Net durante 323 épocas. La curva naranja representa el *accuracy* del conjunto de validación, que comienza con 66% en las primeras épocas con un leve decrecimiento en las épocas 2 y 3 que posterior empieza a aumentar progresivamente, alcanzando 91% en la época 323, y mediante la técnica *early stopping* detiene automáticamente cuando el modelo dejó de mejorar. Además, se observa que hay un buen ajuste del modelo sin indicios de sobreajuste.

En b) representa la evolución del coeficiente de dice en 323 épocas. La curva naranja representa la validación comenzando con 1% durante la primera época, un pequeño decrecimiento en la época 2, 3 pero entre la época 4,5, 6 asciende a 65% y decrece en las épocas 7,8 y 9 en un 55%. Posterior empieza a progresar hasta la época 323 manteniéndose alrededor del 87%, lo que indica que el modelo generaliza bien los datos.

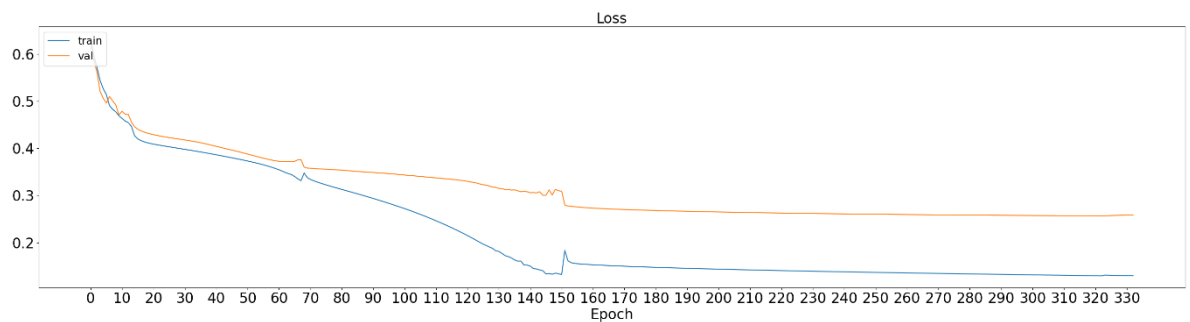
En c) representa la pérdida (*loss*) para el conjunto de datos de validación la curva naranja muestra un comportamiento decreciente en cuanto a la pérdida, ya que a medida que aumentan las épocas, la pérdida de validación también disminuye, alcanzando un 25% para el modelo U-Net. A partir de esa época tiende a aumentar la pérdida y mediante la técnica *early stopping* detiene automáticamente el entrenamiento cuando el modelo dejó de mejorar.



a)



b)



c)

Figura 4. 1 Gráficas: a) Exactitud (*Accuracy*), b) Coeficiente Dice, c) Función de pérdida (*Loss*)

4.1 Resultados de la segmentación

4.1.1 Coeficiente de Sorensen-Dice

En la Figura 4.2 se presenta cualitativamente los resultados de la segmentación del sistema propuesto, en donde a) muestra las imágenes de colposcopia original, b) y c) imágenes Ground-truth (*i.e.*, segmentación manual delimitada en color azul), d) y e) imágenes de la segmentación predicha por el modelo U-Net delimitadas por el color verde, donde pueden existir lesiones o anomalías en el cérvix. Para la métrica Dice se obtuvo un

84% del conjunto de datos de validación que es una de las más utilizadas para evaluar este tipo de problemas en segmentación.

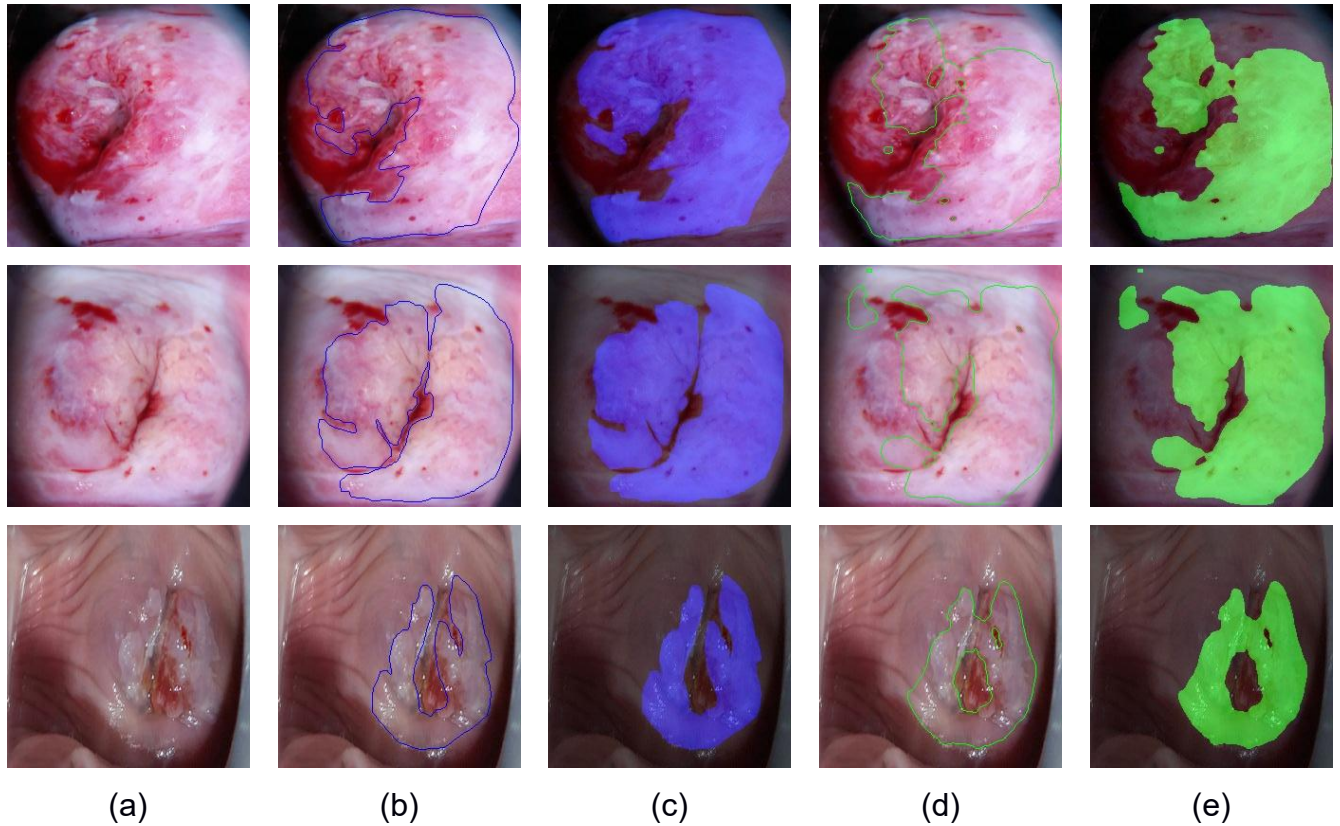


Figura 4. 2 Segmentación: (a) Imágenes originales, (b - c) Imágenes Ground-truth, (d - e) Imágenes U-Net

4.1.2 Intersección Sobre Unión o Índice de Jaccard

En la Figura 4.3 se presenta cualitativamente los resultados de la segmentación del sistema propuesto, en donde a) imágenes de colposcopia original, b) y c) imágenes Ground-truth (*i.e.*, segmentación manual delimitada en color azul), d) y e) imágenes de la segmentación predicha por el modelo U-Net delimitadas por el color verde, donde pueden existir lesiones o anomalías en el cérvix. En cambio, para IoU se obtuvo un 75% del conjunto de datos de validación. Esto se debe a la fórmula de la métrica que al no contar con el factor multiplicador de 2 como en el caso de Dice, penaliza de manera más estricta la forma de las regiones segmentadas.

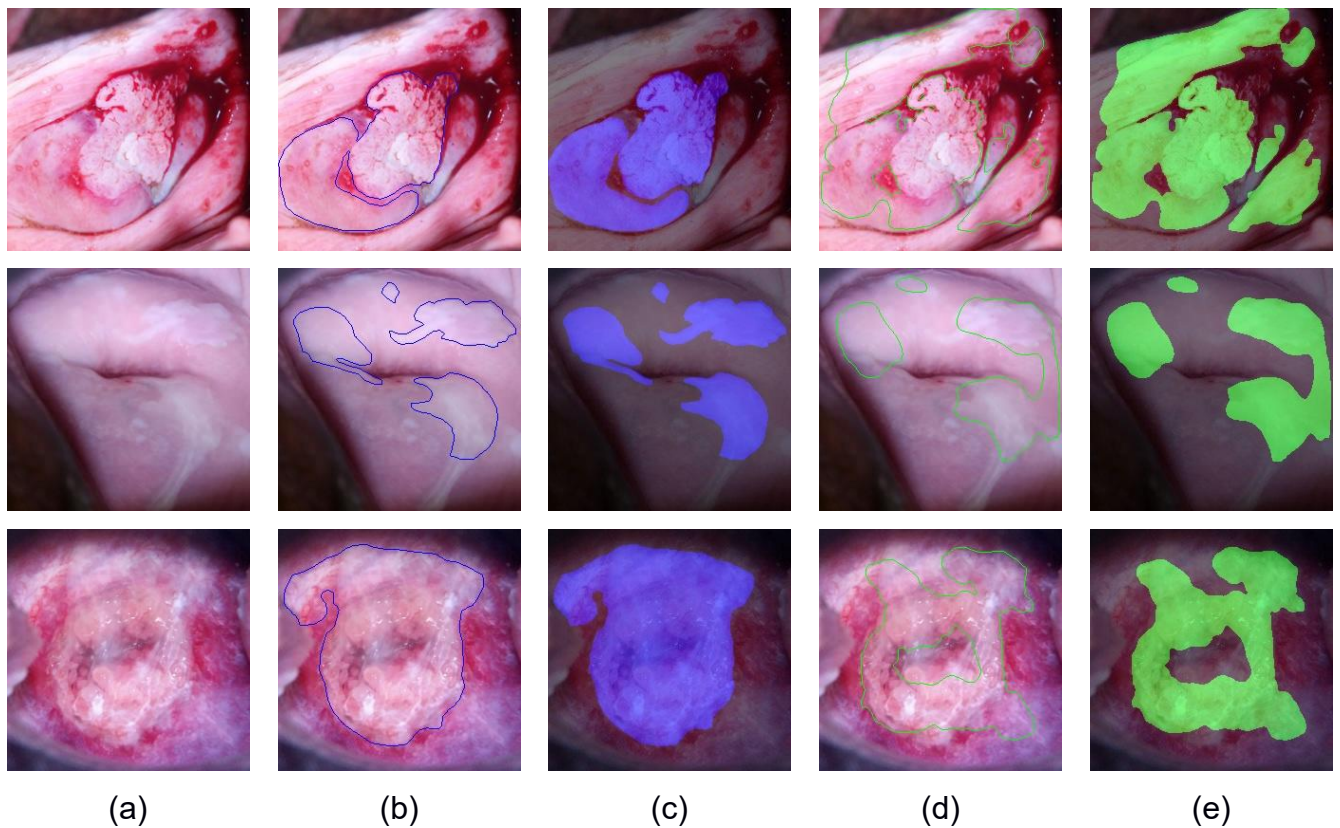


Figura 4. 3 Segmentación: (a) Imágenes originales, (b - c) Imágenes Ground truth, (d - e) Imágenes U-Net

4.2 Análisis Cuantitativo

En la Tabla 4.1 se presentan los resultados promedio obtenidos en el conjunto de validación (160 imágenes de colposcopia). Para el coeficiente de Dice alcanzó un valor del 84%, lo que indica una buena similitud entre las áreas segmentadas por U-Net. Sin embargo, el valor del IoU fue de 75%, lo que señala una menor superposición entre las áreas predichas y las reales del ground-truth.

Los resultados obtenidos para la arquitectura U-Net, se observa que el valor promedio del Dice es superior al del IoU. Esto sugiere que ambas métricas evalúan la calidad de la segmentación de manera diferente. El coeficiente de Dice tiende a ser más alto debido a su fórmula, que tiene un mayor peso a la intersección entre la predicción y el ground-truth, lo que hace menos sensible a errores pequeños en los bordes de la segmentación. Es decir, aunque el modelo tiene una buena capacidad para identificar áreas de superposición, los pequeños errores en los bordes de las regiones predichas afectan menos el valor de Dice. Por otro lado, la

intersección sobre la unión mide la proporción entre las áreas predicha y el ground-truth, y al no contar con el factor multiplicador de 2 como en el caso de Dice, penaliza de manera más estricta la forma de las regiones segmentadas.

Tabla 4. 1 Valor promedio de evaluación

Arquitectura	Dice	IoU
U-Net	84%	75%

4.3 Discusión de los resultados

En la gráfica de accuracy se ilustra que el modelo U-Net presenta un rendimiento de 91% de exactitud en el conjunto de validación. Por otro lado, en la gráfica del coeficiente Dice, el modelo obtiene un 87% en validación. A partir de estos resultados, aún se puede mejorar el aprendizaje de entrenamiento del modelo segmentador.

La principal limitación del sistema de segmentación de lesiones en imágenes de colposcopia utilizando ANN es la dependencia de la calidad y la diversidad del conjunto de datos de entrenamiento. Aunque U-Net ha demostrado ser eficaz en la segmentación en imágenes médicas, el rendimiento puede verse afectado por la calidad de las imágenes como la borrosidad, la iluminación e intensidad de color blanco en las imágenes del cérvix. Estas variaciones limitan la capacidad del modelo para aprender de una manera eficiente.

Para mejorar la segmentación, se podría implementar técnicas adicionales de preprocesamiento de imágenes, como la normalización de la iluminación y la eliminación de ruido mediante filtros. Esto contribuiría a mejorar la calidad de las imágenes de entrada y, por ende, un mejor aprendizaje del modelo para la segmentación. Además, explorar enfoques de segmentación multitarea, en los que se combinen la predicción de la segmentación con la de otros parámetros relevantes, como la forma de las lesiones, para ofrecer un análisis más robusto. Finalmente, considerar incorporar modelos de redes neuronales más avanzados o híbridos para segmentar imágenes médicas.

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

La selección de una base de datos internacional disponible en la plataforma Kaggle fue adecuada para garantizar la calidad del estudio en la segmentación de imágenes de colposcopia. Se logró conformar una base de datos con 200 imágenes de ROIs del cérvix. Esta base de datos pública resultó ideal para asegurar un proceso de segmentación ético, alineado con los estándares legales y con la debida anonimización de los datos. Este enfoque no solo protegió la privacidad de los pacientes, sino que también permitió obtener resultados válidos, ya que el modelo de segmentación alcanzó una exactitud del 91 % en la etapa de validación, sin sobreajuste, lo cual demuestra que los datos utilizados fueron representativos.

La colaboración con el médico especialista en cáncer de cuello uterino fue esencial para alcanzar el objetivo de construir un conjunto de datos confiable, mediante la creación de máscaras binarias que sirvieron como *ground-truth* para entrenar y validar el modelo de segmentación. Esta participación especializada aseguró que las lesiones fueran identificadas y segmentadas según las indicaciones del experto. Como resultado, el modelo U-Net entrenado con las imágenes de ROIs del cérvix y las máscaras binarias alcanzó un coeficiente de Dice del 84% en validación. Este valor indica que el modelo aprendió a identificar las lesiones de cérvix, lo que demuestra un buen desempeño en la predicción de lesiones en imágenes de colposcopia.

El entrenamiento de la red neuronal artificial U-Net logró segmentar lesiones de cérvix en imágenes de colposcopia. A lo largo de las 323 épocas, el modelo mostró un aprendizaje progresivo, alcanzando una exactitud del 91% en el conjunto de validación, lo que refleja un desempeño sólido en la tarea de segmentación.

La implementación de métricas de evaluación como Dice e IoU permitió una valoración cuantitativa del rendimiento del modelo en la tarea de segmentación de lesiones de cuello uterino. Los resultados obtenidos indican un buen desempeño

del modelo con un Dice del 84% y un IoU del 75% en el conjunto de datos de validación, lo que indica una buena similitud entre las lesiones de cérvix en el *ground-truth* y la predicción de la región segmentada por U-Net. De forma cualitativa, las imágenes de colposcopia muestran una segmentación adecuada de las lesiones de cuello uterino.

Recomendaciones

Para asegurar un mejor desempeño del modelo, es necesario contar con imágenes de alta calidad en las que se puedan observar tanto las lesiones como las áreas sin lesión. Esto permitirá obtener mejores resultados durante el entrenamiento del modelo. Aunque el conjunto de datos utilizado consta de 800 imágenes, incluyendo máscaras binarias y técnicas de aumento de datos, un número mayor de imágenes que representen diferentes etapas y tipos de lesiones en el cérvix contribuiría a mejorar la robustez del modelo y a hacerlo más generalizable.

Si bien la arquitectura U-Net ha mostrado buenos resultados, el uso de variantes de redes neuronales más avanzadas como U-Net++, modelos basados en Transformers, entre otros, podrían mejorar aún más el desempeño en la segmentación. Estas redes incorporan bloques diseñados para enfocarse mejor en áreas específicas de interés, lo que contribuiría a una mayor exactitud en la segmentación de lesiones.

Es necesario probar el modelo con imágenes de colposcopia locales para validar su efectividad en la práctica médica. Esto incluye el uso del modelo en imágenes de pacientes que no hayan sido parte del conjunto de datos de validación, para analizar cómo responde el modelo a variabilidad clínica de los datos.

Finalmente, es importante implementar una interfaz intuitiva para los médicos que permita visualizar la segmentación y facilite la interpretación de los resultados. Esta interfaz debería incluir una herramienta de visualización que resalte las áreas sospechosas y sirva como apoyo para que los profesionales puedan tomar decisiones informadas.

BIBLIOGRAFÍA

- [1] Observatorio Global del Cáncer, “Cancer Today.” [Online]. Available: https://gco.iarc.fr/today/en/dataviz/pie?mode=cancer&group_populations=1&sexes=2&types=1. [Accessed: 02-Dec-2024].
- [2] K. S. Aponte Luzuriaga, “Colposcopia como método diagnóstico de patología asociada a infección por el virus del papiloma humano.” Universidad Católica de Cuenca., Cuenca, 2023.
- [3] M. F. C. León *et al.*, “Cáncer de Cérvix y virus del papiloma humano. Diagnóstico y prevención.” *Mediciencias UTA*, vol. 7, no. 2, pp. 2–8, Apr. 2023.
- [4] “Campus Ginecologia Perú,” *Campus Virtual Ginecología*, 2020. [Online]. Available: https://campusvirtualginecologia.org.pe/campus/pluginfile.php/4464/course/overviewfiles/20231125_110318.jpg. [Accessed: 02-Dec-2024].
- [5] L. Hu *et al.*, “An Observational Study of Deep Learning and Automated Evaluation of Cervical Images for Cancer Screening,” *J. Natl. Cancer Inst.*, vol. 111, no. 9, pp. 923–932, Sep. 2019.
- [6] D. Shen, G. Wu, and H. Il Suk, “Deep Learning in Medical Image Analysis,” *Annu. Rev. Biomed. Eng.*, vol. 19, pp. 221–248, Jun. 2017.
- [7] V. Chandran *et al.*, “Diagnosis of Cervical Cancer based on Ensemble Deep Learning Network using Colposcopy Images,” *Biomed Res. Int.*, vol. 2021, no. 1, p. 5584004, Jan. 2021.
- [8] OMS, “Cáncer de cuello uterino,” *Organización Mundial de la Salud*, 2023. [Online]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/cervical-cancer>. [Accessed: 15-Dec-2024].
- [9] S. G. Ramirez Pedro, “Cáncer de cuello uterino,” *MANUAL MSD*, 2023. [Online]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/cervical-cancer>. [Accessed: 10-Dec-2024].
- [10] R. ángels Torné Aureli, Gaba Lydia, Sánchez María, “Cáncer de Cuello de Útero,” *Clínic Barcelona*, 2020. [Online]. Available:

<https://www.clinicbarcelona.org/asistencia/enfermedades/cancer-de-cuello-de-utero>. [Accessed: 10-Dec-2024].

- [11] I. N. del Cáncer, "Definición de cuello uterino," *Cancer.gov*, 2022. [Online]. Available: <https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/def/cuello-uterino>. [Accessed: 10-Dec-2024].
- [12] Bolaños Lucero, "CONOCIMIENTOS Y PRÁCTICAS PREVENTIVAS SOBRE EL CÁNCER DE CUELLO UTERINO EN USUARIAS DE UN CENTRO MÉDICO PARTICULAR DE SAN ISIDRO," UNIVERSIDAD NACIONAL FEDERICO VILLARREAL, LIMA, 2024.
- [13] A. C. Society, "¿Qué es cáncer de cuello uterino (cervical)?," 2020. [Online]. Available: <https://www.cancer.org/es/cancer/tipos/cancer-de-cuello-uterino/acerca/que-es-cancer-de-cuello-uterino.html>. [Accessed: 10-Dec-2024].
- [14] Tenjo Camilo, "Identificación de neoplasia cervical intraepitelial para la predicción de cáncer cervical mediante el uso de aprendizaje profundo," UNIVERSIDAD DEL ROSARIO, BOGOTÁ, 2021.
- [15] S. Sellors J., "Chapter 1: An introduction to the anatomy of the uterine cervix," *Cancer Screening at IARC*, 2003. [Online]. Available: <https://screening.iarc.fr/colpochap.php?lang=1&chap=1.php>. [Accessed: 10-Dec-2024].
- [16] Cuenca Adela, "CÁNCER DE CUELLO UTERINO Y VIRUS DEL PAPILOMA HUMANO: UNA REVISIÓN NARRATIVA," UNIVERSIDAD DE JAÉN, JAÉN, ESPAÑA, 2023.
- [17] V. Lucía, "¿QUÉ ES LA COLPOSCOPIA?," *luciaviorginecologa.org*. [Online]. Available: <https://www.luciaviorginecologa.org/colposcopia-asturias/>. [Accessed: 10-Dec-2024].
- [18] S. MEDICINE, "Cervical Biopsy - Stanford Medicine," *Children's Health*. [Online]. Available: <https://www.stanfordchildrens.org/es/topic/default?id=cervical-biopsy-92-P09281>. [Accessed: 10-Dec-2024].
- [19] I. N. del Cáncer, "Ecografía transvaginal," *Cancer.gov*. [Online]. Available: <https://www.cancer.gov/espanol/buscar/resultados?swKeyword=Ultrasonido+tran>

svaginal. [Accessed: 10-Dec-2024].

- [20] I. N. del Cáncer, "Exploraciones con tomografía computarizada (TC) para el cáncer," *Cancer.gov*. [Online]. Available: <https://www.cancer.gov/espanol/cancer/diagnostico-estadificacion/hoja-informativa-tomografia-computarizada>. [Accessed: 10-Dec-2024].
- [21] K. R. Zand, C. Reinhold, H. Abe, S. Maheshwari, A. Mohamed, and D. Upegui, "Magnetic resonance imaging of the cervix," *Cancer Imaging*, vol. 7, no. 1, p. 69, 2007.
- [22] L. Mansilla and E. Ferrante, "Segmentación multi-atlas de imágenes médicas con selección de atlas inteligente y control de calidad automático," 2018.
- [23] M. R. SÚAREZ, A. E. D; PÉREZ, D. M. & OROZCO, "Diseño de Sistema Automatizado para Detección de Anomalías en Imágenes Digitales de Mama." [Online]. Available: https://www.researchgate.net/publication/338409783_Disenio_de_Sistema_Automatizado_para_Deteccion_de_Anomalias_en_Imagenes_Digitales_de_Mama_Automated_System_Design_to_Detect_Anomalies_in_Digital_Images_of_the_Breast. [Accessed: 24-Jun-2020].
- [24] B. Pérez and M. Aránzazu, "Desarrollo de algoritmos de procesamiento de imagen avanzado para interpretación de imágenes médicas.," Dec. 2016.
- [25] J. F. Vélez, A. B. Moreno Díaz, and Á. Sánchez Calle, *Visión por Computador*. Dykinson y Serv, 2016.
- [26] A. Quintana, "Segmentación de imágenes médicas," 01-Jan-2019. [Online]. Available: https://www.researchgate.net/publication/330183469_Segmentacion_de_imagenes. [Accessed: 19-Feb-2021].
- [27] I. S. Prieto, "Procesado de imágenes médicas en MATLAB," Universidad Politécnica de Madrid, Madrid, 2014.
- [28] P. Pal, S. Bhati, and M. Singhal, "Segmentation Techniques for Skin Lesion: A Comparative Study," Jaipur, 2018.

- [29] S. Boumaraf, X. Liu, C. Ferkous, and X. Ma, "A New Computer-Aided Diagnosis System with Modified Genetic Feature Selection for BI-RADS Classification of Breast Masses in Mammograms," *Biomed Res. Int.*, vol. 2020, pp. 7695207–7695207, May 2020.
- [30] K. R. LastNameLastNameFedorov A., Beichel R., Kalpathy-Cramer J., Finet J., Fillion-Robin J-C., Pujol S., Bauer C., Jennings D., Fennessy F.M., Sonka M., Buatti J., Aylward S.R., Miller J.V., Pieper S., "3D Slicer as an Image Computing Platform for the Quantitative Imaging Network," *Slicer*, 2012. [Online]. Available: <https://www.slicer.org/>. [Accessed: 20-Feb-2025].
- [31] L. K. Paul A. Yushkevich, Jilei Hao ,Sadhana Ravikumar, Alison M. Pouch, Ankush Aggarwal, "ITK-SNAP," 2023. [Online]. Available: <https://www.itksnap.org/pmwiki/pmwiki.php>. [Accessed: 20-Feb-2025].
- [32] D. Sánchez, "Análisis del software ImageJ para el análisis científico de imágenes," E.T.S.I y Sistemas de Telecomunicación (UPM), Madrid, 2014.
- [33] J. Schindelin *et al.*, "Fiji: An open-source platform for biological-image analysis," *Nat. Methods*, vol. 9, no. 7, pp. 676–682, Jul. 2012.
- [34] B. Bai, P. Z. Liu, Y. Z. Du, and Y. M. Luo, "Automatic segmentation of cervical region in colposcopic images using K-means," *Australas. Phys. Eng. Sci. Med.*, vol. 41, no. 4, pp. 1077–1085, Dec. 2018.
- [35] B. Yao and C. Li, "Application Research on Identification of Cervical Lesions Based on Random Forest," *2022 IEEE 2nd Int. Conf. Power, Electron. Comput. Appl. ICPECA 2022*, pp. 1158–1160, 2022.
- [36] Y. Artan and X. Huang, "Combining multiple 2v-SVM classifiers for tissue segmentation," *2008 5th IEEE Int. Symp. Biomed. Imaging From Nano to Macro, Proceedings, ISBI*, pp. 488–491, 2008.
- [37] G. Saranya and C. Sujatha, "A Novel Framework Leveraging Adam Optimization Techniques, Coupled with Mask R-CNN Deep Learning Mechanisms for Cervical Cancer Detection System," *Proc. 3rd Int. Conf. Appl. Artif. Intell. Comput. ICAAIC 2024*, pp. 233–245, 2024.
- [38] Y. C. Lin *et al.*, "Generalizable transfer learning of automated tumor segmentation

from cervical cancers toward a universal model for uterine malignancies in diffusion-weighted MRI,” *Insights Imaging*, vol. 14, no. 1, pp. 1–11, Dec. 2023.

- [39] J. Huang, G. Yang, B. Li, Y. He, and Y. Liang, “Segmentation of Cervical Cell Images Based on Generative Adversarial Networks,” *IEEE Access*, vol. 9, pp. 115415–115428, 2021.
- [40] W. K. y Y. B.-O. BenO, jlJones, Kumar H, Meg Risdal, MRao, Vadim Sherman, Vipul, “ntel & MobileODT Cervical Cancer Screening,” *Kaggle*, 2017. [Online]. Available: <https://www.kaggle.com/competitions/intel-mobileodt-cervical-cancer-screening/rules>. [Accessed: 02-Feb-2024].
- [41] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” *arXiv*, pp. 1–8, 2015.
- [42] T. Esteves, “Segmentación automática de procesos neuronales en microscopía electrónica mediante técnicas de aprendizaje profundo,” Universidad de País Vasco, País Vasco, 2019.
- [43] O. Olalla, “Redes Neuronales artificiales: Qué son y cómo se entrenan ,” *Xeridia*, 16-Sep-2019. [Online]. Available: <https://www.xeridia.com/blog/redes-neuronales-artificiales-que-son-y-como-se-entrenan-parte-i>. [Accessed: 22-Feb-2021].
- [44] ANACONDA, “Anaconda Navigator ,” 2020. [Online]. Available: <https://docs.anaconda.com/anaconda/navigator/>. [Accessed: 01-Mar-2021].
- [45] L. Gonzalez, “INTRODUCCIÓN AL IDE SPYDER ,” 14-Sep-2018. [Online]. Available: <https://www.youtube.com/watch?v=eAaGRRuyg1w>. [Accessed: 01-Mar-2021].
- [46] python, “Bienvenido a Python.org,” *python.org*. [Online]. Available: <https://www.python.org/>. [Accessed: 10-Dec-2024].
- [47] G. Gasca, “Tu primer red neuronal usando Keras ,” *Medium*, 05-Jun-2018. [Online]. Available: https://medium.com/@gogasca_/tu-primer-red-neuronal-usando-keras-72d36130ee6c. [Accessed: 01-Mar-2021].
- [48] Tensorflow, “TensorFlow,” *tensorflow.org*. [Online]. Available: <https://www.tensorflow.org/?hl=es-419>. [Accessed: 10-Dec-2024].

- [49] A. Rosebrock, "Intersection over Union (IoU) for object detection - PyImageSearch," 7 *noviembre*, 2016. [Online]. Available: <https://www.pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>. [Accessed: 22-Jul-2020].
- [50] D. Cournapeau, "API Reference — scikit-learn 0.24.1 documentation," 2020. [Online]. Available: https://scikit-learn.org/stable/modules/classes.html?highlight=sklearn_metrics#module-sklearn.metrics. [Accessed: 04-Feb-2021].
- [51] N. Parsad, "Deep Learning in Medical Imaging V," 2018. [Online]. Available: <https://medium.com/datadriveninvestor/deep-learning-in-medical-imaging-3c1008431aaf>. [Accessed: 04-Feb-2021].