

E.T.S. de Ingeniería Industrial, Informática y de
Telecomunicación

Segmentación de pólipos mediante técnicas de Visión Artificial



Grado en Ingeniería Informática

Trabajo Fin de Grado

Ana Larráyoiz Jimenez

Director: Daniel Paternain Dallo

Codirector: Aranzazu Jurio Munarriz

Pamplona, 26 de enero de 2017

INDICE

1. INTRODUCCIÓN Y OBJETIVOS	3
1.1 Introducción	3
1.2 Objetivos	4
2. PRELIMINARES.....	8
2.1 Umbralización	8
2.2 Morfología matemática	8
2.2.1 Elemento estructurante	9
2.2.2 Dilatación	9
2.2.3 Erosión	10
2.2.4 Apertura.....	10
2.2.5 Cierre	11
3. PREPROCESAMIENTO	12
3.1 Detección de reflejos.....	12
3.1.1 Selección del umbral	13
3.1.2 Mejora de la detección (máscara).....	15
3.2 Eliminación de reflejos mediante inpainting.....	16
3.2.1 Algoritmo de inpainting:	16
3.2.2 Filtro de suavizado	17
3.2.3 Resultados de preprocesamiento (detección e inpainting)	19
4. DETECCIÓN DE VALLES.....	20
4.1 Imagen de valles.....	21
4.1.1 Segunda derivada anisotrópica de la gaussiana.....	21
4.1.2 Resultados de imagen de valles.....	26
4.2 Región de interés (ROI)	27
4.2.1 Ecuilización del histograma.....	29
4.2.2 Erosión de la imagen ecualizada	29
4.2.3 Relleno de formas curvilíneas cerradas.....	29
4.2.4 Apertura.....	30
5. UNION DE REGIONES Y SEGMENTACIÓN.....	31
5.1 Algoritmo de Watershed	31
5.2 Regiones no presentes en ROI	32
5.3 Segmentación final.....	33
6. RESULTADOS.....	34
7. CONCLUSIONES Y LINEAS FUTURAS	37
REFERENCIAS	38

1. INTRODUCCIÓN Y OBJETIVOS

1.1 Introducción

El cáncer colorrectal es un tipo de cáncer que aparece en el intestino grueso en las partes del colon o recto. La mayoría de los cánceres colorrectales son adenocarcinomas (cánceres que empiezan en las células que producen y liberan moco y otros líquidos). El cáncer colorrectal comienza a menudo como un crecimiento, un pólipo, que se puede formar en la pared interna del colon o el recto. Con el tiempo, algunos pólipos se vuelven cancerosos. Al encontrar y eliminar los pólipos se puede prevenir el cáncer colorrectal.

Las muertes por cáncer colorrectal han disminuido con el uso de las colonoscopias y prueba de sangre oculta en la materia fecal.

Según datos estadísticos del SEER [1] (The Surveillance, Epidemiology, and End Results), un programa del NCI (National Cancer Institute) en Estados Unidos, para 2016 se estimaron 134.490 nuevos casos de cáncer colorrectal de los que estimadamente 49.190 morirán de esta enfermedad. Según estos datos es el 4º cáncer más común, produciendo más muertes que el cáncer de mama.

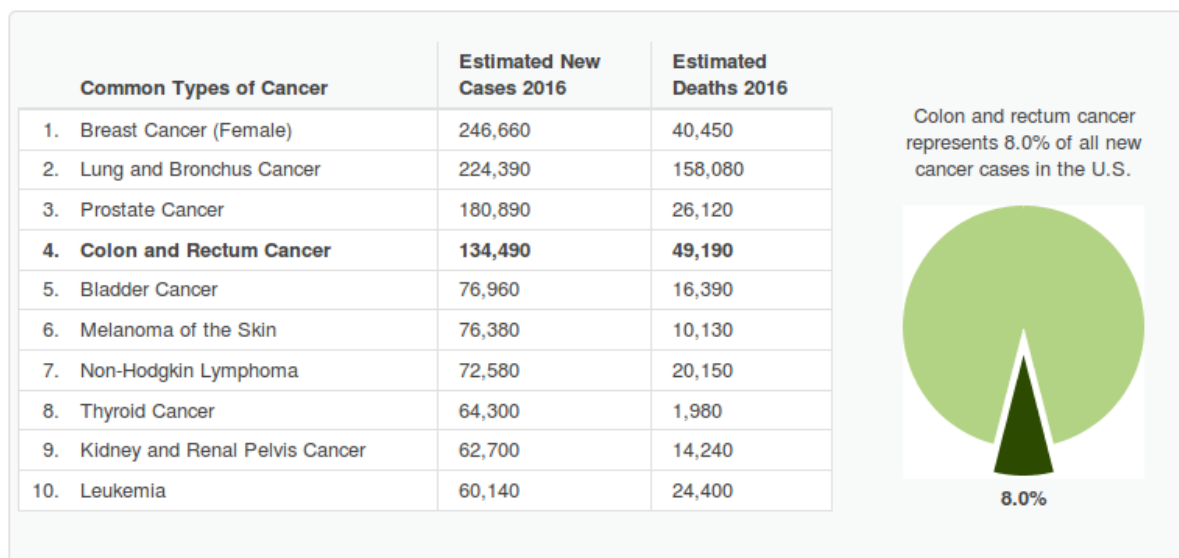


Figura 1: datos estadísticos del cáncer

Debido al elevado número de casos de esta enfermedad, es muy importante seguir avanzando en la investigación para mejorar su detección.

Los métodos más utilizados en exámenes de detección para cáncer colorrectal son:

- Pruebas de sangre oculta en materia fecal de alta sensibilidad (FOBT).
- La prueba de ADN en materia fecal (ADN-FIT)
- Sigmoidoscopia
- Colonoscopia virtual o colonografía por tomografía computarizada (TC)
- Colonoscopia regular (u óptica)

La colonoscopia regular (u óptica) consiste en exploración, por parte del médico, del recto y colon mediante el uso de un tubo iluminado y flexible, llamado colonoscopio, el cual permite observar el interior del recto y colon a través de una lente y extraer tejido en caso de ser necesario. El médico realiza una inspección visual y bajo su criterio determina la presencia o no de pólipos. En este trabajo nos vamos a centrar en este tipo de pruebas. Se quiere complementar, ayudar y mejorar la toma de decisión del médico ya que además de errores de tipo humano, el tamaño del pólipo (en caso de ser muy pequeño), puede dificultar mucho la detección del mismo por parte del médico.

1.2 Objetivos

Detectar cualquier tipo de cáncer en sus primeras fases aumenta significativamente el ratio de supervivencia del paciente. El cáncer colorrectal no es una excepción y la importancia de detectarlo durante las primeras fases mediante técnicas como la colonoscopia es muy elevada. Por esta razón, **el objetivo de este trabajo es diseñar un sistema de visión artificial que sea capaz de automatizar la detección y segmentación de pólipos en imágenes de colonoscopia.**

Para realizar este trabajo de fin de grado hemos trabajado con la base de datos pública CVC-ColonDB [2]. Esta base de datos contiene fotogramas de secuencias de videos de colonoscopias. Este trabajo se va a centrar en aquellos fotogramas de los que, a priori, sabemos que tienen pólipos. Por lo tanto, dado un fotograma concreto, el objetivo de nuestro sistema de visión es identificar y aislar el pólipo.

Evidentemente, la detección de pólipos no es una tarea sencilla ya que en el proceso nos podemos encontrar con problemas: exceso de iluminación en la obtención de la imagen, desfase de información en RGB, movimiento en las imágenes, zonas oscuras... entre otros. Debido a estas dificultades, el sistema de visión artificial que hemos desarrollado consta de varios pasos para poder detectar la presencia de un pólipo:

- **Preprocesamiento:** en esta primera fase necesitamos adaptar y mejorar la imagen a tratar para facilitar la posterior obtención de información. En concreto, se divide el preprocesamiento en dos etapas:

- En la primera etapa, mediante técnicas de umbralización, los reflejos de luz, ocasionados por la iluminación utilizada en el colonoscopio, son detectados.
- En la segunda etapa, vamos a eliminar dichos reflejos para que no afecten en los siguientes pasos (detección y segmentación). Por ello, utilizaremos técnicas de inpainting.

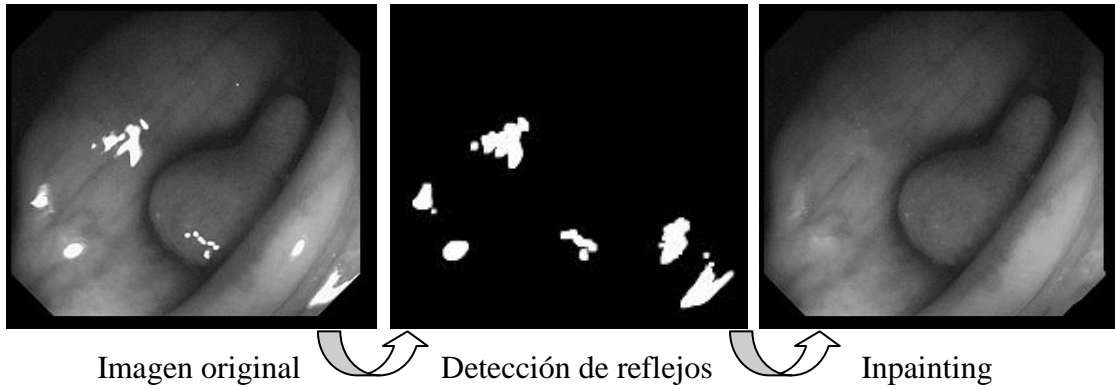
- **Detección de valles:** la luz utilizada en el colonoscopio provoca la aparición de una “sombra” que rodea al pólipo. En este trabajo vamos a utilizar un detector de valles (basado en la segunda derivada anisotrópica gaussiana) para tratar de detectar dicha sombra. Dado que la detección de valles provoca muchos falsos positivos, vamos a obtener una región de interés con el objetivo de reducir el área de búsqueda del pólipo.

- **Segmentación y unión de regiones:** en esta última fase, vamos a aplicar un algoritmo de segmentación (watershed) sobre la imagen de valles. La imagen segmentada contendrá un número muy elevado de zonas. Por ello, el objetivo de esta fase es ir uniendo pequeñas zonas hasta conseguir una única zona que contenga el pólipo. El proceso consta de dos fases:

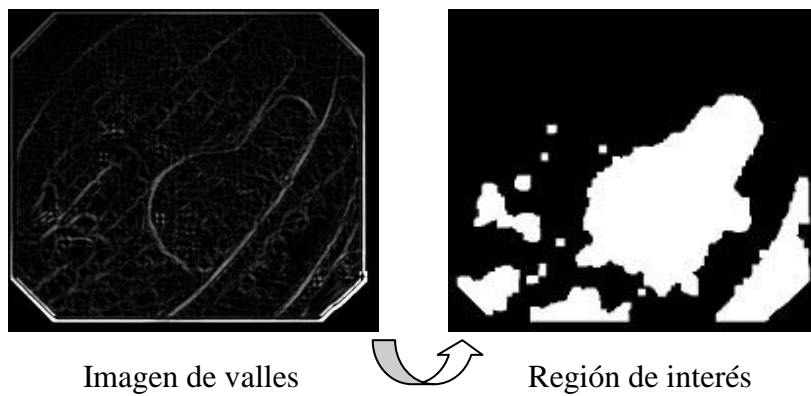
- En la primera, a partir de la región de interés obtenida en la fase anterior, eliminamos todas aquellas regiones del algoritmo de watershed que no estén contenidas en dicha región.
- En la segunda fase, con las regiones restantes, uniremos aquellas regiones que tengan como frontera un valle lo suficientemente débil.

En líneas generales, el esquema del proceso completo que se va a realizar:

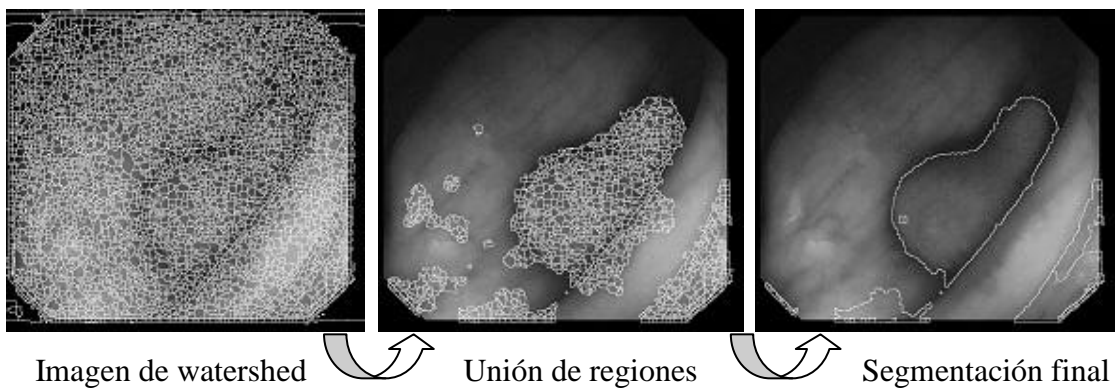
PREPROCESAMIENTO



DETECCIÓN DE VALLES



SEGMENTACIÓN Y UNIÓN DE REGIONES



En las siguientes secciones se explica con detalle los procedimientos realizados para llegar a nuestro objetivo:

Capítulo 2: Preliminares

Introducimos los conceptos teóricos que utilizaremos para desarrollar este trabajo.

Capítulo 3: Preprocesamiento

Adaptación y mejora de la imagen para su posterior procesamiento. Consta de dos fases: detección de reflejos e inpainting de los mismos. Este tratamiento es decisivo para los procedimientos en las siguientes fases.

Capítulo 4: Detección de valles

Se ha observado la presencia de sombras alrededor del pólipo. Vamos a utilizar un detector de valles (basado en la segunda derivada anisotrópica gaussiana). Debido a la cantidad de falsos positivos de la detección de valles, vamos a reducir el área de búsqueda del pólipo, obtendremos una región de interés que contendrá el pólipo y elimina zonas no interesantes de la imagen.

Capítulo 5: Segmentación y unión de regiones

Aplicamos el algoritmo de watershed sobre la imagen de valles para segmentar la imagen en un número elevado de regiones. Para segmentar el pólipo, iremos uniendo dichas regiones hasta llegar al objetivo.

Capítulo 6: Resultados y mejoras

Se presentan algunos de los resultados obtenidos para fotogramas de la base de datos. Proponemos mejoras para que nuestro sistema sea efectivo en más imágenes.

2. PRELIMINARES

En este capítulo recopilamos los conceptos técnicos utilizados a lo largo de todo el trabajo de fin de grado para facilitar la comprensión de los métodos aplicados.

2.1 Umbralización

La umbralización permite segmentar una imagen de manera que, dado un fotograma de niveles de gris resulta en una binaria (separando los objetos del fondo). Para este trabajo, los objetos son los reflejos y el fondo el resto de la imagen.

La imagen binaria o máscara $B(i, j)$ se obtiene a partir de la original $I(i, j)$ y de un umbral U . Este umbral es un nivel de intensidad que determina qué zonas son objeto y cuales son fondo.

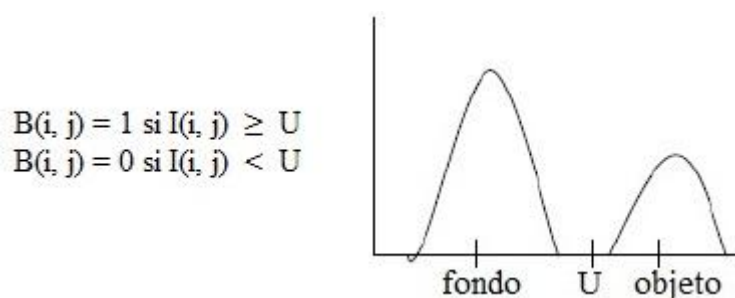


Figura 2: Fórmula y gráfica de umbralización

2.2 Morfología matemática

La morfología matemática estudia las formas y estructura de figuras. Es una herramienta para la extracción de componentes de la imagen útiles para la representación y descripción de las formas de una región. En la literatura es habitual encontrarla en temas relacionados con detección de bordes, esqueletos, valles... Así como en el preprocesamiento y postprocesamiento de sistema de visión por computador.

Una imagen puede considerarse formada por conjuntos de píxeles (objetos) y, por tanto, pueden aplicarse herramientas de la Teoría de conjuntos.

2.2.1 Elemento estructurante

Los operadores morfológicos se crean al hacer interactuar la imagen con un conjunto o subimagen denominado elemento estructurante. Para extraer propiedades de la imágenes, pueden aplicarse diferentes elementos estructurantes (figura izquierda). Sin embargo, debemos hacer que los elementos estructurantes sean rectángulos (figura derecha).



2.2.2 Dilatación

La dilatación es una operación que expande los objetos en una imagen. El tamaño de la dilatación viene dado por el elemento estructural, pequeños conjuntos que se utilizan para extraer propiedades de las imágenes (gráficamente, una matriz de 0s ó 1s). Sea A una imagen y B el elemento estructural (A y B dos conjuntos de Z^2), definimos la dilatación de A por B como:

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\}$$

donde el conjunto de puntos z tales que B, trasladado por z, tiene al menos un punto en común con A.

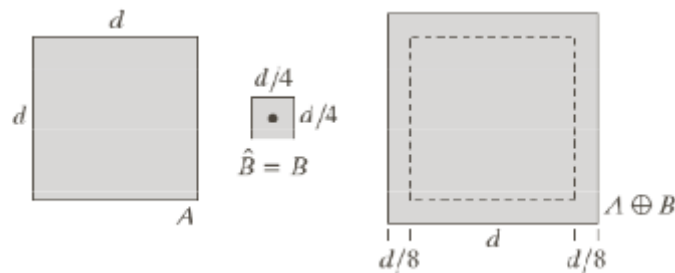


Figura 3: Dilatación de A por B.

2.2.3 Erosión

La erosión es una operación que se emplea para separar regiones débilmente unidas o para eliminar pequeños detalles en una imagen. Sea A una imagen y B el elemento estructural (A y B dos conjuntos de Z^2) definimos la erosión de A por B como:

$$A \ominus B = \{z \mid (B)_z \subseteq A\}$$

donde el conjunto de puntos z tales que B , trasladado por z , está totalmente contenido en A , es decir, el conjunto de puntos en los que B no tienen ningún elemento en común con el fondo de la imagen de A .

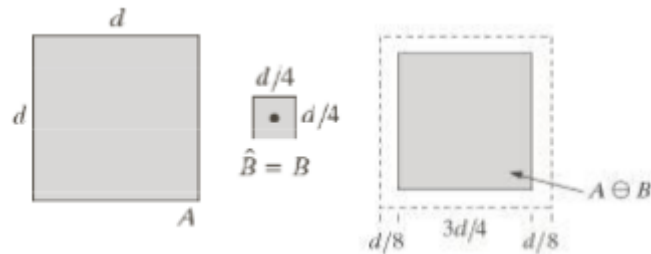


Figura 4: Erosión de A por B .

2.2.4 Apertura

La operación morfológica de apertura suaviza los contornos de una imagen, elimina zonas estrechas que unen dos objetos y pequeñas protuberancias. Utiliza las operaciones explicadas anteriormente de dilatación y erosión. La apertura de A por B es:

$$A \circ B = (A \ominus B) \oplus B$$

Por lo tanto, la apertura de A por B viene dada por la erosión de A por B , seguido de la dilatación del resultado por B .

2.2.5 Cierre

La operación de cierre trata de eliminar agujeros y unir componentes conexas cercanas (rellenar huecos) en la imagen. Utiliza las operaciones de erosión y dilatación. El cierre de A por B es:

$$A \cdot B = (A \oplus B) \ominus B$$

Es decir, el cierre de A por B consiste en la dilatación de A por B, seguido de la erosión del resultado por B.

En la siguiente ilustración vemos el funcionamiento de la combinación de ambas operaciones, (apertura y cierre).

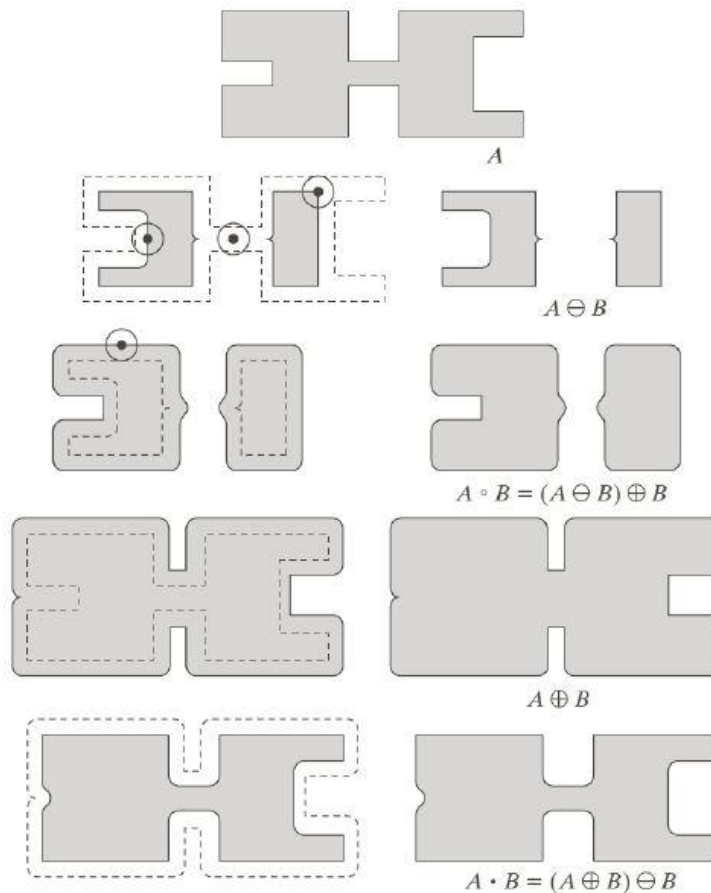


Figura 5: Apertura y cierre

3. PREPROCESAMIENTO

El objetivo de cualquier preprocesamiento es mejorar las imágenes para su posterior procesamiento. En este caso, el objetivo de nuestro preprocesamiento consiste en detectar y eliminar los reflejos de las imágenes ocasionados por la luz emitida por el colonoscopio. Este preprocesamiento es un paso clave para las siguientes fases de nuestro trabajo.

Para ello, vamos a realizar dos pasos:

- Detección de reflejos: basado principalmente en la umbralización con el objetivo de obtener una máscara (separando los reflejos del resto de la imagen) para localizar las zonas que deben ser tratadas con inpainting.
- Eliminación de reflejos: basado en el método de restauración de inpainting. Con este método se obtiene la imagen que tendríamos sin la presencia de reflejos.

3.1 Detección de reflejos

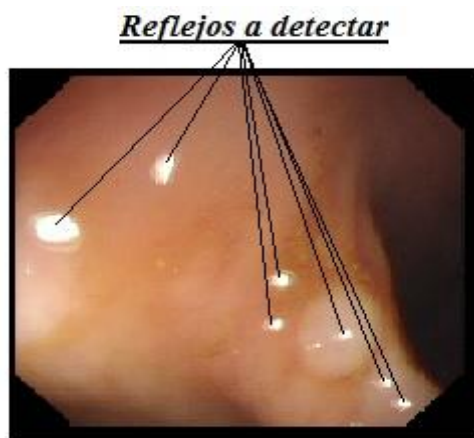


Figura 6: Reflejos

Como se observa en la imagen, pueden aparecer numerosos reflejos que dañan la calidad de la imagen. Dichos reflejos son puntos de intensidad muy alta y suficiente para diferenciarlos del resto de la imagen. Por ello, utilizamos uno de los métodos más básicos: la umbralización.

A partir de un umbral (intensidad), se va a segmentar la imagen de manera que separemos los reflejos del fondo. Como resultado, obtenemos una máscara que utilizaremos posteriormente para eliminar los reflejos.

3.1.1 Selección del umbral

Para determinar qué umbral es más adecuado para segmentar la imagen, se han estudiado distintas intensidades (umbrales). De manera que, para las distintas intensidades, analizaremos cuál de las máscaras obtenidas resulta la más completa o ajustada con la mayoría de fotogramas de la base de datos. A partir de dicho umbral, segmentamos la imagen en dos regiones (fondo-objeto).

Los umbrales estudiados son: 0.6, 0.7, 0.8. La imagen que queremos umbralizar se encuentra en el rango de intensidad [0, 1]. En los siguientes ejemplos vemos los resultados para distintas intensidades.

Ejemplo 1:



Figura 7: imagen original



Figura 8: máscara con umbral 0.6

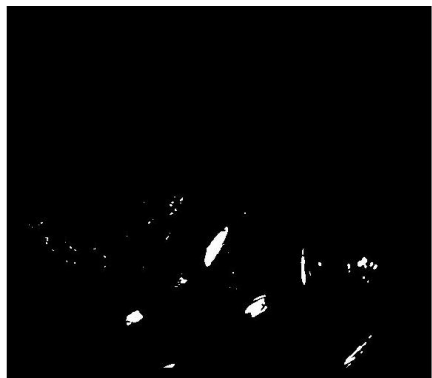


Figura 9: máscara con umbral 0.7

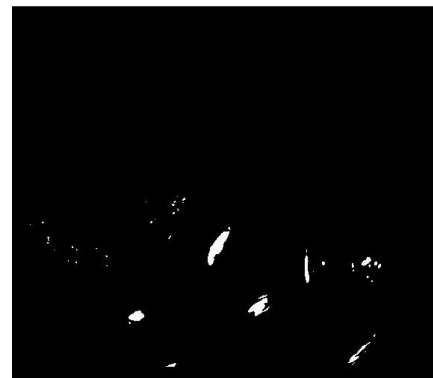


Figura 10: máscara con umbral 0.8

Ejemplo 2



Figura 11: imagen original



Figura 12: máscara con umbral 0.6

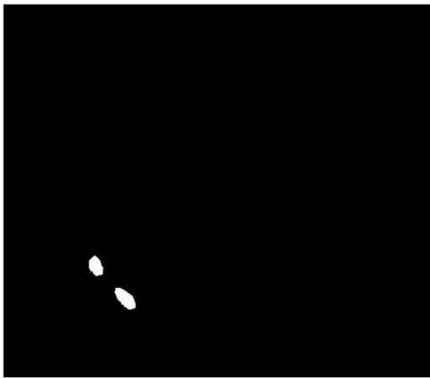


Figura 13: máscara con umbral 0.7

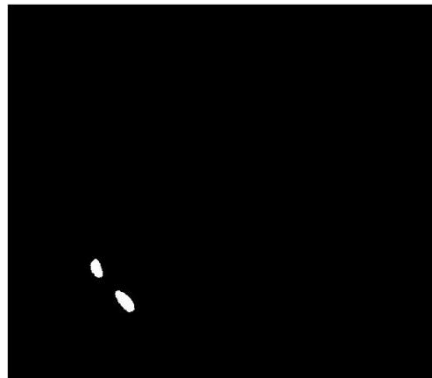


Figura 14: máscara con umbral 0.8

Como se ilustra en ambos ejemplos, con un umbral de 0.6 la segmentación no es buena ya que identifica muchas zonas de la imagen como objeto cuando, claramente, tienen que ser fondo. Los resultados para los umbrales 0.7 y 0.8 son similares. Sin embargo, en varias imágenes, con un umbral de 0.8 se pierden algunos píxeles en los bordes de los brillos. Por lo tanto, para conseguir resultados óptimos, seleccionamos un umbral de 0.7 para umbralizar los reflejos en nuestros fotogramas.

3.1.2 Mejora de la detección (máscara)

Alrededor de algunos reflejos, quedan píxeles con una intensidad ligeramente superior que la del umbral. Esto puede darse debido a zonas borrosas o movidas del fotograma. Eliminaremos dichos píxeles ya que forman parte de los reflejos. Sin embargo, seleccionar un umbral de mayor intensidad, a pesar de que solucionaría este problema, supondría detectar muchas otras zonas no deseadas.

Para asegurar que identificamos todos los píxeles que pertenecen a los reflejos, una opción es aumentar o expandir el área de las zonas previamente identificadas. Para ello, se va a emplear la operación morfológica de dilatación. Dicho método, explicado en detalle en el capítulo 2, consiste en expandir los objetos de una imagen en función de un elemento estructurante. El elemento estructural utilizado definirá el tamaño de la dilatación. En nuestro caso, vamos a utilizar como elemento estructural una matriz [7x7] de unos.



Figura 15: imagen original



Figura 16: máscara con umbral 0.7



Figura 17: Dilatación. Elem. estructurante [7x7]

Como vemos en las ilustraciones, mediante esta técnica aseguramos llegar a todos los píxeles cercanos a los reflejos que queremos eliminar pero no teníamos en la máscara original.

3.2 Eliminación de reflejos mediante inpainting

Hemos visto en la sección anterior que los reflejos detectados son muy problemáticos para la detección y segmentación de pólipos. Por eso, hemos decidido eliminar dichos reflejos mediante técnicas de inpainting.

En esta sección se va a explicar cómo eliminar esos reflejos. El objetivo es hallar un nuevo valor para las zonas detectadas anteriormente que se aproxime al que hubieran tenido si no estuvieran los brillos. Para ello hemos utilizado el método de restauración de imagen de inpainting.

Inpainting (restauración de imagen) es una técnica que permite recuperar una zona deteriorada o seleccionada de una imagen con el objetivo de restaurarla. Este método se utiliza para eliminar objetos, personas o ruido de una imagen y reconstruir las zonas a partir de la información que tenemos alrededor de dichas zonas.

Como se explica a lo largo de la sección, la eliminación de reflejos consta de dos pasos:

1. Inpainting de la imagen con la máscara obtenida anteriormente
2. Suavizado de la imagen obtenida del paso previo. Con esto conseguimos moderar el cambio abrupto de intensidades entre la zona restaurada por inpainting y el resto de la imagen.

3.2.1 Algoritmo de inpainting:

Entrada: imagen I , máscara Ω , tamaño de vecindad n

Salida: imagen I restaurada

PARA cada píxel $(i, j) \in \Omega$ HACER

 Calcular el vecindario V de $n \times n$ píxeles alrededor del píxel (i, j)

 Calcular la intensidad restaurada del píxel (i, j)

$$I(i, j) = \frac{1}{|V \setminus \Omega|} \sum_{(k, l) \in V \setminus \Omega} I(k, l)$$

 Eliminar (i, j) de la máscara Ω

Este algoritmo busca en la imagen los valores de la máscara que representan a los objetos (reflejos) y cuando lo encuentra, empieza a recorrer el objeto “por capas”, es decir, del borde del objeto hacia el interior. De esta manera conseguimos que el nuevo valor sea el más parecido a su homólogo hipotético. Para calcular la intensidad restaurada del píxel, utilizamos el vecindario del píxel a restaurar. Únicamente tendremos en cuenta aquellos píxeles del vecindario donde, según indique la máscara, tenemos información, es decir, fondo y no objeto.

Los resultados son buenos para la mayoría imágenes, pero como se observa en las ilustraciones al final de la sección, puede aparecer un cambio abrupto en las intensidades entre las zonas eliminadas y el resto de la imagen. Esto se debe a que en la imagen original, en los alrededores de los puntos brillantes aparecen “picos” de intensidad que no son suficientemente altos para pertenecer a la máscara pero sí para que se note su presencia. Debido a esto, realizamos un suavizado a las zonas tratadas de la imagen.

3.2.2 Filtro de suavizado

La imagen restaurada, (sección 3.2.1), contiene saltos abruptos de intensidad entre las zonas con nuevo valor y el resto de la imagen. Esto se debe, principalmente, a que la imagen original pueda estar movida o que haya píxeles cercanos a los reflejos sin la intensidad suficiente para ser segmentados (método de umbralización) pero que queremos eliminar.



Figura 18: Fotograma con reflejos no detectados y detalle

Para reducir el cambio abrupto de intensidad, realizamos la convolución de un filtro de suavizado únicamente sobre las zonas de la imagen restauradas (realizarlo sobre toda la imagen supondría emborracharla y perder información necesaria). El filtro consiste en la combinación de las operaciones morfológicas de apertura y cierre (explicadas en detalle en el capítulo 2). El objetivo de la convolución con dicho filtro es suavizar los contornos de las zonas de los reflejos, eliminar pequeñas protuberancias, agujeros y componentes conexas cercanas.

En las siguientes ilustraciones, vemos el proceso realizado hasta ahora. A partir de la imagen original, obtenemos una máscara detectando los reflejos de la imagen que impiden segmentar, posteriormente, el pólipo. Con la máscara y la imagen original eliminamos dichos reflejos (método de inpainting). Como se observa en la figura 21 quedan píxeles que provocan un cambio visible entre la zona restaurada y el resto de la imagen. Por ello, finalmente se aplica un filtro de suavizado a dichas zonas, consiguiendo moderar el cambio de intensidad en el área (figura 22).



Figura 19: imagen original

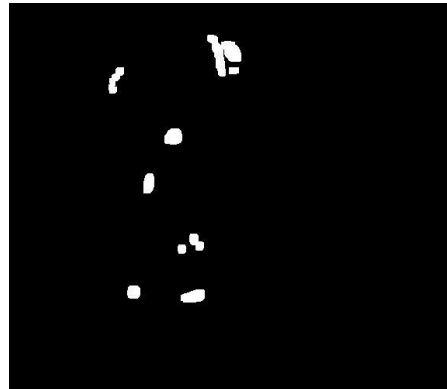


Figura 20: máscara



Cambio abrupto de intensidad



Figura 21: imagen restaurada (izq.), ampliación de detalle recuadro rojo (dcha.)

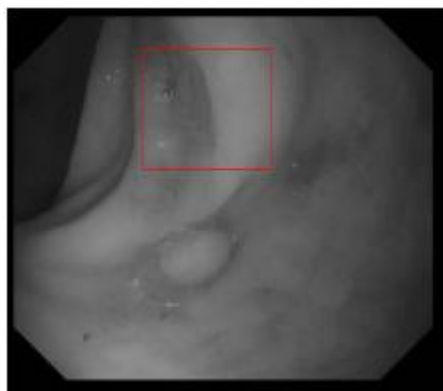


Figura 22: imagen tras filtro de suavizado (izqda.) y ampliación del recuadro (dcha.)

3.2.3 Resultados de preprocesamiento (detección e inpainting)

En los siguientes ejemplos, mostramos el proceso llevado a cabo en este capítulo. A partir de una imagen dada (figuras 23, 26, 29) detectamos los reflejos de dichas imágenes mediante umbralización (sección 3.1) de manera que obtenemos una máscara (figuras 24, 27, 30). Por último, la imagen final es el resultado de aplicar el método de inpainting a la imagen original únicamente en las zonas detectadas en la máscara (figuras 25, 28, 31).

Ejemplo 1:

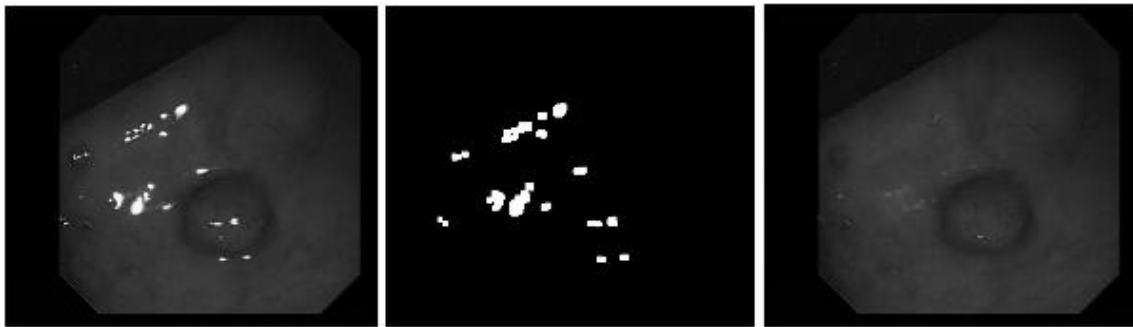


Figura 23: im. original

Figura 24: máscara

Figura 25: im. Restaurada

Ejemplo 2:

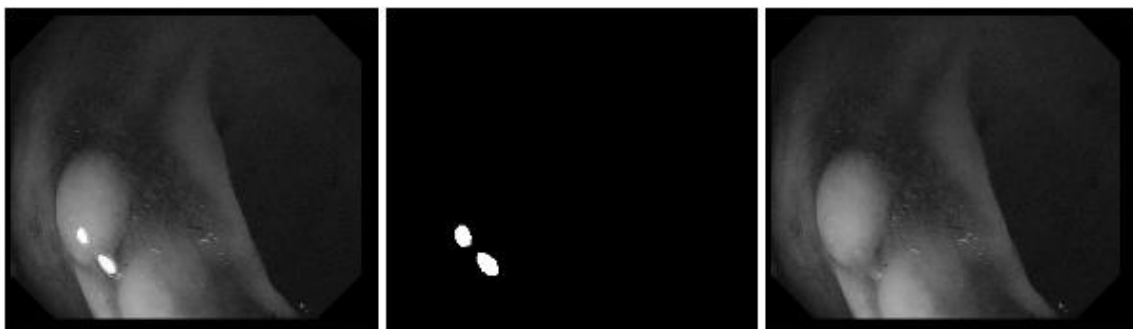


Figura 26: im. original

Figura 27: máscara

Figura 28: im. Restaurada

Ejemplo 3:



Figura 29: im. original

Figura 30: máscara

Figura 31: im. Restaurada

4. DETECCIÓN DE VALLES

La aparición de pólipos en fotogramas de colonoscopia es muy variable. La forma del pólipo no tiene por qué ser uniforme ni seguir un determinado patrón. Sin embargo, en la mayoría de imágenes encontramos un denominador común. Se trata de la presencia de “sombras” alrededor del pólipo debido a la iluminación del colonoscopio (instrumento utilizado en la colonoscopia). Estas sombras, valles en términos de intensidad, van a guiar la detección y segmentación del pólipo (ver figura 32).



Figura 32: sombras alrededor del pólipo

La detección de pólipos consta de dos fases:

- Imagen de valles: se identifica y resalta las sombras que aparecen alrededor de los pólipos. Para ello, utilizamos la segunda derivada anisotrópica gaussiana.
- Región de interés: la imagen de valles, resalta diversas zonas de la imagen, además del pólipo, que no interesan. Por este motivo a partir de dicha imagen reducimos el área a considerar para segmentar el pólipo. De esta manera, facilitamos la última fase de este trabajo, la segmentación (capítulo 5).

4.1 Imagen de valles

El objetivo en esta fase del trabajo es detectar la zona en la que probablemente se encuentre el pólipo. Para detectarla, las sombras o valles que encontramos alrededor de los pólipos servirán de guía. La *segunda derivada anisotrópica gaussiana* resalta dichos valles, aunque además de estos también resalta otras zonas no deseadas de la imagen. Por este motivo, posteriormente reduciremos la zona a segmentar únicamente en la región de interés.

4.1.1 Segunda derivada anisotrópica de la gaussiana

El objetivo de esta fase es identificar y resaltar el valle de intensidad (sombra) que aparece alrededor de los pólipos. La detección de valles (junto con la detección de bordes, líneas, etc) es una de las técnicas de extracción de características de bajo nivel más estudiadas. Sin embargo, existe cierta diferencia entre la detección de bordes y la detección de valles: un valle está compuesto, generalmente, por dos bordes paralelos. Así, mientras que los bordes responden a un máximo/mínimo local de la primera derivada, los valles responden a máximos/mínimos de la segunda derivada.

La convolución con filtros (kernels) gaussianos ha sido muy utilizada en la literatura para la eliminación de ruido, suavizado (orden cero), la detección de bordes (derivada de orden uno) y la detección de valles y crestas (derivada de orden dos). Sin embargo, la utilización de filtros gaussianos simétricos respecto al origen de coordenadas puede no ser la mejor herramienta para la detección de valles, dado que los valles tienen, por lo general, diferentes escalas y orientaciones dentro de una imagen.

Para solucionar estos problemas, en este trabajo proponemos la utilización de filtros gaussianos anisotrópicos de orden dos (segunda derivada de dichos filtros). La idea es generar diferentes filtros (diferentes orientaciones y amplitudes) e intentar capturar la presencia de valles independientemente de su tamaño y orientación.

En la figura 33, vemos gráficamente el resultado de las distintas derivadas de la función de Gauss (derivadas de orden 0, 1 y 2). A continuación se explica cómo calcular la derivada de segundo orden.

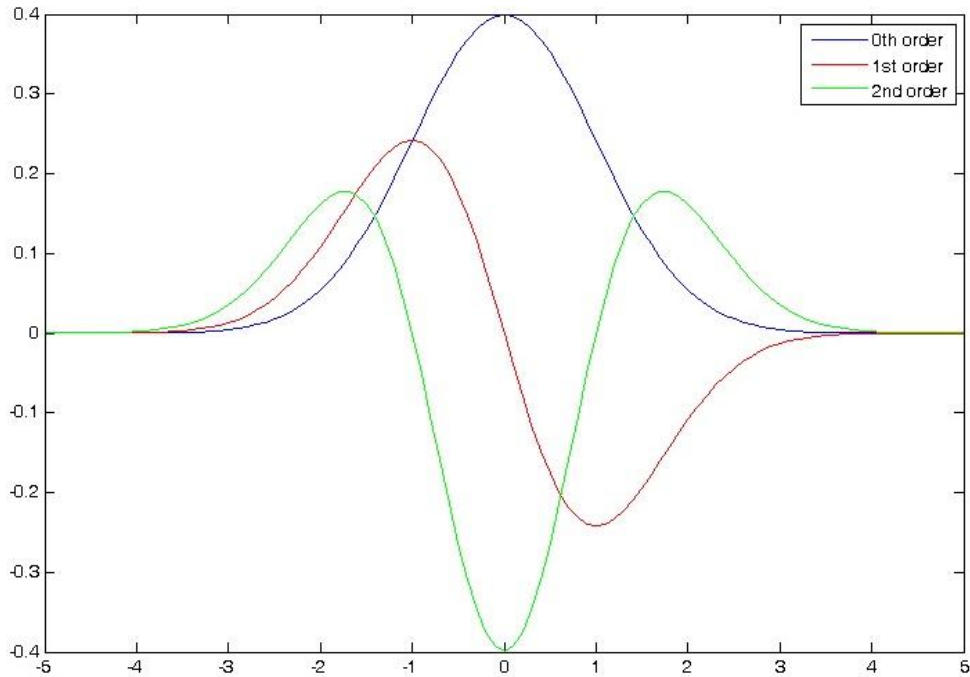


Figura 33: Representación gráfica de las funciones gaussianas

4.1.1.1 Calculo de la segunda derivada anisotrópica de Gauss

Definimos los valores del núcleo gaussiano mediante la siguiente función:

$$G_{(\sigma_x, \sigma_y), \theta} = \frac{1}{(2\pi)\sigma_x\sigma_y} e^{-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)}$$

donde (σ_x, σ_y) son la amplitud de la gaussiana en su correspondiente eje, y θ es el ángulo de rotación del filtro. En función del ángulo de rotación las coordenadas 'x' e 'y' se definen como:

$$\begin{aligned}\tilde{x} &= x \cos \theta + y \sin \theta \\ \tilde{y} &= x \sin \theta - y \cos \theta\end{aligned}$$

Como ya hemos mencionado, se va a utilizar la función gaussiana anisotrópica, esto es con una amplitud de gaussiana $\sigma = \sigma_x = 2\sigma_y$. Por lo tanto la función resultante es,

$$G_{\sigma, \theta} = \frac{1}{(2\pi)2\sigma^2} e^{-\left(\frac{\tilde{x}^2}{2(2\sigma)^2} + \frac{\tilde{y}^2}{2\sigma^2}\right)}$$

Finalmente el núcleo gaussiano empleado para detectar los valles de los pólipos, queda definido como,

$$\partial_{\tilde{y}}^2 G_{\sigma, \theta} = \frac{\tilde{y}^2 - 1}{\sigma^4} G_{\sigma, \theta}$$

Vamos a utilizar 8 orientaciones distintas, para calcular los núcleos, igualmente distribuidas entre $[0, \pi]$. La amplitud de la gaussiana, σ , se ha definido como $\sigma = 4$. Por lo tanto, para cada orientación, se convoluciona la imagen I con el núcleo gaussiano resultante.

$$G_{\sigma, \theta}^N = \partial_y^2 G_{\sigma, \theta} * I$$

Lógicamente, obtendremos 8 imágenes distintas correspondientes a cada una de las orientaciones θ . Tenemos, en cada imagen, el valor concreto de los valles para la orientación correspondiente. La imagen de valles final debe contener el contorno completo del valle. Por este motivo, el resultado final ha de ser una combinación de todas las orientaciones. Para obtener tal combinación de las 8 imágenes, sin perder información, calculamos la máxima intensidad de cada píxel en todas las orientaciones. Como resultado obtendremos la imagen de valles:

$$I_{valles} = \max (G_{\sigma, \theta_j}^N)$$

4.1.1.2 Orientaciones de la segunda derivada anisotrópica gaussiana

En la siguiente secuencia de imágenes, se observan las imágenes obtenidas tras aplicar la convolución de la imagen con el núcleo resultante para cada orientación establecida. Como muestran las imágenes, para cada orientación se “resalta” una zona distinta (correspondiente a la orientación) de manera que dependiendo del valor de θ , resaltamos detalles distintos pero igualmente importantes para conseguir el resultado final. La imagen de valles obtenida es la suma de todos esos detalles, es decir, el máximo de cada píxel para las 8 imágenes. De esta manera se consigue recoger toda la información de todas las orientaciones establecidas en una sola imagen.

Las orientaciones (θ) con las que se trabaja son: $\theta = [0, \pi/6, \pi/4, \pi/3, \pi/2, 2/3\pi, 3/4\pi, \pi]$.

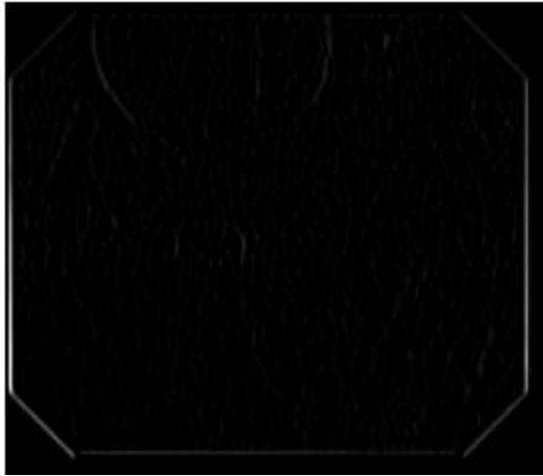


Figura 34: $\theta = 0$



Figura 35: $\theta = \pi/8$

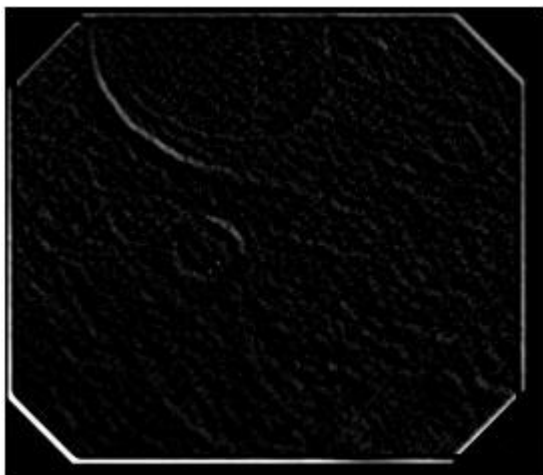


Figura 32: $\theta = \pi/4$

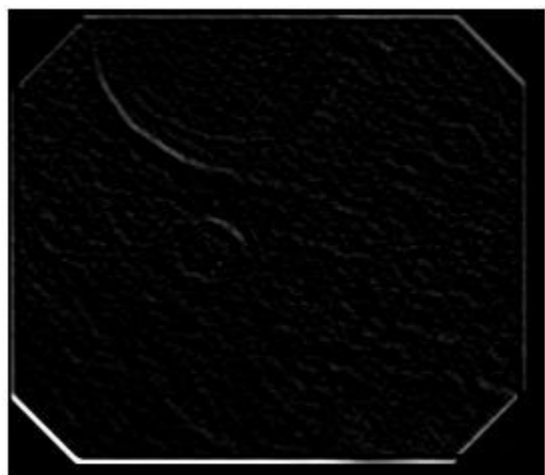


Figura 33: $\theta = 3\pi/8$

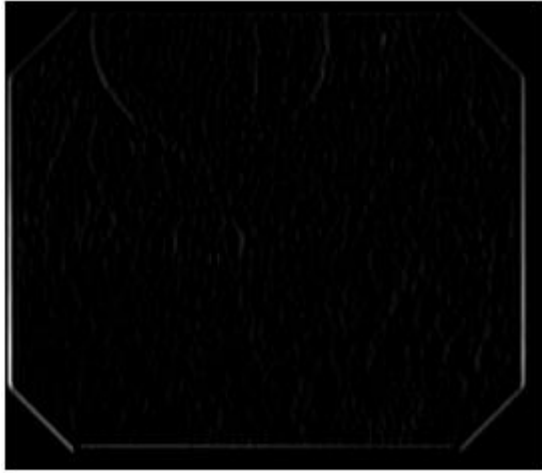


Figura 34: $\theta = \pi/2$



Figura 35: $\theta = 5\pi/8$

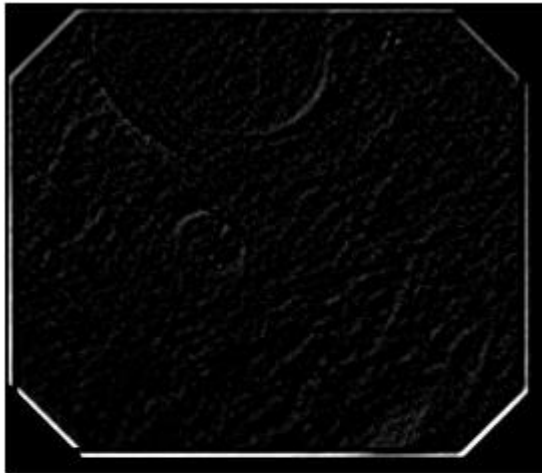


Figura 36: $\theta = 3\pi/4$

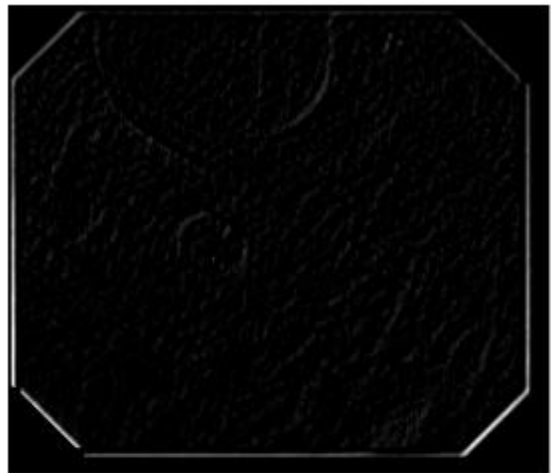


Figura 37: $\theta = 7\pi/8$

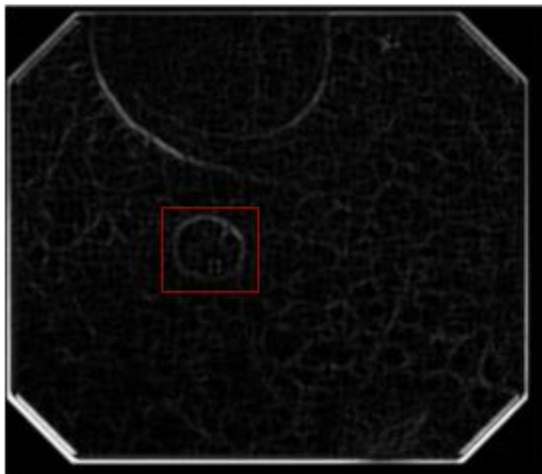


Figura 42: Imagen de valles

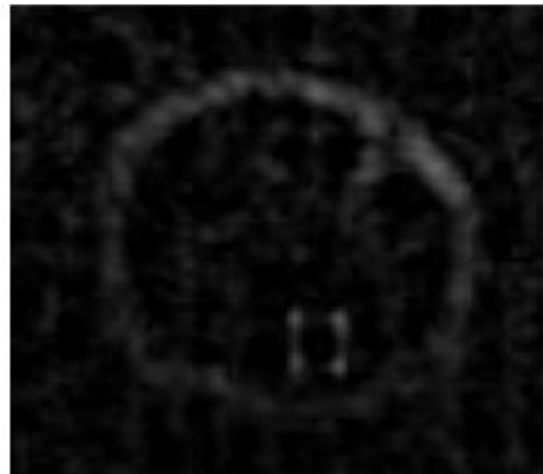


Figura 43: detalle de pólipo

4.1.2 Resultados de imagen de valles

En los siguientes ejemplos, el pólipo bastante definido (objetivo de esta fase del trabajo), a pesar de que como ya se ha comentado, resaltamos zonas no deseadas de la imagen.

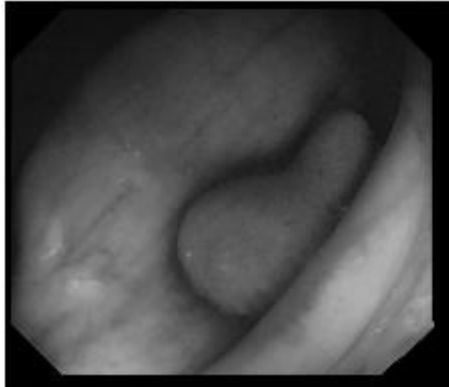


Figura 44: imagen original

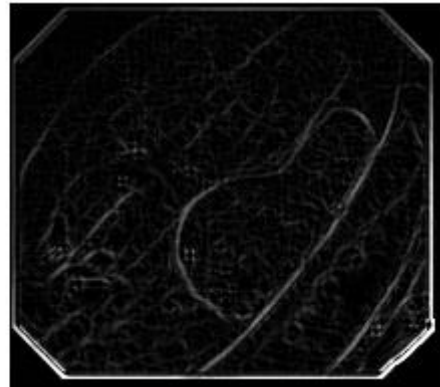


Figura 45: imagen de valles

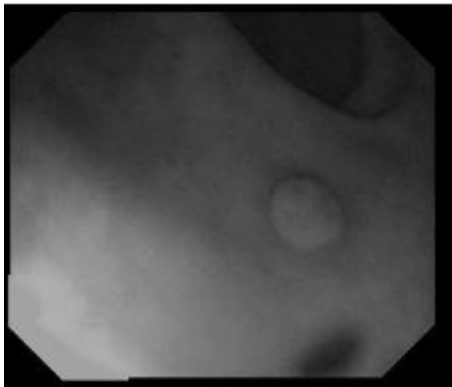


Figura 46: imagen original

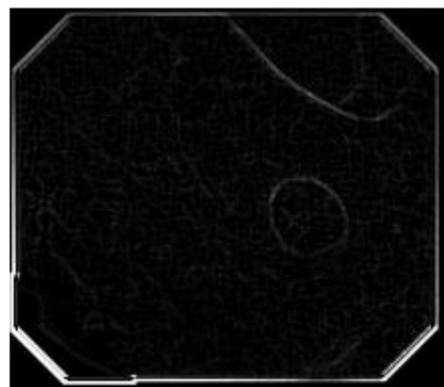


Figura 47: imagen de valles

Sin embargo, no siempre se consigue una detección “suficiente” del pólipo (ver figura 48). En la figura 49, el contorno del pólipo queda difuminado parcialmente. Esto se debe a que en la imagen original no hay una sombra bien definida alrededor del pólipo. En estos casos, la detección falla.

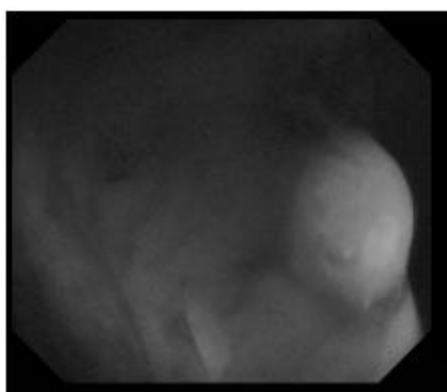


Figura 48: imagen original

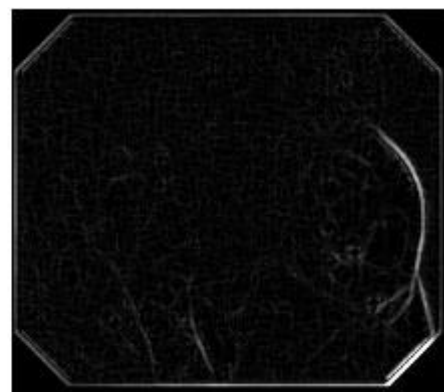


Figura 49: imagen de valles

4.2 Región de interés (ROI)

La imagen de valles contiene un número elevado de falsos positivos, de manera que además del pólipo, se resaltan otras zonas de la imagen no deseadas. Esta sección trata de minimizar esas zonas mediante operaciones morfológicas sobre la imagen de valles. De esta manera, reducimos y nos aproximamos considerablemente al área de búsqueda del pólipo.

En la figura 50, el recuadro representa la zona de mayor importancia en la imagen de valles. Dicha zona contiene en su totalidad al pólipo. En el resto de la imagen, aparece información sobre elementos de la escena que no nos interesan. Por lo tanto, el objetivo es encontrar una región de interés semejante al área descrita por el recuadro mencionado.

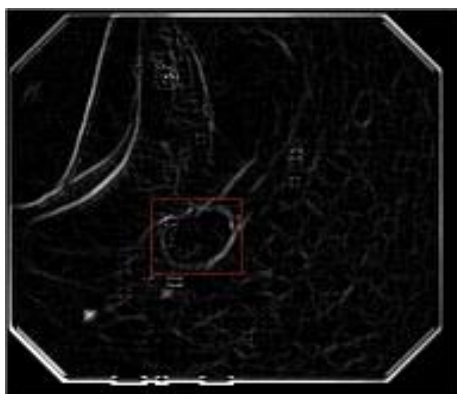


Figura 50: imagen de valles

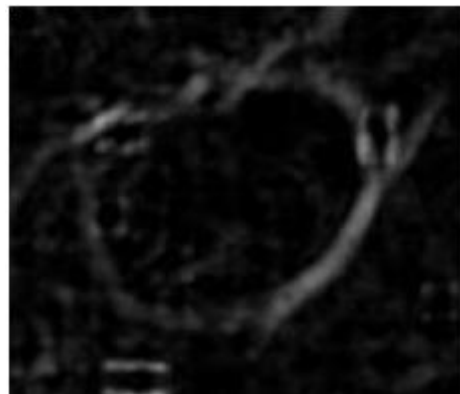


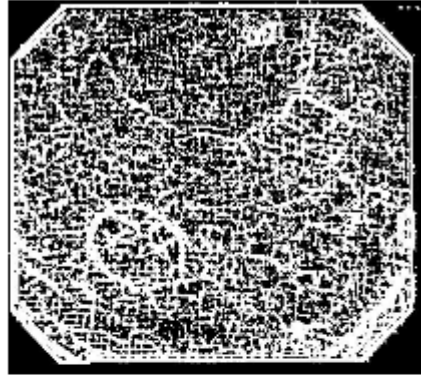
Figura 51: pólipo

La figura 51 muestra en detalle el pólipo que queremos detectar y segmentar. Podemos observar que los valles resultantes en la zona del pólipo, forman una curva cerrada. Este hecho es fundamental para reducir el área de búsqueda del pólipo ya que al formar una curva cerrada, aísla, de alguna manera, el pólipo del resto de la imagen. A partir de la imagen de valles y de la idea de la curva cerrada que forman algunos valles, vamos a realizar varios pasos para identificar la región de interés:

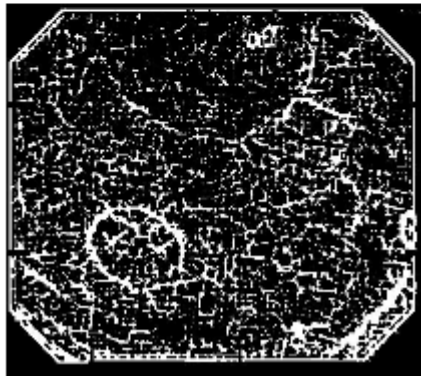
1. Ecuación del histograma de la imagen de valles (binarizar imagen).
2. Operación morfológica de erosión sobre la imagen del paso anterior.
3. Relleno de formas curvilíneas cerradas presentes en la imagen.
4. Operación morfológica de apertura sobre la imagen rellena.

En las siguientes ilustraciones hacemos un seguimiento de los pasos previamente mencionados para un fotograma concreto:

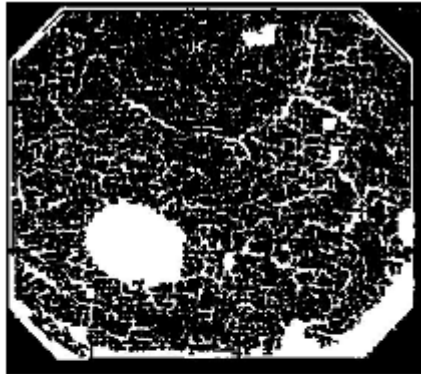
Paso 1 (binarización im. de valles):



Paso 2 (operación de erosión) :



Paso 3 (relleno de curvas cerradas):



Paso 4 (operación de apertura):



4.2.1 Ecuación del histograma

El histograma representa el número de píxeles de la imagen con un determinado nivel de intensidad (cuantización). La ecualización del histograma trata de obtener un nuevo histograma en el que todas las intensidades tengan un número parecido de píxeles. Esto es, obtener un histograma uniforme. En nuestro caso, el nuevo histograma contendrá únicamente dos niveles de intensidad, es decir, vamos a binarizar la imagen.

Al binarizar la imagen, todos los valles detectados tendrán la misma intensidad. Esta representación ecuánime de los valles facilita la búsqueda, en los siguientes pasos, de la silueta de curva cerrada que forma el pólipo.

4.2.2 Erosión de la imagen ecualizada

Los valles detectados pertenecientes a un pólipo sobresalen, lógicamente, del resto de la imagen. Aunque no son los únicos valles “fuertes” presentes en la imagen, podemos observar en la figura 46 que su grosor destaca del resto. Para nuestra región de interés, este hecho es beneficioso ya que diferencia al pólipo del resto de la imagen.

Ya que el objetivo es eliminar zonas irrelevantes de la imagen, sin perder ninguna información del pólipo, vamos a debilitar la presencia de todos los valles. De modo que aquellos valles inicialmente débiles lo sean aún más mientras que aquellos fuertes continúan destacando del resto. Por estos motivos vamos a erosionar la imagen.

4.2.3 Relleno de formas curvilíneas cerradas

Necesitamos asegurar que en la región de interés obtenida esté contenido completamente el pólipo. Para ello, tras las operaciones anteriores realizadas (paso 1 y 2) vamos a rellenar todas aquellas formas curvilíneas cerradas (filling holes) ya que sabemos que una de ellas va a ser el pólipo. En el paso 2 hemos erosionado la imagen. Esta operación elimina la posible presencia de otras formas curvilíneas cerradas que, significativamente, tienen valles más débiles que los del pólipo.

4.2.4 Apertura

Este último paso consiste en eliminar todas aquellas zonas(valles) de la imagen que no han resultado rellenos en el paso anterior. Para ello, vamos a utilizar la operación morfológica de apertura. Dicha operación precisa de un elemento estructural (ver capítulo 2), en nuestro caso ha de ser lo suficientemente grande para suprimir todos los valles restantes sin eliminar elementos del pólipo. En concreto, se ha utilizado un elemento estructural de 11×11 .

En la ilustración del paso 4, representa la región de interés. Como se puede observar, se ha reducido y acotado considerablemente el área de búsqueda para proceder a la segmentación del pólipo.

5. UNION DE REGIONES Y SEGMENTACIÓN

En esta última fase de nuestro trabajo vamos a aplicar un algoritmo de segmentación (watershed) sobre la imagen de valles. La imagen segmentada contendrá un número muy elevado de zonas. Este hecho podría tratarse de un caso de sobre segmentación pero en nuestro trabajo, para ajustarnos a la segmentación final más aproximada del pólipo precisamos de este método. Debido al elevado número de zonas, el objetivo de esta fase es ir uniendo pequeñas regiones hasta conseguir una única zona que contenga al pólipo. El proceso consta de dos fases:

- En la primera, a partir de la región de interés (capítulo 4), eliminamos todas aquellas regiones del algoritmo de watershed que no estén contenidas en dicha región.
- En la segunda fase, con las regiones restantes, uniremos aquellas zonas que tengan como frontera un valle lo suficientemente débil.

5.1 Algoritmo de Watershed

Para proceder con una segmentación inicial, vamos a aplicar el algoritmo de watershed [3]. Este método se basa en el concepto de “línea divisoria”.

En términos geográficos, watershed es una cresta que divide las áreas drenadas por los diferentes sistemas fluviales. La transformación de watershed aplica estas ideas al procesamiento de imágenes de manera que puede usarse para segmentar imágenes. Para resolver problemas de segmentación de imágenes, el concepto clave es seleccionar una imagen inicial cuyas cuencas son los objetos o regiones que queremos identificar.

En este trabajo, siguiendo con la idea de que los objetos (pólipos) están rodeados por valles de intensidad hemos seleccionado como imagen inicial la propia imagen de valles ya que de esta manera el algoritmos se adapta mejor a la estructura de la imagen.

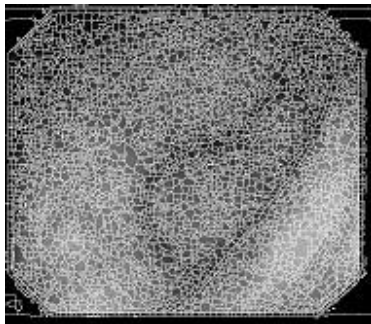


Figura 52: Watershed

5.2 Regiones no presentes en ROI

En este paso previo a la segmentación final, reducimos el elevado número de regiones resultantes del algoritmo de watershed. Para ello, trataremos la región de interés (capítulo 4) como una máscara. Vamos a realizar los siguientes pasos:

1. Segmentación en regiones (algoritmo de watershed).
2. Para cada región, calcular su adyacencia (regiones contiguas)
3. Ordenar las regiones en función de su tamaño (número de elementos).
 - a. $region_to_merge_1 = \min(\text{regiones})$
 - b. $region_to_merge_2 = \min(\text{adyacencia}(region_to_merge_1))$
4. Comprobación de region merging
 - a. frontera = píxeles de separación de las regiones a unir
 - b. Si la frontera no pertenece ROI
 - i. Unir regiones
 - ii. Volver al paso 3
 - [] frontera pertenece ROI
 - i. regiones no unificables
 - ii. Volver al paso 3
5. Fin tras tratar todas las regiones

La forma de proceder a la hora de decidir el orden de unión de las regiones es decisiva para el resultado final. En este trabajo se ha propuesto un algoritmo que selecciona las regiones a unir en función de su tamaño. Primero, elige la región más pequeña de todas ($region_to_merge_1$) como candidata para unificar. La elección de la segunda región ($region_to_merge_2$) parte de las regiones adyacentes o contiguas a la primera. Nuevamente, elegiremos la región más pequeña de las adyacentes para tratar de unificarlas. La decisión de unir o no dichas regiones cae sobre la región de interés. En función de si los bordes que separan a las regiones candidatas forman parte o no de la región de interés, quedarán o no unidas. El algoritmo finaliza cuando hemos tratado todas las regiones.

En las siguientes ilustraciones, vemos el resultado para un fotograma concreto. Como se puede observar, se ha reducido drásticamente el número de regiones iniciales quedando únicamente aquellas con elementos comunes a la máscara (región de interés).



Figura 53: máscara (ROI)

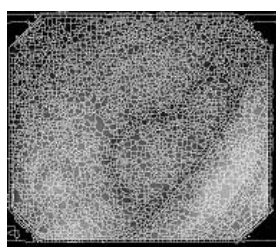


Figura 54: watershed

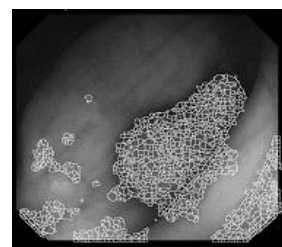


Figura 55: Unión de regiones

5.3 Segmentación final

De manera análoga a lo explicado anteriormente, uniremos las regiones restantes con el objetivo de que en la segmentación definitiva quede delimitado el pólipo. A diferencia de la segmentación previa, unificamos las regiones en función de la información de la imagen de valles. Así pues, partiendo de la segmentación previa, el algoritmo propuesto para la segmentación final es el siguiente:

1. Para cada región, calcular su adyacencia (regiones contiguas)
2. Ordenar las regiones en función de su tamaño (número de elementos)
 - a. $region_to_merge_1 = \min(\text{regiones})$
 - b. $region_to_merge_2 = \min(\text{adyacencia}(region_to_merge_1))$
3. Comprobación de region merging
 - a. frontera = píxeles de separación de las regiones a unir
 - b. Si la frontera no es un valle fuerte
 - i. Unir regiones
 - ii. Volver al paso 2
 - [] frontera es un valle fuerte
 - i. regiones no unificables
 - ii. Volver al paso 2
4. Fin tras tratar todas las regiones.

Como vemos, este algoritmo es similar al anterior. La diferencia es cómo decidimos si unimos o no las regiones. Cuantificar el valor para definir un valle fuerte o débil no es tarea sencilla. En este trabajo se ha definido una intensidad concreta para cada fotograma, tratando de discernir un valle fuerte de uno débil. Siguiendo con el ejemplo del apartado anterior, la intensidad escogida para diferenciar valles es 45 (Ver figura 56).

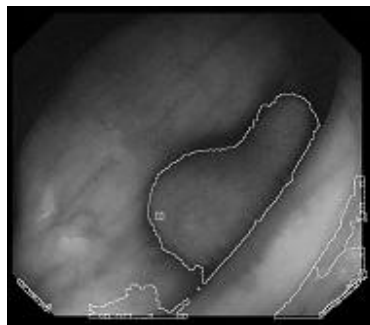


Figura 56: pólipo segmentado

6. RESULTADOS

A continuación se muestran los resultados obtenidos para algunos fotogramas de la base de datos:

Fotograma 1:



Figura 57: Fotograma

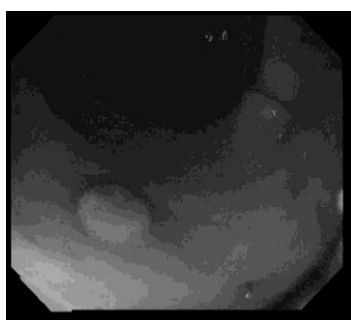


Figura 58: im. restaurada

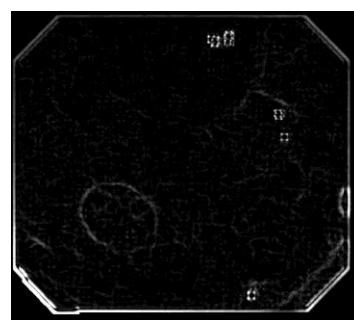


Figura 59: im. de valles



Figura 60: Región de interés

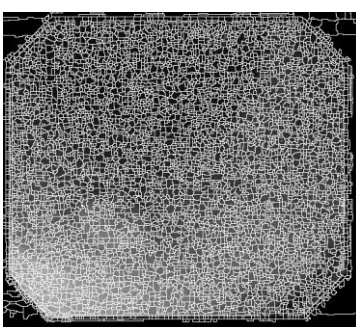


Figura 61: Alg. Watershed

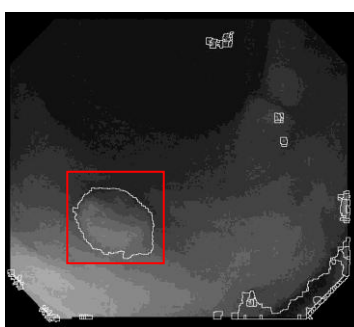


Figura 62: Pólipo segmentado

Fotograma 2:

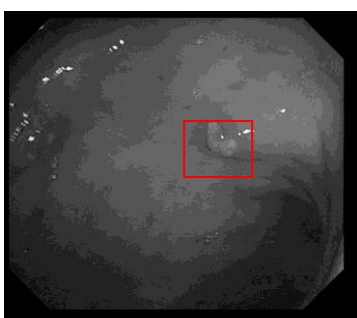


Figura 63: Fotograma

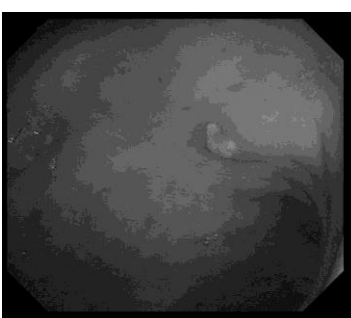


Figura 64: im. restaurada

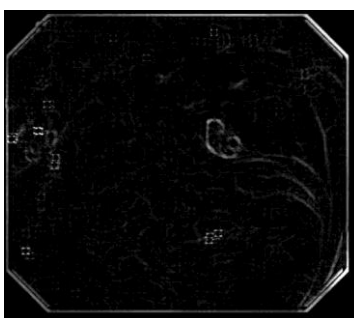


Figura 65: im. de valles



Figura 66: Región de interés

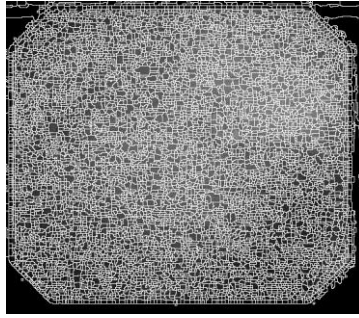


Figura 67: Alg. Watershed

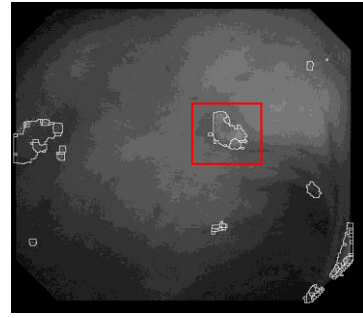


Figura 68: Pólipo segmentado

Fotograma 3:

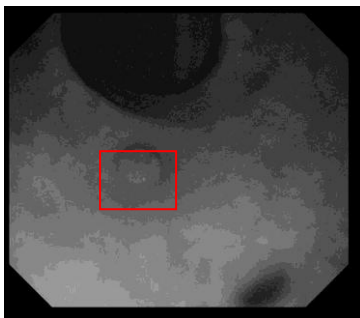


Figura 69: Fotograma

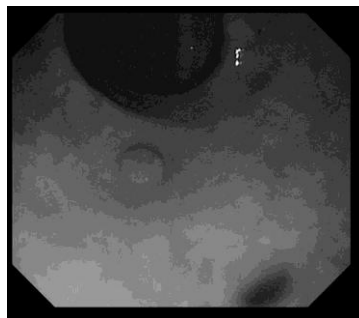


Figura 70: im. restaurada

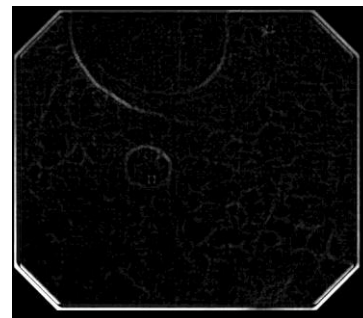


Figura 71: im. de valles



Figura 72: Región de interés

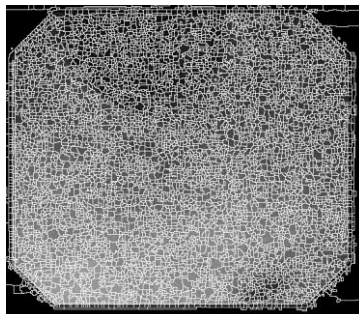


Figura 73: Alg. Watershed

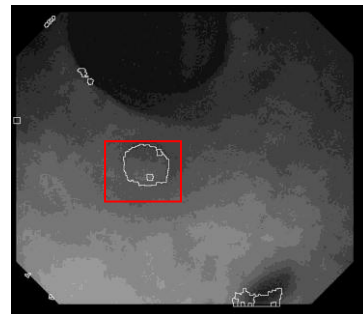


Figura 74: Pólipo segmentado

Analizando los resultados aquí presentados, vemos que en el fotograma 1 la segmentación del pólipo es completa y ajustada. En los dos siguientes fotogramas también queda el pólipo perfectamente segmentado aunque en su interior queda alguna región que no se ha unido.

Ante estos resultados, la primera mejora que podríamos incluir en nuestro sistema de visión artificial es un método de preprocesamiento para eliminar los píxeles aislados de alta intensidad que provocan gran respuesta en la imagen de valles impidiendo unir la región que les representa.

En los siguientes fotogramas, debido al entorno de la imagen no conseguimos segmentar el pólipo:

Fotograma 4:

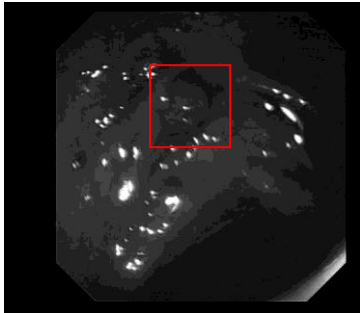


Figura 75: Fotograma

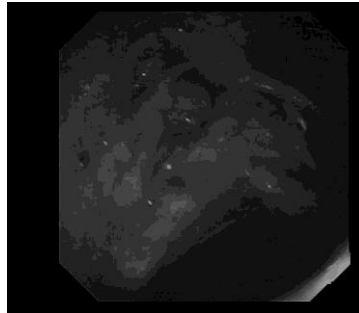


Figura76: im. restaurada

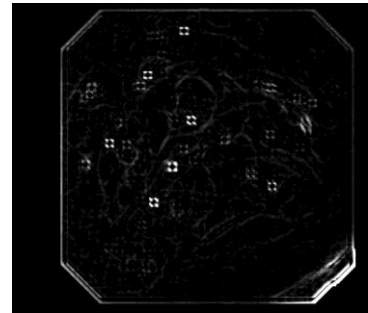


Figura77: im. de valles



Figura 78: Región de interés

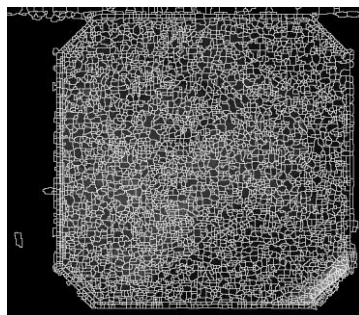


Figura 79: Alg. Watershed

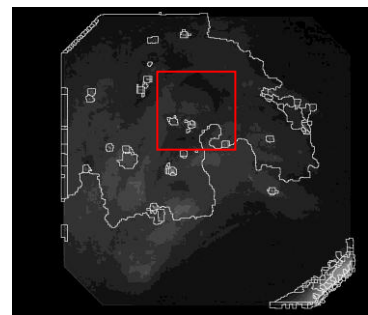


Figura 80: Pólipo segmentado

Fotograma 5:

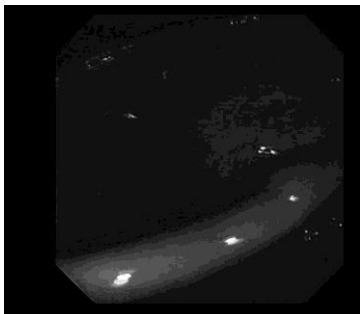


Figura 81: Fotograma

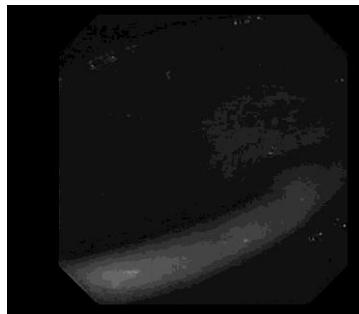


Figura 82: im. restaurada

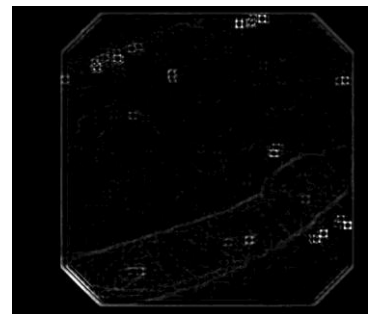


Figura 83: im. de valles



Figura 84: Región de interés

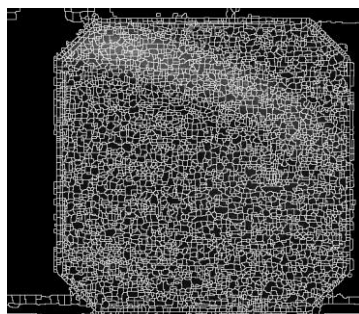


Figura 85: Alg. Watershed

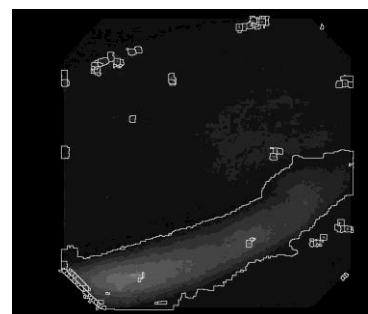


Figura 86: Pólipo segmentado

7. CONCLUSIONES Y LINEAS FUTURAS

En este trabajo se ha diseñado un sistema de visión artificial para automatizar la detección y segmentación de pólipos en imágenes de colonoscopia. El sistema se ha desarrollado principalmente en las siguientes fases:

- **Preprocesamiento:** en esta primera fase se ha adaptado y mejorado la imagen para facilitar el procesamiento posterior.
- **Detección de valles:** la luz utilizada por el colonoscopio provoca la aparición de una “sombra” que rodea al pólipo. Para detectar dicha sombra se ha utilizado un detector de valles (basado en la segunda derivada anisotrópica gaussiana). Dado que la detección de valles provoca muchos falsos positivos, hemos obtenido una región de interés con el objetivo de reducir el área de búsqueda del pólipo.
- **Segmentación y unión de regiones:** en esta última fase, se ha aplicado un algoritmo de segmentación (watershed) sobre la imagen de valles. La imagen segmentada contendrá un número muy elevado de zonas. Por ello, hemos ido uniendo pequeñas zonas hasta conseguir una única zona que contenga el pólipo.

De manera experimental hemos visto en los fotogramas 1, 2 y 3 que el procedimiento desarrollado da buenos resultados para la mayoría de las imágenes. Sin embargo, en los fotogramas 4 y 5 no conseguimos segmentar el pólipo.

Para mitigar los resultados desfavorables en los fotogramas anteriores se propone las siguientes mejoras:

- En ambos fotogramas (4 y 5), los reflejos aparecen movidos provocando una “sombra” de intensidad elevada. El preprocesamiento propuesto en este trabajo no es suficiente para corregir tal evento. Para evitarlo, se propone preprocesar la imagen para eliminar los efectos de imágenes en movimiento.
- En la mayoría de imágenes, se detectan un número elevado de valles debido, en parte, a la presencia de vasos sanguíneos en la escena. Detectarlos y eliminarlos resultaría en una imagen de valles de mayor precisión y claridad. Además, encontramos píxeles de elevada intensidad aislados que provocan gran respuesta en la imagen de valles. Estos píxeles impiden la unión de algunas regiones por lo que eliminarlos mejoraría la precisión en los resultados.
- Recordamos que para la segmentación final (sección 5.3), adaptamos para cada fotograma la intensidad que define el límite entre un valle débil o fuerte. Para poder automatizar todo el proceso sería necesario de un método adaptativo para el mencionado nivel de intensidad.

REFERENCIAS

- [1] National Cancer Institute, SEER program: <https://seer.cancer.gov/>
- [2] CVC-ColonDB: <http://mv.cvc.uab.es/projects/colon-qa/cvccolondb>
- [3] L. Vincent and P. Soille. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE transactions on pattern analysis and machine intelligence*, 13(6):583-598, 1991.
- [4] Jorge Bernal del Nozal. Polyp Localization and Segmentation in Colonoscopy Images by Means of a Model of Appearance for Polyps.
- [5] Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins. Digital Image Processing Using MATLAB, 2nd Ed.
- [6] Rafael C. Gonzalez, Richard E. Woods. Digital Image Processing, 2nd Ed.

