

# SEGMENTACIÓN DEL CUELLO UTERINO EN IMÁGENES DE COLPOSCOPIA MEDIANTE TÉCNICAS DE APRENDIZAJE DE MAQUINA

1° Ana María Bolaños Semanate  
10°semestre Ingeniería electrónica  
Pontificia Universidad Javeriana de Cali  
Cali, Colombia  
anamariabs@javerianacali.edu.co

2° Santiago Hutado Bustos  
10°semestre Ingeniería electrónica  
Pontificia Universidad Javeriana de Cali  
Cali, Colombia  
santiagohurtado11@javerianacali.edu.co

**Abstract**—The human papillomavirus (HPV), a sexually transmitted disease that can cause cervical cancer, is one of the most common neoplasms in women globally and nationally. Although there are health campaigns promoting screening tests to detect the disease, the waiting times for results are high due to deficiencies in laboratory infrastructure, affecting diagnosis.

To address this issue, it is proposed to leverage colposcopy images and apply machine learning techniques to analyze the state of the cervix. A desktop application with unsupervised learning models was developed to improve diagnostic times. This project (CITOBOT), funded by Minciencias and developed by a multidisciplinary team, aims to provide a solution to this significant public health problem.[1][2][3]

**Index Terms**—segmentation, machine learning, cervix, colposcopy imaging, HPV

## I. INTRODUCCIÓN

El cáncer de cuello uterino es una de las principales causas de muerte en mujeres a nivel mundial, convirtiéndolo en uno de los problemas de salud pública con mayor incidencia en Latinoamérica. Muestra de ello es que en Colombia para el año 2020 cerca de 4569 mujeres padecían de esta enfermedad y 2309 fallecieron por su causa. Así lo indica la agencia internacional para la investigación de cáncer de la Organización Mundial de la Salud (OMS). Además, las bases de datos de esta intintución muestran que este tipo de cáncer es el segundo con mayor índice de mortalidad y frecuencia a nivel nacional.[18]

El virus del papiloma humano es una enfermedad de transmisión sexual, pero existen factores que pueden incrementar la probabilidad de contraer esta afección, tales como, sistemas inmunitarios debilitados, contacto con personas o superficies que hayan sido infectadas, uso de duchas comunes o piscinas. La Sociedad Americana de Cáncer, estipula que existen cerca de 150 variaciones relacionadas con el VPH, algunos de ellos pueden generar verrugas que son tumores no cancerosos, sin embargo, hay casos en los que estas lesiones generarán cáncer, algunos de estos son: tipo 6, 11, 16 y 18. [17]

El VPH se clasifica en dos grandes grupos, VPH cutáneo y mucoso, a su vez, el VPH mucoso tiene dos clasificaciones, de alto riesgo y bajo riesgo. Las de alto riesgo [19]

causan síntomas que pueden desaparecer en algunos años, e incluso pueden desencadenar algún tipo de cáncer, debido a alteraciones en las estructuras de las células.[20]

Se ha identificado que el proceso de diagnóstico de cáncer de cuello uterino puede presentar falsos positivos, así como también, lapsos de espera superiores a los cinco días en la obtención de los resultados. Los diagnósticos equivocados se generarán por los siguientes factores: Recolección inadecuada de las muestras, generando que la cantidad de células recolectadas sea insuficiente para el análisis [21]. El segundo motivo se relaciona con los tiempos de respuesta[22][23] en los laboratorios para evaluar las muestras debido a que en algunos hospitales hay carencia de equipos especializados para esta labor.

Trabajos previos han tratado esta problemática desde diferentes perspectivas, tal como, "segmentación de la región acetoblancas en colposcopías del cérvix / Aceto - white", el cuál propone estrategias para procesar imágenes colposcópicas, automatizando la detección y segmentación del área cervical y la región acetoblancas. Se emplean algoritmos de agrupamiento y morfología matemática en el espacio de colores Lab, como k-means, Minibatch k-means, modelos de mezclas Gaussianas y modelos de mezclas Gaussianas por Inferencia Bayesiana. El preprocesamiento aborda peculiaridades de las imágenes, como regiones brillantes y ruido. Se destaca la importancia de detectar la Región de Interés (ROI), definida como la región cervical, antes de aplicar algoritmos de agrupación para la segmentación de la región acetoblancas,region segmentation in cervix colposcopies.[24]

Otro ejemplo es, utilizar técnicas de Clustering para mejorar la reconstrucción de modelos anatómicos 3D a partir de imágenes médicas DICOM. Se emplean dos métodos de clustering, k-means y Fuzzy K-means. Tras analizar algoritmos y realizar experimentos, se desarrolló un módulo de segmentación que utiliza la diferencia de color como indicador de disimilitud y la heurística de análisis de frecuencia de color para la inicialización de centroides. Se identificaron vías para mejorar las técnicas de segmentación, como la búsqueda de procedimientos avanzados de inicialización de centroides y el

uso de algoritmos que buscan el óptimo global. El módulo presenta un diseño extensible y fue probado en sistemas operativos Ubuntu 9.10 y Windows XP, demostrando alta fiabilidad y robustez.[25]

También se propone el uso de redes neuronales artificiales (RNA) para diagnosticar lesiones precancerosas. Las RNA han demostrado altos porcentajes de aciertos en aplicaciones médicas, como el reconocimiento de células con un 87.4% de precisión. Se desarrolló una aplicación basada en RNA para diagnosticar lesiones precancerosas en cuellos uterinos. Se utilizaron imágenes digitales y diagnósticos del centro piloto Corposalud Aragua. Destacan dos módulos: adquisición de imágenes limitadas a cuellos uterinos sin exéresis y segmentación para eliminar información irrelevante, definida por un protocolo de captación.[26]

Con base a lo anterior, durante este proyecto se emplearán algoritmos de aprendizaje supervisado tal como Segnet, Unet y FPN. Y algoritmos no supervisados, al igual que, Watershed, RGB, Otsu, K-means, Rango Umbral y mezclas Gaussianas.

Este trabajo destaca por implementar la segmentación automática del cuello uterino en imágenes de colposcopia mediante técnicas de aprendizaje de máquina. Su aporte fundamental radica en mejorar la identificación de etapas de cáncer cervical de manera eficiente. La evaluación del rendimiento se realiza con métricas estándar, como el coeficiente Dice, el Índice de Jaccard (IoU) y la exactitud (Accuracy), proporcionando resultados cuantitativos fiables. En conjunto, este enfoque innovador tiene un impacto directo en el diagnóstico médico, ofreciendo avances significativos para la sociedad. Presentando resultados favorables para los algoritmos supervisados.

## II. MATERIALES Y MÉTODOS

Este trabajo es de carácter exploratorio, puesto que se desea demostrar que la segmentación de imágenes de colposcopia mejora los índices de identificación de cáncer de cuello uterino. Para cumplir con ello, es necesario realizar una búsqueda de información para identificar la manera en como otras personas que han desarrollado proyectos relacionados, han actuado. Posterior a ello, se buscan métodos de aprendizaje de máquina para entenderlos y para su posterior implementación. Después, se organiza la base de datos que contienen las imágenes de colposcopia y se construyen máscaras binarias para construir un marco de referencia de la segmentación. Seguidamente, se implementaron 7 métodos de aprendizaje no supervisado, como son: Otsu, Rango Umbral, RGB, Grabcut, K-means, Wathersed y Mezclas Gaussianas, y 3 modelos de aprendizaje supervisado como son Unet, FPN y SegNet. Una vez implementados, se evalúan los resultados en relación al gold standar para determinar el rendimiento, para ello, se hizo uso del coeficiente Dice, IoU y Acurracy, siendo este último específico para los modelos supervisados. Finalmente son seleccionados para el aplicativo final.

### A. Base de datos

Las imágenes con las cuales se entrenarán los modelos pertenecen a una base de datos de la Organización Mundial de la Salud y el Instituto de Cáncer de Estados Unidos. Es importante recalcar que antes de usar la base de datos, se realizará una preselección y segmentación manual de los datos que se usarán para que la información sea uniforme.

### B. Algoritmos no supervisados

1) *K-means*: El algoritmo k-means es un método de agrupamiento que trabaja con conjuntos de datos que son descritos por valores numéricos y es uno de los algoritmos más populares en minería de datos. Su objetivo es encontrar una partición de ese conjunto, la cual consiste en una serie de grupos que son representados por un centro. Para determinar el número de grupos a crear por el algoritmo, el usuario tiene que definir previamente el valor del parámetro "k", indicando cuantos grupos se formarán en la partición resultante del conjunto de datos. La idea del algoritmo es ir afinando la posición de los centros representativos en el espacio de objetos, es decir, encontrar a los grupos que integren a los individuos más entre ellos. Para llevar a cabo el proceso del algoritmo k-means se define a priori el número de grupos a formar. Los pasos para desarrollar este algoritmo se explican en la figura 1.[4]

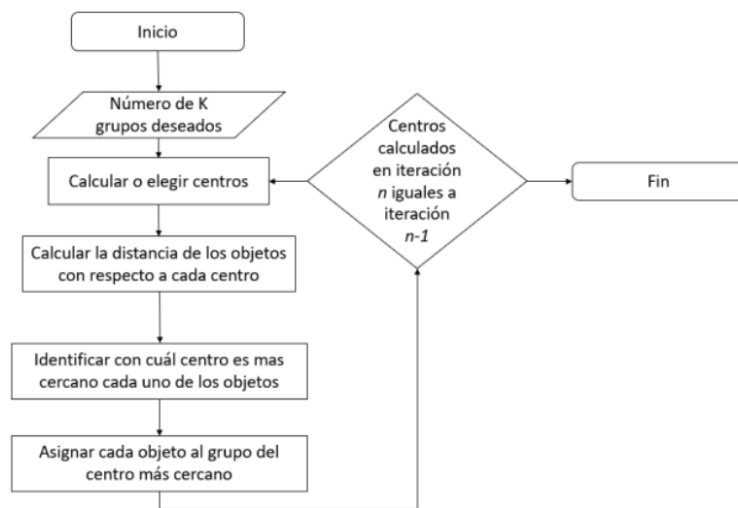


Fig. 1. Algoritmo k-means

2) *Otsu*: Este es un modelo estadístico que busca separar los objetos de una imagen basados en la intensidad de los píxeles que componen la imagen. El modelo de umbralización extrae objetos de su fondo basado en la selección de umbral óptimo global que maximiza la varianza entre clases de una imagen en escala de grises, de manera que el píxel cuyo nivel de gris sea menor que el umbral se asignará al fondo, en caso contrario, al segmento de objeto.[5]

3) *Rango Umbral*: El modelo de segmentación por rango umbral es un método estadístico que, al igual que el modelo

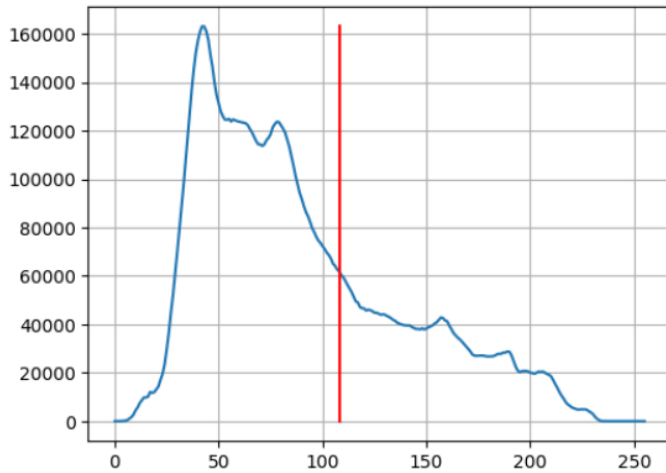


Fig. 2. Histograma Otsu

Otsu, busca diferenciar el cuello uterino del fondo de una imagen. Para ello, se calcula el tono de gris que tiene mayor frecuencia en la imagen, este tono de gris se utiliza como punto de partida para definir un rango inferior y superior. El rango inferior se obtiene restando cincuenta al tono de gris de mayor frecuencia, y el rango superior se obtiene sumando cincuenta.[6]

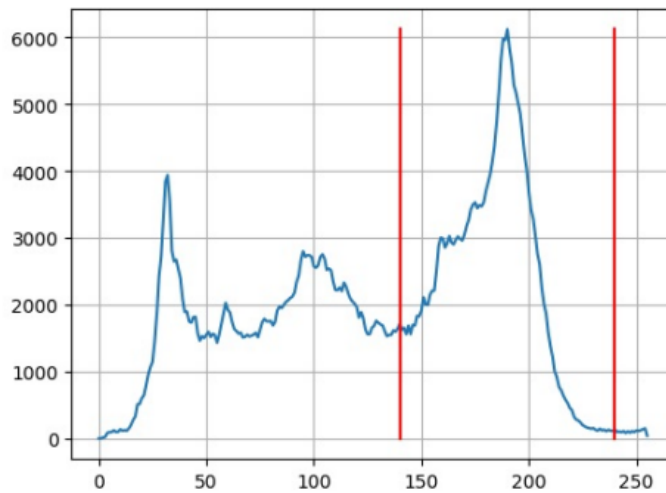


Fig. 3. Histograma rango umbral

4) *RGB*: El modelo RGB segmenta una imagen de color asignando un segmento a cada píxel. Para ello, el modelo selecciona una región central de la imagen y separa los canales RGB de la sección. Luego, calcula el valor mínimo y máximo para cada canal, estos valores definen un rango. Finalmente, el modelo verifica que píxeles tienen un tono que está en el rango especificado, aquellos que si lo están se les asigna un valor de uno y aquellos que no un cero, de esta manera se construye la máscara binaria del método. Máscara que después

se multiplica por la imagen original para finalmente obtener la imagen segmentada.[7]

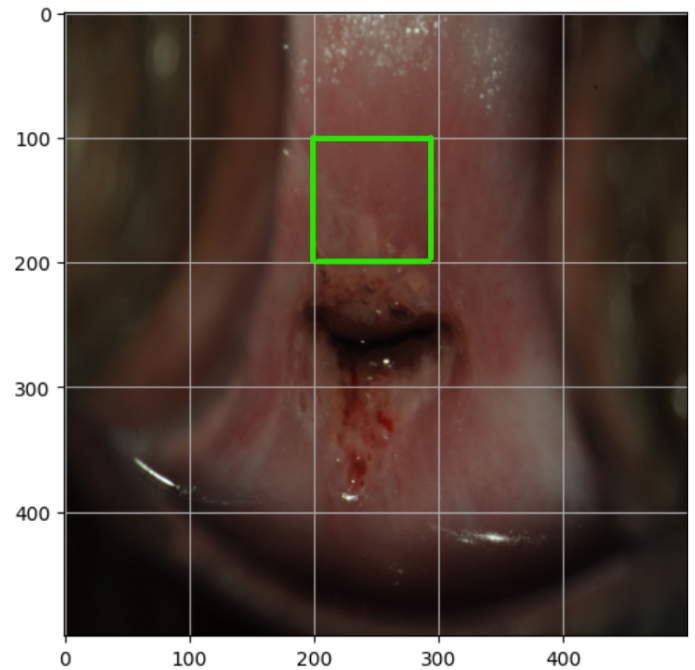


Fig. 4. Sección rgb para segmentar imágenes de colposcopia

5) *Watershed*: El algoritmo de watershed es una técnica de segmentación basada en morfología matemática, que permite extraer las fronteras de las regiones que hay en una imagen. A la vez, se considera una técnica de segmentación basada en regiones, debido a que clasifica los píxeles debido a su proximidad espacial, el gradiente de sus niveles de gris y la homogeneidad de sus texturas. Por ello se toma como un método de detección de contornos y crecimiento de regiones al mismo tiempo. El algoritmo puede considerar una imagen en escala de grises como la imagen topográfica de un relieve terrestre, en donde a cada píxel se le asocia como valor de "altura" su nivel de gris correspondiente. En este sentido se puede pensar que las intensidades de gris de mayor amplitud corresponden con montañas, mientras que las intensidades de menor valor son valles o ríos. La técnica además incorpora un proceso de inundación de los valles desde los niveles más bajos de altura (mínimos locales) hasta los más altos. Las zonas de baja intensidad de gris también se conocen como "basins" por donde fluiría el agua e inundaría toda la topografía de la imagen. Es decir, el agua circulará por todas las "basins" identificadas. Este proceso continúa hasta que las aguas de cuencas contiguas se unan, formando líneas[8]

6) *Mezclas Gaussianas*: El modelo de mezclas gaussianas (GMM) es un modelo probabilístico que retorna centroides y límites de grupos que describen distribuciones gaussianas. Para la implementación del modelo se empleó la librería `sklearn.mixture` que hace uso del algoritmo `expectation-maximization (EM)`, el cual consiste en un proceso iterativo para estimar los parámetros de un modelo de mezclas gaussianas

para ajustar un modelo mixto a un conjunto de datos. Si bien las mascarar binarias permiten diferenciar el fondo del elemento, la manera en la que el modelo asigna la región del fondo y del elemento es aleatoria, por este motivo, en algunos casos la predicción del modelo podría ser buena pero la mascara binaria estaba al revés, además este fenómeno no tiene un patrón establecido para tanto no fue posible hacer una corrección [9][10]

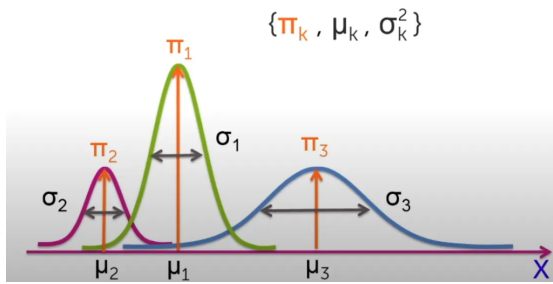


Fig. 5. Distribución Gaussiana en 1D con tres clusters

7) *Grabcut*: Grabcut es una técnica de segmentación de imágenes donde se requiere poca intervención del usuario. Inicialmente el usuario debe seleccionar un recuadro alrededor de la zona de interés y luego la segmentación se realiza de manera automática. Posteriormente el usuario puede seleccionar manualmente ciertas áreas de la imagen para mejorar el resultado obtenido[11]

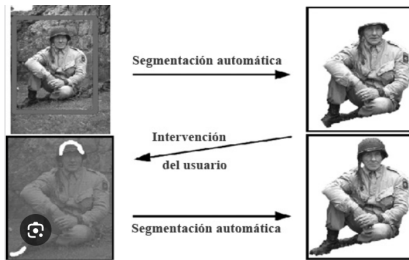


Fig. 6. Pasos segmentación Grabcut

### C. Algoritmos supervisados

1) *Unet*: La arquitectura U-net es un modelo de red neuronal usado en tareas de visión artificial, más específicamente en problemas de segmentación semántica de imágenes biomédicas, de manera que se le asigna una etiqueta una clase a cada píxel, permitiendo un análisis detallado de la composición de la imagen. Las redes neuronales se componen de codificador y un decodificador, para el caso de U-net, el codificador son una serie de capas convolucionales o Down Convolutions, que reducen el tamaño de las imágenes, pero se incrementan el número de canales, de manera que se extraen el mayor número de características, identificando la información más relevante, en cuanto al decodificador se hacen una serie de capas de deconvolución o Up Convolutions, siendo esta una estructura simétrica al codificador, el decodificador ayuda a

producir una salida de las mismas dimensiones de la entrada, pero en la medida que se ejecutan las capas de Up Convolution se fusiona la imagen con las características extraídas en el codificador. Algo que distingue a este método es que se hace un proceso de Copy and Drop, en el cual se concatenan las salidas de cada capa de la Down Convolution con cada capa del decodificador, de esta manera se puede conservar detalles finos.[12][13][14]

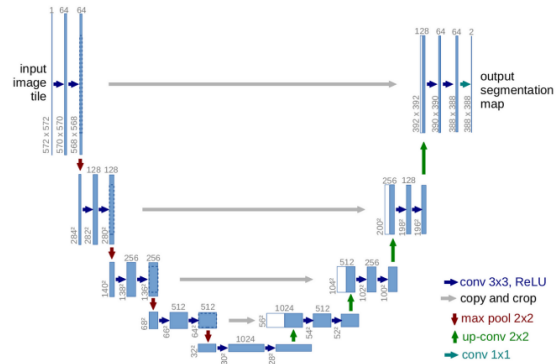


Fig. 7. Arquitectura Unet

2) *SegNet*: Segnet es una arquitectura para realizar la segmentación de imágenes, esta conformada por una red de codificación y una red de decodificación, seguida por una capa final de clasificación de píxeles. La red de codificación consta de 13 capas convolucionales que corresponden a las 13 primeras capas convolucionales en la red VGG16, diseñada para la clasificación de objetos. Por lo tanto, podemos inicializar el proceso de entrenamiento a partir de pesos previamente entrenados para la clasificación de grandes volúmenes de datos. También se puede descartar las capas totalmente conectadas en favor de retener mapas de características de alta resolución en las salidas más profundas de la red de codificación. Esto también reduce el número de parámetros en la red de codificación de Segnet. Cada capa de codificación tiene una capa de decodificación correspondiente, por lo tanto, la red de decodificación consta de 13 capas. La salida final del decodificador se alimenta a un clasificador softmax multiclase para producir probabilidades de clase para cada píxel de manera independiente. [15]

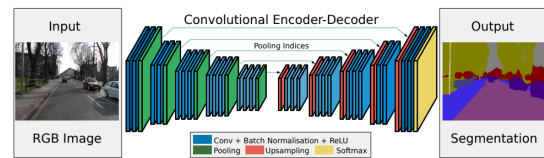


Fig. 8. Arquitectura Segnet

3) *FPN*: El modelo FPN o Feature Pyramid Network es una arquitectura empleada en aplicaciones de visión por computadora, como la detección de objetos o segmentación de imágenes, esto es posible mediante la representación de características multidimensionales para cualquier tamaño de

imagen. Esta extracción de características se hace mediante una red neuronal convolucional en forma de pirámide a partir de una imagen. La estructura FPN está basada en una red pre-entrenada y en la medida que los datos de la imagen pasa por la red se construye una pirámide de características mediante fusiones ascendentes, de esta manera se aumenta la resolución de las características mediante la interpolación y la suma de mapas característicos de resoluciones inferiores, y descendentes, donde se concatenan las mapas de resoluciones superiores con mapas de resolución inferior, de esta manera es posible mejorar la precisión y robustez del modelo de segmentación semántica.[16]

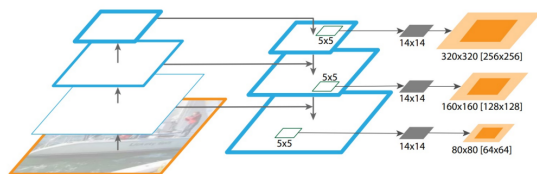


Fig. 9. Arquitectura FPN

### III. RESULTADOS

Se implementaron seis modelos no supervisados para la segmentación de imágenes de cuello uterino: Otsu, Rango umbral, K-means, Mezclas Gaussianas, Grabcut, RGB y Watershed. Para evaluar la precisión de los modelos, se construyó un dataset de pruebas de 1050 imágenes, estas imágenes fueron extraídas de la base de datos de la OMS, NHS y kaggle, este dataset incluye imágenes de diferentes calidades, lo que permite evaluar el rendimiento de los modelos en una variedad de condiciones. Los modelos se evaluaron segmentando las imágenes del dataset de pruebas, generando máscaras binarias para cada imagen y estas se compararon con las máscaras binarias generadas manualmente, para determinar la precisión se utilizaron los parámetros IoU y Dice.

#### A. Modelos no supervisados

1) *Modelo Otsu*: Se segmentaron 1050 imágenes con el modelo Otsu, el parámetro IoU se calculó para cada imagen resultante. Los resultados se presentan a continuación en un histograma. Los resultados del parámetro Dice para el modelo Otsu tienen una distribución con un pico central alto. Dado el valor central, es posible afirmar que los resultados varían mucho. Adicional a ello el rango con mayor frecuencia es (0,463 - 0,540], lo que indica que la segmentación que el modelo tiene una baja precisión en la predicción para la aplicación de segmentación de imágenes de colposcopia. Por otra parte, los resultados del parámetro IoU aunque tienen una distribución de datos diferente la interpretación de los datos es la misma, el modelo en la mayoría de los casos no puede diferenciar por completo el cuello uterino del fondo.

2) *Rango umbral*: Los resultados obtenidos para el modelo rango umbral de la prueba de segmentación para calcular los parámetros Dice e IoU se presentan a continuación. El desempeño del método rango umbral es deficiente, en los

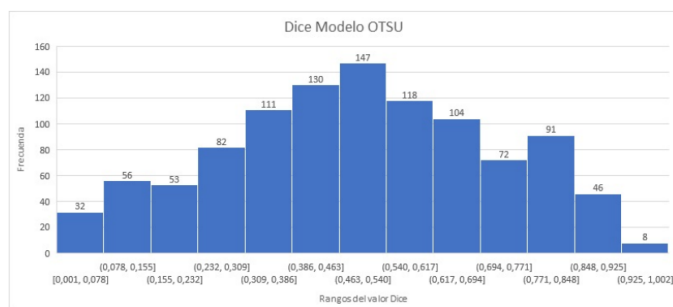


Fig. 10. Resultado parámetro Dice para modelo Otsu

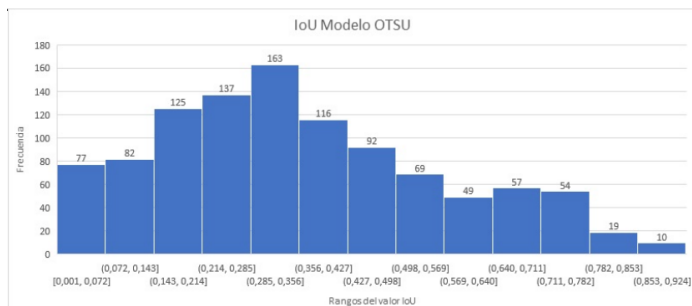


Fig. 11. Resultado parámetro IoU para modelo Otsu

histogramas se observa que en los dos casos de prueba más del 50% de las imágenes evaluadas tienen una precisión entre 0 y 0,1, indicando así que el modelo no es recomendable para la aplicación dado el desempeño.

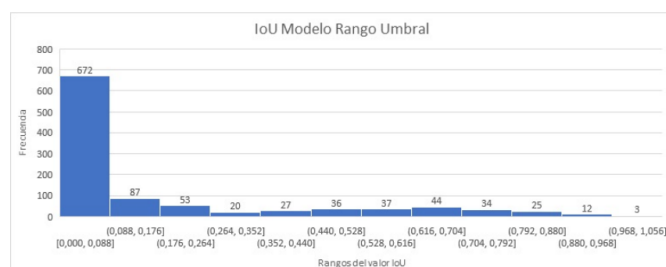


Fig. 12. Resultado parámetro IoU para rango umbral

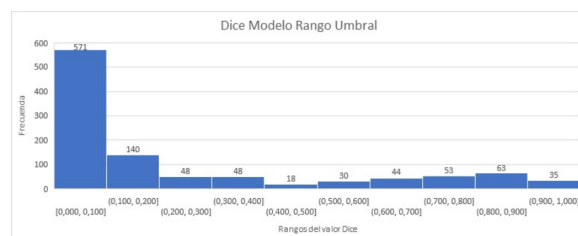


Fig. 13. Resultado parámetro Dice para rango umbral

3) *K-means*: A continuación, se presentan los resultados de los parámetros IoU y Dice después de efectuar la prueba de segmentación. Los resultados consignados en el histograma

muestra que el 81.51% de las imágenes segmentadas tienen una precisión inferior a 0.5, además los valores de precisión con mayor frecuencia están en el rango de 0.115 a 0.215.

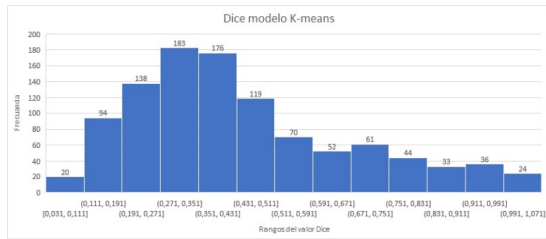


Fig. 14. Resultado parámetro Dice para K-means

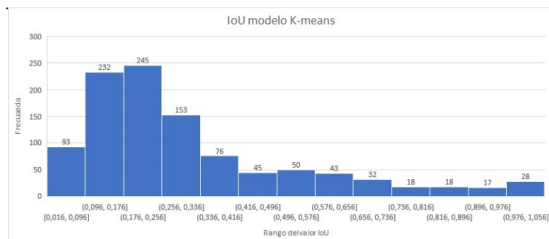


Fig. 15. Resultado parámetro IoU para K-means

4) *Mezclas Gaussianas*: Los resultados obtenidos de la prueba para este modelo son los siguientes, donde cabe resaltar que el 72% de las imágenes presentan un parámetro Dice inferior o igual a 0.5. Mientras que el 91% de las imágenes presentan un parámetro IoU inferior o igual a 0.5

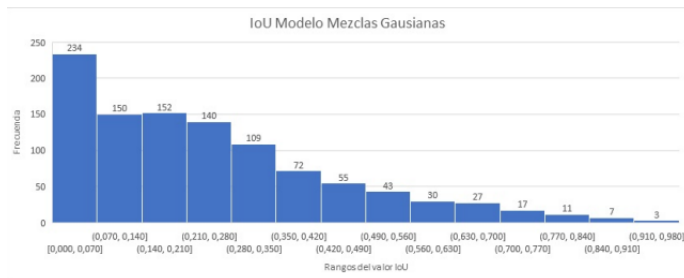


Fig. 16. Resultado parámetro IoU para mezclas gaussianas

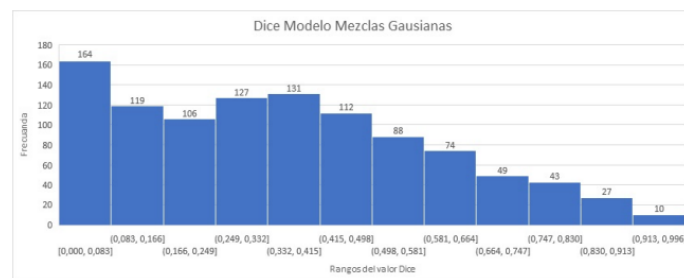


Fig. 17. Resultado parámetro Dice para mezclas gaussianas

5) *Grabcut*: Como se observa en los histogramas, la distribución de los datos es diferente, para el parámetro Dice se observa una tendencia de sesgo negativo, pues la mayoría de los datos se encuentra en el extremo derecho del histograma, lo que indica que los valores más bajos son menos frecuentes. Por el contrario, para el parámetro IoU se observa que los datos tienen mayor tendencia a los rangos centrales.

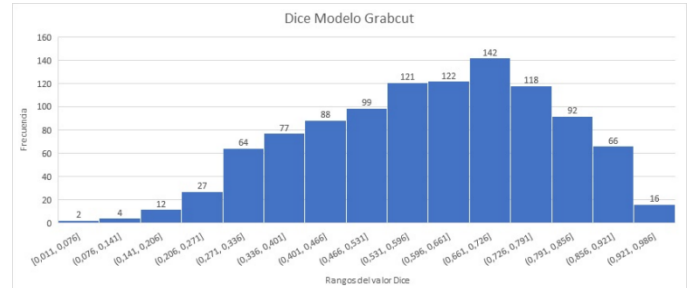


Fig. 18. Resultado parámetro Dice para Grabcut

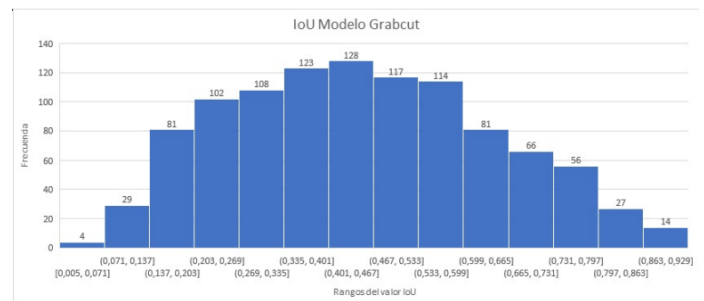


Fig. 19. Resultado parámetro IoU para Grabcut

6) *RGB*: Después de segmentar 1050 imágenes, los resultados de los parámetros Dice e IoU indican que el modelo de mezclas gaussianas es relativamente preciso, aunque algunas imágenes presentan dificultades en la segmentación. Esta discrepancia puede deberse a la complejidad intrínseca o al bajo contraste de ciertas imágenes. Es esencial reconocer que la efectividad del modelo no es uniforme, destacando la necesidad de abordar específicamente las imágenes más desafiantes. Este análisis subraya la importancia de evaluar y ajustar continuamente los modelos para optimizar la segmentación en imágenes médicas.

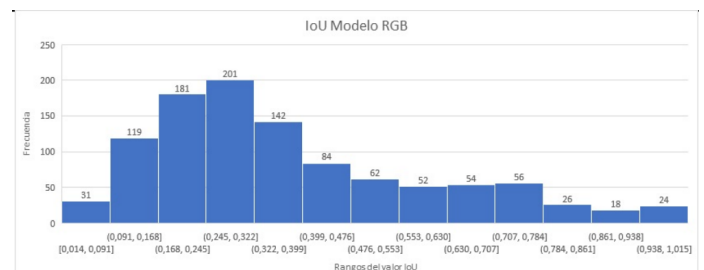


Fig. 20. Resultado parámetro IoU para RGB

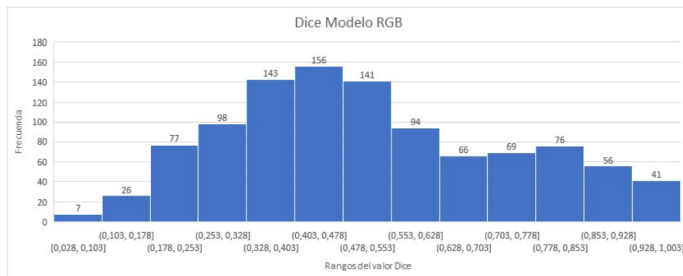


Fig. 21. Resultado parámetro Dice para RGB

7) *Watershed*: Los resultados de los parámetros de precisión del modelo se presentan a continuación. Cabe resaltar las imágenes que obtuvieron valores entre 0.85 y 0.9 para el parámetro Dice.

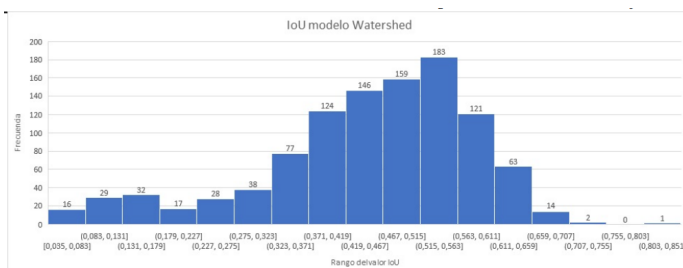


Fig. 22. Resultado parámetro IoU para watershed

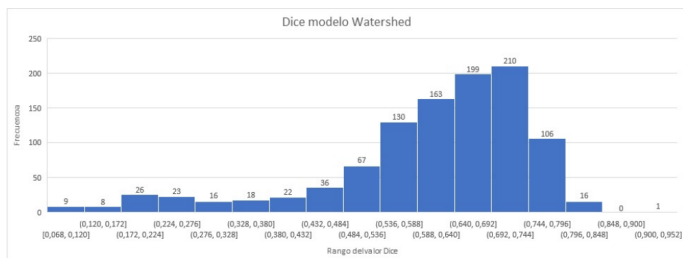


Fig. 23. Resultado parámetro Dice para watershed

## B. Modelos supervisados

1) *Segnet*: El modelo genera las siguientes métricas, accuracy, loss. Obteniendo valores superiores a 0,7 e inferiores a 0.3 respectivamente al finalizar el entrenamiento.

2) *Unet*: Después de implementar y entrenar el modelo con 741 imágenes, con cinco épocas, los resultados se muestran a continuación por medio de las imágenes segmentadas o datos de salida del modelo mostrados en la figura 24.

Resultados del entrenamiento inicial de Unet:

- *Accuracy*: 0.8364.
- *Dice*: 0.7467.
- *IoU*: 0.5958.

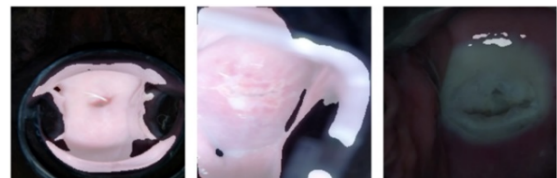


Fig. 24. Entrenamiento inicial Unet

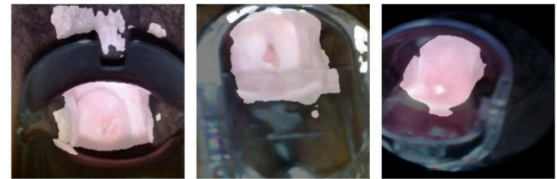


Fig. 25. Reentrenamiento 1 modelo Unet

Resultados del reentrenamiento del modelo Unet:

- *Accuracy*: 0.8517.
- *Dice*: 0.7465.
- *IoU*: 0.5955.

A partir de las métricas obtenidas se puede evidenciar un leve incremento en el parámetro accuracy, sin embargo, el desempeño del modelo puede ser mejor, por esta razón se realizó un reentrenamiento por segunda vez, con imágenes 1050 y 10 épocas.

3) *FPN*: El entrenamiento inicial del modelo supervisado Feature Pyramid Network se realizó con 741 imágenes y 5 épocas, con estas condiciones se obtuvieron resultados en los que visualmente se puede identificar el cuello uterino, sin embargo, con el objetivo de determinar la calidad de la segmentación se añadieron los parámetros accuracy, IoU y Dice, para establecer la precisión del modelo al final del entrenamiento.

- *Accuracy*: 0.9588.
- *Dice*: 0.9354.
- *IoU*: 0.8787.

Con estos resultados es posible afirmar que este modelo es capaz de segmentar cualquier imagen de colposcopia con una precisión aproximada del 90%.



Fig. 26. Ejemplos de resultados para arquitectura FPN

## IV. DISCUSIÓN

### A. Modelos no supervisados

Para establecer que modelo es mejor, se realizaron comparaciones entre la imagen original y la máscara binaria generadas por los modelos, es posible evidenciar que aquellas imágenes que alcanzan un IoU superior o igual 0.50, Dice superior o igual 0.66, ya que, a partir de estos valores se tienen un nivel de precisión suficiente para diferenciar el cuello uterino de los demás componentes de la imagen.

Modelo	Dice > 0.66	IoU > 0.50
Watershed	47,6%	41,4%
RGB	41,5%	40,0%
Otsu	25,5%	24,0%
K-means	19,9%	19,2%
Rango Umbral	16,0%	15,8%
Mezclas Gaussianas	12,8%	12,3%

Después de analizar los resultados obtenidos para cada modelo no supervisado con los criterios mencionados anteriormente, IoU mayor o igual a 0.5 y Dice mayor o igual a 0.66, se obtuvo que los modelos con mejor rendimiento son Watershed y Grabcut, pues la predicción de las máscaras binarias para estos modelos tienen resultados notoriamente mejores respecto a los otros.

### B. Modelos supervisados

Después de establecer los criterios de precisión aceptables para los modelos segmentación aplicados en imágenes de cuello uterino se presentan los resultados globales para definir cual se adapta en mayor medida a los parámetros de selección.

Modelo	Accuracy	IoU	Dice
SegNet	0.9605	-	-
FPN	0.9588	0.8787	0.9354
Unet	0.9303	0.7885	0.8817

El modelo que tiene un valor de desempeño más alto es SegNet, sin embargo, como se mostró anteriormente con valores mayores o iguales a 0.5 y 0.66, para IoU y Dice, se obtienen resultados favorables. En este caso todos los modelos tienen un desempeño cercano a 1, por lo tanto los tres pueden ser seleccionados.

### C. Aplicación

La aplicación se diseñó empleando la librería Tkinter debido a que es utilizada para realizar el diseño de interfaces gráficas de usuario o GUI. La biblioteca proporciona diferentes tipos de widgets, tal como botones, menús desplegables, campos de texto, entre otros. En el panel izquierdo está el recuadro "Imagen original" donde carga la imagen de cuello uterino sin ningún tipo de retoque, es decir, sin modificación de bordes, color, brillo, contraste, nitidez, etc. Para hacerlo, se debe hacer click sobre el botón "cargar imagen", posterior a esto la aplicación tendrá acceso a las carpetas donde se han guardado previamente las imágenes a segmentar. Seleccionamos la foto de interés para cargar los datos. Posteriormente se debe seleccionar el método a emplear, para esto debe elegir la opción

no supervisados en el panel central, luego escoger los algoritmos de segmentación disponibles, algunos son: K-means, Watershed, mezclas Gaussianas, etc. Una vez seleccionado el método, se debe dar click en el botón (Aceptar) para aplicar el método a la imagen cargada previamente. En la parte posterior del panel central. Finalmente la imagen segmentada carga en el recuadro derecho.

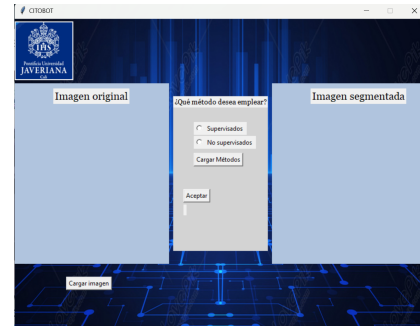


Fig. 27. Interfaz principal de la GUI

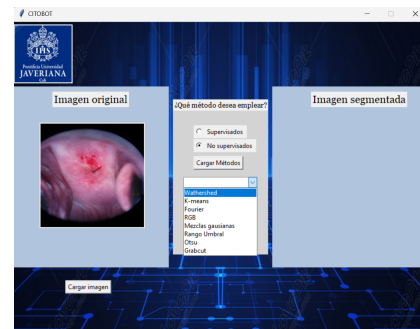


Fig. 28. Cargar la imagen a la API

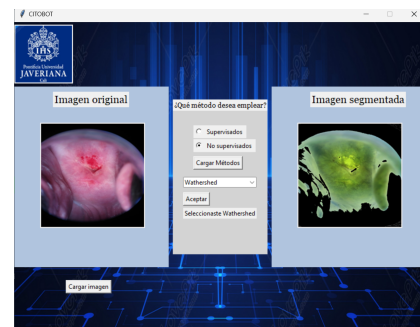


Fig. 29. Algoritmo watershed

## V. CONCLUSIONES

En este trabajo de grado, se abordó la implementación de la segmentación automática del cuello uterino en imágenes de colposcopia mediante técnicas de aprendizaje de máquina para la identificación de etapas de cáncer. Las actividades principales incluyeron la organización de una base de datos para identificar la sección del cuello uterino y la implementación

de algoritmos de aprendizaje de máquina, tanto supervisados (Segnet, FPN, Unet) como no supervisados (grabcut, otsu, k-means, mezclas gaussianas). Estos algoritmos generaron máscaras, representaciones gráficas que resaltan áreas específicas de las imágenes.

Tras la segmentación, se emplearon métricas como IoU, Dice, Accuracy y Loss para evaluar el rendimiento de los modelos. Los modelos no supervisados, especialmente grabcut y Watershed, destacaron en rendimiento, mientras que los supervisados mostraron un rendimiento superior a 0.9, haciéndolos adecuados para la aplicación en el proyecto CITOBOT.

Como parte del proyecto, se desarrolló una aplicación que permite el preprocesamiento de las imágenes de colposcopia con cada uno de los métodos implementados. Para trabajos futuros, se propone el entrenamiento de modelos supervisados con imágenes de los centros de salud colaboradores y la unificación de las etapas de preprocesamiento y clasificación para obtener una herramienta robusta que apoye el diagnóstico del cáncer de cuello uterino en sus diversas etapas.

## VI. REFERENCIAS

### REFERENCES

- [1] J. Guerrero, J. Rivas, P. Linares, and J. Rodríguez, Red neuronal artificial para detectar lesiones precancerosas en el cuello uterino, vol. 15, Apr. 2008. doi:<https://www.redalyc.org/pdf/707/70712315003.pdf>
- [2] “Prueba de Vph,” American Cancer Society, <https://www.cancer.org/es/cancer/tipos/cancer-de-cuello-uterino/deteccion-diagnostico-clasificacion-por-etapas/pruebas-de-deteccion/prueba-de-vph.html> (accessed Dec. 15, 2023).
- [3] “Diccionario de Cáncer del NCI,” Instituto Nacional del Cáncer, <https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-cancer/def/colposcopia> (accessed Dec. 15, 2023).
- [4] A. Franco-Arcega et al., Sistema de enseñanza para la técnica de agrupamiento k-means Teaching system for the k-means clustering technique, 2021. doi:<https://repository.uaeh.edu.mx/revistas/index.php/icbi/article/view/7384/8278>
- [5] Otsu method and K-means — IEEE conference publication - IEEE Xplore, <https://ieeexplore.ieee.org/document/5254345/> (accessed Dec. 15, 2023).
- [6] Region of adaptive threshold segmentation between mean ... - IEEE xplore, <https://ieeexplore.ieee.org/abstract/document/6914204/> (accessed Dec. 15, 2023).
- [7] Unsupervised RGB-D image segmentation by multi-layer clustering — IEEE ..., <https://ieeexplore.ieee.org/document/7760095/> (accessed Dec. 15, 2023).
- [8] N. L. S. Palomino, Watershed: un algoritmo eficiente y flexible para segmentación de imágenes de geles 2-DE.
- [9] “Segmentación de objetos basada en el modelo de mezcla gaussiana y campos aleatorios condicionales,” IEE, pp. 900–904, 2016.
- [10] R. Alberto, “Clustering con gaussian mixture model en sklearn y sus parámetros Aplicado al marketing.,” Medium, <https://rubialesalberto.medium.com/clustering-con-gaussian-mixture-model-en-sklearn-y-sus-par%C3%A1metros-aplicado-al-marketing-74b9d8454b86> (accessed Dec. 15, 2023).
- [11] Esmitt Ramírez J, David Martínez R, and Rhadamés Carmona S., SEGMENTACIÓN DE IMÁGENES A COLOR BASADA EN EL ALGORITMO DE GRABCUT, Apr. 2012.
- [12] G. T. Y. L. Juntao Jiang, Xiyu Chen, “Vig-unet: Vision graph neural networks for medical image segmentation,” IEEE, 2023.
- [13] V. C. S. S. Aswathy A. L., “Cascaded 3d unet architecture for segmenting the covid-19 infection from lung ct volume,” Scientific Reports, 2022.
- [14] J. Zhang. (2019) Unet — line by line explanation.
- [15] Vijay Badrinarayanan, Alex Kendall, Roberto Cipolla, and Senior Member, SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation.
- [16] T.-Y. Lin et al., “Feature Pyramid Networks for Object Detection,” arXiv.org, <https://arxiv.org/abs/1612.03144> (accessed Dec. 15, 2023).
- [17] “Infección por vph,” Mayo Clinic, <https://www.mayoclinic.org/es/diseases-conditions/hpv-infection/symptoms-causes/syc-20351596> (accessed Jan. 21, 2024).
- [18] “Cancer Today,” Global Cancer Observatory, (accessed Jan. 21, 2024).
- [19] “Prueba de Vph,” American Cancer Society, <https://www.cancer.org/es/cancer/tipos/cancer-de-cuello-uterino/deteccion-diagnostico-clasificacion-por-etapas/pruebas-de-deteccion/prueba-de-vph.html> (accessed Jan. 21, 2024).
- [20] S. de Salud, “Cáncer de Cuello uterino,” gob.mx, <https://www.gob.mx/salud/acciones-y-programas/cancer-de-cuello-uterino> (accessed Jan. 21, 2024).
- [21] MinisterioDeSalud. Ministerio de la protección social instituto nacional de cancerología. [https://minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/IA/INCA/Guia-tamizacion-cuello-uterino.pdf.\(2007\)](https://minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/IA/INCA/Guia-tamizacion-cuello-uterino.pdf.(2007))
- [22] L. Bailey. Comprender los resultados anormales en las pruebas de detección del cáncer cervical. <https://es.familydoctor.org/comprender-los-resultados-anormales-en-las-pruebas-de-deteccion-del-cancer-cervical/> (Abril 2022)
- [23] RedJurista. (18 Mayo 2016) ¿cuánto tiempo se puede demorar una eps en asignar una cita con un médico general y un especialista? [https://www.redjurista.com/N\\_ewsPaper/35/salud/1011/cuanto-tiempo-se-puede-demorar-una-eps-en-asignar-una-cita-con-un-medico-general-y-un-especialista](https://www.redjurista.com/N_ewsPaper/35/salud/1011/cuanto-tiempo-se-puede-demorar-una-eps-en-asignar-una-cita-con-un-medico-general-y-un-especialista)
- [24] D. V. A. S. Saul Olivera, Marta Bager, “Segmentacion de la región acetoblancas en colposcopías del cervix/ acetowhite region segmentation in cervix colposcopies,” Ciencias Matematicas, vol. 31, no. 1, pp. 171–181, 2017
- [25] O. P. Gustavo Lorca, José Arzola, “Segmentación de imágenes médicas digitales mediante técnicas de clustering,” Aporte Santiaguino, vol. 3, no. 1, pp. 108–116, 2010.