



Desarrollo de un método para la detección y tratamiento de artefactos presentes en imágenes endoscópicas por medio de procesamiento digital de imágenes.

Daniela Díaz Palacio

Johana Nataly Tinoco Gómez

Universidad El Bosque
Facultad de ingeniería, Programa de Bioingeniería
Bogota D.C., Colombia
2019

Desarrollo de un método para la detección y tratamiento de artefactos presentes en imágenes endoscópicas por medio de procesamiento digital de imágenes.

Daniela Díaz Palacio

Johana Nataly Tinoco Gómez

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:
Bioingeniero

Director(a):
Jhonathan Tarquino González

Bogotá D.C., Colombia
2019

Contenido

Lista de figuras	vii
Lista de tablas	xI
Lista de abreviaturas	xIII
Resumen	1
Introducción	2
1 Planteamiento del problema	4
2 Justificación	6
3 Objetivos	7
4 Marco referencial	8
4.1 Estado del arte	8
4.2 Marco teórico	11
4.2.1 Sistema digestivo	11
4.2.2 Características visuales de endoscopia asociadas a padecimientos de sistema gastrointestinal	17
4.2.3 Artefactos	20
4.2.4 Imágenes digitales	23
4.2.5 Visión por computadora	24
4.2.6 Características	27
4.2.7 Métodos de identificación	35
4.2.8 Métodos de corrección	37
4.2.9 Métricas de evaluación	40
4.3 Marco normativo	42
5 Metodología y Resultados	43
5.1 Base de datos	43
5.2 Pre-procesamiento	47
5.2.1 Normalización	47

5.2.2	Ecualización	56
5.2.3	Filtrado	59
5.2.4	Espacio de color	62
5.3	Región de Interés (ROI)	65
5.4	Segmentación y extracción de candidatos a artefacto	69
5.4.1	Umbralización	69
5.4.2	Separación o extracción de candidatos	75
5.5	Procesamiento y extracción de características	76
5.5.1	Balanceo de clases:	77
5.5.2	Evaluación de características:	78
5.6	Método de identificación	87
5.7	Método de corrección de RE	99
5.8	Interfaz de presentación de resultados	105
6	Resultados finales y discusión	110
7	Conclusiones y recomendaciones	115
8	Anexos	116
8.1	Diagrama de flujo de la normalización	116
8.2	Diagrama de flujo de la ecualización	117
8.3	Diagrama de flujo de la evaluación de las fases del preprocesamiento	118
8.4	Diagrama de flujo sobre la región de interés	119
8.5	Diagrama de flujo sobre la segmentación y extracción de posibles artefactos .	120
8.6	Diagrama de flujo de la extracción de características	121
8.7	Diagrama de flujo de la evaluación de separabilidad de características	122
8.8	Diagrama de flujo del método de identificación	123
8.9	Diagrama de flujo del método de corrección	124
8.10	Diagrama de flujo de la interfaz	125
Bibliografía		126

Lista de Figuras

4-1. Organización del sistema digestivo [Tortora and Derrickson, 2018]	12
4-2. Anatomía de la cavidad bucal [Tortora and Derrickson, 2018]	12
4-3. Subdivisiones de la faringe [Latarjet et al., 2019]	13
4-4. Fase esofágica de la deglución [Latarjet et al., 2019]	13
4-5. Vista anterior de las regiones del estómago [Tortora and Derrickson, 2018] . .	14
4-6. Vista anterior de la anatomía externa del intestino delgado [Latarjet et al., 2019]	14
4-7. Vista anterior de las principales regiones del intestino grueso [Tortora and Derrickson, 2018]	15
4-8. Imágenes endoscópicas de gastritis [Vakil, 2019a]	17
4-9. Úlceras presentes en imágenes endoscópicas [de vídeo endoscopia Gastrointestinal de El Salvador,], [Vakil, 2019a]	18
4-10. Pólipos presentes en imágenes endoscópicas [Murra Saca, 2018]	18
4-11. Acalasia presente en imágenes endoscópicas [Gaudó Perales and Revnic,] . .	19
4-12. Imágenes endoscópicas con presencia de cáncer gástrico, (a) precoz y (b) avanzado [Grávalos and González Flores, 2017]	19
4-13. Imágenes endoscópicas con presencia de enfermedad de Crohn [Vera, 2017] . .	20
4-14. Diferentes artefactos [Atasoy et al., 2010]	21
4-15. Imagen de endoscopia con la presencia de varios artefactos, en el que se destaca la saturación [Ali and Zhou, 2019]	21
4-16. Imágenes endoscópicas del (a) Colon y (b) Esófago con reflejos especulares [Stehle, 2006]	22
4-17. Imágenes endoscópicas con presencia de burbujas [Yoo et al., 2016]	22
4-18. Imágenes endoscópicas con artefacto de contraste [Ali and Zhou, 2019]	22
4-19. Imágenes colonoscópicas con artefacto de movimiento [Kaçmaz et al., 2018] . .	23
4-20. Imagen endoscópica con presencia de un instrumento [Ali and Zhou, 2019] . .	23
4-21. Representación de una imagen binaria y los valores de sus píxeles a través de una matriz de 6x6 [Marques, 2011]	24
4-22. Etapas de un sistema de visión por computadora	24
4-23. Realce de contraste de una imagen mediante normalización [Ortiz et al., 2017]	25
4-24. Imagen original con sus versiones ecualizadas [of Edinburgh, 2010]	25
4-25. Segmentación de una imagen [Li et al., 2018]	26
4-26. Extracción de características a partir del área [Molina Cabello et al., 2018] . .	26

4-27. Resultado de un modelo de clasificación binaria o multiclase [Mishra, 2018]	27
4-28. Filtro Gaussiano, A) operación de convolución [Kabbai et al., 2016], B) ejemplo de mascara 5x5 para un filtro gaussiano [R. Fisher and Wolfart., 2003]	28
4-29. Histéresis de umbral [OpenCV, 2015]	29
4-30. Ejemplo cálculo matriz de co-ocurrencia de nivel de gris [Löfstedt et al., 2019]	30
4-31. Histograma típico del canal de luminancia de una imagen endoscópica [Stehle, 2006]	33
4-32. Método de agrupamiento K-means [Iris et al.,]	35
4-33. Representación de un hiperplano (línea azul) [Betancourt, 2005]	36
4-34. Proceso de restauración de una imagen [Dey, 2018]	37
4-35. Proceso de restauración de una imagen con RE [Stehle, 2006]	38
4-36.(a) Imagen con una zona afectada . (b) Bloque de la imagen afectada DIB , Bloque de la imagen IB y Bloque del píxel PB [Stehle, 2006]	39
4-37. Dice similarity coefficient (DSC) [Zou et al., 2004]	40
5-1. Fases de desarrollo del proyecto de grado	43
5-2. Número de Bounding box por artefacto	45
5-3. Anotaciones en forma de bbox, rojo (RE) y azul (DM)	46
5-4. Segmentaciones entregadas por la base de datos	47
5-5. Normalizaciones aplicadas sobre las imágenes endoscópicas	47
5-6. Normalización de Min-Max	48
5-7. Normalización RGB	49
5-8. Normalización de contraste local	50
5-9. Normalización de iluminación	51
5-10. Estandarización	52
5-11. Normalización Logarítmica	52
5-12. Diferencia entre el fondo y la región con artefactos	54
5-13. Imágenes ecualizadas	56
5-14. Ecualización global	57
5-15. Ecualización Adaptativa	57
5-16. Estiramiento del contraste	58
5-17. Imágenes filtradas	60
5-18. A) Imagen con bbox de RE, B) Imagen con filtro de media con máscara de 20x20	61
5-19. Distintos espacios de color con sus componentes para RE	62
5-20. Distintos espacios de color con sus componentes para DM	63
5-21. Media de medidas de correlación y distancia en diferentes espacios de color	64
5-22. Desviación estándar de medidas de correlación y distancia en diferentes espacios de color	64
5-23. Diferentes regiones de interés	65

5-24. Primera técnica para la detección de la región de interés	66
5-25. Segunda técnica para la detección de la región de interés	67
5-26. Tercera técnica para la detección de la región de interés	68
5-27. Diferentes técnicas de umbralización para RE	69
5-28. Metodología técnica ₁	70
5-29. Metodología técnica ₂	72
5-30. Metodología técnica ₃	73
5-31. Técnica ₃ con segmentaciones de saturación (bbox verde)	75
5-32. Representación de como se separan y extraen los candidatos a RE	75
5-33. Imágenes excluidas de la base de datos	76
5-34. A) Imagen con anotación de RE (bbox rojo), B) Imagen sin anotación de RE	77
5-35. Diagramas de dispersión entre características (ventana de 300x300) de DM .	79
5-36. Boxplot, características no escogidas	79
5-37. Boxplot, con evidencia de separabilidad entre la clase 0 del resto	80
5-38. Diagramas de dispersión entre características de RE	84
5-39. Boxplot, con características no separables.	85
5-40. Boxplot, con evidencia de separabilidad entre la clase 1 del resto.	85
5-41. Métodos de identificación binario (A. para identificación de DM y B. para identificación de RE)	88
5-42. ROC del clasificador elegido para identificación de DM	94
5-43. ROC del clasificador elegido para identificación de RE	95
5-44. Método para simular RE	100
5-45. Técnica de suavizado	101
5-46. Corrección de RE mediate <i>Inpainting 1</i>	102
5-47. Medidas de evaluación	105
5-48. Interfaz de presentación de resultados	106
5-49. Interfaz cuando se detecta RE	107
5-50. Interfaz cuando se detecta DM	108
5-51. Interfaz cuando se detecta RE y DM	108
5-52. Interfaz cuando no se detecta RE ni DM	109
6-1. Método de ubicación e identificación del tipo de artefacto (DM y/o RE) . . .	111
6-2. Algoritmo para la corrección o eliminación de imágenes endoscópicas . . .	113
8-1. Diagrama de flujo de la normalización	116
8-2. Diagrama de flujo de ecualización	117
8-3. Diagrama de flujo de la evaluación del preprocesamiento	118
8-4. Diagrama de flujo de la región de interés (ROI)	119
8-5. Diagrama de flujo de la segmentación y extracción de candidatos	120
8-6. Diagrama de flujo de la extracción de características	121
8-7. Diagrama de flujo de la evaluación de separabilidad de características . . .	122

8-8. Diagrama de flujo método de identificación	123
8-9. Diagrama de flujo del método de corrección	124
8-10. Diagrama de flujo de la GUI	125

Lista de Tablas

5-1. Características de los datos de entrenamiento de la base de datos EAD 2019	45
5-2. Cantidad de imágenes endoscópicas con RE-DM	46
5-3. Rangos de intensidad	53
5-4. Medidas de evaluación normalización	55
5-5. Niveles de contraste	59
5-6. Medidas de evaluación ecualización	59
5-7. Medidas de evaluación filtrado	61
5-8. Medidas de similitud ROI ₁	66
5-9. Medidas de similitud ROI ₂	67
5-10. Medidas de similitud ROI ₃	68
5-11. Medidas de similitud con diferentes técnicas de detección de ROI _s	69
5-12. Medidas de similitud de la técnica ₁ con diferentes kernels	71
5-13. Medidas de similitud de la técnica ₂ con diferentes kernels	72
5-14. Medidas de similitud de la técnica ₃ con diferentes kernels	74
5-15. Medidas de similitud técnicas de umbralización	74
5-16. Lista de características para detectar DM y RE	76
5-17. Cantidades de bbox para detección de DM por clase (0:DM, 1:RE, 2:Otro) después de la depuración de la base de datos	77
5-18. Cantidades de bbox para detección de RE por clase (0:DM, 1:RE, 2:Otro) después de la depuración de la base de datos	77
5-19. Cantidad de bbox para detección de DM por clase (0:DM, 1:RE, 2:Otro) después del balanceo de clases	78
5-20. Cantidad de bbox para detección de RE por clase (0:DM, 1:RE, 2:Otro) des- pués del balanceo de clases	78
5-21. Percentiles 25 y 75 de los boxplots visualmente separables	80
5-22. Resultados de la prueba T-Student	83
5-23. Resultados de la prueba de suma de rango de Wilcoxon	83
5-24. Percentiles 25 y 75 de los boxplots visualmente separables para identificación de RE	85
5-25. Resultados de la prueba T-Student para identificación de RE (clase 0: DM, clase 1: RE y clase 2: Ninguno)	87
5-26. Resultados de la prueba de suma de rango Wilcoxon para identificación de RE	87

5-27. Cantidad de características para identificación de DM por clase (0: DM, 1: RE +Otro) y para identificación de RE por clase (0: DM + Otro, 1: RE)	88
5-28. Resultados métricas de desempeño métodos de identificación de DM y RE con partición hold out	91
5-29. Resultados métricas de desempeño métodos de identificación de DM con partición k-fold	92
5-30. Resultados métricas de desempeño métodos de identificación de RE con partición k-fold	92
5-31. Resultados métricas de desempeño métodos de identificación de DM y RE con partición leave one out	93
5-32. Resultados métricas de desempeño para los métodos de identificación de DM .	94
5-33. Comparación de los resultados de las métricas de desempeño para los métodos de identificación elegido de DM contra resultados expuestos en la literatura .	94
5-34. Resultados métricas de desempeño métodos de identificación de RE	95
5-35. Comparación de los resultados de las métricas de desempeño para los métodos de identificación elegido de RE contra resultados expuestos en la literatura .	96
5-36. Correlación entre características de DM	97
5-37. Resultados métricas de desempeño en imágenes de entrenamiento con la unión de características DM	97
5-38. Resultados métricas de desempeño con imágenes de prueba DM	98
5-39. Correlación entre características de RE	98
5-40. Resultados métricas de desempeño en imágenes de entrenamiento con la unión de características RE	99
5-41. Resultados métricas de desempeño con imágenes de prueba RE	99
5-42. Resultados métricas de similitud método corrección por suavizado	102
5-43. Resultados métricas de similitud método corrección por Inpaiting_1	102
5-44. Aplicación de la ecuación de Navier-Stokes en el método de Inpainting	103
5-45. Resultados métricas de similitud método corrección por Inpaiting_2	103
5-46. Resultados métricas de similitud método corrección por Inpaiting_3	104
6-1. Resultados de la evaluación de separabilidad de las características seleccionadas	110
6-2. Resultados unión de clasificadores binarios	112
6-3. Resultados por fase metodológica del método	114

Lista de abreviaturas

Abreviatura Término

<i>AUC</i>	Área bajo la curva
<i>bbox</i>	Bounding boxes
<i>BD</i>	Base de datos
<i>CGAN</i>	Redes generativas antagónicas
<i>CLAHE</i>	Ecualización adaptativa del histograma limitada por contraste
<i>DAO</i>	Diagnóstico asistido por computador
<i>DFT</i>	Transformada discreta de Fourier
<i>DM</i>	Desenfoque por movimiento
<i>DSC</i>	Coeficiente de similitud de datos
<i>DWT</i>	Transformada Wavelet discreta
<i>EAD2019</i>	Endoscopic Artefact detection Challenge 2019
<i>FDC</i>	Función de distribución acumulativa
<i>FMM</i>	Método de marcha rápida
<i>GI</i>	Gastrointestinal
<i>GLCM</i>	Matriz de co-ocurrencia de nivel de gris
<i>GT</i>	Ground truth
<i>GUI</i>	Interfaz gráfica de usuario
<i>ID</i>	Intestino delgado
<i>IG</i>	Intestino grueso
<i>IPR</i>	Porcentaje de píxeles aislados
<i>IP</i>	Píxeles aislados
<i>KLT</i>	Kanade Lucas Tomasi
<i>LCN</i>	Normalización de contraste local
<i>mDM</i>	Metodología para DM
<i>mRE</i>	Metodología para RE
<i>MSE</i>	Error cuadrático medio
<i>PSNR</i>	Relación señal a ruido de pico
<i>RE</i>	Reflejos especulares
<i>RF</i>	Random forest
<i>RN</i>	Red Neuronal
<i>ROC</i>	Característica Operativa del Receptor

Abreviatura Término

<i>ROI</i>	Región de interés
<i>SSIM</i>	Índice de similitud estructural
<i>SVM</i>	Máquina de soporte vectorial

Resumen

Las enfermedades digestivas están dentro de las primeras veinte causas de muerte en hombres y mujeres en Colombia [Rodríguez García et al., 2017], sin embargo, es importante mencionar que muchos de estos padecimientos hoy en día pueden ser diagnosticados a tiempo mediante el uso de la endoscopia [Triana et al., 2017], [American Cancer Society, 2015]. Para lograr lo anterior, el médico especialista debe analizar las imágenes endoscópicas, tarea con alto grado de complejidad, no solo por el volumen de información sino también porque hay diferentes factores que afectan la calidad de las imágenes [Arnold et al., 2011].

Es por esto que se han venido desarrollando herramientas de apoyo al diagnóstico hecho por el especialista a través de métodos de procesamiento de imágenes endoscópicas, no obstante, dichas imágenes presentan diferentes artefactos que comprometen el desempeño de las herramientas [Moccia et al., 2018]; razón por la que este trabajo presenta el desarrollo de un método de identificación, corrección o eliminación de artefactos (reflejos especulares y desenfoque por movimiento).

Lo anterior se logró aplicando dos metodologías diferentes para la identificación de RE y DM mediante el análisis y procesamiento de imágenes endoscópicas, lo cual permitió extraer características específicas (**SODMS, Entropía, brillo, contraste y desviación de intensidades**) para ser introducidas en dos clasificadores binarios complementarios (Random Forest), donde se obtuvo una precisión del 96 % para la identificación de RE y de 76 % para DM, datos comparables con lo encontrado en el estado del arte. Dependiendo del tipo de artefacto que se identifique, el método concluye con la corrección de RE mediante la técnica de inpainting o la eliminación de la imagen desenfocada. Finalmente, el método completo de tratamiento de artefactos presenta una precisión de 82 % con un MSE de ubicación de artefactos de 4517.

Introducción

La endoscopia es un procedimiento clínico mínimamente invasivo para “mirar dentro” del cuerpo humano, en este se introduce un dispositivo (endoscopio) para visualizar el interior de un órgano hueco o una cavidad corporal, el cual cuenta con una fuente de luz y una cámara que transmite las imágenes a una pantalla de computador [Society, 2019], [Münzer et al., 2018]. A partir de la información visual contenida en dichas imágenes se pueden realizar diagnósticos de diferentes enfermedades asociadas al sistema gastrointestinal (GI) [American Cancer Society, 2015]. Para lograr lo anterior, el médico especialista comúnmente debe analizar la secuencia de imágenes que son adquiridas durante los videos de endoscopia, tarea que no es para nada trivial, pues cada cuadro de la secuencia de imágenes generada contiene un fragmento de la superficie recorrida durante el examen, es decir, que la cantidad de imágenes es considerablemente alta y sumado a esto, hay diferentes factores que afectan la calidad de las mismas [Arnold et al., 2011].

Por lo anterior, se ha venido investigando sobre herramientas computacionales que apoyen el diagnóstico hecho por el especialista a través de métodos de procesamiento de imágenes endoscópicas, un ejemplo de ellas son los algoritmos de reconstrucción 3D, de diagnóstico asistido por computador y de mosaico [Moccia et al., 2018], [Kaçmaz et al., 2018]. No obstante, el desempeño de estas herramientas se ve comprometido por la presencia de artefactos, dado que estos se convierten en una fuente directa de error en aspectos como la segmentación, el reconocimiento de objetos, la reconstrucción de formas a partir de sombreados, entre otros [Gao et al., 2017].

Según la literatura los errores expuestos anteriormente pueden disminuirse si las imágenes con artefactos son corregidas o eliminadas antes de entrar a la herramientas de apoyo [Moccia et al., 2018], [Kaçmaz et al., 2018]. Por esta razón, se decidió desarrollar un método automático que permita identificar dos de los tipos artefactos más comunes y hacer su posible corrección o eliminación.

Por lo tanto, la estructura de este documento se encuentra dividida de la siguiente forma: en el cuarto capítulo se expondrá un marco referencial compuesto por estado del arte y marco teórico, el primero expondrá lo que se ha realizado hasta al momento respecto a la identificación y corrección de artefactos en imágenes endoscopias, el segundo por otro lado, explicará los elementos técnicos y teóricos usados en el documento para facilitar su entendimiento.

miento y se dan ejemplos de su uso y funcionamiento.

En el quinto capítulo se encuentra la metodología usada para el desarrollo del presente trabajo de grado, en el que se evidencian los resultados y discusión de la misma, por otro lado, en el sexto capítulo se muestran los resultados finales y su discusión. Por último, se expone el séptimo capítulo en el que se encuentra las conclusiones y recomendaciones del proyecto.

1 Planteamiento del problema

Según el estudio “Estimación de la carga de enfermedad global para Colombia” publicado por la revista de salud pública, las enfermedades digestivas están dentro de las primeras veinte causas de muerte en hombres y mujeres, generando 10.938 en el año 2010 y 11123 en el 2012 [Rodríguez García et al., 2017]. Sin embargo, es importante mencionar que el diagnóstico temprano de algunos de estos padecimientos asociados al sistema gastrointestinal (GI) como cáncer gástrico y de estómago puede aumentar la tasa de supervivencia en un 70 %, según evidencias de algunos estudios a nivel internacional [Triana et al., 2017]. Este proceso de diagnóstico puede realizarse mediante pruebas de laboratorio como análisis de sangre y/o heces, o por imágenes de endoscopía, las cuales se caracterizan por contener información visual altamente relacionada con las patologías asociadas al sistema GI [American Cancer Society, 2015], [U.S. National Library of Medicine, 2018].

Cabe mencionar que para realizar dicho diagnóstico se deben analizar una gran cantidad de imágenes que comprenden el vídeo de endoscopia, por ejemplo, en una colonoscopia de 10 minutos tomada a 25 cuadros por segundo se obtienen 15000 cuadros [Arnold et al., 2009]. En efecto, esta tarea resulta difícil por el número de cuadros a procesar y es aún más compleja debido a que dichas imágenes son susceptibles a presentar diferentes tipos de artefactos, como la saturación de píxeles (brillos en zonas), reflejos especulares (iluminaciones provenientes del reflejo de la luz de la sonda), burbujas, residuos y desenfoque por movimiento [Arnold et al., 2011]; estos pueden comprometer entre 30 al 40 % de los cuadros del vídeo [Armin et al., 2016], por ejemplo, el desenfoque por movimiento puede estar presente en aproximadamente el 25 % de los cuadros [Münzer et al., 2018]. Dichos artefactos afectan características de textura del tejido y de manera general la imagen, dificultando el proceso de interpretación y diagnóstico que se hace a partir de ellas [Kim et al., 2016], [Kaçmaz et al., 2018].

Es por esto que la problemática a abordar es la presencia de diferentes artefactos en las imágenes endoscópicas, dado que afectan el desempeño de las herramientas automáticas de apoyo al diagnóstico, convirtiéndose en una fuente directa de error en aspectos como la segmentación, el reconocimiento de objetos, la detección de regiones de interés (ROI), la reconstrucción de formas, etc [Gao et al., 2017]; lo cual limita la calidad de la información de entrenamiento y representa un mayor procesamiento (diez veces más) con un menor rendimiento para las herramientas existentes como lo son los algoritmos de diagnóstico asistido

por computador (DAO) o los de reconstrucción 3D [Moccia et al., 2018]. Sumado a esto, es importante mencionar que los artefactos dificultan la visualización del tejido subyacente generando el descarte de imágenes que pueden tener información relevante para el desarrollo de herramientas de apoyo, como los métodos de mosaico que requieren al menos un 60 % de los cuadros sucesivos de un vídeo de endoscopia para que sean exitosos [Kaçmaz et al., 2018], [Ali and Zhou, 2019], [Ali et al., 2019].

2 Justificación

El cáncer colorrectal y de estómago encabezan las listas de carcinomas más frecuentes en Colombia [López et al., 2017]. Donde se registraron 15744 casos en el 2018 y 4118 en el 2017, respectivamente. Razón por la que la endoscopia cobra gran sentido, pues es uno de los exámenes diagnósticos y terapéuticos más comunes en Colombia para este tipo de padecimientos, esto se debe principalmente a que permite la detección temprana de estas y otras patologías asociadas al aparato digestivo, sin causar un daño o ionizar el paciente [Ortiz et al., 2017]. Teniendo en cuenta lo anterior, es de gran importancia realizar aportes, desde diferentes áreas del conocimiento, que apoyen al análisis del registro de imágenes endoscópicas.

Uno de los aportes en los que actualmente se ha venido investigando es sobre herramientas computacionales, las cuales apoyan el diagnóstico a través de métodos de procesamiento de imágenes de endoscopia; sin embargo, el desarrollo de herramientas efectivas y confiables se ve afectado por la presencia de artefactos, dado que estos desencadenan posibles errores en segmentación y clasificación [Saint-Pierre et al., 2011], [Liedlgruber and Uhl, 2011], [Moccia et al., 2018], [Ali and Zhou, 2019]. Por lo anterior, resulta oportuno proveer un método automático que sea un insumo para dichas herramientas, con el que se logre limitar el efecto negativo de la presencia de los artefactos más comunes en imágenes de endoscopia (reflejos especulares y desenfoque por movimiento) [JungHwan Oh, 2004], [Münzer et al., 2018], mediante su detección y posible corrección o eliminación completa del cuadro.

Cabe resaltar que para llevar a cabo lo anterior, se hace necesaria la asociación de conocimientos en procesamiento de imágenes junto con saberes propios de la fisiología humana que permitan relacionar variables biológicas a los patrones encontrados en las imágenes. Dichos requerimientos conceptuales evidencian la pertinencia de que la solución sea desarrollada por un Bioingeniero, pues este profesional se forma para usar los principios de la biología y las herramientas de la ingeniería en el diseño y desarrollo de productos que mejoren la calidad de vida de los individuos (Universidad El Bosque, 2018). Dicho lo anterior, el producto a desarrollar se enmarca dentro de uno de los cuatro ejes temáticos del programa, el cual es tecnología para la salud de entes biológicos.

3 Objetivos

Dado lo anterior, se plantea desarrollar un método automático para la detección de dos artefactos en imágenes endoscópicas y su posible corrección o eliminación completa del cuadro por medio del procesamiento digital de imágenes como objetivo general y los siguientes como objetivos específicos:

- Evaluar la capacidad de separación de características encontradas a partir del procesamiento de imágenes de endoscopía para la identificación de artefactos presentes en la base de datos Endoscopic Artifact Detection Challenge (EAD) 2019.
- Implementar un método de ubicación e identificación del tipo de artefacto, basado en las características encontradas en las imágenes de la base de datos EAD 2019.
- Desarrollar un algoritmo para la corrección o eliminación de imágenes con presencia de uno de los artefactos identificados a través de la metodología de clasificación propuesta.
- Evaluar el rendimiento del método propuesto de identificación y corrección de artefactos en las imágenes de endoscopia de la base de datos EAD 2019.

4 Marco referencial

4.1. Estado del arte

La endoscopia es un procedimiento mínimamente invasivo, ampliamente utilizado actualmente en el campo médico debido a que permite diagnosticar y tratar enfermedades asociadas al tracto digestivo [Liedlgruber and Uhl, 2011]. Sin embargo, el análisis de las imágenes adquiridas mediante este examen se ve afectado por la presencia de artefactos, dado que estos generan dificultades en la visualización del tejido subyacente durante el procedimiento, afectando el diagnóstico hecho por el especialista o provocando errores en las herramientas DAO basadas en análisis de imágenes [Arnold et al., 2011].

Es por esto que desde hace más de una década se han venido reportando diferentes métodos en la literatura para la detección, restauración y/o eliminación de artefactos en imágenes de endoscopia. Los primeros trabajos intentaban atacar de manera individual los artefactos en endoscopia, algunos (2004) con el desarrollo de un algoritmo para detectar y eliminar fotogramas desenfocados dentro del vídeo [JungHwan Oh, 2004]. En este trabajo se propuso una medida basada en la información de los bordes detectados por el algoritmo de Canny; a partir de esta información se calculó el porcentaje de píxeles aislados (IPR - Isolated Pixel Ratio) de cada cuadro, el cual indica la relación entre los píxeles aislados (IP - Isolated Pixel) con el número total de píxeles de borde obtenidos en el cuadro. Este índice permite clasificar las imágenes en: claras (cuando el IPR es menor al umbral más bajo), desenfocadas (cuando el IPR es mayor al umbral más alto) y ambiguas (cuando el IPR está entre los dos umbrales); finalmente, este método demostró una precisión superior al 90 % en la detección de imágenes claras, en consecuencia, se logró eliminar las imágenes desenfocadas y ambiguas, manteniendo una representación visual adecuada de todo el vídeo de endoscopia. Cabe aclarar que este trabajo se basa únicamente en el desenfoque de los frames para la eliminación de los mismos [JungHwan Oh, 2004].

Atacando otros tipos de artefactos, a mediados del año 2006 se propuso un método para remover y reemplazar reflejos especulares, donde se detectó el artefacto mediante la segmentación por umbral en el espacio de color YUV y se reemplazó por píxeles corregidos mediante un algoritmo de deconvolución espectral que aprovecha los picos de frecuencia en el histograma; sin embargo, este método presentó problemas de segmentación aumentando el número de falsos positivos [Stehle, 2006]. En un estudio posterior, se demostró que la de-

tección de reflejos especulares aumentaba la precisión al clasificar marcos informativos de los no informativos (marcos desenfocados); lo anterior se evidenció tras comparar el desempeño de dos métodos de clasificación según las imágenes de entrada (que contenían o no reflejos especulares), donde el primero, clasificación basada en bordes, pasó de tener una precisión del 89 % al 94 % al eliminar dicho artefacto; asimismo, el desempeño del clasificador basado en técnicas de agrupamiento (k-means) y transformada discreta de Fourier (DFT) aumentó alrededor de un 3 %, obteniendo como resultado un porcentaje de precisión del 96 %. Cabe resaltar que para la detección, los reflejos especulares se dividieron en dos tipos de áreas, una de brillo absoluto (zona con píxeles absolutamente brillantes) para la que se empleó la técnica de umbrales múltiples y otra de brillo relativo (zona con píxeles relativamente brillantes) en la que se utilizó la detección de valores atípicos [Stehle, 2006], [Oh et al., 2007].

Un avance en la inclusión de nuevas fuentes de ruido en las imágenes, como lo son los generados cuando la cámara está cubierta de líquido, está cerca de la mucosa o la está tocando directamente, se evidenció en el año 2009; en el cual se propuso un sistema automático de clasificación bayesiano basado en las características de la transformada Wavelet discreta (DWT), para distinguir los marcos informativos de los que presentaban las fuentes de ruido anteriormente nombradas (marcos no informativos). Además, compararon el rendimiento de este método con otro que fue usado en estudios anteriores, más específicamente un sistema de clasificación por agrupamiento (k-means) basado en características de la transformada discreta de Fourier (DFT). Este estudio demostró que el área bajo la curva ROC (AUROC) de las características de la DWT son significativamente más alta con un nivel de confianza del 95 % que el de las características de DFT, además la extracción de las características de DWT se realizó a 62.2 marcos por segundo, lo que es 9.5 veces más rápido que la extracción de características de DFT, que se realizó a 6.5 frames por segundo. Sin embargo, este estudio solo propone el descarte de marcos no informativos y no plantea ninguna estrategia, por ejemplo el inpainting, para su corrección [Oh et al., 2007], [Arnold et al., 2009].

Con ánimos de mejorar la falencia de trabajos anteriores, en los que no se generaban estrategias para la corrección de artefactos, se presentó un estudio en el año 2011 con el objetivo de extraer y restaurar automáticamente reflejos especulares presentes en imágenes colposcópicas. La detección de dichas regiones se logró en el espacio de color CIE-XYZ, en el que se comparó la componente de luminancia y (crominancia) del espacio CIE-xyY normalizado y la luminancia Y del espacio CIE-XYZ. Posteriormente, los píxeles de las regiones encontradas fueron reemplazados por el promedio de bloques (particiones de la imagen inicial) con un tamaño determinado o con píxeles adyacentes, a esto se le conoce como la técnica de inpainting. Finalmente, se aplica una ecualización de histograma a cada bloque restaurado en el paso anterior. Este proceso permite homogeneizar regiones de bloques restaurados con regiones vecinas y reducir el efecto causado por la partición de la imagen en varios bloques [Arnold et al., 2009], [Meslouhi et al., 2011].

No obstante, en el año 2014 los métodos basados en inpainting no fueron recomendados para la reducción de reflejos especulares en tiempo real, pues estos tienen una alta demanda computacional, en consecuencia su funcionamiento durante el examen de endoscopia es bastante lento, además también se mencionó que pueden ser engañosos para el especialista y provocar un error quirúrgico. Es por esto, que se propuso un método automático de dos umbrales de resta de intensidad para disminuir el deslumbramiento (imagen compuesta por color real y reflexiones especulares); este se caracterizó por mantener adecuado el contraste de la imagen, dado que se restó el mismo valor RGB a cada píxel de las regiones definidas con deslumbramiento. Sumado a esto, se propuso también una técnica de eliminación de artefactos presentes en los límites de las imágenes, para el cual realizaron una máscara binaria, con la que convertían a cero (negro) todos los píxeles que se encontraran fuera de ella. Por último, determinaron que el método propuesto no es computacionalmente exigente, por lo que se puede tener en cuenta en estudios futuros y probarlo a tiempo real [Meslouhi et al., 2011], [Abel et al., 2014].

Hasta este año toda investigación relacionada con la detección, clasificación y/o restauración de imágenes de endoscopia con presencia de artefactos, se había limitado en su mayoría a la detección y reemplazo de reflejos especulares o clasificación de marcos desenfocados. Sin embargo en el 2015, un estudio propuso un método para distinguir entre marcos informativos de los no informativos, estos últimos incluyen imágenes con presencia de 4 tipos de artefactos, los cuales son: desenfoque por movimiento, chorro de agua, contacto directo de la cámara con la mucosa (indistintos) y burbujas. Dicho método empleó un clasificador Random forest basado en el cálculo de tres clases de características, estas son: las extraídas por el método Kanade Lucas Tomasi (KLT), la desviación estándar y la media del espacio de color HSV, así como, la característica de porcentaje de píxeles de bordes. Cabe mencionar, que se comparó los resultados de las características anteriormente nombradas con las extraídas a partir de la matriz de coexistencia de nivel de gris (GLCM) y el IPR, donde las primeras obtuvieron 94 % de exactitud con 86 % de precisión y las otras un 92 % con un 79 %, respectivamente [Abel et al., 2014], [Armin et al., 2016].

La clasificación de marcos (informativos y no informativos) en los videos de endoscopia sigue siendo un tema de continua investigación, un ejemplo de ello se evidencia en la estrategia presentada en el año 2018 para seleccionar automáticamente los marcos informativos dentro del video, teniendo en cuenta que para el presente estudio los marcos no informativos eran marcos: subexpuestos, con saliva o reflejos especulares y/o desenfocados. Para lograr dicha clasificación, primero se extrae el conjunto de características de los marcos, las cuales estaban basadas en el coeficiente de luminancia, distribución frecuencial (varianza de Laplace en la imagen), magnitud del gradiente de la imagen, entropía, intensidad, histograma y número de puntos clave en la escala de gris. Posteriormente, se emplea el algoritmo de clasi-

ficación máquina de soporte vectorial (SVM), el cual usa un Kernel Gaussiano y el esquema de clasificación multiclas (4 tipos de marco) uno-vs-uno. El método propuesto tuvo una exactitud del 91 % en la clasificación de marcos informativos [Armin et al., 2016], [Moccia et al., 2018].

Todo lo anterior demuestra que la clasificación de marcos no informativos (presencia de artefactos) ha sido un tema de continua investigación, sin embargo, en ninguno de estos estudios se evidencia tanto la detección como la corrección de múltiples artefactos. No obstante, en abril de este año se publicó un estudio donde propusieron un método en tiempo real, en el cual detectaron y clasificaron 6 tipos de artefactos (desenfoque por movimiento, burbujas, reflejos especulares, instrumento y saturación de píxeles) mediante una red neuronal convolucional multiclas; por otro lado, para la corrección primero hicieron la evaluación de la calidad del frame de acuerdo a la afectación dada por la ubicación y tipo de los artefactos, si el cuadro estaba “levemente” afectado procedieron a hacer la corrección mediante redes generativas antagónicas (CGAN). Cabe resaltar que este enfoque conserva, en promedio, el 68.7 % de los cuadros de un vídeo, y obtuvieron una puntuación media (0-10) dada por dos especialistas para la corrección de desenfoque por movimiento (DM) de 7.87, para reflejos especulares de 7.7 y para saturación de píxeles de 1.5 [Ali et al., 2019].

Es por esto que este trabajo de grado busca una nueva aproximación al problema presentado; el cual consiste en la ubicación e identificación de reflejos especulares y desenfoque por movimiento que son unos de los artefactos más comunes que se presentan en las imágenes endoscópicas, así mismo su posible corrección o la eliminación completa del cuadro.

4.2. Marco teórico

4.2.1. Sistema digestivo

Los alimentos constituyen la fuente de una gran variedad de nutrientes necesarios para el correcto funcionamiento del cuerpo; sin embargo, la mayor parte de estos nutrientes se encuentran en moléculas demasiado grandes para ser usadas por la células del cuerpo. Por consiguiente, el aparato digestivo debe degradar, mediante un grupo de órganos, los alimentos hasta el tamaño de moléculas que puedan ser absorbidas y usadas por las células, entre otros, para construir nuevos tejidos corporales y reparar aquellos dañados [Tortora and Derrickson, 2018].

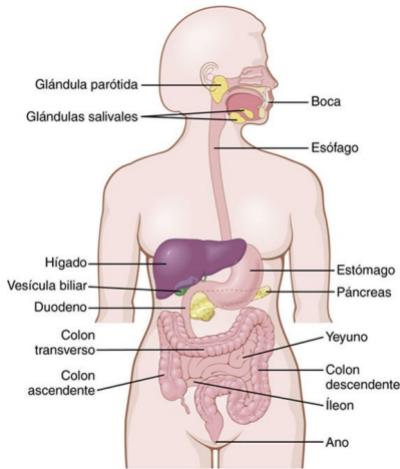


Figura 4-1: Organización del sistema digestivo [Tortora and Derrickson, 2018]

El sistema digestivo está formado por un tubo tapizado de epitelio que, como se observa en la figura 4-1, empieza en la cavidad bucal y se extiende hasta el conducto anal, comprendiendo seis segmentos [Latarjet et al., 2019]:

Boca: en la figura 4-2 se puede apreciar que la boca se compone de paladar, dientes, encías, glándulas salivales y lengua para desarrollar funciones de succión, masticación, deglución y fonación.



Figura 4-2: Anatomía de la cavidad bucal [Tortora and Derrickson, 2018]

La mucosa del paladar, mejillas, lengua y labios contiene numerosas glándulas salivales que secretan directamente en la cavidad bucal para ayudar en la digestión, reblandecimiento y deglución de los alimentos, por ejemplo, la lipasa lingual, secretada por las glándulas serosas de la lengua, se mezcla con la saliva y comienza la digestión de las grasas (Hansen, 2015). Asimismo, la lengua es un receptor de las sensaciones gustativas, esto es posible por los tres nervios que sostienen los músculos de este órgano: el nervio lingual, nasofaríngeo y laríngeo superior. Estos nervios transmiten las sensaciones de contacto, temperatura, dolor, presión, así como las imágenes gustativas, que permiten apreciar la calidad y sabor de las sustancias sólidas y líquidas ingeridas [Latarjet et al., 2019].

Posteriormente se encuentra la faringe, que es un canal compuesto por músculo esquelético y revestido por una mucosa. Está dispuesto verticalmente desde las coanas (orificios internos de la cavidad nasal) hasta el esófago y la laringe en dirección anterior, pues este conducto está compartido por las vías respiratorias y digestivas. En la figura 4-3 se puede observar que la faringe se divide en tres partes: nasofaringe, la cual funciona solo en la respiración y tanto la orofaringe como la laringofaringe tienen funciones digestivas y respiratorias a la vez. Estas dos estructuras toman los alimentos deglutidos y los impulsan al esófago y posteriormente al estómago mediante contracciones musculares [Tortora and Derrickson, 2018].

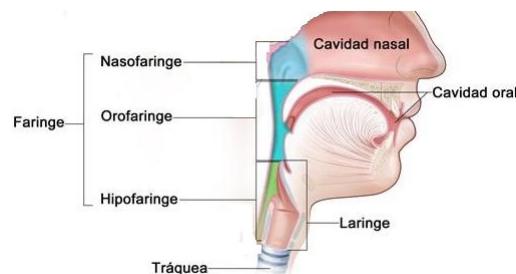


Figura 4-3: Subdivisiones de la faringe [Latarjet et al., 2019]

Las contracciones musculares son trabajo de los músculos presentes a cada lado de la faringe, tres constrictores y tres elevadores. Estos son músculos estriados que obedecen a un comando voluntario [Latarjet et al., 2019].

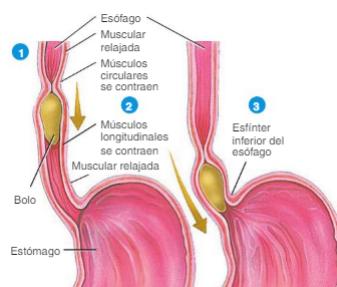


Figura 4-4: Fase esofágica de la deglución [Latarjet et al., 2019]

El esófago es un tubo tapizado de mucosa que transporta los alimentos desde la faringe (desde el borde inferior del músculo constrictor) hacia el estómago mediante la deglución, proceso que ocurre en tres pasos: voluntario, en la que el bolo pasa a la orofaringe; faríngeo, el bolo pasa involuntariamente a través de la faringe para ingresar al esófago; esofágica, paso involuntario en el que actúa, como se observa en la figura 4-4, el peristaltismo (progresión de contracciones y relajaciones coordinadas de las capas musculares) para mover el bolo a través del esófago hasta llegar al estómago [Tortora and Derrickson, 2018]. Ya en el estómago

(bolsa muscular) se mezcla el bolo macerado de alimento mediante las capas de músculo liso, pues sus capas están orientadas en diferentes direcciones entre sí, como se observa en la figura 4-5.

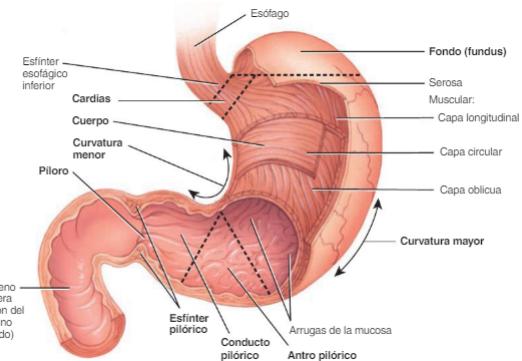


Figura 4-5: Vista anterior de las regiones del estómago [Tortora and Derrickson, 2018]

La mucosa del estómago está dispuesta en gruesos pliegues gástricos y en miles de pliegues microscópicos y fositas gástricas tapizadas con un epitelio de renovación. En la base de la fosita gástrica se encuentran las glándulas que aportan a la digestión enzimática del alimento hasta dejarlo en una mezcla líquida denominada quimo; estas glándulas contienen cuatro tipos de células, las cuales secretan: moco para proteger el revestimiento del estómago; pepsinógeno, que se convierte en pepsina una vez hace contacto con el jugo gástrico y ayuda a la digestión de las proteínas; ácido clorhídrico (HCl) y factor intrínseco; hormonas que regulan la digestión [Hansen and Carlos, 2015].

El quimo pasa al intestino delgado (ID), donde ocurre el 90 % de absorción de los nutrientes y el agua que lo atraviesan. Esta estructura empieza en el esfínter pilórico, se enrolla a través de la parte central de la cavidad abdominal y finalmente se abre en el intestino grueso. Se divide en tres partes: duodeno, yeyuno e íleon, como se observa en la figura 4-6.

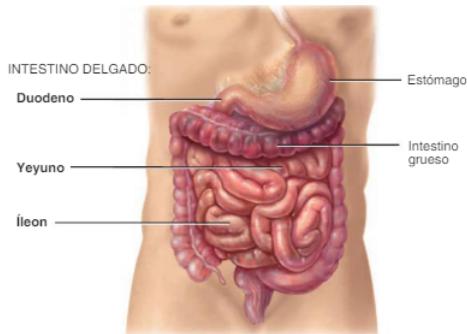


Figura 4-6: Vista anterior de la anatomía externa del intestino delgado [Latarjet et al., 2019]

El ID tiene segmentaciones que mezclan el quimo (se mueve a través del intestino por la peristalsis) con los jugos digestivos y permiten que los alimentos entren en contacto con la mucosa para su absorción. También completa la digestión de los carbohidratos, proteínas y lípidos [Tortora and Derrickson, 2018].

Finalmente, en el intestino grueso (IG) se completa la absorción de nutrientes, además se lleva a cabo la producción de vitaminas, formación de heces y la expulsión de estas fuera del cuerpo.

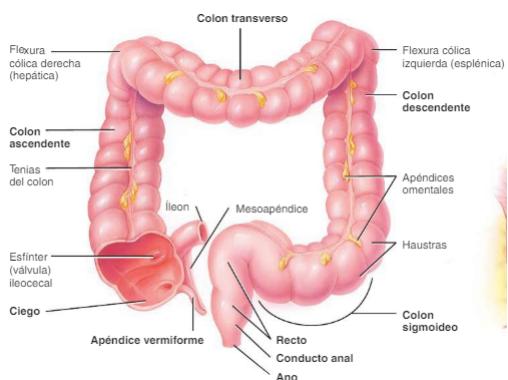


Figura 4-7: Vista anterior de las principales regiones del intestino grueso [Tortora and Derrickson, 2018]

Las regiones del IG, tal como se aprecia en la figura 4-7 son: ciego, colon (ascendente, transverso, descendente y sigmoide), recto y conducto anal. En ellas ocurren los principales procesos del IG [Tortora and Derrickson, 2018]:

- El batido de las haustras y el peristaltismo en masa impulsan el contenido del colon hacia el recto.
- Las bacterias del IG convierten proteínas en aminoácidos, los degradan y producen vitaminas del grupo B y K.
- Se lleva a cabo la absorción de agua, iones y vitaminas.
- Se expulsan las heces (vaciamiento del recto).

El sistema digestivo también incluye glándulas asociadas como: hígado, vesícula biliar y páncreas, cuyas características y funciones se describen a continuación.

El hígado es la glándula más grande del cuerpo y funcionalmente se divide en dos lóbulos: derecho e izquierdo. Es una glándula importante debido a que recibe el drenaje venoso del tracto GI, sus órganos accesorios y el bazo a través de la vena porta hepática [Hansen and Carlos, 2015]. El hígado realiza una serie de funciones importantes como: almacenar fuentes

de energía, producir combustibles celulares, producir ácidos biliares, hacer el metabolismo de los fármacos y toxinas y excreción de bilirrubina. La bilis secretada por el hígado es de alrededor 900 ml/día, la cual es almacenada en la vesícula biliar. Este es un líquido que ayuda en la digestión y a las enzimas para descomponer las grasas en ácidos grasos [Hansen and Carlos, 2015].

Por otra parte, el páncreas es una glándula retroperitoneal que consiste en cabeza, cuerpo y cola. Las células acinares del páncreas exocrino secretan una serie de enzimas que son necesarias para la digestión de proteínas, almidones y grasas. Las células ductales pancreáticas secretan líquido con un alto contenido de bicarbonato que neutraliza el ácido que entra en el duodeno desde el estómago [Hansen and Carlos, 2015].

Proceso diagnóstico de enfermedades gastrointestinales

Teniendo en cuenta la variedad de tejidos, órganos y funciones que componen el sistema digestivo, este es susceptible a diferentes padecimientos. Por lo cual en el ejercicio médico, diferentes entidades nacionales como la Asociación Colombiana de Gastroenterología e internacionales como Ministerio de salud pública de Santo Domingo proponen un protocolo de diagnóstico que consiste en:

- Anamnesis: se tiene en cuenta la historia clínica del paciente y se le interroga sobre síntomas específicos, según la patología gastrointestinal estos pueden ser: dolor abdominal moderado o severo, náuseas, vómitos, característica de las heces ó pérdida de apetito. También se hacen preguntas sobre síntomas generales asociados como: fiebre, pérdida de peso y fatiga [de Salud Pública, 2016].
- Examen físico: se evalúa el aspecto general (hidratación, coloración de piel y mucosas), estado nutricional (peso y talla), abdomen (distensión, ruidos intestinales, dolor a la palpación y masas), aftas orales, iritis (enrojecimiento del ojo), artritis [de Salud Pública, 2016].
- Pruebas de laboratorio: estas pruebas incluyen exámenes coprológicos seriados y coprocultivos obligatorios; hemograma completo; velocidad de eritrosedimentación, proteína C reactiva; electrolitos; urocultivo [de Salud Pública, 2016].
Posteriormente se realizan pruebas complementarias para confirmar el diagnóstico o en ocasiones tratar alteraciones específicas, por ejemplo:
 - Endoscopia: es una prueba diagnóstica por imágenes mínimamente invasiva muy utilizada actualmente, dado que permite la visualización de una gran parte del aparato digestivo y otras partes del cuerpo. En este examen se introduce un endoscopio al interior de una estructura o cavidad corporal, el cual cuenta con una fuente de luz y una cámara que transmite las imágenes a una pantalla de computador [Society, 2019], [Münzer et al., 2018].

- Tomografía Computarizada (TC): es un examen no invasivo de diagnóstico por imágenes que se utiliza para ayudar a detectar enfermedades del ID, del colon y de otros órganos del cuerpo, tales como: apendicitis, colitis ulcerativa (enfermedad de Crohn), pancreatitis, cirrosis hepática, cálculos y cáncer de: hígado, riñón, páncreas, ovarios y vejiga. Sin embargo, se debe tener en cuenta que este es un examen costoso, lo cual dificulta su acceso [Info, 2018].
- Resonancia Magnética: es un examen imagenológico no invasivo, que a través del uso de imanes, ondas potentes y una computadora genera imágenes nítidas que permiten la visualización de todas las estructuras internas del cuerpo, sin emplear radiación ionizante; se usa para detectar enfermedades como: cáncer de estómago o de esofago, enfermedad inflamatoria intestinal, entre otras [Gallego Ojea, 2019].

4.2.2. Características visuales de endoscopia asociadas a padecimientos de sistema gastrointestinal

A nivel de endoscopia, que es el tema central del presente trabajo de grado, las enfermedades son diagnosticadas bajo características visuales asociadas a cambios fisiológicos que resultan ser importantes para la propuesta de trabajo de grado, sobre todo para limitar la posibilidad de que estos patrones sean confundidos con artefactos. Algunos de estos descriptores relacionados a enfermedades se presentan a continuación:

- Para la **gastritis**: como se observa en la figura 4-8 la mucosa del estómago con presencia de gastritis se visualiza como un enrojecimiento del tejido (inflamación), junto con erosiones de las células superficiales, que se ven como pequeños puntos o placas rojas, y en algunos casos pequeñas lesiones que sangran [Murra Saca, 2018].

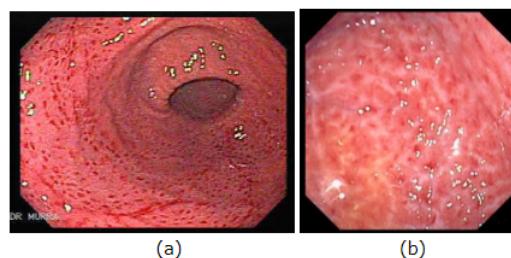


Figura 4-8: Imágenes endoscópicas de gastritis [Vakil, 2019a]

Donde en la figura 4-8 la parte (a) representa una gastritis aguda, donde se presenta en forma de puntos rojos pequeños y dispersos son áreas inflamadas de la mucosa; (b) gastritis crónica, la cual se presenta en forma de placas de color rojo y rosa oscuro que son las áreas de inflamación crónica y profunda de la mucosa gástrica; imágenes adaptadas de Manual MSD [Vakil, 2019a] y el atlas gastrointestinal [de vídeo endoscopia Gastrointestinal de El Salvador,].

- Para las **úlceras**: como se observa en la figura 4-9, las úlceras se visualizan como un agujero o perforación en la mucosa, pero también se puede ver como un tipo de hundimiento de la mucosa (bache) [Vakil, 2019b].

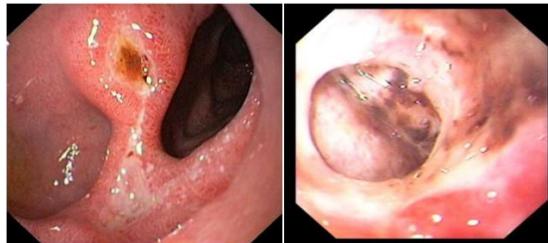


Figura 4-9: Úlceras presentes en imágenes endoscópicas [de vídeo endoscopia Gastrointestinal de El Salvador,], [Vakil, 2019a]

- Para **pólipos**: como se observa en el figura 4-10, los pólipos se ven como pequeñas protuberancias de diferentes diámetros que se ubican en la mucosa, suelen tener coloración eritematosa (enrojecimiento) comparado con la mucosa adyacente [Macías-Rodríguez et al., 2014]. Además dichas protuberancias pueden crecer como “tallos” y parecer honros, que son más conocidos como: pólipos pediculados; otros crecen sin cuello (pólipos sésiles) y existen otros que son planos (pólipos planos) [Lizarzábal García, 2016].

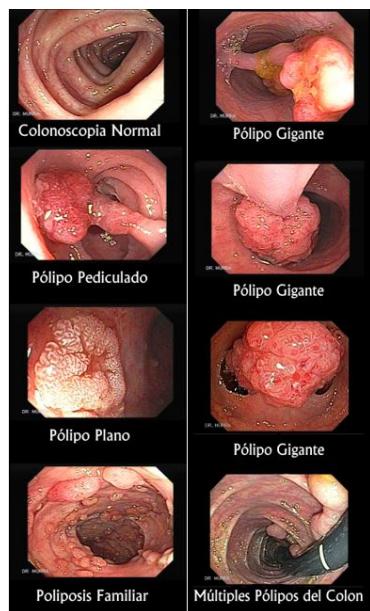


Figura 4-10: Pólipos presentes en imágenes endoscópicas [Murra Saca, 2018]

- Para la **acalasia**: en el examen endoscópico para esta enfermedad se suele visualizar los restos alimenticios a nivel esofágico, así como, la dilatación del mismo y la resistencia que se produce ante el paso del endoscopio hacia el estómago [Gaudó Perales

and Revnic,]. En las imágenes endoscópicas si bien no se visualizan cambios en la mucosa, se pueden observar las alteraciones de las contracciones del tubo digestivo y los estrechamientos de la unión gastroesofágica como se ve en la figura 4-11 [Bigot and Alvarado, 2009].



Figura 4-11: Acalasia presente en imágenes endoscópicas [Gaudó Perales and Revnic,]

En la figura 4-11 se observa la unión gastroesofágica muy cerrada por espasmos del cardias [Gaudó Perales and Revnic,].

- Para **cáncer gástrico**: Como se observa en la figura 4-12, el cáncer precoz se caracteriza por la presencia de un borde bien marcado y un patrón irregular de color [Yao et al., 2013]; por otro lado, en la figura 12b se ve que el cáncer más avanzado es una unión anormal de los pliegues, sangrado espontáneo, ulceraciones o erosiones, y cambios de color de la mucosa [Lee et al., 2017].

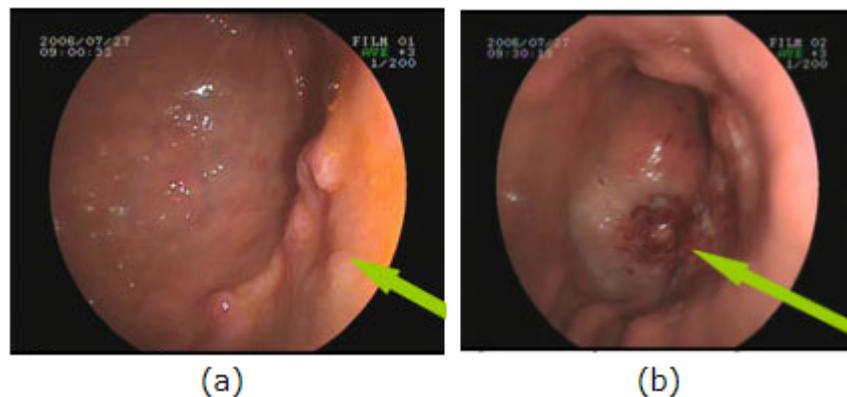


Figura 4-12: Imágenes endoscópicas con presencia de cáncer gástrico, (a) precoz y (b) avanzado [Grávalos and González Flores, 2017]

- Para la **enfermedad de Crohn (EC)**: en fases iniciales, como se observa en la figura (a) 4-13, se encuentran úlceras aftosas, caracterizadas por ser lesiones redondeadas,

superficiales, de pequeño tamaño, con fondo recubierto de fibrina, rodeadas de mucosa de aspecto normal. Cuando la enfermedad avanza estas úlceras superficiales tienden a aumentar de tamaño, a hacerse más profundas o juntarse entre ellas dando lugar a úlceras lineales y estrelladas (figura (b) 4-13).

En las fases más avanzadas de la enfermedad, aparecen ulceraciones lineales y serpiginosas profundas, que se caracterizan por seguir el eje del intestino, tal como se aprecia en la figura (c) 4-13 [Vera, 2017].

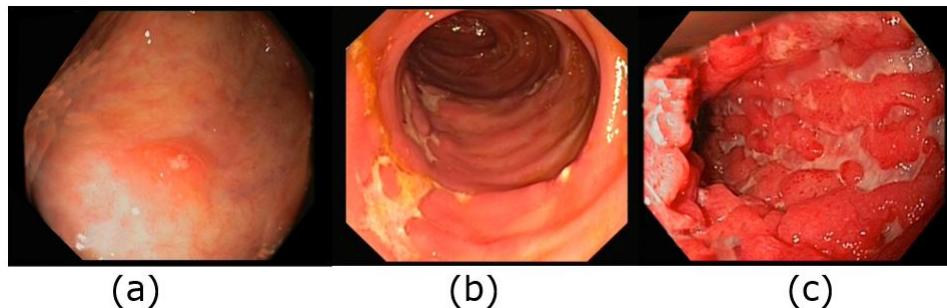


Figura 4-13: Imágenes endoscópicas con presencia de enfermedad de Crohn [Vera, 2017]

En la figura 4-13 podemos observar (a) úlcera aftosa, (b) úlceras de bordes irregulares o estrellados, (c) úlceras serpiginosa [Vera, 2017].

A pesar de que la endoscopia es un examen ampliamente usado en el campo médico, es importante mencionar que los cuadros generados en el video presentan inconvenientes, pues la mayoría están contaminados con múltiples artefactos que impiden la correcta observación de las características visuales anteriormente mencionadas [Ali and Zhou, 2019].

4.2.3. Artefactos

Los artefactos son definidos como la distorsión, adición o error en una imagen que no tiene correlación con la región anatómica estudiada; los cuales etimológicamente proceden de las raíces artis (artificial) y actum (efecto), lo que indica que los artefactos son el efecto artificial que altera la calidad y fidelidad de la imagen, donde puede llegar a ocultar alguna información relevante (ej. patología) o puede generar hallazgos falsos [Sartori et al., 2015].

A continuación se explicarán los artefactos que se presentan con mayor frecuencia en los marcos del video de endoscopia, como lo son: saturación de píxeles, reflejos especulares, burbujas, residuos y desenfoque por movimiento [Ali and Zhou, 2019].

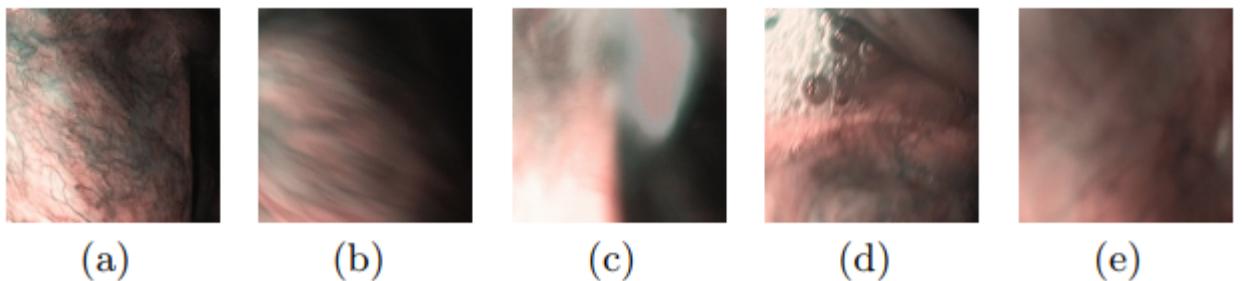


Figura 4-14: Diferentes artefactos [Atasoy et al., 2010]

En la figura 4-14 se puede observar: (a) Imagen endoscópica gastrointestinal ideal; las otras imágenes contienen artefactos como: (b) desenfoque por movimiento, (c) reflejos especulares, (d) burbujas, (e) indistinto [Atasoy et al., 2010].

- Saturación de píxeles (brillos en zonas): en las imágenes de endoscopia se pueden presentar zonas en las que hay un alto nivel de saturación, como se observa en la figura 4-15. Por ejemplo, en una imagen (canal RGB - cuantificación 8 bits), un valor de píxel R, G o B está saturado cuando toma su valor máximo, en este caso 255 [Zhang and Brainard, 2004].

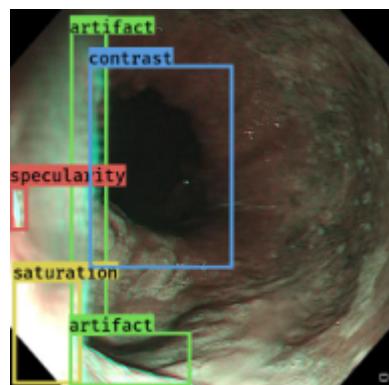


Figura 4-15: Imagen de endoscopia con la presencia de varios artefactos, en el que se destaca la saturación [Ali and Zhou, 2019]

- Reflejos especulares: un fenómeno en que los rayos de luz (en este caso, provenientes de la sonda) que inciden sobre la superficie de la mucosa se reflejan de tal manera que el ángulo de reflexión es igual al ángulo de incidencia, como se ve en la figura 4-16, [Queiroz and Ren, 2019]. Este tipo de artefactos afecta la calidad de la imagen desde el punto de vista del diagnóstico y puede ser una fuente de error para muchos algoritmos de diagnóstico, puesto que pueden generar regiones falsas, así como, errores

en segmentación, reconocimiento de objetos, reconstrucción de formas, etc [Gao et al., 2017].

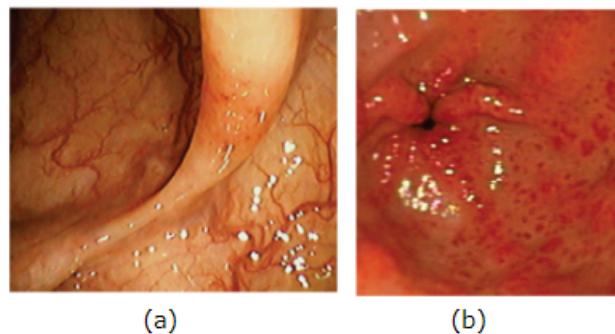


Figura 4-16: Imágenes endoscópicas del (a) Colon y (b) Esófago con reflejos especulares [Stehle, 2006]

- Burbujas: como se puede ver en la figura 4-17 son la presencia de burbujas de aire y espuma en alguna parte del tracto digestivo, lo que limita la visualización de la mucosa [Ahsan et al., 2011].



Figura 4-17: Imágenes endoscópicas con presencia de burbujas [Yoo et al., 2016]

- Contraste: se define como la diferencia relativa de intensidades de la imagen entre regiones con detalles adyacentes [Luna-Sánchez et al., 2011], como se observa en la figura 4-18.

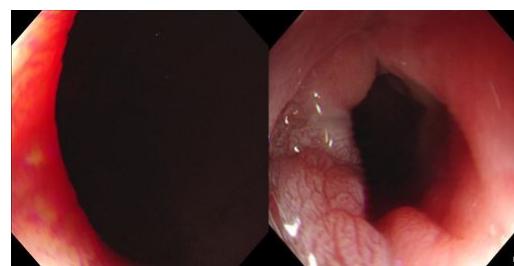


Figura 4-18: Imágenes endoscópicas con artefacto de contraste [Ali and Zhou, 2019]

- Desenfoque por movimiento: puede deberse a un movimiento rápido de la cámara [Kaçmaz et al., 2018], o a los cambios frecuentes de enfoque de la cámara mientras se mueve a lo largo del tracto digestivo, pues en muchos casos se usan endoscopios equipados sólo con un lente gran angular que puede dejar imágenes fuera de foco, como se observa en la figura 4-19 [JungHwan Oh, 2004].



Figura 4-19: Imágenes colonoscópicas con artefacto de movimiento [Kaçmaz et al., 2018]

- Instrumentos: ocurre cuando el tubo del endoscopio queda registrado en la imagen (marco), obstruyendo la visualización del tejido, como se observa en la figura 4-20.

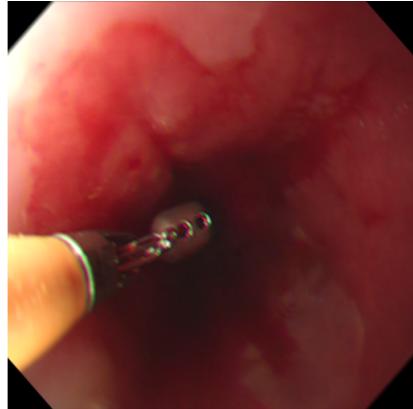


Figura 4-20: Imagen endoscópica con presencia de un instrumento [Ali and Zhou, 2019]

4.2.4. Imágenes digitales

Considerando que hasta el momento se ha hablado de imágenes endoscópicas y estas serán la entrada del producto a desarrollar, se torna relevante entender qué es una imagen digital y cómo es guardada la información en estas. Las cámaras digitales captan longitudes de onda mediante sensores (datos analógicos), estos datos pasan por un proceso de digitalización y son almacenados de forma ordenada en matrices con coordenadas espaciales x,y (filas y columnas), donde el valor correspondiente a cada elemento de la matriz que representa el nivel de intensidad de luz es ese punto, es conocido como píxel [Grávalos and González Flores, 2017], [Luis Gerardo de la Fraga, 2011], como se observa en la figura 4-21.

Las imágenes pueden ser binarias, en escala de grises y a color, donde las primeras son imágenes de un solo bit, es decir, solo puede tener dos tonos (blanco y negro); en el caso de las imágenes en escala de grises, suelen tener 8 bits, lo que significa que pueden tener 256 tonos diferentes de grises; por último, las imágenes a color se crean con parámetros RGB ((R) rojo, (G) verde, (B) azul), donde cada uno de estos se almacena en una matriz diferente y cada una generalmente es de 8 bits, es decir, que por los tres canales son 24 bits, por lo que una imagen a color con estos bits mostrará 16.7 millones de colores, cabe resaltar que este tipo de imágenes permite cambiar de espacio de color [Luis Gerardo de la Fraga, 2011], [Ortiz et al., 2017].

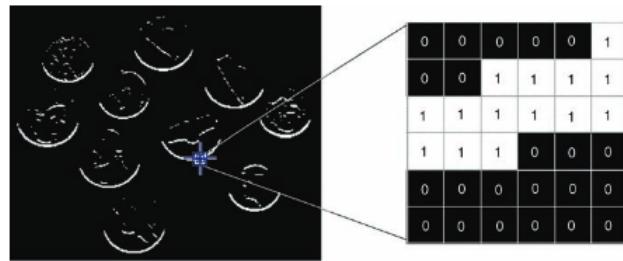


Figura 4-21: Representación de una imagen binaria y los valores de sus píxeles a través de una matriz de 6x6 [Marques, 2011]

4.2.5. Visión por computadora

Estas imágenes deben ser procesadas y analizadas mediante una herramienta conocida como visión por computadora, la cual busca establecer relaciones entre el mundo tridimensional y sus vistas bidimensionales tomadas de él [Ortiz et al., 2017], [Mery, 2004] emulando el proceso de asociación de conceptos cuando el lóbulo occipital de la corteza visual capta imágenes [Mery, 2004], [Watson et al., 2010].

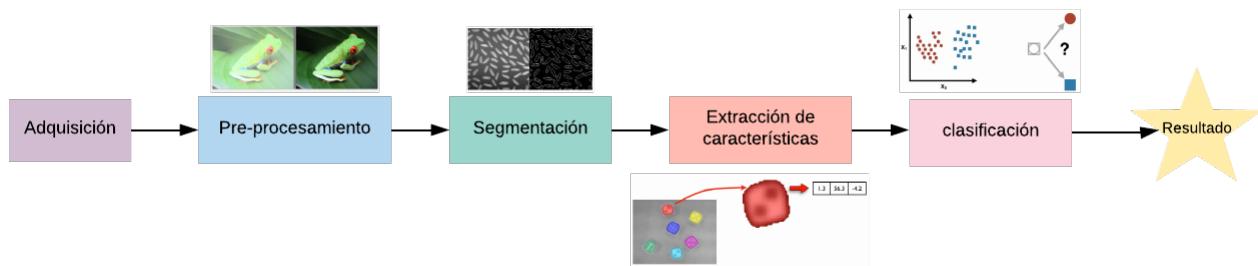


Figura 4-22: Etapas de un sistema de visión por computadora

Como se observa en la Figura 4-22, los principales componentes de un sistema de visión por computadora o visión de máquina son:

1. Adquirir información del medio físico (imágenes digitales).
2. Etapa de preprocesamiento: se hacen transformaciones de la imagen dirigidas a mejorarla para el proceso posterior, algunas de estas transformaciones pueden consistir en:
 - Transformación de intensidad: se modifica el brillo y el contraste; en el primero se reduce el contraste y la pérdida de detalle; y en el segundo se realza el contraste para mejorar ciertas características visuales de la imagen como se puede observar en la figura 4-23 y 4-24. Algunas técnicas para ello son: operaciones puntuales de transformación, ecualización, binarización, entre otras [Watson et al., 2010].

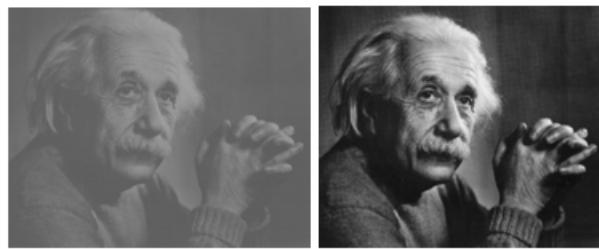


Figura 4-23: Realce de contraste de una imagen mediante normalización [Ortiz et al., 2017]

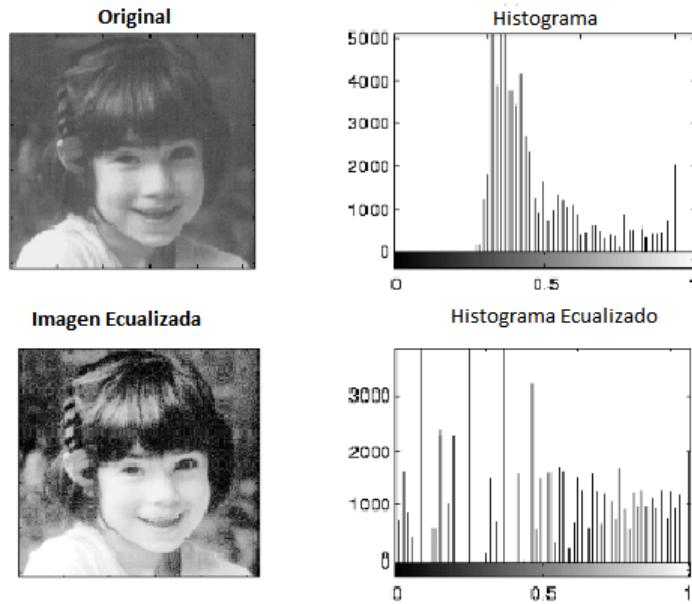


Figura 4-24: Imagen original con sus versiones ecualizadas [of Edinburgh, 2010]

- Normalización: lleva a un tamaño estándar las dimensiones de la imagen sin generar en ella una distorsión significativa [Marques, 2011]

3. Segmentación: divide la imagen en sus componentes principales, es decir separa los objetos de interés del fondo como se ve en la figura 4-25, basándose en características que permiten distinguir un objeto del fondo y objetos entre sí. Tras aplicar algoritmos como discontinuidad, similaridad o conectividad, se obtiene una imagen binaria [Universidad De Jaén, 2017], [Molina Cabello et al., 2018]. Este proceso se puede llevar a cabo mediante diferentes métodos como: basado en intensidad (umbralización), basado en regiones (región en crecimiento, dividir y fusionar) y otros (basados en bordes, textura, movimiento) [Marques, 2011].



Figura 4-25: Segmentación de una imagen [Li et al., 2018]

4. Extracción de características: proceso en el que se detectan ciertas características de una imagen. Esta representación resultante es la entrada a las técnicas de clasificación y reconocimiento, las cuales etiquetan, clasifican o reconocen contenido de una imagen o sus objetos como se ve en la figura 4-26. Las técnicas para encontrar los vectores característicos son muchas, por ejemplo: área, centroide, excentricidad, transformada de Fourier, Transformada Wavelet, etc [Molina Cabello et al., 2018], [Marques, 2011].

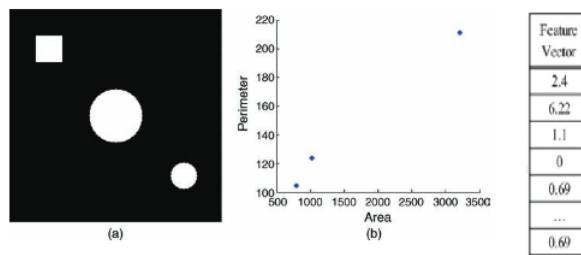


Figura 4-26: Extracción de características a partir del área [Molina Cabello et al., 2018]

5. Clasificación: se encarga de asignar una clase a un patrón desconocido a partir del conocimiento adquirido previamente acerca de los objetos y las clases a las que pertenecen (entrenamiento del modelo de clasificación con las características encontradas en la etapa anterior) [Marques, 2011], como se observa en la figura 4-27

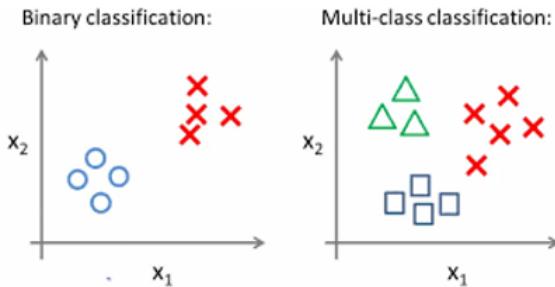


Figura 4-27: Resultado de un modelo de clasificación binaria o multiclas [Mishra, 2018]

4.2.6. Características

A continuación se exponen las características que se han empleado (ver sección 4.1) y las que se extrajeron en el presente trabajo de grado para la detección de RE y DM.

4.2.6.1. Detección de DM

- **IPR:** se calcula a partir de los bordes de una imagen endoscópica, más específicamente calcula el porcentaje de píxeles IP (píxeles de borde que no están conectados con otro píxel), mediante la siguiente ecuación 4-1.

$$IPR = \frac{IPs}{Total\ de\ pixeles\ de\ borde} \cdot 100\% \quad (4-1)$$

Dado que la característica IPR necesita de bordes, se hace relevante explicar cómo se obtienen dichos bordes; estos utilizan técnicas de detección de discontinuidades en el brillo con las que se encuentran los límites de un objeto dentro de una imagen; para esto se han desarrollado una variedad de algoritmos como lo son: Sobel, Canny, Pre-witt, Roberts y métodos de lógica difusa [Rebaza, 2007], [MathWorks,].

Por ejemplo, el algoritmo de **Canny** se basa en los cambios bruscos de la primera derivada, más específicamente este consiste en tres grandes pasos [Rebaza, 2007]:

1. Obtención del gradiente: para este paso lo primero que se hace es aplicar un filtro gaussiano para el suavizado de la imagen como se ve en la figura 4-28 , sin embargo, cabe resaltar que no se debe hacer un suavizado excesivo dado que se pueden perder algunos bordes de la imagen [Rebaza, 2007].

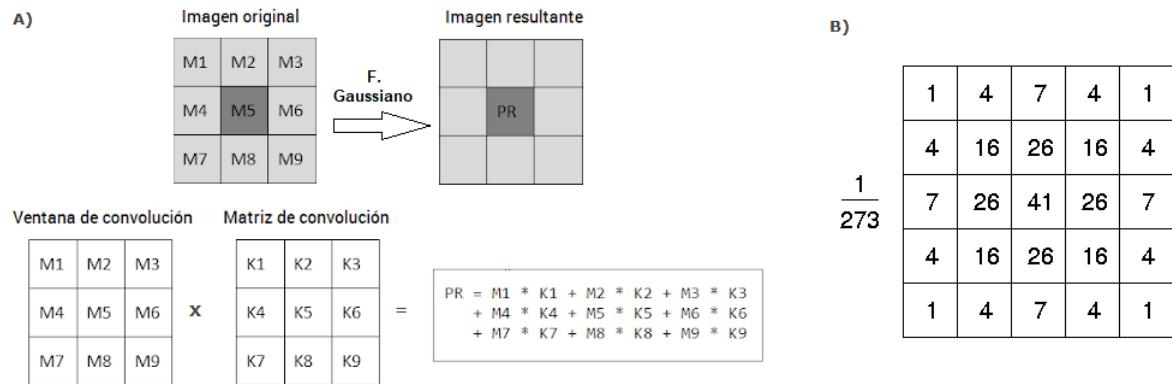


Figura 4-28: Filtro Gaussiano, **A)** operación de convolución [Kabbai et al., 2016], **B)** ejemplo de mascara 5x5 para un filtro gaussiano [R. Fisher and Wolfart., 2003]

Posteriormente, se calcula la magnitud (ecuación 4-3) y orientación (ecuación 4-4) del gradiente en cada píxel. Donde, el gradiente $f(x, y)$ es un vector bidimensional en un píxel (x, y) , definido por la ecuación 4-2.

$$G[f(x, y)] = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x} f(x, y) \\ \frac{\partial}{\partial y} f(x, y) \end{bmatrix} \quad (4-2)$$

$$|G| = \sqrt{G_x^2 + G_y^2} = |G_x| + |G_y| \quad (4-3)$$

$$\nabla f = \tan^{-1} \left(\frac{G_y}{G_x} \right) \quad (4-4)$$

2. Supresión no máxima: tras obtener el gradiente (magnitud y orientación), se hace el adelgazamiento del ancho de los bordes, es decir, se eliminan los píxeles no deseados (no son parte del borde); esto se hace verificando en cada píxel, si este es un máximo local en dirección del gradiente. Si no lo es se asigna un valor de 0 al píxel (se suprime), en caso contrario se asigna el valor que tenga la magnitud del gradiente [Rebaza, 2007], [OpenCV, 2015].
3. Histéresis de umbral: en este paso se busca reducir la posibilidad de aparición de bordes falsos, para esto se aplica una función de histéresis basada en dos umbrales (`minVal` y `maxVal`). Si el borde es mayor que `maxVal` se considera un borde seguro, si es menor que `minVal` no es un borde (se elimina), por otro lado, si el borde se encuentra entre estos dos umbrales se clasifican como bordes o no según su conectividad (ver figura 4-29) [Rebaza, 2007], [OpenCV, 2015].

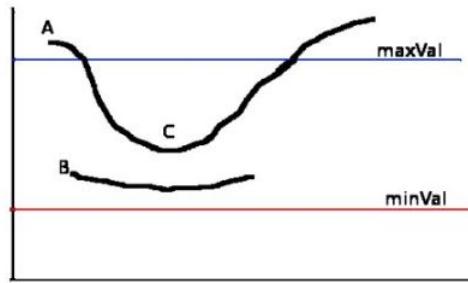


Figura 4-29: Histéresis de umbral [OpenCV, 2015]

En la figura 4-29, A es un "borde seguro" dado que es mayor que maxVal; C aunque es menor que maxVal está conectado al borde A, por lo que también se considera un borde válido; por otro lado, el borde B esta en entre los dos umbrales pero no está conectado, por ende no es considerado como borde [OpenCV, 2015].

- **Matriz de co-ocurrencia de nivel de gris (GLCM):** Es una matriz que cuenta la frecuencia con la que un píxel con cierto nivel de gris tiene una relación de espacio específica con otro píxel, donde estos están separados por alguna distancia en un determinado ángulo; esta matriz revela ciertas propiedades sobre la distribución de los grises en la textura de la imagen [Ríos-Díaz et al., 2009], [Pladellorens Mallofré, 2010].

En el caso de [Oh et al., 2007], hicieron un análisis de textura del espectro de frecuencia, para esto lo primero que realizaron fue convertir las imágenes a escala de grises y luego la transformaron al espectro de frecuencias gracias a la transformada de Fourier (DFT).

El espectro de frecuencias muestra la distribución de las frecuencias en una imagen, generando diferentes patrones; generalmente las características más notorias de la transformada son que: las altas frecuencias dan información de los bordes y límites de la imagen, y que las bajas frecuencias corresponden a los componentes de variaciones lentas. En el caso de los marcos endoscópicos con desenfoque, el grafico en 2D de la transformada de Fourier no contiene información clara, excepto que en sus esquinas tiene bordes fuertes; por otro lado, el marco endoscópico sin artefacto si muestra información clara [Oh et al., 2007].

Tras obtener el gráfico 2D de la transformada hicieron el análisis de textura de este con la matriz GLCM, donde para calcularla se consideran la relación espacial entre los píxeles, uno de referencia (i) y su vecino (j), donde cada elemento de la matriz C_{ij} es el número de veces que en la imagen donde se pasa del píxel i al píxel j para una distancia y ángulo determinado (como se ve la figura 4-30), este ángulo puede variar entre 0, 45, 90 o 135 grados [Pladellorens Mallofré, 2010], [Haralick et al., 1973].

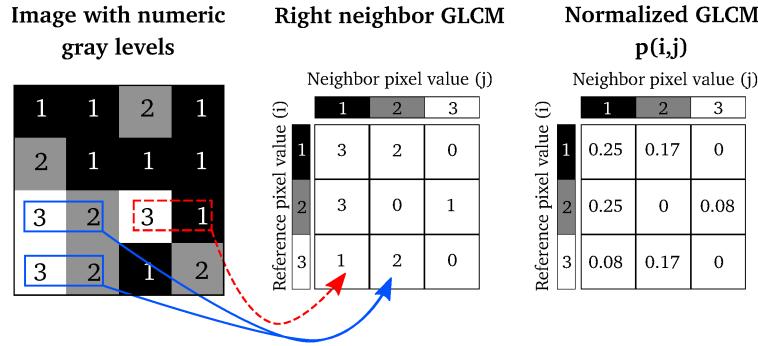


Figura 4-30: Ejemplo cálculo matriz de co-ocurrencia de nivel de gris [Löfstedt et al., 2019]

En la figura **4-30** se presenta una imagen (4x4) con tres niveles de gris (1,2y3), en este ejemplo solo se mira la relación del píxel de referencia con el vecino de la derecha, donde el píxel de i determina la columna del GLCM y el j determina la fila. En este ejemplo, hay 2 casos en el que el píxel de $i=3$ “co-ocurre” con un píxel vecino $j = 2$ (línea azul), por otro lado, la línea roja indica un ejemplo de un píxel $i = 3$ con un vecino de 1. En esta figura también se observa un GLCM normalizado, el cual representa la probabilidad estimada (P_{ij}) de que esa combinación ocurra en la imagen [Pladellörens Mallofré, 2010], [Haralick et al., 1973].

Al tener el GLCM calculado se procede a calcular algunas o todas las características de textura propuestas por Haralick [Haralick et al., 1973], de la siguiente manera:

$$\text{Segundo Momento Angular (ASM)} : \sum_i \sum_j P(i,j)^2 \quad (4-5)$$

$$\text{Contraste} : \sum_i \sum_j (i,j)^2 \cdot P(i,j) \quad (4-6)$$

$$\text{Correlación} : \sum_i \sum_j \frac{(i - \mu_x) \cdot (j - \mu_y) \cdot P(i,j)}{\sigma_x \sigma_y} \quad (4-7)$$

$$\mu_x : \sum_i \sum_j i \cdot P(i,j) \quad (4-8)$$

$$\mu_y : \sum_i \sum_j j \cdot P(i,j) \quad (4-9)$$

$$\sigma_x = \sqrt{\sum_i \sum_j (i - \mu_x)^2 \cdot P(i, j)} \quad (4-10)$$

$$\sigma_y = \sqrt{\sum_i \sum_j (j - \mu_y)^2 \cdot P(i, j)} \quad (4-11)$$

$$Disimilitud: \sum_i \sum_j P(i, j) \cdot |i - j| \quad (4-12)$$

$$\text{Entropía: } \sum_i \sum_j P(i, j) \cdot \ln P(i, j) \quad (4-13)$$

$$\text{Energía: } \sqrt{ASM} \quad (4-14)$$

$$Uniformidad: \sum_i \sum_j \frac{P(i, j)}{1 - |i - j|} \quad (4-15)$$

$$Sumadecuadrados: \sum_i \sum_j (i - \mu)^2 \cdot P(i, j) \quad (4-16)$$

$$Momentodediferenciainversa: \sum_i \sum_j \frac{1}{1 + (i + j)^2} \cdot P(i, j) \quad (4-17)$$

$$Sumapromedio: \sum_{i=2}^{2N_g} iP_{x+y}(i) \quad (4-18)$$

- **Evaluador de calidad espacial de imágenes sin referencia (BRISQUE):** Es una métrica holística con la que se evalúa la calidad de la imagen sin necesidad de una referencia, dicha medida se calcula en el dominio espacial. Esta medida utiliza coeficientes de luminancia normalizado para cuantificar posibles pérdidas de 'naturalidad' en la imagen por la presencia de distorsiones [Moccia et al., 2018].

- **Varianza (laplaciana):** Es una medida basada en el laplaciano (L) de la imagen (I) de tamaño $M \times N$, se calcula de la siguiente manera [Moccia et al., 2018]:

$$\Delta VAR : \sum_m^M \sum_m^N ((L(m, n)) - \bar{L})^2 \quad (4-19)$$

$$\bar{L} = \frac{1}{MN} \sum_m^M \sum_m^N |L(m, n)| \quad (4-20)$$

Donde L se obtiene por la convolución de la imagen con el núcleo laplaciano (K_L).

$$K_L : \frac{1}{6} \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (4-21)$$

- **Función de evaluación de enfoque Sobel–Tenengrad (TEN):** Esta medida se basa en el valor del gradiente en dos direcciones G_x, G_y , usando los operadores sobel para convertir la imagen y sumando al cuadrado todas las magnitudes mayores al umbral (T), definiéndose como:

$$TEN : \sum_m^M \sum_m^N [S(m, n)]^2, \quad \text{para } S(m, n) > T \quad (4-22)$$

$$S(m, n) = \sqrt{[G_x(m, n)]^2 + [G_y(m, n)]^2} \quad (4-23)$$

$$G_x : \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \quad G_y \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (4-24)$$

4.2.6.2. Detección de RE

- **Luminancia o luma (Y):** Es la medida de la cantidad de energía que un observador percibe procedente de una fuente luminosa, provee la información en blanco y negro de una imagen, es decir, provee la imagen en escala de grises; esta se puede encontrar separada del color en espacios como: YUV, YIQ o YCrCb [Urueña et al., 2011]. El cálculo del canal de luminancia se hace de la siguiente manera:

$$Y = 0,3R + 0,59G + 0,11B \quad (4-25)$$

- En el espacio de color CIE-XYZ la luminancia se calcula como se muestra a continuación:

$$Y = 0,21R + 0,71G + 0,07B \quad (4-26)$$

Donde, R,G,B = Corresponden a los canales de una imagen en RBG. En el caso de una imagen endoscópica, las áreas con RE tendrán un valor de luminancia elevado; estos niveles se observan mejor en un histograma.

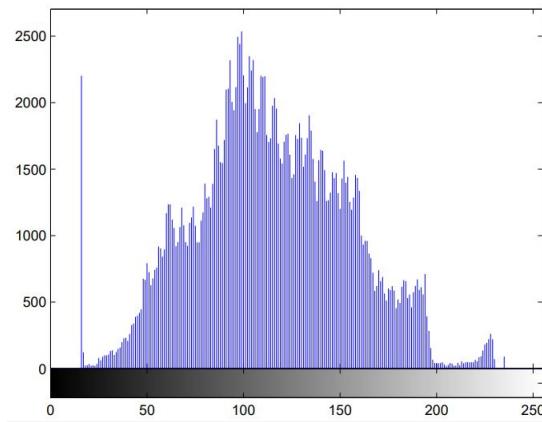


Figura 4-31: Histograma típico del canal de luminancia de una imagen endoscópica [Stehle, 2006]

En la figura 4-31 se observa un pequeño pico que corresponde a píxeles que contienen RE. Para hallar la posición de dicho pico se puede pasar un filtro por el histograma para encontrar el punto donde la derivada tiene un cambio significativo de un valor positivo a un valor cercano a cero [Stehle, 2006].

- **Brillo:** El brillo se refiere a la claridad general u oscuridad de la imagen, y se calcula como se ve en la ecuación 4-27 [Morgand and Tamaazousti, 2014].

$$Brillo = \frac{\sqrt{0,241R^2 + 0,691G^2 + 0,068B^2}}{W * H} \quad (4-27)$$

Donde, R,G,B hacen referencia a cada uno de los canales del espacio de color RGB, W y H las dimensiones de la imagen.

- **Varianza global de intensidad (GVAR):** Se calcula mediante la siguiente ecuación [Moccia et al., 2018]:

$$GVAR : \frac{1}{MN} \sum_m^M \sum_n^N [I(m, n) - \bar{I}]^2 \quad (4-28)$$

Donde, \bar{I} : es la media de intensidad de todos los píxeles en la imagen.

- **Variación local de la intensidad del canal de luminancia:** Esta medida se define como [Moccia et al., 2018]:

$$Varl : \frac{1}{MN} \sum_m^M \sum_m^N [lv(m, n) - \bar{lv}]^2 \quad (4-29)$$

Donde,

$$lv(m, n) : \frac{1}{w_x w_y} \sum_i^{w_x} \sum_j^{w_y} [I(m + i, n + j) - \bar{I}_w]^2 \quad (4-30)$$

$$\bar{I}_w : \frac{1}{MN} \sum_m^M \sum_m^N lv(m, n) \quad (4-31)$$

\bar{I}_w : Valor de intensidad media en la ventana

- **Número de puntos clave (NP):** Son los puntos significativos en la imagen en escala de grises [Moccia et al., 2018].

Como se observa a continuación, también fueron calculados diferentes parámetros estadísticos para la detección de los dos artefactos.

- **Media:** es el promedio de un conjunto de valores, o su distribución, definida como se muestra a continuación

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i \quad (4-32)$$

- **Desviación estándar:** es una medida que indica qué tan dispersos están los datos con respecto a la media, como se observa en la ecuación 4-33. Esto significa que mientras mayor sea la desviación estándar más dispersos estarán los datos. En el caso de imágenes es una característica importante porque generalmente los cuadros desenfocados tienen una distribución de color no uniforme. [Armin et al., 2016].

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (4-33)$$

4.2.7. Métodos de identificación

4.2.7.1. K-means

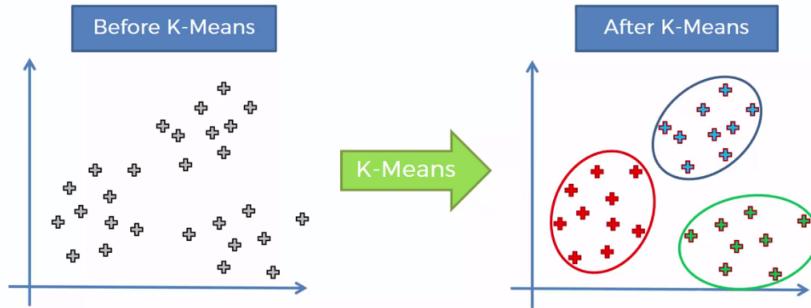


Figura 4-32: Método de agrupamiento K-means [Iris et al.,]

Es un método de agrupamiento con aprendizaje no supervisado, donde se utilizan datos no etiquetados. Este método tiene como objetivo la partición de n observaciones en k (cluster) grupos, donde se agrupa por medio de la similitud de la suma de las distancias entre cada objeto y el centroide (el medio de los elementos que componen el cluster) de su grupo [Ramos, 2016], [Zhang, 2012]. Dicho algoritmo se divide en las siguientes etapas:

1. Se elige al azar los objetos que forman así los K grupos iniciales, estos tienen un valor inicial del centroide (C_i), siendo estos el único objeto dentro de los grupos [Ramos, 2016].
2. Se reasignan los objetos en los grupos mediante la medida de la distancia [Ramos, 2016].
3. Se calculan los centroides [Ramos, 2016].
4. Repetir los pasos 2 y 3 hasta que no se hagan más reasignaciones [Ramos, 2016].

El resultado de la aplicación de este algoritmo depende de la escogencia inicial de los centroides [Zhang, 2012].

- Evaluación de desempeño: La medida más común es la suma de error cuadrado (SSE), donde el error es la distancia al grupo más cercano y para obtener el SSE se cuadran estos errores y se suman como se ve en la ecuación 4-34 [Zhang, 2012].

$$SSE : \sum_{i=1}^k \sum_{x \in C_i} dist^2(m_i, x) \quad (4-34)$$

Donde, x es un punto de datos en el clúster C_i y m_i es el punto representativo del clúster [Zhang, 2012].

4.2.7.2. Random Forest(RF)

Es un algoritmo de aprendizaje supervisado que crea árboles decisión sobre muestras (variables) aleatoriamente seleccionadas, las cuales son empleadas para tomar la decisión de cada árbol y así seleccionar la mejor opción mediante el promedio del cálculo de la clase más votada. Para llevar a cabo lo anterior, en el k -ésimo árbol se genera un vector aleatorio θ_k independiente de los últimos vectores aleatorios $(\theta_1, \dots, \theta_{k-1})$ pero con la misma distribución; dicho lo anterior un árbol tiene como parámetros de entrada el conjunto de entrenamiento y $\theta_k \mathbf{h}(\mathbf{x}, \theta_k)$.

Como se mencionaba anteriormente, RF se basa en un conjunto de árboles de decisión, por lo tanto la muestra que entra a un árbol es sometida a una serie de test binarios en cada nodo, llamados *split*, hasta llegar a la hoja en la que se encuentra la respuesta [Medina-Merino and Nique-Chacón, 2017]. A partir de las muestras de entrenamiento, el algoritmo busca optimizar los parámetros de las funciones de *split*. Para ello se utiliza la siguiente función de ganancia de información:

$$I_j = H(j) - \sum_{i \in \epsilon_{1,2}} \frac{|S_j^i|}{|S_j|} \cdot H(S_j^i) \quad (4-35)$$

4.2.7.3. Máquina de soporte vectorial (SVM)

Este algoritmo se puede emplear para clasificación binaria o regresión, donde primero se transforman los datos de entrada en un espacio de características de una dimensión mayor, es decir, si la entrada está en el espacio R^2 se transforma a un espacio R^3 y se encuentra un hiperplano óptimo que sea capaz de clasificar las características como se ve en la figura 4-33 [Betancourt, 2005], [Vargas et al., 2012].

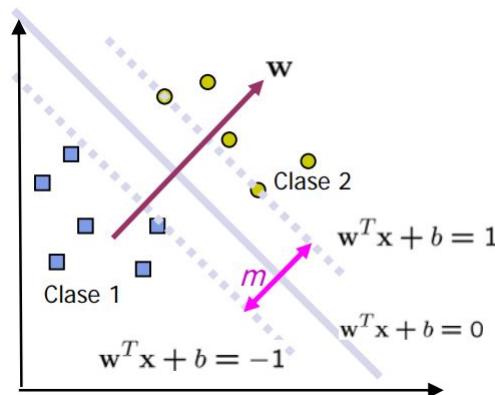


Figura 4-33: Representación de un hiperplano (línea azul) [Betancourt, 2005]

Un conjunto de entrenamiento de tamaño N compuesto por datos tipo (x_i, y_i) , debe tener una ecuación de un hiperplano que divida dicho conjunto, de manera que los datos con la

misma etiqueta queden en un mismo lado del hiperplano [HENAO et al., 2005], [Vargas et al., 2012], este se calcula con la siguiente formula matemática:

$$y_i(\omega^T x_i + b) > 0, i : 1, \dots, N \quad (4-36)$$

Cuando los datos son linealmente separables es porque existe un hiperplano que satisfaga la ecuación 4-36, cabe resaltar que existen diferentes tipos de hiperplanos, pero aquel cuya distancia al punto más cercano es máxima se denomina "óptimo hiperplano de separación (OSH)" [HENAO et al., 2005].

Por otro lado, cuando los datos no son linealmente separables a la ecuación 4-36 se le puede agregar variables de relajación (ξ) [Vargas et al., 2012], de la siguiente manera:

$$y_i(\omega^T x_i + b) \geq 1 + \xi_i, \forall i \quad (4-37)$$

4.2.8. Métodos de corrección

- **Deconvolución espectral:** También conocida como desenfoque de la imagen, es una técnica computacional que permite compensar en parte la distorsión de la imagen causada por los artefactos. La mejora puede ser significativa tanto en términos de atenuación de la luz desenfocada y aumento de la resolución espacial. Dicho proceso consta de la reconstrucción o estimación de la imagen real a partir de la degradada como se ve en la figura 4-34, existen dos tipos de clasificaciones de estos algoritmos, la primera los divide en: métodos de filtrado espectral y métodos iterativos; por otro lado, la otra clasificación se divide en los algoritmos en métodos que no requieren ninguna información sobre el artefacto (también conocido como deconvolución ciega) y los métodos que necesitan esa información [Thakur and Datar, 2014], [Dey, 2018].

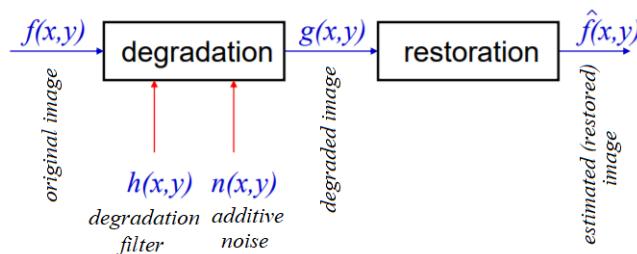


Figura 4-34: Proceso de restauración de una imagen [Dey, 2018]

La ecuación para expresar la imagen degradada se modela de la siguiente manera:

$$g = f * h + n \quad (4-38)$$

Donde, f es la imagen original y g es la versión que se ha degradado, debido a la convolución de la imagen con el filtro de degradación (h) y la adición de un ruido aleatorio n [Thakur and Datar, 2014].

Lo que se busca con la deconvolución es despejar f , es decir encontrar la imagen original o una aproximación a esta. Para despejarla, lo primero que se realiza es pasar la ecuación 4-38 a Fourier (ecuación 4-39).

$$G(u, v) = F(u, v) \cdot H(u, v) + N(u, V) \quad (4-39)$$

Hay diferentes formas de estimar la función de degradación (H) para poder restaurar las imágenes que tienen presente el artefacto de desenfoque por movimiento. Dicha estimación se puede hacer por modelado matemático que incluye métodos como el filtrado inverso, el filtrado de Wiener, el filtrado de mínimos cuadrados, entre otros [Thakur and Datar, 2014].

Por otro lado, para restaurar imágenes con reflejos especulares [Stehle, 2006] propuso un método para hacer una interpolación basada en la deconvolución espectral como se ve en la figura 4-35.

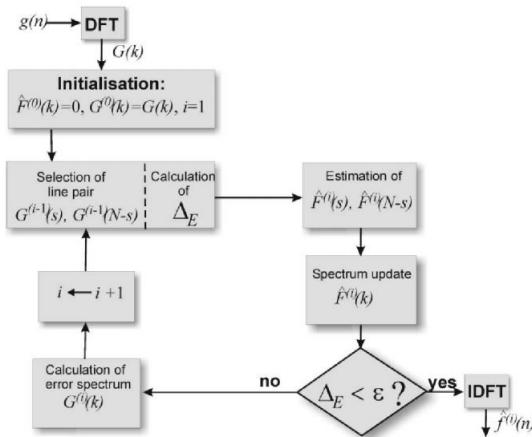


Figura 4-35: Proceso de restauración de una imagen con RE [Stehle, 2006]

Cabe resaltar que f se obtiene con la transformada inversa.

- **Inpainting:** Es una técnica que usa información espacial o frecuencial para quitar objetos no deseados o restaurar porciones dañadas de una imagen o vídeo. Existen diferentes métodos de inpainting y en la mayoría de casos la función requiere como entrada la

imagen y la máscara de las regiones que deben ser corregidas. [Shih and Rong-Chi Chang, 2005]

En el caso de [Meslouhi et al., 2011] se empleó un método de inpainting multi-resolución propuesto por [Shih and Rong-Chi Chang, 2005], el cual analiza los detalles de la sección afectada para así decidir qué información vecina debe utilizar. El nivel de detalles puede ser indicado por la varianza en la distribución del color en una porción de la imagen.

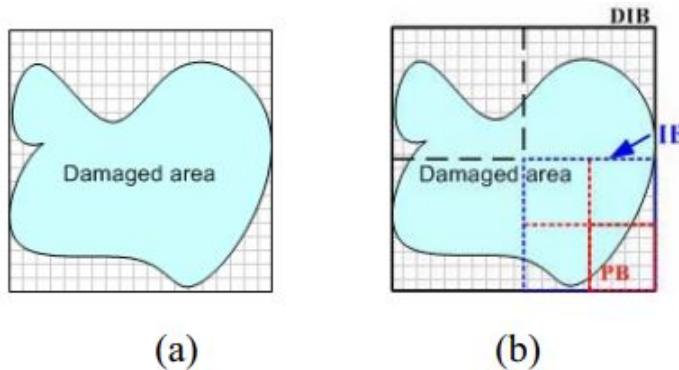


Figura 4-36: (a) Imagen con una zona afectada .(b) Bloque de la imagen afectada **DIB**, Bloque de la imagen **IB** y Bloque del píxel **PB** [Stehle, 2006]

Como se observa en la figura 4-36 la imagen afectada se va a denominar **DIB** (Damage Image Block), la cual será procesada por algoritmo como se explica a continuación:

1. Subdividir DIB en bloques de $n \times n$, los cuales se llamarán **IB** (Image Blocks) (figura 4-36).
2. Calcular la varianza (var) en cada bloque IB con la ecuación 4-40

$$\bar{x} = \frac{\sum_{\forall i} \sum_{\forall j} x_{ij}}{i \times j} \quad var = \frac{\sqrt{\sum_{\forall i} \sum_{\forall j} (x_{ij} - \bar{x})^2}}{i \times j - 1} \quad (4-40)$$

3. Establecer el valor de los umbrales
 - Varianza de los colores del píxel (α). El cual determina si se requiere aplicar el método de inpainting por multi-resolución. De acuerdo a los experimentos realizados por [Shih and Rong-Chi Chang, 2005] éste valor corresponde al 80 %.
 - Umbrales del porcentaje de píxeles afectados en **PB** (β_1 y β_2), donde ($\beta_1 < \beta_2$)

4. Si el valor de la *var* de cada **IB** calculada en el paso 2 es < al umbral α entonces se subdivide IB en **PB**, como se observa en la figura 4-36-b
5. Calcular el porcentaje de píxeles afectados en **PB** y llevar a cabo el proceso de corrección.
 - Si dicho valor es $> \alpha \beta_2$ se utiliza el color medio de IB para hacer la corrección.
 - Si dicho valor es $> \alpha \beta_1$ se utiliza el color medio de PB para hacer la corrección.
 - Si dicho valor es $< \alpha \beta_1$ se utilizan los píxeles vecinos para hacer la corrección.

4.2.9. Métricas de evaluación

- **Dice similarity coefficient (DSC):** en español coeficiente de similitud de datos, es un índice de superposición espacial y una métrica de validación de reproducibilidad; si el valor DICE da cercano a cero indica que no hay superposición espacial entre dos imágenes binarias, por otro lado, si es cercano a uno, indica una superposición completa (ver figura 4-37) [Zou et al., 2004], [Shamir et al., 2019]; este índice se calcula siguiendo la ecuación 4-41.

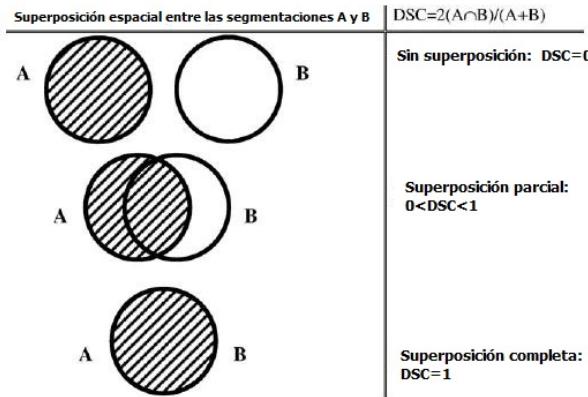


Figura 4-37: Dice similarity coefficient (DSC) [Zou et al., 2004]

$$2 \cdot \frac{|A \cap B|}{|A| + |B|} \quad (4-41)$$

Donde A representa el *ground truth* y B es la segmentación calculada, esta medida se calcula en cada uno de los píxeles de la imagen (a_i, b_i), de la siguiente manera:

$$|A \cap B| = \sum_i a_i b_i \quad (4-42)$$

$$|A| = \sum_i a_i \quad (4-43)$$

$$|B| = \sum_i b_i \quad (4-44)$$

- **Relación Señal a Ruido de Pico o PSNR:** es uno de los métodos más comunes para medir la calidad de las imágenes, en el que se expresa la relación entre el valor máximo posible (potencia) de una señal/imagen y la potencia de ruido de distorsión que afecta la calidad de su representación. Para definirla, primero se debe tener en cuenta el **error cuadrático medio (MSE. por sus siglas en inglés)** [Hore and Ziou, 2010]:

$$MSE(f, g) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (f(ij) - g(ij))^2 \quad (4-45)$$

Donde: (f) es la imagen de referencia y (g) es la imagen de prueba, ambas de un tamaño $M \times N$

Así, el PSNR se define como:

$$PSNR = 10 \cdot \log_{10} \frac{255^2}{MSE} \quad (4-46)$$

La ecuación 4-46 demuestra que un valor de PSNR alto representa una mayor calidad en la imagen.

- **Índice de similitud estructural (SSIM):** evalúa una imagen de prueba (X) con respeto a una imagen de referencia (Y) para cuantificar la similitud visual entre ellas, donde 0 indica que son completamente diferentes y 1 imágenes idénticas. En este sentido el SSIM cuantifica la calidad de la imagen X con respecto a Y mediante el cálculo de un índice espacial local que se define de la siguiente manera:

$$SSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [r(x, y)]^\gamma \quad (4-47)$$

Dónde: $l(x, y)$ es el índice de relacionado con las diferencias de luminancia $c(x, y)$ con las diferencias de contraste y $r(x, y)$ con las variaciones de estructura entre x y y.

4.3. Marco normativo

- Normatividad que respalda la calidad del algoritmo implementado para el análisis cualitativo y cuantitativo de las imágenes como la ISO/IEC 14598-5 y la ISO/IEC 25000
- Se tomó en cuenta la normatividad que respalda la calidad de las imágenes.

5 Metodología y Resultados

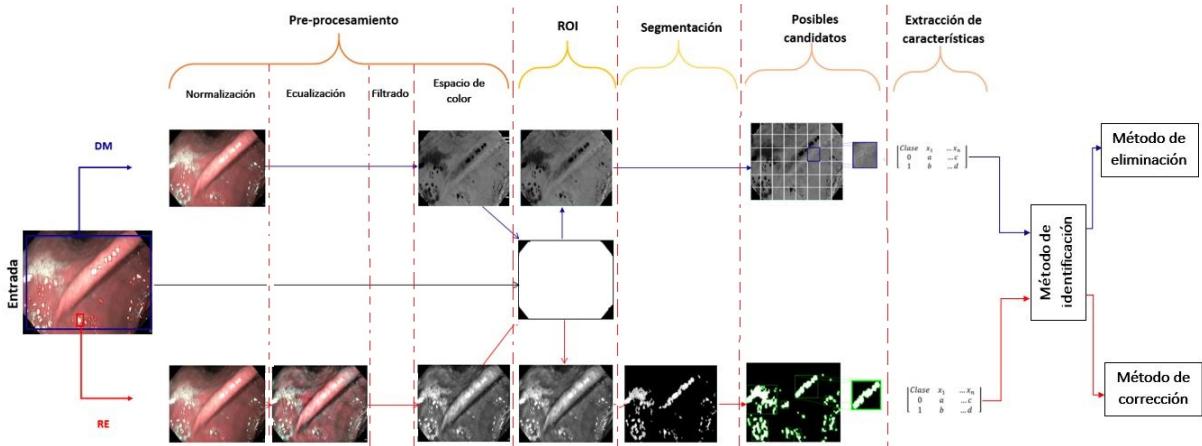


Figura 5-1: Fases de desarrollo del proyecto de grado

El proceso de desarrollo del método propuesto para la identificación y tratamiento de artefactos presentes en imágenes endoscópicas, contempló distintas fases como se evidencia en la figura 5-1. Este, inició con la descripción y caracterización de la base de datos, así como la asignación de etiquetas. Posteriormente, se efectuó una tarea de pre-procesamiento para selección de color, homogeneización de intensidades, la corrección de ruido, seguido de la disminución del área a analizar, es decir, la detección de la región de interés; después, se realizaron tareas de segmentación que permitieron obtener una imagen binaria preliminar que resalta los posibles artefactos. A dichos candidatos se les realizó el proceso de extracción de características que fueron la entrada del método de identificación.

Al tener la ubicación e identificación del artefacto se procedió a realizar la corrección o eliminación del marco, en caso de que la imagen tuviera RE se realizó la corrección, por el contrario si fue DM se procedió a eliminarla. Por ultimo, todo el desarrollo del método incluyó la evaluación de cada uno de los procesos descritos anteriormente.

5.1. Base de datos

Teniendo en cuenta que el objetivo general de este proyecto se centra en la detección de artefactos más no en la adquisición de las imágenes endoscópicas, y que la obtención de

estas junto con marcaciones de artefactos verificadas por especialistas requiere una cantidad de tiempo considerable, se optó por hacer uso de la base de datos Endoscopic Artefact detection Challenge (EAD) 2019, la cual contiene marcos de video endoscópico multi-tejido (gastroscopia, cistoscopia, endoscopia superior, colonoscopia) y multi-modal (luz blanca e imágenes de banda estrecha). Por ende, cabe resaltar que el desarrollo de este proyecto se vio limitado por las características y marcaciones de las imágenes que la componen.

Dichas imágenes fueron adquiridas en centros de salud de Francia, Italia y Suiza bajo el protocolo de imágenes estándar, sin embargo dado que fueron empleados diferentes tipos de resolución, las dimensiones de las imágenes son variables; además fueron etiquetadas en forma de cuadros delimitadores (bounding boxes - bbox) por dos especialistas (endoscopistas clínicos), quienes propusieron 7 tipos de artefactos (**reflejos especulares**, saturación, artefacto, **desenfoque por movimiento**, contraste, burbujas e instrumento). Es importante resaltar que de los siete artefactos enunciados, este trabajo de grado trabaja únicamente con RE y DM, que fueron marcados mediante el siguiente protocolo:

- Para una misma región se anotaron múltiples bbox si la región pertenecía a más de una clase.
- El tamaño mínimo de los bbox fueron usados para describir la región del artefacto, por ejemplo: si hay muchos reflejos especulares presentes en una imagen en vez de poner un bbox muy grande se ponen pequeños bbox para capturar el tamaño natural del artefacto.
- Cada tipo de artefacto es diferenciable y en general todos los conjuntos de datos de endoscopia.

La base de datos se divide en imágenes de entrenamiento y prueba, las primeras se dividen así:

- Imágenes de entrenamiento:

Como se menciona en la tabla **5-1**, en la fase de entrenamiento hay 9352 cuadros delimitadores o bbox, dichas anotaciones están en un formato .txt en el que se especifican las coordenadas x,y, altura y ancho del cuadro; además se indica la clase a la que pertenece. En la figura **5-2** se especifica cuántos bbox hay por artefacto.

Número frames	887
Instituciones de datos	<ul style="list-style-type: none"> -John Radcliffe Hospital -Lorraine Institute of Oncology -Ambroise Paré Hospital of Boulogne-Billancourt -Istituto Oncologico Veneto -University Hospital Vaudois -Botkin Clinical City Hospital
Modalidades	<ul style="list-style-type: none"> -Luz blanca -Fluorescencia -Imágenes de banda estrecha
Resolución	<ul style="list-style-type: none"> -Estándar HD. -Full HD. -Gran aumento.
Tipos de tejido	<ul style="list-style-type: none"> -Esófago -Estómago -Píloro -Recto -Ano -Colon -Vejiga
Clases de artefactos	<ul style="list-style-type: none"> -Reflejos especulares - Saturación -Artefacto -Desenfoque por movimiento -Contraste -Burbujas -Instrumento
Bounding Boxes	9.352

Tabla 5-1: Características de los datos de entrenamiento de la base de datos EAD 2019

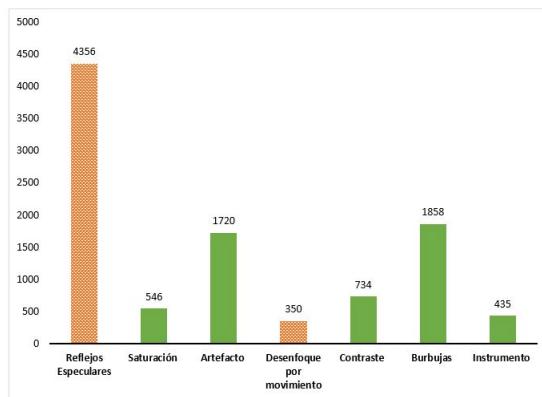


Figura 5-2: Número de Bounding box por artefacto

Las anotaciones para los bbox vienen dadas de la siguiente manera:

$< object-class >< x >< y >< width >< height >$ donde:

- $< object-class >$, corresponde al tipo de artefacto, se representa con un número entero de 0 a 6 (ver figura 5-2, 0: reflejos especulares 6: instrumento).
- $< x >< y >$, corresponde a las coordenadas del centro del bbox.
- $< width >$, corresponde al ancho normalizado del bbox.
- $< height >$, corresponde a la altura normalizada del bbox.

■ Imágenes de prueba:

- Número de cuadros de video / imágenes : 1.306.
- Instituciones de datos: JRC, UK/ICL, France/ APHP, France/ UHV, Switzerland/Botkin Clinical City Hospital, Russia.

■ Imágenes con los artefactos a trabajar:

Es importante tener en cuenta que la información descrita anteriormente de la base de datos contempla todos los 7 artefactos, sin embargo, para este proyecto de grado solo fueron relevantes las imágenes que contenían RE y DM; cuya cantidad puede ser observada en la tabla 5-2. Igualmente cabe aclarar que dichas imágenes contaron con 4706 anotaciones en forma de bbox (ver figura 5-3), donde la cantidad específica por clase puede ser consultada en la figura 5-2; sumado a esto para RE se contó con imágenes segmentadas (ground truth) como se visualiza en la figura 5-4.

	Entrenamiento	Prueba
DM	251	298
RE	352	343
RE y/o DM	533	553

Tabla 5-2: Cantidad de imágenes endoscópicas con RE-DM

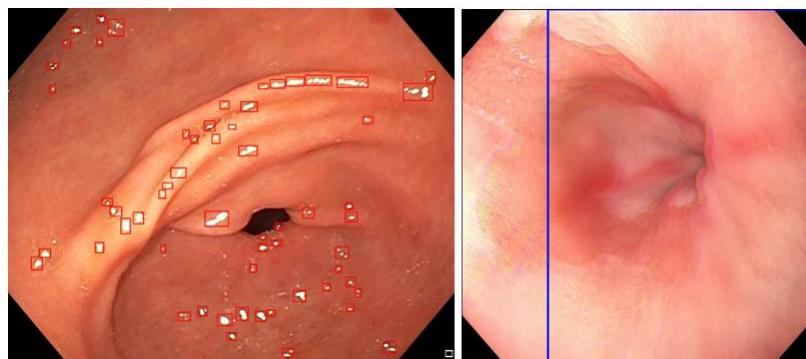


Figura 5-3: Anotaciones en forma de bbox, rojo (RE) y azul (DM)

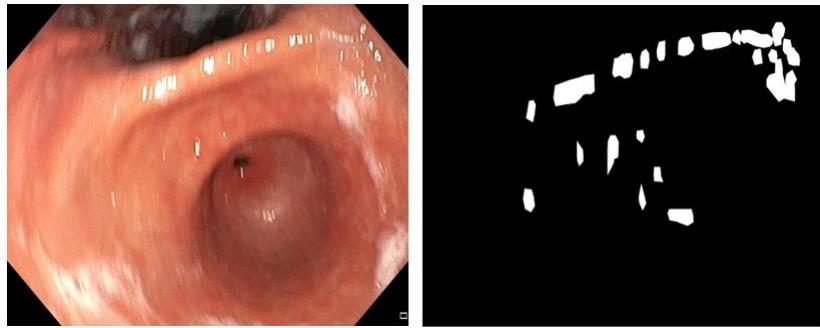


Figura 5-4: Segmentaciones entregadas por la base de datos

5.2. Pre-procesamiento

Una vez caracterizada la base de datos se llevó a cabo el preprocesamiento de las imágenes para mejorar la calidad de las mismas, lo cual incluyó técnicas de normalización, ecualización, filtrado y determinación del espacio de color. Es importante mencionar que, según la literatura consignada en la subsección 4.1, generalmente no se realiza la etapa de pre-procesamiento o se limita a la elección del espacio de color. No obstante, las técnicas inicialmente mencionadas se aplicaron y evaluaron puesto que hacen parte de la metodología usualmente empleada para el procesamiento de imágenes [Marques, 2011].

5.2.1. Normalización

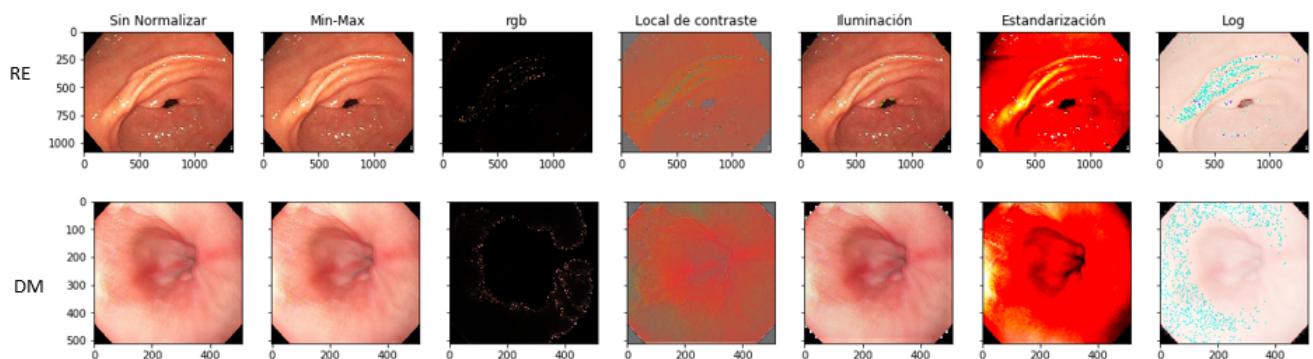


Figura 5-5: Normalizaciones aplicadas sobre las imágenes endoscópicas

Con la etapa de normalización se buscó cambiar diferentes parámetros de las imágenes a unos más convenientes y así disminuir el efecto de la variabilidad; para esto se probaron diferentes técnicas (ver figura 5-5), como lo son: la normalización min-max, de canales RGB, local de contraste, de iluminación, estandarización y logarítmica.

- **Min-Max:** este método de normalización lineal se caracteriza por ajustar los datos a un límite predefinido, en este caso a un rango de (0, 255), manteniendo la relación entre los datos originales y los deseados [Patro and Sahu, 2015]. A diferencia del proceso que se lleva a cabo en la ecualización del histograma, este método no tiene en cuenta la función de distribución de intensidades, por lo tanto ajusta los valores al límite definido sin depender del número de píxeles en cada nivel de gris del histograma. En la figura 5-6 se puede observar los resultados de la ejecución de dicha normalización, más específicamente se visualiza una imagen más clara; aunque el histograma no es la base conceptual en este método si es posible observar el cambio de distribución de intensidades, puesto que pasó de tener un valor máximo de 188 a uno normalizado de 255.

Dicha normalización transforma el rango de intensidad (*Min, Max*), es decir, transforma la intensidad más baja (*Min*) y la más alta (*Max*) presente en una imagen a nuevos valores más convenientes (*Numin, Numax*), mediante la siguiente ecuación:

$$I_N(i, i) = \frac{(I(i, j) - Min)}{(Max - Min)} \cdot (Numax - Numin) + Numin \quad (5-1)$$

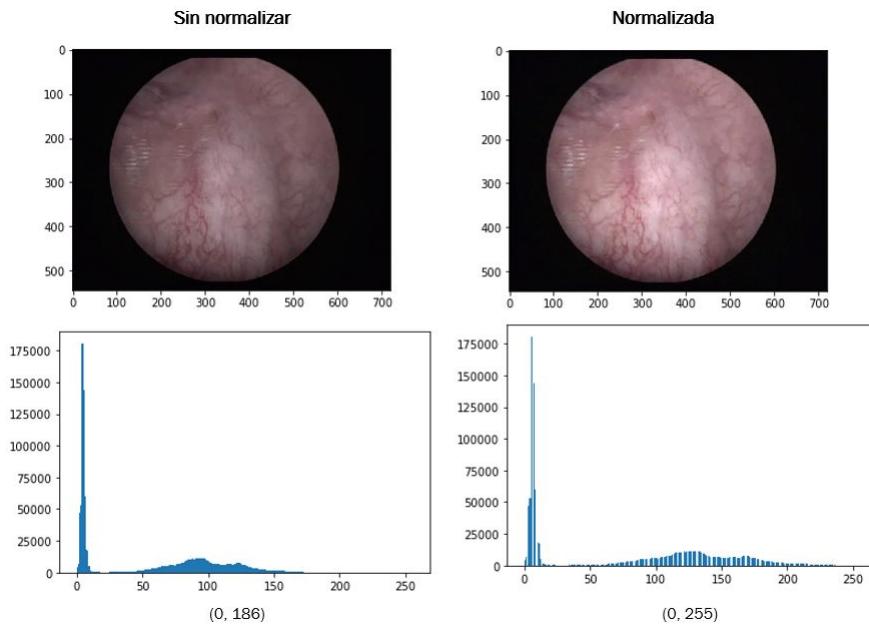


Figura 5-6: Normalización de Min-Max

- **Canales rgb:** por otro lado, se aplicó una normalización de canales RGB (R',G',B') siguiendo las ecuaciones 5-2, 5-3 y 5-4, dado que según literatura los canales normalizados (r,g,b) son menos sensibles a los cambios de iluminación, disminuyendo así efectos

de luz en la imagen [Kundu et al., 2018]. En la figura 5-7 se puede observar los resultados de la ejecución de dicha normalización, donde el máximo rango de intensidad del histograma disminuyó aproximadamente un 42 %, es decir que la mayor parte de los niveles de intensidad de la imagen se encuentran en tonalidades oscuras.

$$r(i,j) = \frac{R(i,j)}{R(i,j) + G(i,j) + B(i,j)} \quad (5-2)$$

$$g(i,j) = \frac{G(i,j)}{R(i,j) + G(i,j) + B(i,j)} \quad (5-3)$$

$$b(i,j) = \frac{B(i,j)}{R(i,j) + G(i,j) + B(i,j)} \quad (5-4)$$

La ecuación 5-4 también puede ser expresada como se muestra en la ecuación 5-5, por lo tanto B' se puede obtener solo con el porcentaje de los componentes R' y G' . Esta pérdida de información puede representar la información de color real de una imagen sin dependencia de brillo [Cheng et al., 2001]:

$$B' = 1 - R' - G' \quad (5-5)$$

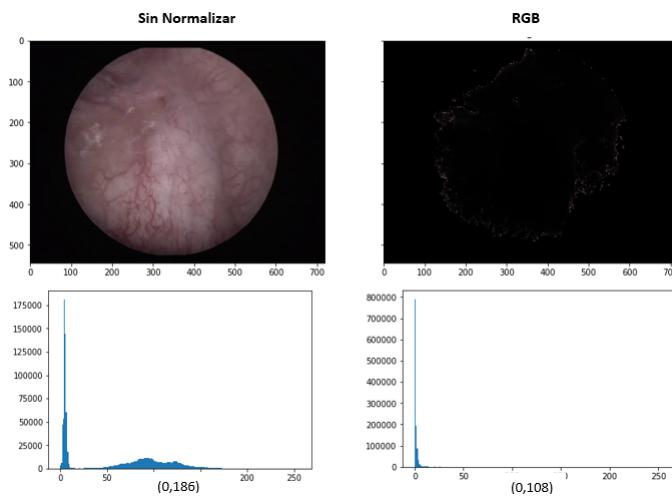


Figura 5-7: Normalización RGB

- **Contraste local (LCN):** también se aplicó un método de normalización de contraste no lineal, conocido como normalización de contraste local (LCN), el cual tiene como

propósito hacer que las características de la imagen sean más significativas; además permite corregir la iluminación o sombreado no uniforme [Jarrett et al., 2009]. En la figura 5-8 se puede observar que la imagen normalizada tiene un tono más uniforme y un rango de intensidad ajustado entre valores de 0 a 1.

Éste método realiza dos tipos de normalizaciones: sustractiva y divisoria. En la primera, para un punto dado X_{ijk} se calcula la ecuación 5-6.

$$v_{ijk} = x_{ijk} - \sum_{ipq} w_{pq} \cdot x_{i,j+p,k+q} \quad (5-6)$$

Dónde, w_{pq} es una ventana de ponderación Gaussiana (vecindad de 9x9), normalizada para que $\sum_{ipk} w_{pq} = 1$. En el caso de la divisoria, se calcula la ecuación 5-7.

$$y_{ijk} = \frac{v_{ijk}}{\max(c, \sigma_{jk})} \quad (5-7)$$

Dónde, $\sigma_{jk} = \sqrt{(\sum_{ipk} w_{pk} \cdot v_{i,j+p,k+q}^2)}$ y para cada muestra, la constante c está dada por la media(σ_{jk}).

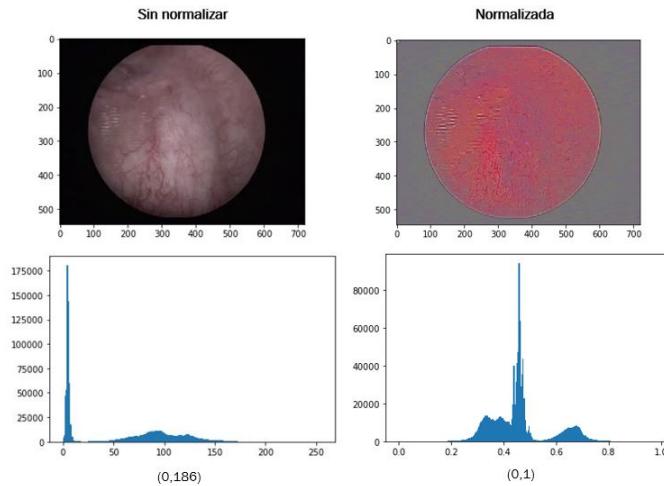


Figura 5-8: Normalización de contraste local

Sin embargo, tras la primera evaluación cualitativa de los resultados de ésta normalización se reconsideró la pertinencia de una imagen uniforme, dado que era posible que aumentara la dificultad para distinguir los artefactos del fondo.

- **Iluminación:** según el estudio realizado por [Sánchez-González and Soto, 2017], este tipo de normalización reduce los efectos de la iluminación desigual de una imagen. Para lograrlo se realizó una normalización local del brillo de la imagen tomando ventanas

de 20x20 píxeles y en cada una se calculó la media de cada canal multiplicada por un peso, en este caso de 0.05. Los resultados se pueden observar en la figura 5-9, en la cual no hay un cambio visual significativo, sin embargo el rango de intensidades se vio modificado, aumentando el valor máximo de 186 a 255.

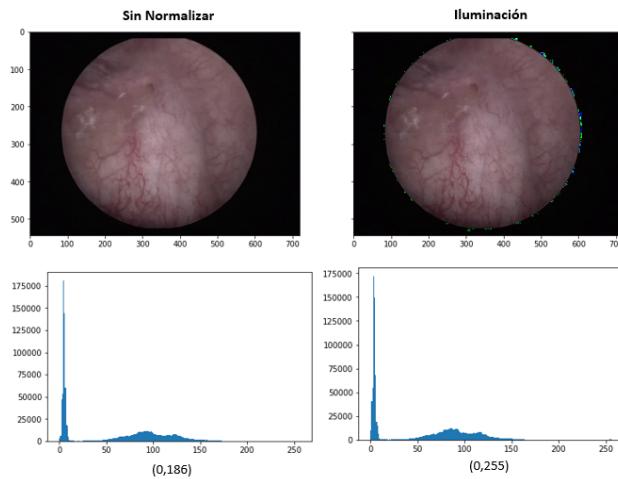


Figura 5-9: Normalización de iluminación

- **Estandarización:** es una técnica de escalado que transforma la distribución de los datos de intensidad de la imagen a una distribución gaussiana, donde se obtiene una media igual a uno y una desviación estándar de cero [Lincoln University, 2014], por medio de la ecuación 5-8. El resultado tras aplicar esta normalización se puede observar en la figura 5-10, donde todas las intensidades de los píxeles se vieron afectadas obteniendo un cambio drástico en el rango de intensidades, puesto que pasó de 0 a 186 a un rango de -2.93 a 5.66.

$$E = \frac{I - \bar{I}}{\sigma} \quad (5-8)$$

Dónde, la estandarización (E) es la resta de cada píxel de la imagen (I) con la media (\bar{I}), dividido por la desviación estándar (σ).

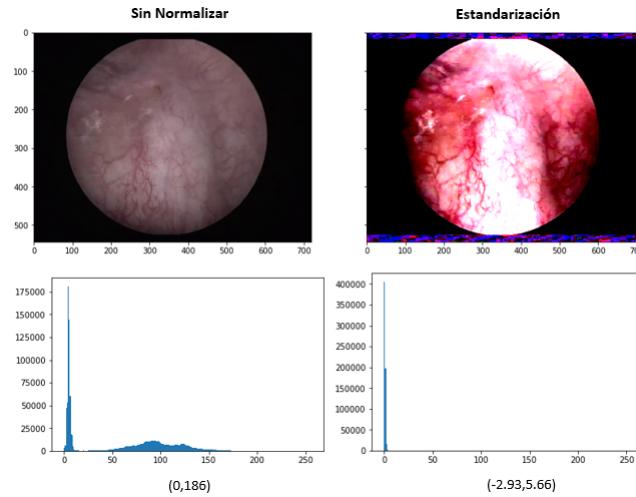


Figura 5-10: Estandarización

- **Transformación logarítmica:** el operador logarítmico es una transformación en dónde la función de mapeo es una curva logarítmica. En otras palabras, cada píxel es reemplazado según la ecuación 5-9. En la figura 5-11 se puede observar como los píxeles de baja intensidad son transformados a valores de intensidad más altos, debido a la transformación logarítmica de cada píxel. Esto también puede ser observado en el histograma, dado que la distribución de intensidades en la imagen normalizada está en su mayoría en la zona clara (cercano a 255).

$$Q(i, j) = c \log(|P(i, j)|) \quad (5-9)$$

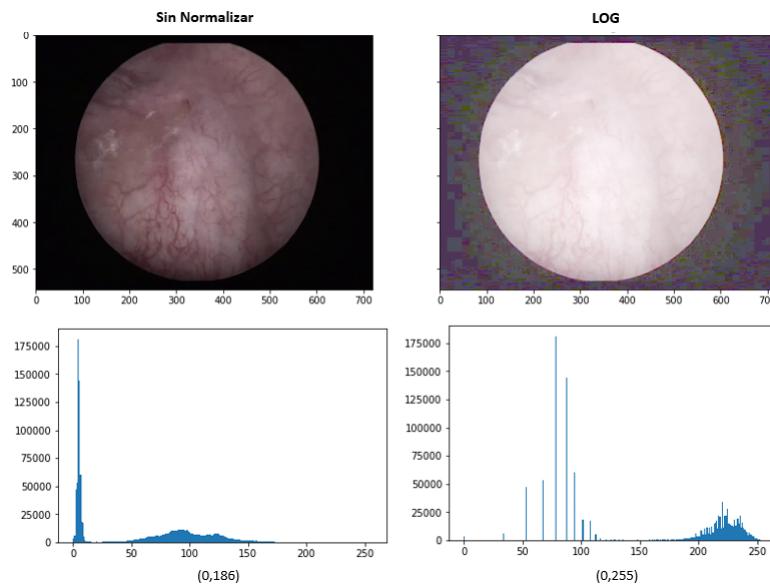


Figura 5-11: Normalización Logarítmica

Evaluación de resultados

Tras la evaluación cualitativa de cada normalización, se esperó que las técnicas más pertinentes para aplicar en el presente trabajo fueran **normalización min-max**, **iluminación** y **transformación logarítmica**. Sin embargo, teniendo en cuenta que ésta es una evaluación subjetiva se optó por realizar una evaluación cuantitativa y así identificar la técnica que mejor eliminó la variabilidad entre datos; cuyos resultados de media y desviación estándar de los rangos de intensidad se evidencian en la tabla 5-3. De modo que, para DM las técnicas que obtuvieron un mejor desempeño (menor desviación) fueron: Min-Max e iluminación; por otro lado, para RE, todas las técnicas excepto rgb, alcanzaron una desviación estándar de 0.

Rangos de intensidad			
Media ± Desviación Estándar			
	RE	DM	
	Min	Max	Min
Sin Normalizar	0 ± 0	255 ± 0	0 ± 0
Min-Max	0 ± 0	255 ± 0	0 ± 0
rgb	0 ± 0	$247,4 \pm 12,5$	0 ± 0
LCN	0 ± 0	1 ± 0	0 ± 0
Iluminación	0 ± 0	255 ± 0	0 ± 0
Estandarización	0 ± 0	255 ± 0	$-2,7 \pm 1$
Log	0 ± 0	244 ± 0	0 ± 0
			$254,3 \pm 0,3$

Tabla 5-3: Rangos de intensidad

Teniendo en cuenta que los resultados obtenidos anteriormente son bastante cercanos, se decidió evaluar cada normalización que obtuvo la menor variabilidad en los rangos de intensidad en términos de similitud, por consiguiente se calculó: correlación, distancia Bhattacharyya y distancia Euclíadiana de los histogramas de la imagen normalizada con artefacto y la misma imagen normalizada pero sin artefacto (ver figura 5-12). Con esto se buscó encontrar la técnica que más **resaltara** las regiones con artefactos, es decir, la normalización que menor correlación y mayor distancia arrojara es la que mejor permitía diferenciar entre la región con artefactos del fondo.

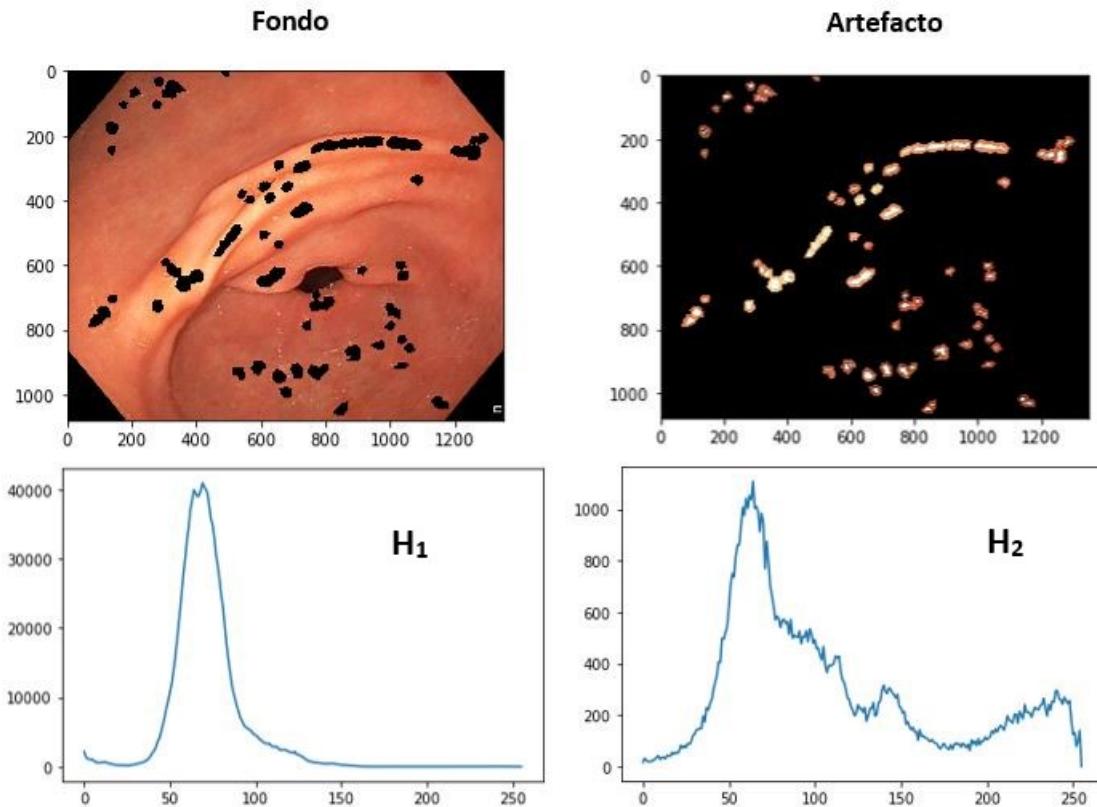


Figura 5-12: Diferencia entre el fondo y la región con artefactos

Las medidas anteriormente mencionadas se calcularon de la siguiente manera:

- Correlación entre histogramas: ésta es una métrica que describe la dependencia lineal entre un histograma y otro, en este caso entre el histograma del fondo (H_1) y el del artefacto (H_2). El coeficiente de correlación está dado por la ecuación 5-10 y su rango oscila entre 0 y 1, donde 0 indica que no existe relación lineal entre los dos histogramas [Briñez de León et al., 2012].

$$r = \frac{\sum_I (H_1(I) - \bar{H}_1)(H_2(I) - \bar{H}_2)}{\sqrt{(\sum_I (H_1(I) - \bar{H}_1)^2)(\sum_I (H_2(I) - \bar{H}_2)^2)}} \quad (5-10)$$

Donde H_1 y H_2 son los histogramas a comparar y \bar{H}_1 y \bar{H}_2 son las medias de dichos histogramas

- Distancia Bhattacharyya: evalúa la distancia (similitud) entre los dos histogramas de las imágenes que están siendo comparadas. Su valor está dado por la ecuación 5-11 y se

encuentra en un rango de 0 a 1, dónde 1 representa la distancia más lejana [Briñez de León et al., 2012].

$$d = \sqrt{1 - \frac{1}{\sqrt{H_1 H_2 N^2}} \sum_I \sqrt{H_1(I) H_2(I)}} \quad (5-11)$$

- Distancia Euclíadiana: generalmente, ésta es una expresión que se emplea para evaluar distancias entre espacios numéricos. En éste caso, compara los histogramas en cada plano R,G,B de la imagen [Briñez de León et al., 2012] y su valor está dado por la expresión 5-12.

$$L = \sqrt{\sum_m^M (H_1 - H_2)^2} \quad (5-12)$$

A partir de los resultados presentados en la tabla 5-3 y 5-4 se escogió la técnica de normalización más conveniente, es decir la técnica con menor variabilidad en los rangos de intensidad y así mismo, como un valor agregado, la que menor correlación y mayor distancia tuviera; que en éste caso para los dos artefactos fue la **normalización min-max**.

	Medidas de distancia					
	Media ± Desviación Estándar					
	Correlación		Bhattacharyya		Euclíadiana	
	RE	DM	RE	DM	RE	DM
Sin Normalizar	0,26 ± 0,35	0,44 ± 0,29	0,66 ± 0,19	0,56 ± 0,17	120507 ± 99050	340106 ± 334514
Min-Max	0,26 ± 0,35	0,43 ± 0,29	0,66 ± 0,19	0,56 ± 0,17	120507 ± 99050	340106 ± 334514
Iluminación	0,26 ± 0,38	0,44 ± 0,29	0,66 ± 0,19	0,55 ± 0,17	118609 ± 89861	339881 ± 335423
Log	0,48 ± 0,32	0,57 ± 0,28	0,53 ± 0,19	0,50 ± 0,18	157377 ± 105505	338498 ± 327944

Tabla 5-4: Medidas de evaluación normalización

Una vez determinada la técnica de normalización más conveniente para cada artefacto, se consideró importante destacarlos aún más del fondo, por lo tanto se decidió aumentar el contraste mediante la aplicación de diferentes tipos de ecualización, como se muestra a continuación.

Cabe señalar que, todos los resultados de las métricas de similitud y rangos de intensidad de todas las normalizaciones se encuentran en la carpeta: anexos/tablas/Normalizacion.xlsx.

5.2.2. Ecualización

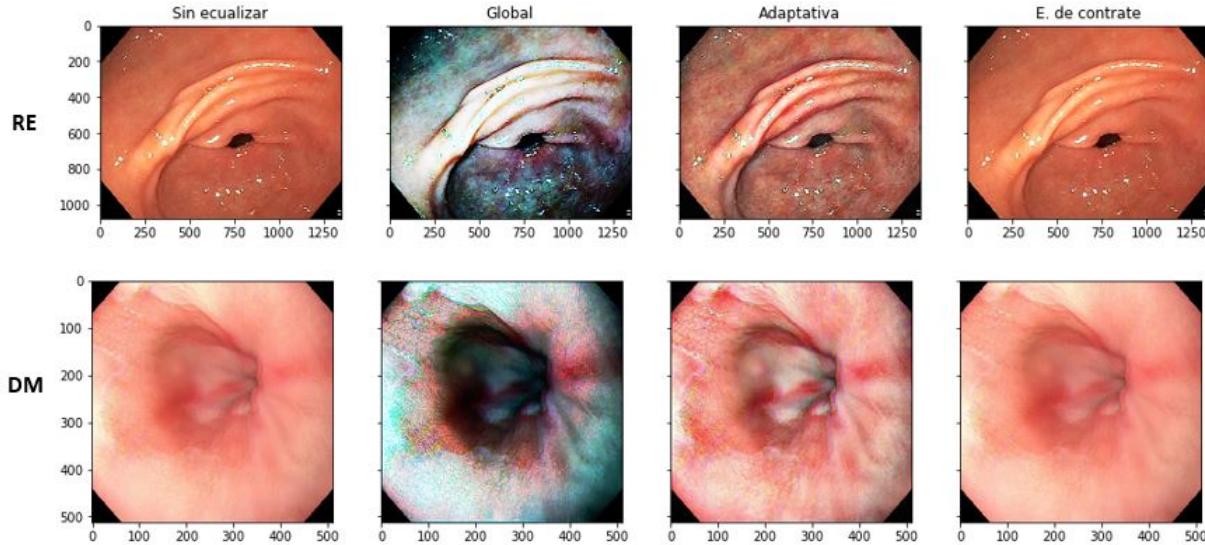


Figura 5-13: Imágenes ecualizadas

En ésta fase se aplicaron tres tipos de ecualización: ecualización global del histograma, ecualización adaptativa y estiramiento del histograma (Contrast Stretching, por su nombre en inglés); las cuales fueron evaluadas de acuerdo al contraste y las métricas de similitud utilizadas para la elección de la normalización, esto para determinar la técnica (o ninguna) más conveniente para este proyecto de grado.

- **Ecualización global del histograma:** se realiza trabajando sobre la función de distribución acumulativa (FDC), la cual está dada por la ecuación 5-13. Una vez calculada la función de mapeo (FDC), ésta es aplicada a todos los píxeles de la imagen de entrada. En la figura 5-14 se observa que los píxeles están agrupados en su mayoría en los niveles de intensidad más alto, esto puede deberse a que la ecualización global no se adapta a las características de brillo local de la imagen [Marques, 2011]; lo cual indica que no es la mejor técnica para resaltar los artefactos del fondo.

$$s_k = \sum_{j=0}^k p(r_j) \quad (5-13)$$

Donde, (s_k) es el nuevo nivel de gris calculado para todos los píxeles, cuyo nivel de gris solía ser (r_k).

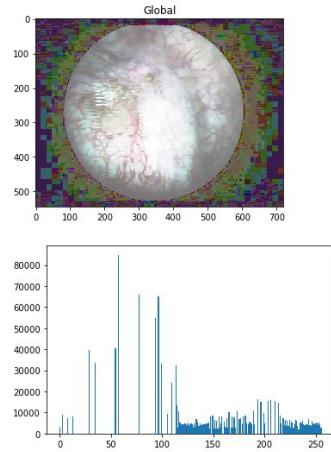


Figura 5-14: Ecualización global

- **Ecualización local o ecualización adaptativa del histograma limitada por contraste (CLAHE):** tomando en cuenta que la técnica anterior no arrojó cualitativamente los resultados esperados, se aplicó CLAHE, pues esta es sensible a las características de brillo local de la imagen. Dicha técnica consiste en tomar una ventana rectangular (conocida como *tile*) que recorre toda la imagen. En *python*, el tamaño por defecto de ésta ventana es de 8x8 y es la que delimita la región vecina en la que se calcula el histograma.

Para cada píxel en la imagen (alineado con el centro de la ventana) se calculó el histograma y se normalizó (principal diferencia con la ecualización global), posteriormente se calculó la **FDC** con la ecuación 5-13 y el píxel de referencia toma el nuevo valor determinado por la pendiente (m) de la función de transformación.

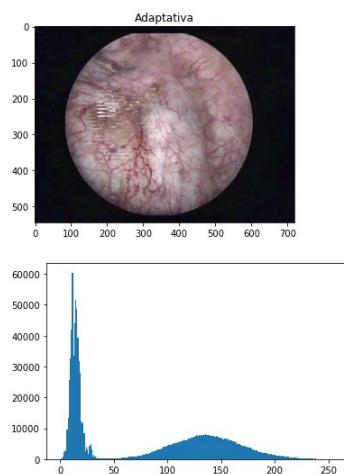


Figura 5-15: Ecualización Adaptativa

En la figura 5-15 se puede observar los resultados tras aplicar CLAHE, que en comparación con la ecualización global evitó el sobre-brillo y conservó la distribución de las intensidades, como se observa en el histograma.

- **Estiramiento del contraste o ecualización del histograma:** ésta técnica se aplicó para ampliar el rango de intensidades y así abarcar todos los píxeles presentes en las imágenes (transformación lineal) [Singh and Patel, 2017]. Para éste caso, se empleó un método en el que la imagen es re-escalada para que incluya todas las intensidades que estén dentro del percentil 2 y 98.

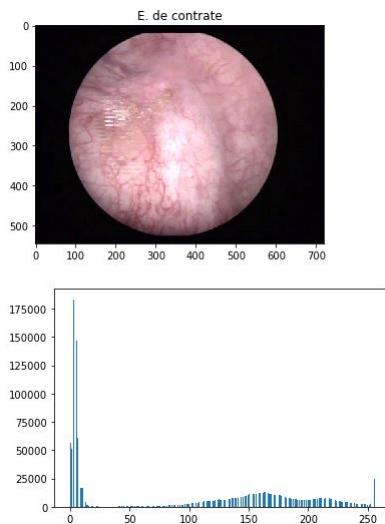


Figura 5-16: Estiramiento del contraste

En la figura 5-16 se puede observar que tras aplicar el estiramiento de contraste la distribución de intensidades se encuentra en la región más clara del histograma para aumentar así el contraste. Además, visualmente, es la imagen con menor cambio respecto a la imagen de entrada, esto se debe a que ésta técnica aplica una transformación con una relación uno a uno.

Evaluación de resultados

Debido a que con las diferentes técnicas de ecualización aplicadas se buscó mejorar la calidad de la imagen en términos de contraste, se evaluó este ítem tanto en la imagen original como en las ecualizadas, tras aplicar cada técnica se obtuvieron los resultados observados en la tabla 5-5, de la que se resalta la ecualización global como posible candidato a la técnica que mejor desempeño tiene (en términos de contraste).

	Niveles de contraste	
	Media ± Desviación Estándar	
	RE	DM
Sin ecualizar	9,33 ± 1,03	8,83 ± 1,43
E.Global	11,05 ± 0,30	10,97 ± 0,56
Adaptativa	9,90 ± 0,86	9,51 ± 1,21
E. contraste	9,53 ± 1,01	9,64 ± 1,38

Tabla 5-5: Niveles de contraste

Sin embargo, también es importante evaluar que éstas técnicas no homogenicen los artefactos de interés presentes en las imágenes. Para llevar a cabo lo anterior, se evaluó cada técnica en términos de similitud (correlación, distancia Bhattacharyya, distancia Euclíadiana), cuyos resultados pueden ser evidenciados en la tabla **5-6** en dónde se muestra la media y desviación estándar de cada métrica de similitud empleada.

	Medidas de distancia					
	Media ± Desviación Estándar					
	Correlación		Bhattacharyya		Euclíadiana	
	RE	DM	RE	DM	RE	DM
Sin ecualizar	0,26 ± 0,35	0,44 ± 0,29	0,66 ± 0,19	0,56 ± 0,17	120483 ± 99036	340106 ± 334514
Global	0,47 ± 0,20	0,43 ± 0,28	0,62 ± 0,16	0,55 ± 0,18	122468 ± 99450	45133 ± 40070
Adaptativa	0,16 ± 0,31	0,41 ± 0,26	0,66 ± 0,19	0,55 ± 0,16	115969 ± 59441	55915 ± 99994
E. Contraste	0,25 ± 0,34	0,40 ± 0,23	0,66 ± 0,19	0,57 ± 0,20	122223 ± 99075	45509 ± 38762

Tabla 5-6: Medidas de evaluación ecualización

Según los datos observados en la tabla **5-5**, la ecualización global es la que mayor contraste tiene; **sin embargo** esta aun no es la decisión final pues al evaluar las medidas de similitud entre el artefacto y el fondo, es posible observar en la tabla **5-6** que los resultados de correlación para esta ecualización son los más altos, tanto en RE como en DM, es decir que se homogenizó la imagen; en otras palabras, la región con artefactos ya no resaltaba del fondo. De manera general, para DM tras aplicar las técnicas de ecualización hay un aumento en la correlación con respecto a la imagen original, por ende, no se escogió ningún tipo de ecualización para dicho artefacto y con esto lograr evitar la pérdida de características importantes que permitan la identificación de DM. Por el contrario, para RE la correlación disminuyó y el valor de las distancias aumentó con la **ecualización adaptativa**, por lo cual fue la escogida. Cabe señalar que, todos los resultados de las métricas de similitud y contraste de todas las ecualizaciones se encuentran en la carpeta: **anexos/tablas/Ecuallizacion.xlsx**.

5.2.3. Filtrado

Cabe resaltar que en ninguno de los artículos consignados en la sección 4.1 se aplicó alguna técnica de filtrado. Sin embargo, era importante determinar si esta etapa aportaba algún

beneficio en términos de separabilidad entre el artefacto y el fondo, por lo mismo, como se observa en la figura 5-17 se probaron las siguientes técnicas de filtrado: auto nivel, mediana, mínimo, Gaussiano y media.

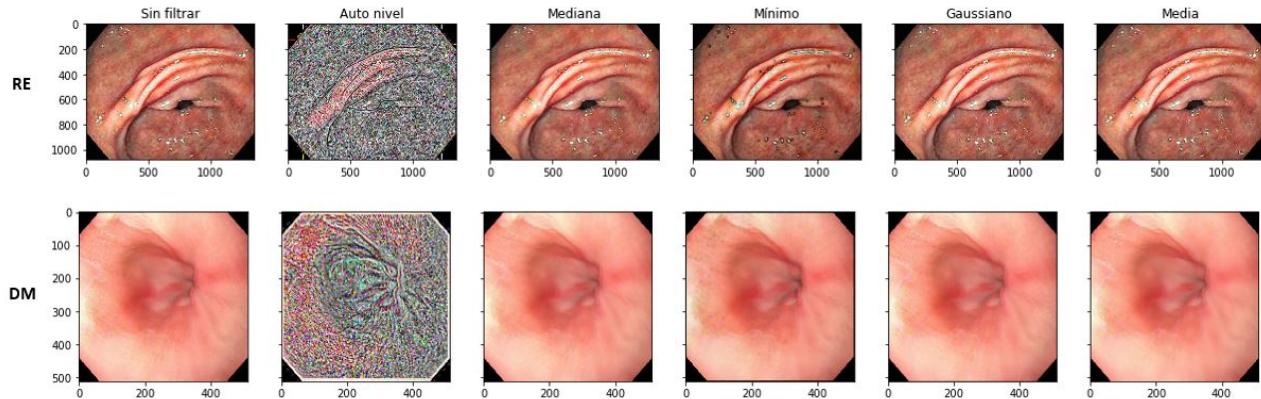


Figura 5-17: Imágenes filtradas

- **Nivelación automática local (auto nivel):** buscando resaltar los valores atípicos de las imágenes, en este caso resaltar las áreas con reflejos especulares, se probó el filtro de nivelación automática local, el cual es un filtro de rango no lineal que se calcula a partir del histograma local de nivel de gris y es muy sensible a resaltar los valores atípicos de la imagen [Soille, 2002]. Sin embargo, como se observa en la figura 5-17 el filtro no resalta ningún artefacto de interés para éste trabajo de grado, por el contrario homogeniza la imagen.
- **Mediana:** ésta es una técnica de filtrado no lineal que se suele emplear para eliminar ruido y bajo ciertas condiciones preserva los bordes de la imagen [Gupta and chandel, 2013]. Para ello se reemplazó un nivel de gris de un punto por la mediana de los niveles de gris de una vecindad de 3×3 .
- **Mínimo:** filtro morfológico que se empleó para reemplazar el píxel del centro con el más oscuro de la siguiente ventana (neighborhood) [Gupta and chandel, 2013]. En la figura 5-17 se observa que tras aplicar éste filtro se resaltan los RE con un color verde oscuro, lo cual fue un primer indicio cualitativo de que esto ayudaría a diferenciar este artefacto del fondo.
- **Media:** reduce la variación de intensidad entre píxeles vecinos al tomar cada píxel de la imagen y reemplazarlo por la media de los píxeles vecinos [Gupta and chandel, 2013], este filtro operó mediante la convolución de las imágenes con una máscara de 5×5 , dando como resultado una imagen suavizada.
- **Gaussiano:** es similar al filtro de media, pero se usó una máscara de 3×3 modelada con una función Gaussiana, como se muestra en la ecuación 5-14 [Gupta and chandel,

2013], lo cual dio como resultado una imagen suavizada.

$$G(x, y) = \frac{1}{\sqrt{(2\pi)\sigma}} e^{\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \quad (5-14)$$

Cabe resaltar que los filtros anteriores se probaron con tamaños de máscaras pequeñas, pues al aplicar tamaños grandes se homogenizaba la imagen, es decir, que los RE señalados por los bbox rojos en la figura 5-18-A ya no resaltaban del fondo tras aplicar el filtro de media, como se ve en la figura 5-18-B.

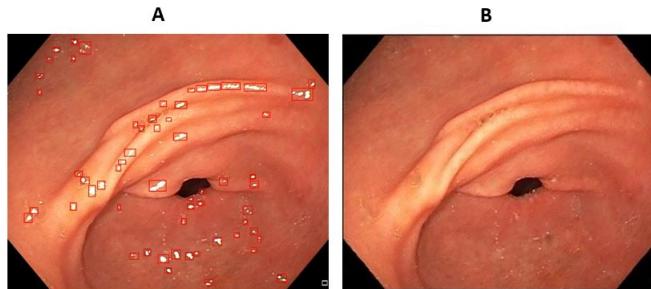


Figura 5-18: A) Imagen con bbox de RE, B) Imagen con filtro de media con máscara de 20x20

Evaluación de resultados

Los filtros probados se evaluaron con las medidas de similitud empleadas para la elección de la técnica de normalización y ecualización. Los resultados en la tabla 5-7 demostraron que tras aplicar los filtros la correlación aumentó y las distancia disminuyó, por ejemplo el filtro mínimo, que se esperaba obtuviera buenos resultados, para RE tuvo una correlación más alta que la imagen sin filtrar, lo cual indicó que de utilizar éste filtro se dificultaría la separación entre el artefacto del fondo. Cabe destacar que éste resultado fue repetitivo en todas las técnicas anteriormente empleadas, en consecuencia se decidió eliminar esta etapa del método.

Medidas de distancia						
Media ± Desviación Estándar						
Correlación		Bhattacharyya		Euclidianas		
RE	DM	RE	DM	RE	DM	
Sin filtrar	0,16 ± 0,31	0,44 ± 0,29	0,66 ± 0,19	0,56 ± 0,17	115969 ± 59441	340106 ± 334514
Auto nivel	1,00 ± 0,10	0,95 ± 0,18	0,40 ± 0,23	0,33 ± 0,38	84642 ± 39047	51883 ± 40672
Mediana	0,18 ± 0,32	0,44 ± 0,28	0,66 ± 0,20	0,55 ± 0,18	116536 ± 59435	44702 ± 37228
Mínimo	0,20 ± 0,33	0,44 ± 0,29	0,63 ± 0,16	0,56 ± 0,18	120081 ± 61356	44361 ± 36899
Gaussiano	0,18 ± 0,32	0,44 ± 0,28	0,65 ± 0,19	0,55 ± 0,17	111225 ± 57958	44503 ± 37119
Media	0,18 ± 0,43	0,45 ± 0,28	0,66 ± 0,20	0,55 ± 0,17	111900 ± 57395	44876 ± 36999

Tabla 5-7: Medidas de evaluación filtrado

Tras identificar que ninguna técnica de filtrado fue conveniente para el desarrollo del método, se procedió a realizar la última etapa del preprocesamiento propuesto, que es la elección del espacio de color, como se muestra a continuación.

Cabe señalar que, todos los resultados de las métricas de similitud de todas los filtros probados se encuentran en la carpeta: **anexos/tablas/Filtrado.xlsx**.

5.2.4. Espacio de color

Para la determinación del espacio de color sobre el cual se trabajarán las imágenes, se evaluaron los espacios de color que fueron usados en el estado del arte (RGB, YUV, HSV, CIE-XYZ) como se observa en las figuras **5-19** y **5-20**, donde se buscó encontrar el canal en que el artefacto resaltaba más del fondo.

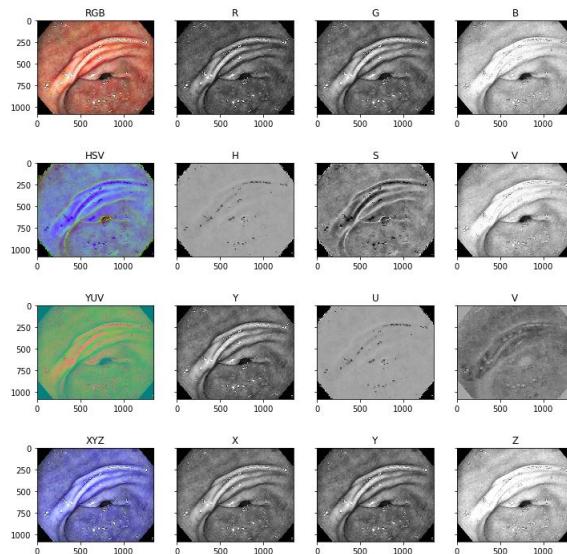


Figura 5-19: Distintos espacios de color con sus componentes para RE

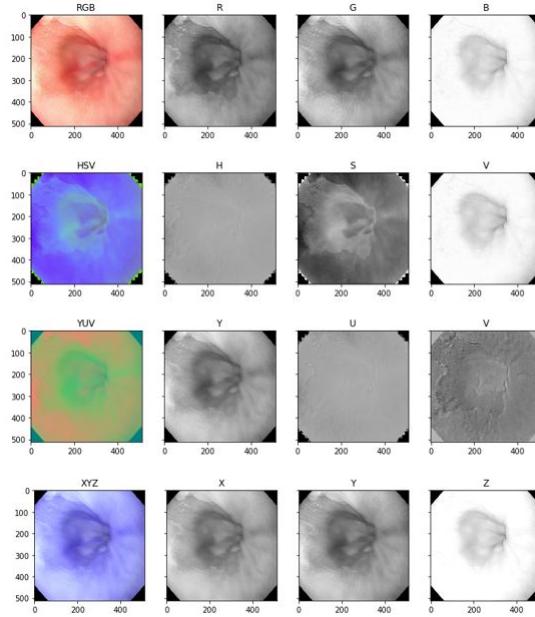


Figura 5-20: Distintos espacios de color con sus componentes para DM

Evaluación de resultados

En cada canal se evaluó la similitud entre el artefacto y el fondo mediante las métricas empleadas en etapas anteriores (correlación y distancia Bhattacharyya). De la figura 5-21-RE se puede observar que los canales con menor correlación y mayor distancia (media) entre fondo y los artefactos, son el HSV, B y G, sin embargo, los resultados mostrados en la figura 5-22-RE muestran que el canal con menor desviación estándar es el B y el G.

En el caso de DM, de acuerdo a la figura 5-21-DM los canales con menor correlación son el B y HSV. Por otra parte, en la figura 5-22-DM se puede observar que las desviaciones estándar están muy cercanas, por lo tanto se decidió mantener estos canales como posibles candidatos y seleccionarlos de acuerdo al desempeño de cada uno en la evaluación de separabilidad de características.

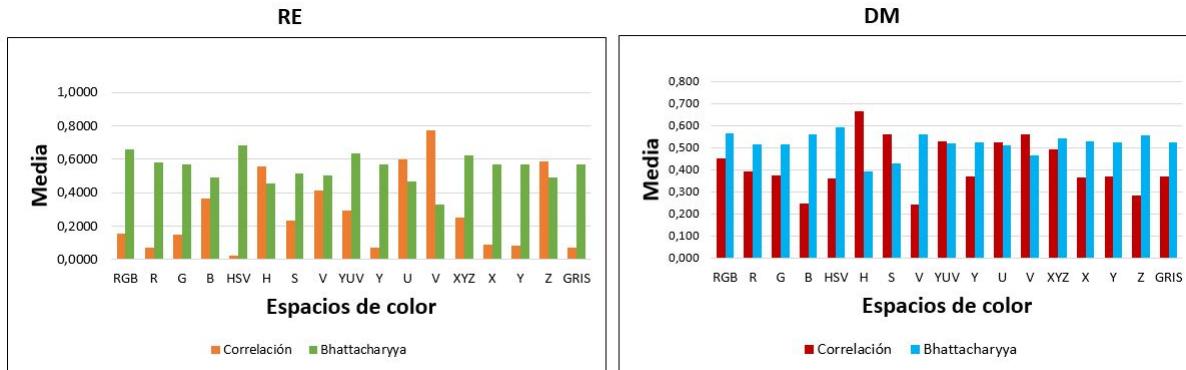


Figura 5-21: Media de medidas de correlación y distancia en diferentes espacios de color

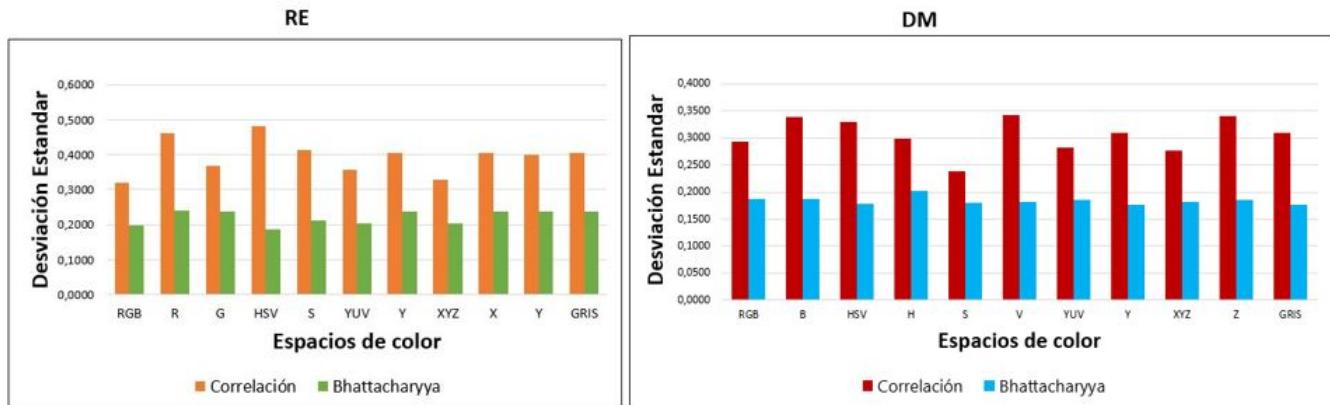


Figura 5-22: Desviación estándar de medidas de correlación y distancia en diferentes espacios de color

Con la evaluación del espacio de color se da por terminada la etapa de preprocesamiento, que para el desarrollo de éste método consistió en:

- RE: normalización min-max, ecualización CLAHE, canal G y B del espacio de color RGB.
- DM: normalización min-max, y espacio de color pendiente por definir.

Cabe señalar que, todos los resultados de las métricas de similitud de todas los espacios de color probados se encuentran en la carpeta: **anexos/tablas/Espacio de color.xlsx**.

5.3. Región de Interés (ROI)

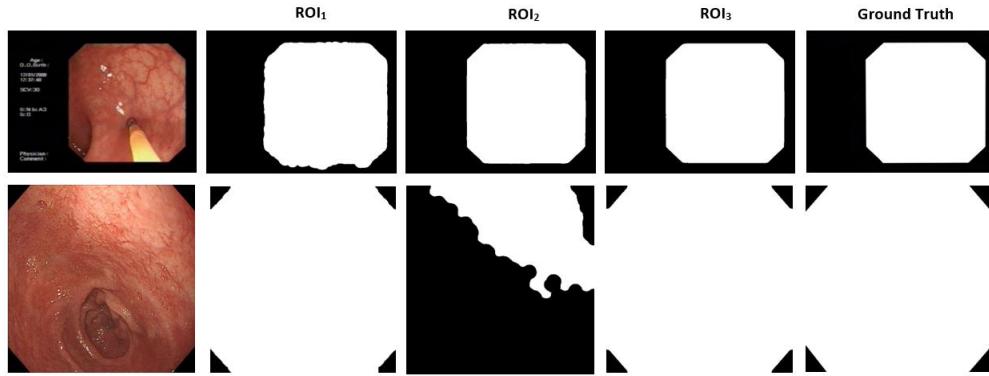


Figura 5-23: Diferentes regiones de interés

Con la región de interés (ROI) se buscó reducir las zonas a analizar en la imagen, asimismo eliminar ciertas etiquetas que suelen estar presentes en los bordes de algunas imágenes endoscópicas como se observa en la figura 5-23. Para la detección automática de dicha ROI se probaron diferentes métodos en tres espacios de color (RGB, HSV, XYZ) como se ve a continuación, y posteriormente, se evaluó la similitud entre las imágenes binarias obtenidas y los *ground truth*, dicha evaluación se realizó con el cálculo del coeficiente de similitud *DICE*, el índice de similitud estructural (*SSIM*) y el error cuadrático medio (*MSE*) (técnicas explicadas en la sección 4.3). Es preciso mencionar que, las imágenes control (*ground truth*) fueron realizadas manualmente dado que la base de datos EAD2019 no las proporcionaba, se realizaron en total 306 imágenes control.

Todos los métodos probados para la de detección de la ROI implementarán una umbralización simple por niveles de intensidad de la imagen (I), que permitió crear una máscara binaria (m) a partir de un umbral (T), como se observa en la ecuación 5-15.

$$m(i, j) = \begin{cases} 1 & \text{si } I(i, j) > T \\ 0 & \text{si } I(i, j) \leq T \end{cases} \quad (5-15)$$

A continuación se explican tres métodos diferentes para calcular el valor de T :

- **ROI₁:** la detección de esta ROI se basó en el artículo [Sánchez-González and Soto, 2017], en el que primero se realizó cada transformación de color mencionada anteriormente (RGB, HSV y XYZ), seguido de la aplicación de un umbral fijo de 0.03 en el canal B, V y Z, respectivamente. Sin embargo, como se observa en la imagen original de la figura 5-24, algunos bordes contenían información que permanecía visible tras aplicar la umbralización, para eliminarlos se empleó una transformación morfológica

de erosión con un kernel de 5x5. Tras obtener la mascara binaria, esta se multiplicó con cada uno de los canales de la imagen original (RGB) para así obtener la región de interés. Las medidas de similitud obtenidas en cada canal se muestran en la tabla 5-8.

	Medidas		
	DICE	SSIM	MSE
Canal B (RGB)	0,97±0,09	0,97±0,05	1139±2694
Canal V (HSV)	0,97±0,08	0,97±0,05	1424±3153
Canal Z (XYZ)	0,98±0,06	0,97±0,05	1190±2863

Tabla 5-8: Medidas de similitud ROI₁

Cabe resaltar que en la figura 5-24 se puede observar la metodología que se siguió para obtener la ROI a partir del canal Z, que según los datos consignados en la tabla 5-8 es el canal con los mejores resultados de las métricas de similitud, puesto que el valor de DICE y SSIM fue el más cercano a 1 con la menor desviación estándar; e igualmente obtuvo el menor MSE.

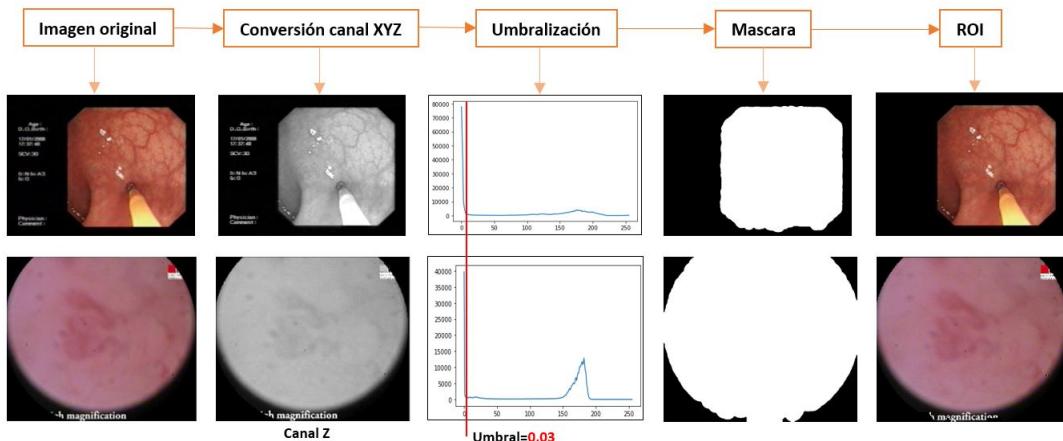


Figura 5-24: Primera técnica para la detección de la región de interés

- **ROI₂:** detección basada en el algoritmo interactivo para un umbral mínimo de entropía cruzada, conocido también como umbral de Li [Li and Tam, 1998]; para este método se seleccionó un umbral (T) con el que se consiguió la menor entropía cruzada a partir del histograma de cada imagen. Al tener el valor de T se obtuvo una máscara binaria con la ecuación 5-15, a esta se le aplicó dos transformaciones morfológicas, una de apertura y otra de cierre, con kernel de 20x20 y 30x30, respectivamente. Tras obtener la máscara binaria, esta se multiplicó con cada uno de los canales de la imagen original (RGB) para así obtener la región de interés. Las medidas de similitud obtenidas en cada canal se muestran en la tabla 5-8. Posteriormente se evaluó la similitud entre dicha máscara y su *ground truth*, cuyos resultados se pueden ver en la tabla 5-9, los cuales indican

que la detección de la ROI hecha a partir del canal V (del espacio de color HSV) tiene los mejores resultados.

	Medidas		
	DICE	SSIM	MSE
Canal B (RGB)	0,80±0,19	0,70±0,24	17922±15277
Canal V (HSV)	0,80±0,17	0,70±0,22	17615±14329
Canal Z (XYZ)	0,79±0,18	0,70±0,23	18161±15053

Tabla 5-9: Medidas de similitud ROI₂

Cabe resaltar que en la figura 5-25 se puede observar la metodología que se siguió para obtener la ROI a partir del canal V, donde la imagen original en RGB se transformó al espacio de color HSV y posteriormente se calculó el umbral Li en el canal V, cuyos valores de umbral variaron entre 17.8 y 179.2.

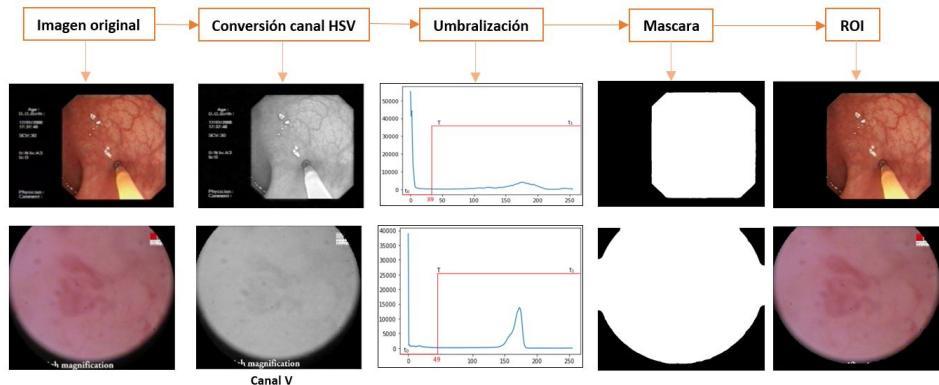


Figura 5-25: Segunda técnica para la detección de la región de interés

- **ROI₃:** ésta detección se basó en el cálculo del umbral más apropiado para cada imagen dependiendo de su histograma, más específicamente se encontró el rango de intensidades en este con mayor información (ver figura 5-26) y se escogió el 30 % del valor mínimo de dicho rango como umbral. Al tener el valor de T se obtuvo una máscara binaria con la ecuación 5-15, a la que se le aplicó dos transformaciones morfológicas, una de apertura y otra de cierre, con kernel de 27x27 y 37x37, respectivamente. Tras obtener la máscara binaria, esta se multiplicó con cada uno de los canales de la imagen original (RGB) para así obtener la región de interés. Posteriormente se evaluó la similitud entre dicha máscara y su ground truth, cuyos resultados se pueden ver en la tabla 5-10, los cuales indican que la detección de la ROI hecha a partir del canal Z (del espacio de color XYZ) tiene los mejores resultados.

	Medidas		
	DICE	SSIM	MSE
Canal B (RGB)	0,94±0,07	0,88±0,12	6812±7601
Canal V (HSV)	0,99±0,03	0,97±0,05	1082±33028
Canal Z (XYZ)	0,97±0,03	0,97±0,05	752±1566

Tabla 5-10: Medidas de similitud ROI_3

Cabe resaltar que en la figura 5-26 se puede observar la metodología que se siguió para obtener la ROI a partir del canal Z, donde la imagen original en RGB se transformó al espacio de color XYZ y posteriormente se calculó el umbral con el método propuesto anteriormente, cuyos valores de umbral variaron entre 3 y 74.

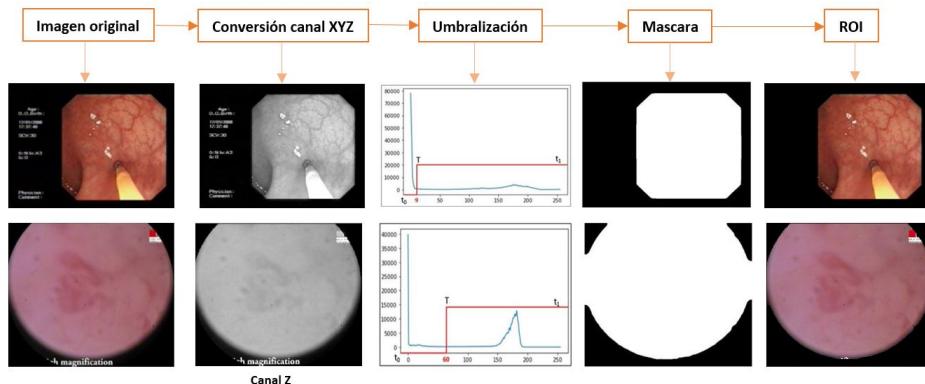


Figura 5-26: Tercera técnica para la detección de la región de interés

Evaluación de resultados

Por cada método descrito anteriormente se eligió el canal con los mejores resultados, es decir, en el caso del coeficiente DICE y el índice SSIM el resultado más cercano a uno con la menor desviación estándar, por otro lado, el valor MSE más cercano a cero con la menor desviación estándar. Para la ROI_1 y ROI_3 se escogió el canal Z y para la ROI_2 el canal V. Finalmente, de acuerdo a la evaluación cualitativa de las máscaras resultantes de cada una de las tres técnicas expuestas anteriormente (ver en la figura 5-23 un ejemplo de ellas) las técnicas que mejor desempeño se esperaba que tuvieron fue la ROI_1 y la ROI_3 pues en la mayor parte de las máscaras resultantes de la ROI_2 no se alcanza a tomar la región esperada de acuerdo al ground truth.

Según las métricas de similitud obtenidas en la tabla 5-11 se escogió la ROI_3 , puesto que es la técnica con mejor desempeño, lo que reafirma la hipótesis contemplada en la evaluación cualitativa.

	Medidas		
	DICE	SSIM	MSE
ROI₁	0,98±0,06	0,97±0,04	1190±2863
ROI₂	0,80±0,17	0,70±0,22	17615±14329
ROI₃	0,97±0,03	0,97±0,05	752±1566

Tabla 5-11: Medidas de similitud con diferentes técnicas de detección de ROI_s

Cabe señalar que, todos los resultados de las métricas de similitud y las técnicas de ROI se encuentran en la carpeta: **anexos/tablas/Region de interes (ROI).xlsx**.

5.4. Segmentación y extracción de candidatos a artefacto

Una vez obtenidas las regiones de interés de las imágenes preprocesadas se llevó a cabo la segmentación de los posibles candidatos a RE obteniendo así su ubicación y así mismo disminuyendo el área a analizar. Por el contrario, no se realizó segmentación a los posibles candidatos a DM, dado que la mayoría de las imágenes entregadas por la base de datos presentaban un desenfoque global, es decir en toda la imagen. Por ende, se optó por dividir la imagen en ventanas, las cuales son las posibles regiones con dicho artefacto.

5.4.1. Umbralización

Una vez determinada la ROI se probaron diferentes técnicas de umbralización para RE, con el fin de obtener una discriminación completa del fondo y los objetos de interés, asimismo, obtener la ubicación de los posibles reflejos especulares, como se observa en la figura 5-27.

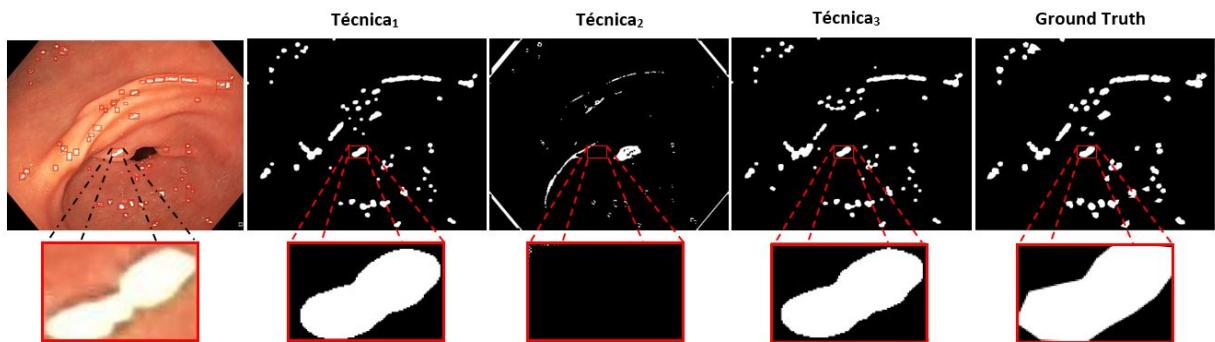


Figura 5-27: Diferentes técnicas de umbralización para RE

Según la literatura para la separación entre reflejos especulares y el fondo son ampliamente utilizadas las técnicas de umbral [Karapetyan and Sarukhanyan, 2013], por esto, se probaron

tres diferentes técnicas como se ve a continuación. Antes de explicar en que consistieron, es importante resaltar que, para la evaluación del desempeño de las técnicas de umbralización se contó con 135 imágenes con segmentaciones de RE (*ground truth*) proporcionadas por la base de datos EAD2019, de las que se obtuvieron 1120 bbox control.

- **Técnica₁ (Umbral global):** esta se basó en el **umbral adaptativo** descrito en el artículo [Oh et al., 2007], que tomó como referencia el hecho de que los RE suelen tener un valor de brillo (V) alto y una baja saturación (S) [Gross, 2014]. De acuerdo a lo anterior, [Oh et al., 2007] se encontró mediante pruebas experimentales que los umbrales en el espacio de color HSV que mejor detectaban las regiones con RE son: TH_s de 89.25, que como se observa en el histograma del canal S en la figura 5-28 actúa como un filtro pasa-bajos y un TH_v de 191.25, cuya representación en el histograma del canal V, actúa como un filtro pasa-alto. Lo anterior también es descrito por la ecuación 5-16. En este caso la aplicabilidad de la técnica descrita siguió la metodología propuesta en la figura 5-28.

$$m(i) = \begin{cases} 1 & \text{Si } S(i) < TH_s \text{ y } V(i) > TH_v \\ 0 & \text{Si } \text{Otros casos} \end{cases} \quad (5-16)$$

Donde, i es el píxel correspondiente en la máscara binaria resultante tras la aplicación de la técnica, los valores de umbral son TH_s y TH_v para los canales de saturación S y brillo V del espacio de color HSV, respectivamente.

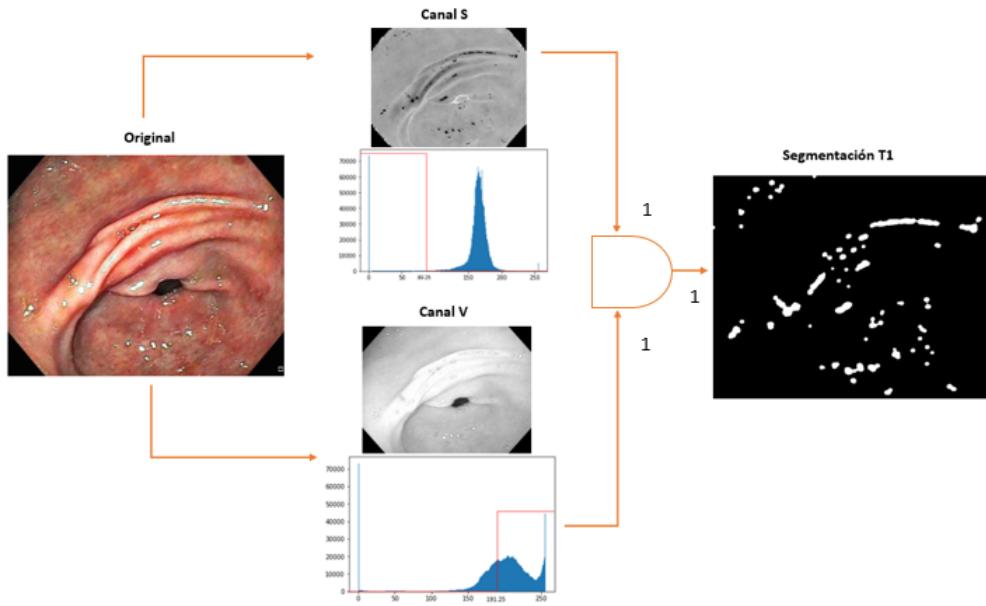


Figura 5-28: Metodología técnica₁

En la figura 5-28 se puede observar que la entrada de esta técnica fue la región de interés de la imagen preprocesada hasta la ecualización para RE (ver subsección 5.2.2), la cual se transformó al espacio de color HSV y se le aplicó la ecuación 5-16, dando como resultado una máscara binaria (imagen con extracción del fondo), en la que los candidatos a reflejos especulares tomaron un valor de 1 y los que no un valor de 0.

Posteriormente, a la máscara se le aplicó una dilatación y un cierre morfológico, buscando ampliar la región y llenar vacíos de los posibles candidatos. Para ésta tarea se probaron kernels en forma de disco con diámetro de 5, 10 y 15 píxeles, cabe señalar que la elección del kernel se basó en los resultados de la evaluación de similitud (DICE y MSE). La mencionada selección fue hecha a partir de la comparación de los bbox con RE de las imágenes *ground truth* (bbox control) y los bbox en la misma ubicación que estos sacados de las máscaras resultantes. Los resultados de esta evaluación se pueden observar en la tabla 5-12.

	Medidas de similitud	
	DICE	MSE
kernel 5x5	0,70±0,17	21061±8524
kernel 10x10	0,70±0,17	21061±8524
kernel 15x15	0,70±0,17	21059±8523

Tabla 5-12: Medidas de similitud de la técnica₁ con diferentes kernels

Como se puede observar en la tabla 5-12, la medida de DICE es igual para cada Kernel aplicado, por lo tanto la elección del más conveniente se basó en el que obtuvo el valor más pequeño de MSE, que en éste caso fue el Kernel de 15x15. Sumado a esto, es importante resaltar el buen desempeño de los valores obtenidos en la métrica DICE, lo cual supone que la mayoría de RE que deben ser segmentados (de acuerdo al *ground truth* de cada imagen) están presentes en las máscaras resultantes tras aplicar ésta técnica.

- **Técnica₂ (Derivada global):** ésta técnica se basó en el análisis de intensidades de una imagen mediante su histograma, descrita en el artículo [Stehle, 2006]. En éste caso se siguió la metodología propuesta en la figura 5-29.

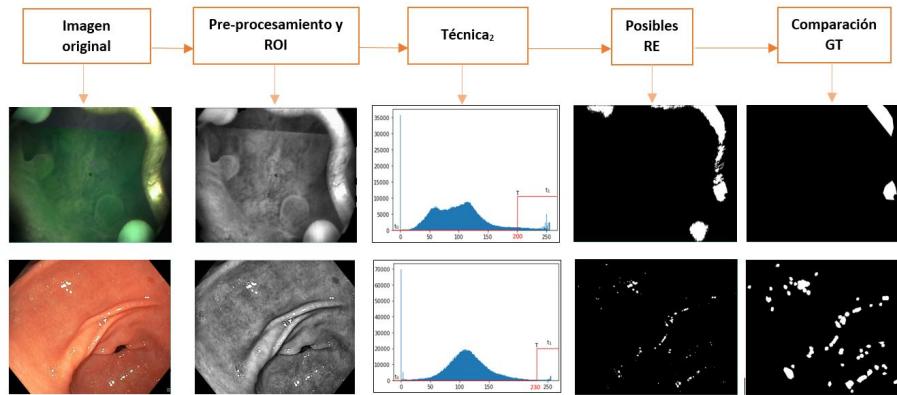


Figura 5-29: Metodología técnica₂

En la figura 5-29 se puede observar que la entrada para esta técnica fue la región de interés de la imagen preprocesada (normalización Min-Max, ecualización adaptativa, canal G del espacio de color RGB), de la que se seleccionó un umbral, cuyo valor fue la posición del pico presente en la región más clara del histograma, dado que los RE son muy brillantes y los píxeles que los contienen tienden a tener un valor alto de intensidad [Stehle, 2006]. Cabe resaltar que, dicha posición se encontró calculando el punto donde la derivada del histograma tenía un cambio significativo de un valor positivo a un valor cercano a cero; en este caso, los valores de umbral variaron entre 200 y 255. Seguido a esto se obtuvo una máscara binaria siguiendo la ecuación 5-15, en la que se le asignó un valor de 1 a los píxeles que contenían posibles candidatos a reflejos especulares y un valor de 0 a los que no.

Posteriormente, a la máscara se le aplicó una dilatación y un cierre morfológico, buscando ampliar la región y llenar vacíos del posible candidato. Para esta tarea se probaron kernels en forma de disco con diámetro de 5, 10 y 15 píxeles, cabe señalar que la elección del kernel se basó en los resultados de la evaluación de similitud (DICE y MSE) hecha a partir de la comparación de los bbox con RE de las imágenes *ground truth* (bbox control) y los bbox en la misma ubicación que éstos, sacados de las máscaras resultantes. Los resultados de esta evaluación se pueden observar en la tabla 5-13.

Medidas de similitud		
	DICE	MSE
kernel 5x5	0,31±0,33	34229±11407
kernel 10x10	0,31±0,33	34231±11407
kernel 15x15	0,31±0,33	34232±11406

Tabla 5-13: Medidas de similitud de la técnica₂ con diferentes kernels

De acuerdo con los resultados en la tabla 5-13, la medida de DICE es igual para cada Kernel aplicado, por lo tanto la elección del más conveniente se basó en el que obtuvo el valor más pequeño de MSE, que en este caso fue el Kernel de 5x5. Sumado a esto, es importante resaltar el bajo desempeño de los valores obtenidos en la métrica DICE, lo cual supone que la mayoría de RE que deben ser segmentados (de acuerdo al *ground truth* de cada imagen) NO están presentes en las máscaras resultantes tras aplicar esta técnica, lo que también puede ser observado en la figura 5-29.

- **Técnica₃ (Local adaptativa):** esta sigue la misma metodología que la técnica anterior, solo que en vez de calcular el histograma de toda la imagen, calcula el histograma por parches de la imagen como se ve en la siguiente figura.

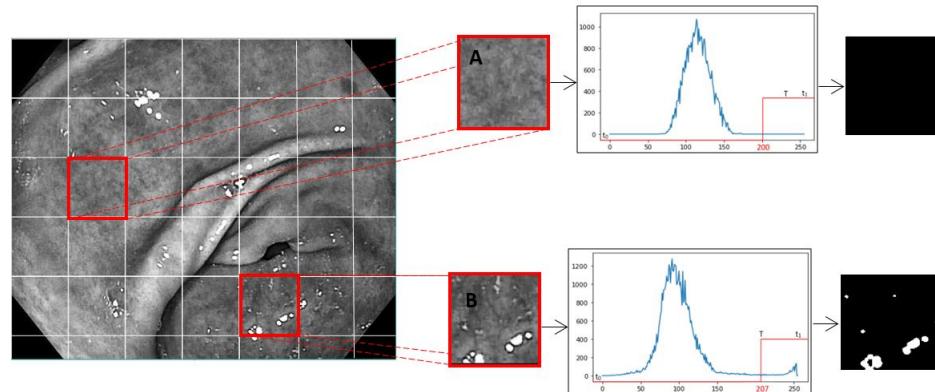


Figura 5-30: Metodología técnica₃

Como se ve en la figura 5-30 se utilizó una ventana deslizante de 200x200 píxeles para que recorriera la imagen, y de cada una se calculó el histograma para la selección del valor del umbral como se describió en la técnica₂. Por cada ventana se obtuvo una máscara binaria siguiendo la ecuación 5-15, en la que se le asignó un valor de 1 a los píxeles que contenían posibles candidatos a reflejos especulares (ver figura 5-30 -B) y un valor de 0 a los que no (ver figura 5-30 -A).

Posteriormente, a la máscara se le aplicó una dilatación y un cierre morfológico, buscando ampliar la región y llenar vacíos del posible candidato. Para esta tarea se probaron kernels en forma de disco con diámetro de 5, 10 y 15 píxeles, cabe señalar que la elección del kernel se basó en los resultados de la evaluación de similitud (DICE y MSE) hecha a partir de la comparación de los bbox con RE de las imágenes *ground truth* (bbox control) y los bbox en la misma ubicación que estos, sacados de las máscaras resultantes. Los resultados de esta evaluación se pueden observar en la tabla 5-14.

Medidas de similitud		
	DICE	MSE
kernel 5x5	0,69±0,19	20961±8798
kernel 10x10	0,69±0,19	20960±8798
kernel 15x15	0,69±0,19	20959±8797

Tabla 5-14: Medidas de similitud de la técnica₃ con diferentes kernels

Según la tabla **5-14**, la medida de DICE es igual para cada Kernel aplicado, por lo tanto la elección del más conveniente se basó en el que obtuvo el valor más pequeño de MSE, que en éste caso fue el Kernel de 15x15. Sumado a esto, es importante resaltar el buen desempeño de los valores obtenidos en la métrica DICE, lo cual supone que la mayoría de RE (verdaderos), que deben ser segmentados (de acuerdo al *ground truth* de cada imagen) están presentes en las máscaras resultantes tras aplicar ésta técnica.

Evaluación de resultados

Por cada técnica descrita anteriormente se eligió el tamaño de kernel con los mejores resultados, es decir, con mayor valor de DICE y menor MSE junto con la menor desviación estándar. Para la técnica₁ y la técnica₃ se escogió un kernel de 15x15 píxeles y para la técnica₂ el kernel de 5x5. Finalmente, de acuerdo a una evaluación cualitativa hecha de las máscaras resultantes de cada una de las técnicas expuestas anteriormente (ver en la figura **5-27** un ejemplo de ellas) se esperaba que la técnica₂ obtuviera el peor desempeño y que la elección de la umbralización para este método fuera entre las otras dos técnicas.

Efectivamente, después de comparar los resultados de la evaluación de similitud de los bbox control y los bbox sacados de las máscaras resultantes tras la aplicación de cada una de las técnicas, se observó que la técnica dos obtuvo los peores resultados, y que por el contrario, las técnicas uno y tres obtuvieron un muy buen desempeño; sin embargo, la diferencia entre los resultados de estas dos no es evidente, por lo que se decidió comparar la imagen completa de *ground truth* con las máscaras resultantes. Cuyos resultados se pueden observar en la tabla **5-15**

	Medidas		
	DICE	SSIM	MSE
Técnica ₁	0,25±0,20	0,86±0,07	7040±4862
Técnica ₃	0,29±0,17	0,92±0,04	3103±1965

Tabla 5-15: Medidas de similitud técnicas de umbralización

Como se puede observar en la tabla **5-15**, la medida DICE para las dos técnicas es baja, esto se debe a que las técnicas no solo segmentan verdaderos RE sino que también segmentan

regiones con características de brillo similares a dicho artefacto, como lo son las regiones con saturación (ver figura 5-31). Sin embargo se optó por escoger la técnica tres que fue la que mejores resultados obtuvo.

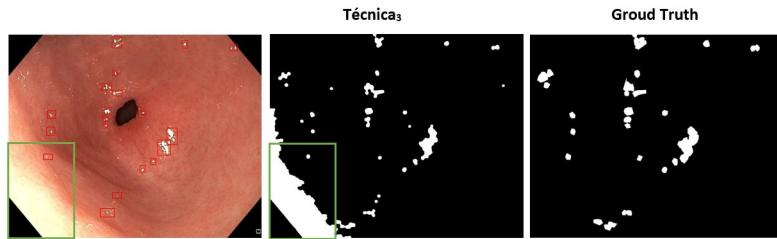


Figura 5-31: Técnica₃ con segmentaciones de saturación (bbox verde)

Una vez adquiridas las imágenes umbralizadas en el canal G se aislaron los componentes de interés como se muestra a continuación.

Cabe señalar que, todos los resultados de las métricas de similitud entre las imágenes de control y las obtenidas tras la aplicación de las técnicas de umbralización se encuentran en la carpeta: anexos/tablas/Umbrealizacion.xlsx.

5.4.2. Separación o extracción de candidatos

Para dar cumplimiento parcial al segundo objetivo específico, se procedió a ubicar los posibles candidatos a RE obtenidos en la segmentación, como se muestra a continuación.

Dado que la extracción de características se va a realizar sobre los posibles candidatos, se hizo necesario la separación y extracción de los mismos a partir de la umbralización anteriormente obtenida.

Dicha extracción consistió en encontrar los contornos de los posibles candidatos dentro de la imagen y asimismo la posición del bounding box que los contenía como se observa en la figura 5-32.

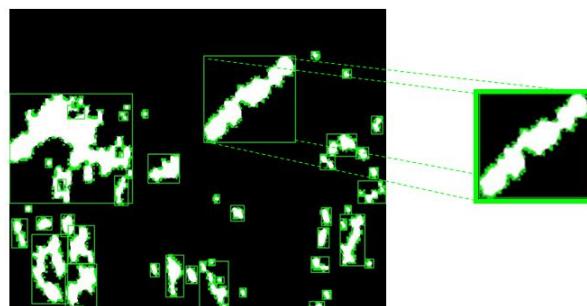


Figura 5-32: Representación de como se separan y extraen los candidatos a RE

5.5. Procesamiento y extracción de características

Características	
DM	RE
Análisis de textura usando la matriz de co-ocurrencia de nivel de gris(GLCM)	
Coeficientes de la transformada Wawelet	Brillo
Descomposición de valores singulares (SVD)	Luminancia
Pendiente espectro de potencias	Características de forma
Varianza laplaciana	Características a partir del histograma
Histograma de orientación de gradientes (HOG)	
Transformada discreta de Walsh-Hadamard (DWHT _s)	
Histograma de intensidades	
Momentos estadísticos	
Distribuciones de magnitud del gradiente	

Tabla 5-16: Lista de características para detectar DM y RE

Antes de comenzar con esta sección es importante resaltar que con los datos entregados por la base EAD2019 se obtuvieron tres tipos de anotaciones por bbox, donde el primer tipo es desenfoque por movimiento (clase 0), el segundo reflejos especulares (clase 1) y el ultimo son las imágenes que no tienen ninguno de estos dos artefactos (clase 2).

De modo que una vez las imágenes fueron pasadas por los procesos nombrados en las secciones anteriores, se trajeron diversas características como se ve en la tabla **5-16** que de acuerdo a la literatura se han utilizado para identificar DM y RE (ver subsección 4.2.6); sin embargo, anterior a esto se decidió realizar una depuración a la base de datos dado que en esta se encontraban imágenes que no correspondían como tal a imágenes endoscópicas con tejido subyacente como se ve en la figura **5-33**, además, se presentaban datos repetidos y que omitían algunas anotaciones, es decir, había imágenes que contenían algún artefacto de la clase cero o uno con anotaciones erróneas o nulas, como se puede observar en la figura **5-34**; por ende se decidió excluir dichas imágenes.



Figura 5-33: Imágenes excluidas de la base de datos

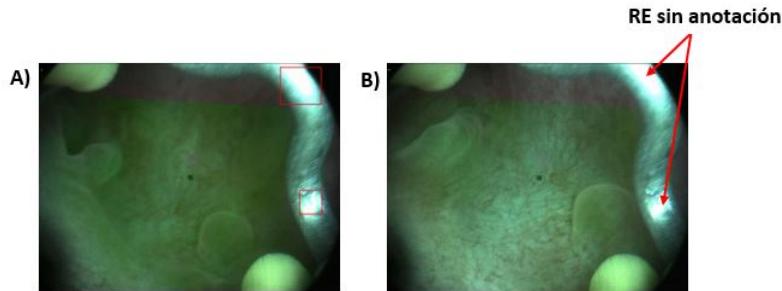


Figura 5-34: A) Imagen con anotación de RE (bbox rojo), B) Imagen sin anotación de RE

Por consiguiente, la cantidad de bbox con los que se contaban por cada clase cambio. En las tablas 5-17 y 5-18 se puede observar el número de bbox con los que se contaron para el entrenamiento del método de identificación de DM y RE, respectivamente; de igual manera se puede apreciar que la cantidad de datos por clase no es igual, por lo que se procedió a balancearlas, como se muestra en la siguiente subsección.

Clase	Número de bbox por clase								
	0			1			2		
Tamaño bbox	50x50	150x150	300x300	50x50	150x150	300x300	50x50	150x150	300x300
Cantidad	42.752	4.539	884	64.833	8.336	1.972	9.929	1.376	309

Tabla 5-17: Cantidad de bbox para detección de DM por clase (0:DM, 1:RE, 2:Otro) después de la depuración de la base de datos

Clase	Número de bbox por clase		
	0	1	2
Cantidad	46.451	3.862	15.696

Tabla 5-18: Cantidad de bbox para detección de RE por clase (0:DM, 1:RE, 2:Otro) después de la depuración de la base de datos

5.5.1. Balanceo de clases:

Como se expuso anteriormente la cantidad de datos con los que se contaba para cada clase no era igual, por lo que se optó por realizar un balanceo de estas, pues un modelo estimado sobre la base de datos completa (sin balanceo) tiene menos oportunidad de reconocer diferencias que sobre una base de datos balanceada [Valdovinos Rosas et al., 2006].

Debido a que se contaba con una cantidad considerable de datos, el tipo de balanceo que se realizó fue el submuestreo aleatorio, dado que disminuye el sesgo implícito al manipular los datos, dicho balanceo consistió en eliminar datos de la clase mayoritaria con el fin de igualar las clases. Es importante tener en cuenta que la base de datos EAD 2019 proporcionaba

imágenes tomadas con diferentes tipos de modalidades, resoluciones y tejidos (ver tabla 5-1), por lo tanto fue necesario asegurarse de incluir en el submuestreo aleatorio imágenes de cada tipo, con la finalidad de que el método de identificación fuera general. El número de bbox resultantes por clase es mostrado en las tablas 5-19 y 5-20

Número de bbox por clase			
Clase	0,1,2		
Tamaño bbox	50x50	150x150	300x300
Cantidad	9.929	1.376	309

Tabla 5-19: Cantidad de bbox para detección de DM por clase (0:DM, 1:RE, 2:Otro) después del balanceo de clases

Número de bbox por clase	
Clase	0,1,2
Cantidad	3.862

Tabla 5-20: Cantidad de bbox para detección de RE por clase (0:DM, 1:RE, 2:Otro) después del balanceo de clases

5.5.2. Evaluación de características:

Para dar cumplimiento al primer objetivo específico, se procedió a evaluar la separabilidad de características encontradas como se muestra a continuación.

Para identificación de DM:

Una vez extraídas las características de la tabla 5-16 para DM en cada uno de los tamaños de ventana propuestos (50x50, 150x150 y 300x300) y en diferentes espacios de color (S, V y B), se procedió a evaluar la capacidad de separabilidad entre clases, inicialmente, de manera cualitativa mediante diagramas de dispersión, como se puede observar en la figura 5-35, donde se muestra la relación entre algunas de las características extraídas anteriormente. De acuerdo a dichos diagramas no es posible identificar con claridad separabilidad entre clases o al menos entre DM (clase 0) y el resto (RE y Otros), pues las características se encuentran solapadas. Por lo tanto, se procedió a evaluar la separabilidad mediante boxplots, herramienta que permite hacer no solo un análisis cualitativo sino también cuantitativo en términos de su distribución en cuartiles.



Figura 5-35: Diagramas de dispersión entre características (ventana de 300x300) de DM

Como se observa en las figuras 5-36 y 5-37, los boxplots indican por percentil la distribución de los datos entre las clases 0, 1 y 2. En este caso se tuvo en cuenta los percentiles 25 y 75, pues estos representan la mayoría de los datos, lo cual es relevante porque para el análisis de separabilidad los percentiles de cada clase no deberían sobreponerse, ya que de hacerlo se indica que los valores de una clase como de otra pueden pertenecer a los tres grupos en un punto dado. Esta situación se presentó en la figura 5-36, pues como se observa las características extraídas a partir de la transformada Wavelet, varianza Laplaciana y HOG se solapan entre las clases, por lo tanto no fueron escogidas.

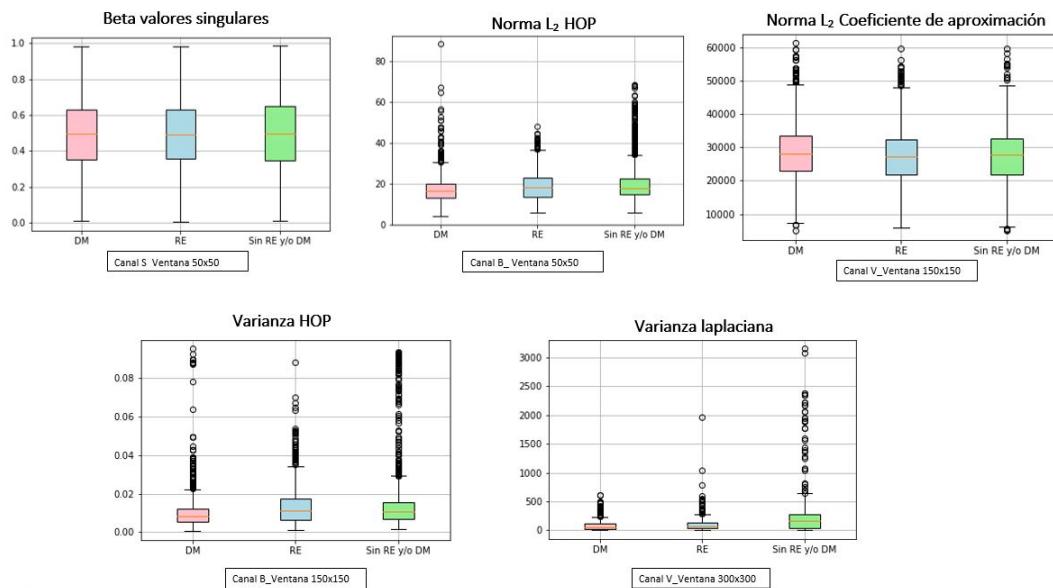


Figura 5-36: Boxplot, características no escogidas

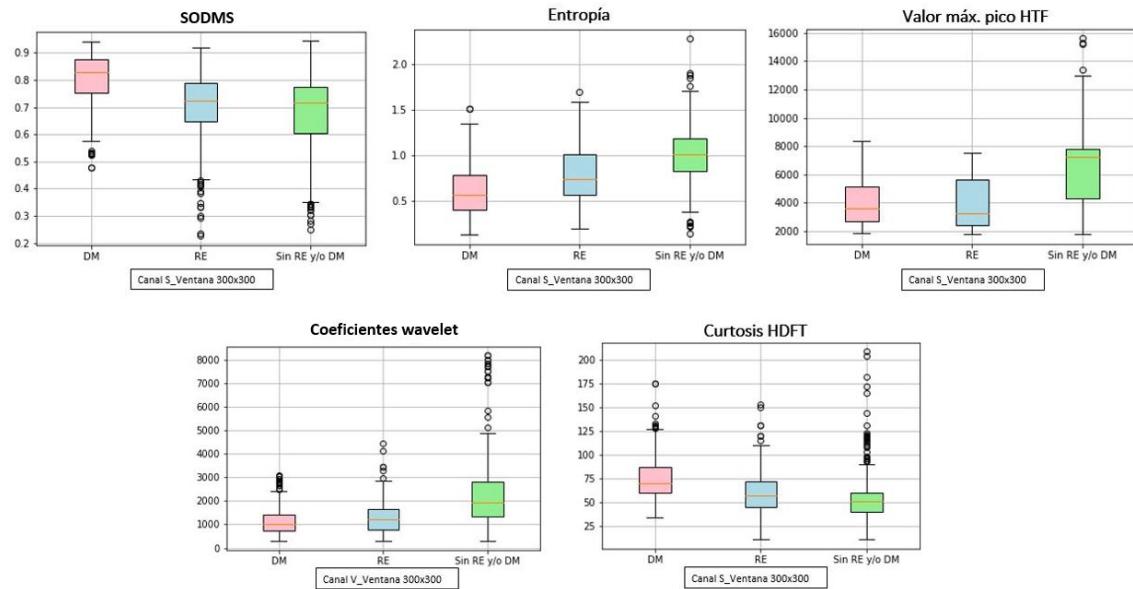


Figura 5-37: Boxplot, con evidencia de separabilidad entre la clase 0 del resto

Clases	SODMS			Entropía			Valor máximo pico HTF			Norma L ₂			Curtosis HDFT		
	0	1	2	0	1	2	0	1	2	0	1	2	0	1	2
Percentil (p25-p35)	0.75-0.87	0.64-0.78	0.60-0.77	0.39-0.78	0.55-1	0.82-1.18	2646-5167	2384-5647	4243-7780	1554-4104	1963-4162	3351-5481	59-86	45-71	39-60

Tabla 5-21: Percentiles 25 y 75 de los boxplots visualmente separables

Por otra parte, de acuerdo a la evaluación cualitativa basada en la distancia de los diferentes boxplots presentes en la figura 5-37, y el análisis de los percentiles 25 y 75 de estos (ver tabla 5-21), se encontró evidencia de que la separabilidad entre la clase 0 (DM) y la 2 es la más notoria con las siguientes características; sin embargo las métricas para todas las características extraídas para DM se encuentran la carpeta: anexos/tablas/Características_DM.xlsx.

- SODMS: Esta es la similitud del parche original del posible candidato a DM con el parche suavizado, para el cálculo de esta característica, lo primero que se hizo fue suavizar una ventana de la imagen (possible candidato a DM) con un filtro de mediana de 15x15. Posteriormente se evaluó la similitud entre la ventana sin difuminar con la difuminada usando la métrica SSIM, con lo que se esperaba obtener un valor cercano a uno cuando la ventana de entrada pertenecía a la clase DM y 0 cuando no, tal como se expresa en la ecuación 5-17.

Lo anterior puede ser explicado dado que si la ventana de entrada es una imagen clara y nítida (perteneciente a la clase 1 y 2) perderá muchos detalles al ser suavizada y se volverá obviamente diferente a la ventana original. Por el contrario, si la ventana de entrada está desenfocada perderá pocos detalles a través de la misma operación de

suavizado [Wang et al., 2019], esto se confirmó con la figura **5-37**, ya que los datos de desenfoque se encuentran más cercanos a 1 que las otras dos clases.

$$SSIM = \begin{cases} \approx 1 & Si \quad w(i,j) \sim w'(i,j) \\ \approx 0 & Si \quad w(i,j) \neq w'(i,j) \end{cases} \quad (5-17)$$

Dónde $w(i,j)$ corresponde a la venta sin difuminar y $w'(i,j)$ a la ventana difuminada de la imagen $W(i,j)$.

- Entropía: esta es una de las características de textura propuesta por Haralick calculada a partir de la matriz GLCM (para mayor información de la matriz, dirigirse a la subsección 4.2.6.1), como se observa en la ecuación 5-18 del posible candidato a DM. Con ella se esperaba que la ventana desenfocada obtuviera una menor entropía que el resto de las clases, ya que las imágenes desenfocadas proveen menos información que las otras clases [Tsai et al., 2008], debido a que hay pérdida de bordes y características visuales del tejido. Esto se comprobó en la figura **5-37**, pues los datos de desenfoque son menores al resto de las clases.

$$\text{Entropía} : \sum_i \sum_j P(i,j) \cdot \ln P(i,j) \quad (5-18)$$

- Valor del máximo pico de intensidad del histograma de la transformada de Fourier: en esta se calculó la transformada de Fourier al posible candidato a DM y posteriormente se halló su histograma de intensidades, del que se obtuvo el valor del máximo pico de intensidad.
- Coeficientes Wavelet: para obtener ésta característica se calcularon los coeficientes de la transformada Wavelet biortogonal de la ventana con el posible candidato a DM; los cuales fueron de aproximación (LL) y coeficientes de detalle: vertical (HL), diagonal(HH) y horizontal (LH). Tomando en cuenta que según el artículo [Jain et al., 2011] las altas frecuencias son representadas por los coeficientes de detalle y las bajas por los de aproximación se calculó la norma l_2 de las altas frecuencias. Con lo que se esperaba que para DM el valor de la norma fuera menor que el de las otras clases, dado que dichas imágenes se caracterizan por la ausencia de bordes (altas frecuencias). Sin embargo en la figura **5-37** se puede observar que la norma arrojó resultados muy cercanos entre la clase 0 y 1 , pero si se evidencia la posible separabilidad de la clase 0 con la clase 2.
- Curtosis del histograma de la transformada discreta de Fourier (DFT): para ésta primero se calculó la DFT de la ventana con el posible candidato a DM, posteriormente

se halló su histograma de intensidades, del que se obtuvo la curtosis calculada a partir de la ecuación 5-19

$$Curtosis(X) = \frac{\sum_{i=1}^T (X - \bar{X})^4}{\sigma^4} \quad (5-19)$$

Donde, X es el vector de intensidades a partir del histograma, \bar{x} su media y σ es su desviación estándar.

Evaluación cuantitativa

Dado que la evaluación cualitativa hecha mediante los boxplots no evidenció una diferenciación significativa entre las clases, se continuó con la evaluación estadística de la separabilidad encontrada en los boxplots. Para esto, se evaluó con dos métodos estadísticos los cuales fueron T-student y la prueba de suma de rango Wilcoxon.

- **T-student:**

La T-student proporciona una medida de nulidad en la media de una distribución de densidad de probabilidad que se asume normal. Sin embargo, dado que esta herramienta estocástica esta pensada para una sola muestra estadística y su evaluación de centralidad, se decidió usar como entrada a la misma, la diferencia aritmética entre las características pertenecientes a dos de las clases. Este método indicó si la media de los datos de las distribuciones de las clases eran iguales o diferentes, demostrando con esta última la independencia de los valores de cada característica para cada una de las clases. Dicho estudio estadístico se calcula según la fórmula 5-20.

$$p = \frac{\bar{X}_1 - \bar{X}_2}{S_{x_1x_2} \times \sqrt{\frac{2}{n}}} \quad (5-20)$$

Donde \bar{X}_1 es la media muestral de la clase 0 (DM) y \bar{X}_2 de la clase con la que fue comparada (1 o 2), $S_{x_1x_2}$ es la desviación estándar combinada y n es el tamaño de la muestra.

Tras calcular la t-student se obtuvieron dos valores importantes, el valor de p y la aprobación (0) o rechazo (1) de la hipótesis, la cual es que las dos medias de las distribuciones de los dos conjuntos de datos comparados son iguales. En este caso, los resultados esperados fueron el rechazo de la hipótesis y un valor de p menor a 0.05, dado que esto significa que las distribuciones de probabilidad de los conjuntos de datos comparados son estadísticamente separables. En la tabla 5-22 se puede observar los resultados de la prueba T-student.

Clases comparadas	SODMS		Entropía		Valor máximo pico HTF		Coeficiente wavelet		Curtosis HDFT	
	0 vs 1	0 vs 2	0 vs 1	0 vs 2	0 vs 1	0 vs 2	0 vs 1	0 vs 2	0 vs 1	0 vs 2
Valor de p	$6,57 \times 10^{-23}$	$7,56 \times 10^{-28}$	$1,25 \times 10^{-19}$	$3,58 \times 10^{-50}$	0,03	$1,63 \times 10^{-29}$	$1,02 \times 10^{-7}$	$3,69 \times 10^{-10}$	$6,09 \times 10^{-15}$	$2,33 \times 10^{-20}$
Valor de h	1	1	1	1	1	1	1	1	1	1

Tabla 5-22: Resultados de la prueba T-Student

Como se puede ver en la tabla **5-22** los resultados de la prueba T-student evidencian que todas las características rechazan la hipótesis inicial, es decir, todos obtuvieron un valor de p menor a 0.05. Además según estos resultados, las mejores características para diferenciar la clase 0 de las demás clases son *SODMS* y *Entropía*.

■ Prueba de rango de Wilcoxon:

No obstante, no se optó por elegir aun estas características, puesto que se prefirió corroborar dichos resultados con la prueba de suma de rango de Wilcoxon (que se calcula como se ve en la ecuación 5-21), la cual indicó si las medianas de las distribuciones de los datos de las clase eran iguales (hipótesis aprobada=1) o diferentes (hipótesis rechazada=0).

$$z = \frac{R - \mu_R}{\sigma_R} \quad (5-21)$$

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2} \quad (5-22)$$

$$\sigma_R = \frac{n_1 n_2 (n_1 + n_2 + 1)}{N} \quad (5-23)$$

Donde, R son las sumas de rangos para muestras de menor tamaño (n_1) y n_2 es el mayor tamaño de la muestra.

Clases comparadas	SODMS		Entropía		Valor máximo pico HTF		Coeficiente wavelet		Curtosis HDFT	
	0 vs 1	0 vs 2	0 vs 1	0 vs 2	0 vs 1	0 vs 2	0 vs 1	0 vs 2	0 vs 1	0 vs 2
Valor de p	$7,05 \times 10^{-20}$	$5,11 \times 10^{-22}$	$2,40 \times 10^{-12}$	$1,69 \times 10^{-33}$	0,07	$2,05 \times 10^{-23}$	0,003	$4,09 \times 10^{-27}$	$3,92 \times 10^{-8}$	$4,96 \times 10^{-5}$
Valor de h	1	1	1	1	0	1	1	1	1	1

Tabla 5-23: Resultados de la prueba de suma de rango de Wilcoxon

Como se puede ver en la tabla **5-23** los resultados de la prueba Wilcoxon evidencian que la característica del valor máximo del pico del histograma de la transformada

de Fourier no es útil para diferenciar la clase 0 de la clase 2, puesto que el valor de p fue mayor a 0.05 es decir los dos conjuntos de datos tienen la misma mediana y probablemente una distribución similar. Además, sumado al análisis cualitativo, los datos de p en la tabla ratificaron que las mejores características para diferenciar la clase 0 de las demás clases son *SODMS* y *Entropía*.

Para la identificación de RE:

Una vez extraídas las características de la tabla **5-16** para RE en el cnal G y B (teniendo en cuenta cada una de las etapas de la sección 5.2), se procedió a evaluar la capacidad de separabilidad entre clases, inicialmente, de manera cualitativa mediante diagramas de dispersión como se puede observar en la figura **5-38**, donde se muestra la relación entre algunas de las características extraídas anteriormente. De acuerdo a dichos diagramas no es posible identificar con claridad separabilidad entre clases o al menos entre RE (clase 1) y el resto (DM y Otros), pues las características se encuentran solapadas. Por lo tanto, se procedió a evaluar la separabilidad mediante boxplots, herramienta que permite hacer no solo un análisis cualitativo sino también cuantitativo mediante la distribución de cuartiles.

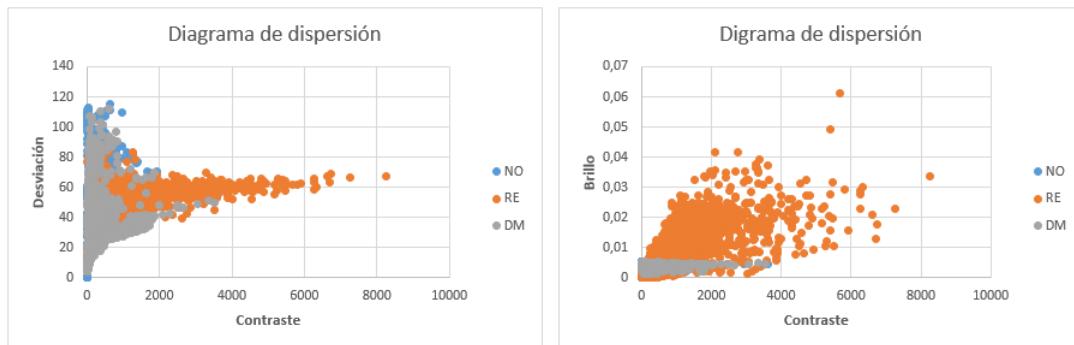


Figura 5-38: Diagramas de dispersión entre características de RE

Para el análisis de separabilidad mediante boxplots, como se explicó en la etapa anterior, los percentiles 25 y 75 son los de mayor interés, pues ellos contienen gran parte de la distribución de los datos. Teniendo en cuenta esto se explica la razón por la cuál no se escogieron las características presentes en la figura **5-39**, pues cada uno de los boxplot de la características *máximo pico de intensidad* y *norma euclídea del brillo* se sobreponen por clase, es decir, no son separables.

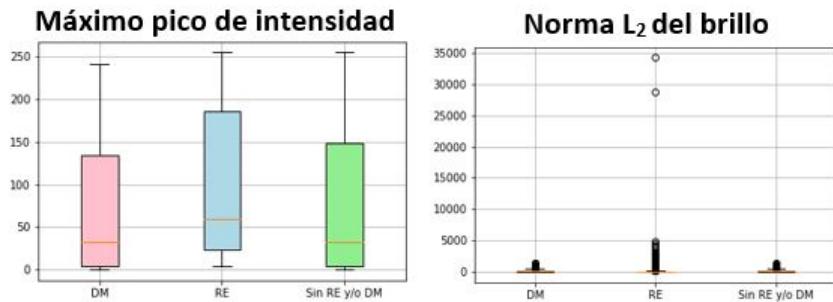


Figura 5-39: Boxplot, con características no separables.

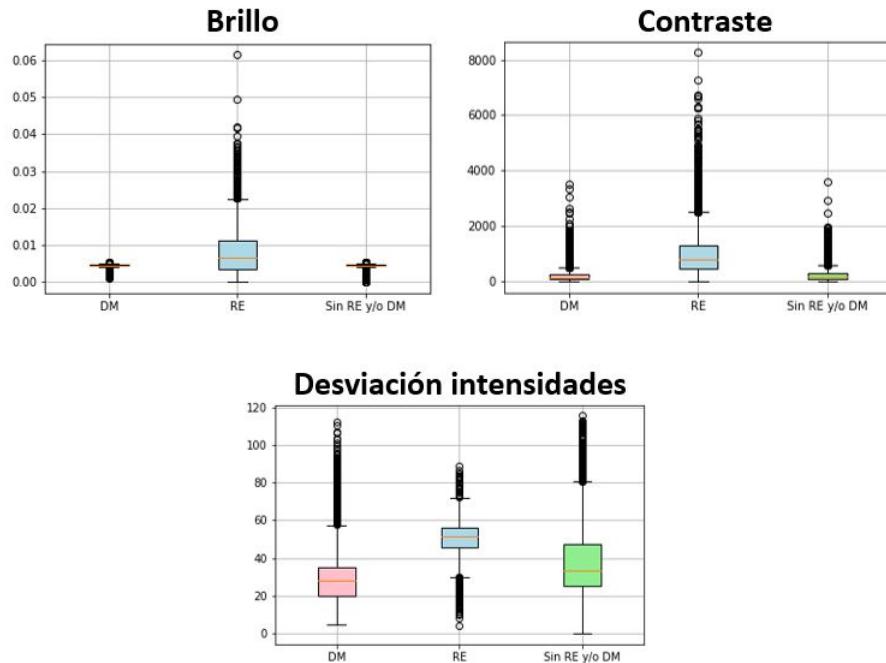


Figura 5-40: Boxplot, con evidencia de separabilidad entre la clase 1 del resto.

Clases	Brillo			Contraste			Desviación de intensidades		
	0	1	2	0	1	2	0	1	2
Percentil (p25-p35)	0.0043-0.0046	0.003-0.011	0.0043-0.0046	79-242	482-1295	87-294	20-35	45-56	25-47

Tabla 5-24: Percentiles 25 y 75 de los boxplots visualmente separables para identificación de RE

Por otro lado, en la figura 5-40 se evidencia tras la observación de los boxplots y la comparación de los percentiles 25 y 75 presentes en la tabla 5-24 la existencia de separabilidad entre clases con las siguientes características, sin embargo las métricas para todas las características extraídas para RE se encuentran la carpeta: anexos/tablas/Características_RE.xlsx:

- Brillo: esta se refiere a la claridad general u oscuridad de la imagen, con esta se esperaba que la clase 1 (RE) fuera mayor que las demás clases, dado que estos se caracterizan por contar con la mayoría de sus distribuciones de intensidad en la parte más clara del histograma [Morgand and Tamaazousti, 2014]. Para esta, se calculó el promedio de brillo obtenido con la siguiente ecuación.

$$\text{Brillo} = \frac{\sqrt{0,241R^2 + 0,691G^2 + 0,068B^2}}{W * H} \quad (5-24)$$

Donde R,G y B son el canal rojo, verde y azul del espacio de color RGB, respectivamente; W y H son las dimensiones del bbox.

- Contraste: este es una de las características propuestas por Haralick calculada a partir de la matriz GLCM de cada uno de los posibles candidatos a RE. Con esta se esperaba que la clase 1 fuera mayor que las demás. Cabe resaltar que esta se calculó con la siguiente ecuación.

$$\text{Contraste} : \sum_i \sum_j (i, j)^2 \cdot P(i, j) \quad (5-25)$$

- Desviación de las intensidades: esta se refiere a que tan dispersos están los valores de intensidad de la imagen respecto a su media [Armin et al., 2016], es decir, que mientras mayor sea la desviación estándar más dispersos estarán las intensidades del posible candidato a RE. Se calculó con la siguiente ecuación.

$$\sigma = \sqrt{\frac{1}{N - 1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (5-26)$$

Evaluación cuantitativa:

Dado que la evaluación cualitativa hecha mediante los boxplots y los diagramas de dispersión no evidencian una diferenciación significativa entre las clases, se continuó con la evaluación estadística de la separabilidad de las características mencionadas anteriormente entre clases mediante el uso de dos métodos estadísticos los cuales fueron T-student y la prueba de suma de rango Wilcoxon.

- **T-student:**

Al igual que en el análisis cuantitativo anterior, realizado para las características que describen DM, se comenzó con evaluar la T-student; cuyos resultados se pueden observar en la tabla 5-25.

Clases comparadas	Brillo		Contraste		Desviación de intensidades	
	1 vs 0	1 vs 2	1 vs 0	1 vs 2	1 vs 0	1 vs 2
Valor de p	$2,673 \times 10^{-293}$	$2,4006 \times 10^{-305}$	0	0	0	$9,8757 \times 10^{-232}$
Valor de h	1	1	1	1	1	1

Tabla 5-25: Resultados de la prueba T-Student para identificación de RE (clase 0: DM, clase 1: RE y clase 2: Ninguno)

Como se puede ver en la tabla 5-25, las tres características evaluadas con la prueba estadística t-student rechazaron la hipótesis nula, es decir, tiene un valor de p menor a 0.05, por ende la media de sus distribuciones son diferentes.

■ Prueba de rango de Wilcoxon:

No obstante, al igual que la evaluación de separabilidad para la identificación de desenfoque, se procedió a evaluar la prueba de suma de rango Wilcoxon, y así confirmar la separabilidad de las características propuestas. Los resultados de esta prueba se pueden observar en la tabla 5-26

Clases comparadas	Brillo		Contraste		Desviación de intensidades	
	1 vs 0	1 vs 2	1 vs 0	1 vs 2	1 vs 0	1 vs 2
Valor de p	$2,7112 \times 10^{-272}$	$3,4115 \times 10^{-282}$	0	0	0	$1,6146 \times 10^{-216}$
Valor de h	1	1	1	1	1	1

Tabla 5-26: Resultados de la prueba de suma de rango Wilcoxon para identificación de RE

Como se puede observar en la tabla 5-26, con la prueba Wilcoxon se corroboró la separabilidad de las características propuestas, puesto que para este test también se rechazo la hipótesis de que las medianas de las distribuciones de los conjuntos de datos son iguales.

A modo de conclusión, el análisis de separabilidad de características arrojó que para la identificación de desenfoque las características más separables son **SODMS y la entropía**, por otro lado, para RE se encontró que las características mas separables son **brillo, contraste y desviación de intensidades**.

Sumado a esto, tras esta evaluación es muy posible que sea necesario usar dos clasificadores binarios, pues solo se obtuvo separabilidad significativa entre una clase y el resto.

5.6. Método de identificación

Tras elegir las características más separables tanto para identificación de DM como para RE, se optó por usar dos clasificadores binarios complementarios como se observa en la figura

5-41, es decir, el primero que clasifique DM vs los Otros (RE y Otro) y el segundo que clasifique RE vs los Otros (DM y Otro); puesto que en la evaluación de separabilidad de características no se encontró ninguna capaz de diferenciar las tres clases.

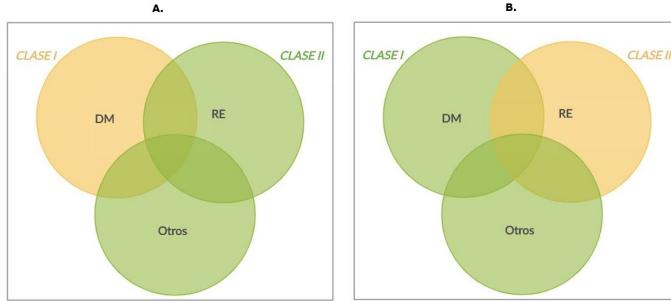


Figura 5-41: Métodos de identificación binario (**A.** para identificación de DM y **B.** para identificación de RE)

Para llevar a cabo lo anterior, se volvieron a balancear las clases debido a que para lograr entrenar los métodos de identificación binarios era necesario tener dos clases y hasta el momento se contaban con tres clases (0:DM, 1:RE y 2:Otro), a continuación se observa el proceso de balanceo.

Balanceo de clases

Como se expuso anteriormente se decidió balancear las clases para pasar de tener tres a dos clases, con el fin de poder usar dos clasificadores binarios complementarios. Se hicieron dos balanceos de submuestreo aleatorio, dado que se contaba con datos para identificación de DM y RE. Para el primer clasificador las clases quedaron organizadas de la siguiente manera: clase 0 (clase I en la figura 5-41) corresponde a DM y la clase 1 (clase II en la figura 5-41) corresponde a la clase Otros, es decir, datos de RE más datos sin ninguno de los artefactos en cuestión (Otro); por su parte para el segundo clasificador binario la clase 0 corresponde a la clase Otros (datos con DM más datos sin DM ni RE) y la clase 1 corresponde a los datos con RE. Por consiguiente, fue necesario asegurarse de incluir en el submuestreo aleatorio características de cada una de las clases, con la finalidad de que el método de identificación fuera general. El número de características resultantes por clase es mostrado en la tabla 5-27.

Número de características por clase			
	1 ^{er} clasificador binario	2 ^{do} clasificador binario	
	Para identificación de DM	Para identificación de RE	
Clase	0 (DM)	1 (RE + Otro)	0 (DM + Otro)
Cantidad	309	309	3852
			1 (RE)

Tabla 5-27: Cantidad de características para identificación de DM por clase (0: DM, 1: RE +Otro) y para identificación de RE por clase (0: DM + Otro, 1: RE)

Posteriormente, se probaron tres métodos de partición de datos (hold out, k-fold, leave one out) con algunos clasificadores usados en el estado del arte (Regresión logística, Máquina de soporte vectorial, Ramdon Forest, Bayes y Red neuronal). Cabe resaltar, que para algunos de dichos clasificadores se variaron diferentes parámetros, como se muestra a continuación:

Clasificadores probados:

- **Regresión logística:** se probó un clasificador de regresión logística dado que según literatura una de sus principales aplicaciones son los problemas de clasificación binaria [Ruiz, 2018]. En este caso se usó este clasificador con la técnica OvR (Uno contra todos) y con el algoritmo de optimización de descenso coordinado conocido como liblinear.
- **Máquina de soporte vectorial (SVM):** Se decidió probar este algoritmo de clasificación dado que se utilizó en el estado del arte y se caracteriza por su capacidad de trabajar con características linealmente no separables. Dicho método evaluó las características previamente seleccionadas e hizo uso de un hiper plano que permitió separarlas y luego clasificarlas, el proceso del cálculo del hiper plano se evidencia en la subsección 4.2.7.

Cabe resaltar que, se usó este sistema de clasificación con diferentes kernels y diferentes valores de optimización, con el fin de evaluar con qué parámetros se presentaban los mejores resultados de las métricas de evaluación. El coeficiente de tolerancia denotado con la letra C es un parámetro empírico de penalización del término de error, que se utiliza para controlar la tolerancia de los valores atípicos, se suele usar un valor de $C = 1$ [Liu et al., 2014], sin embargo, se probaron valores más bajos y más altos con los que se esperaba cambiara el margen del hiper plano, por otro lado, los kernels que se probaron fueron: función de base radial (RBF), lineal, polinomial y sigmoideo.

Con los kernels anteriores se generaron diferentes hiper planos que se esperaba arrojaran diferentes resultados de desempeño, estos se calcularon con las siguientes ecuaciones de la función del kernel $K(x_1, x_2)$.

El kernel RBF y sigmoideo se representan de la siguiente forma, donde x_1 y x_2 son los valores de las clases de entrada y α es la anchura del kernel [Hsu et al., 2003].

$$K(x_1, x_2) = \exp - \frac{\|x_1 - x_2\|^2}{2\alpha^2} \quad (5-27)$$

Para el kernel lineal,

$$K(x_1, x_2) = x_1^T x_2 \quad (5-28)$$

Para el kernel polinómico donde ρ es el tamaño mínimo del polinomio.

$$K(x_1, x_2) = (x_1^T x_2 + 1)^\rho \quad (5-29)$$

- **Random forest (RF):** Se probó este algoritmo dado que fue uno con los mejores resultados de clasificación que muestra el estado del arte. Dicho algoritmo necesitó dos parámetros: el parámetro m (número de variables utilizadas al azar para cada división o número de variables predictivas) y el parámetro k (número de árboles de clasificación). En este caso se variaron la cantidad de árboles de decisión que conforman el clasificador, para elegir el valor de k que mejores resultados de las métricas de clasificación arrojara; el proceso de la toma de decisión de este clasificador se evidencia en la subsección 4.2.7.
- **Bayes:** también se probó el clasificador bayesiano dado que es uno de los que se han utilizado para la clasificación de artefactos en imágenes endoscópicas, la configuración usada para este clasificador fue el algoritmo de Naive Bayes Gaussiano.
- **Red neuronal (RN):** se implementó un algoritmo de perceptrón multicapa (MLP) que se entrenó usando backpropagation cuyo proceso de toma de decisión se evidencia en la subsección 4.2.7. En este caso, se utilizó una RN de 4 capas ocultas con 30 nodos, y de la cual se variaron algunos parámetros. Estos fueron: la función de activación para la capa oculta y el solver para la optimización del peso de la red de forma iterativa en función de los datos de entrenamiento, los optimizadores que se probaron incluyeron algoritmos como los basados en métodos cuasi-Newton (lbfgs) y los basados en el descenso de gradiente estocástico (sgd y adam).

Por otro lado, las funciones de activación con las que se probaron son dadas por las siguientes ecuaciones, donde x son los datos de entrada [Nwankpa et al., 2018].

Para la función de activación no operativa o identidad,

$$f(x) = x \quad (5-30)$$

La función de unidad lineal rectificada (ReLU),

$$f(x) = \max(0, x) = \begin{cases} x_i & Si \quad x_i \geq 0 \\ 0 & Si \quad x_i < 0 \end{cases} \quad (5-31)$$

La función tangente hiperbólica (Tanh),

$$f(x) = \tanh(x) = \left(\frac{e^x - e^{-x}}{e^x + e^{-x}} \right) \quad (5-32)$$

Y para la Función sigmoidea o función logística,

$$f(x) = \left(\frac{1}{1-e^{-x}} \right) \quad (5-33)$$

Los clasificadores con los parámetros anteriormente mencionados fueron los que se probaron con cada uno de los métodos de partición; a continuación se puede observar los clasificadores con mejor desempeño por cada método de partición; cabe resaltar que, no se muestran los resultados de los clasificadores con todos los parámetros que se variaron, solo se muestran los parámetros con los que el clasificador obtuvo mejor desempeño, sin embargo, todos los resultados de las pruebas realizadas se podrán encontrar en la carpeta: anexos/tablas/(metodos_identificacion_DM.xlsx y metodos_identificacion_RE.xlsx).

Métodos de partición de datos:

Hold out

Es uno de los métodos más simples de validación cruzada, donde el conjunto de datos se dividió en dos, en 70 % y 30 % para entrenamiento y prueba, respectivamente. Este se probó dado que se caracteriza por arrojar resultados estables y ser usado para bases de datos grandes, que en este caso es apropiado principalmente para RE por el número de datos con los que se contaban. En la tabla 5-28 se pueden observar los resultados de las métricas de desempeño (ver subsección 4.3) de los clasificadores con mejores resultados al usar este método de partición; donde cuyos resultados arrojaron que tanto para la identificación de DM como para la de RE el clasificador que obtuvo mejor desempeño fue el RF con 30 estimadores.

	Exactitud		Precisión		Sensibilidad		Especificidad		Score	
	RE	DM	RE	DM	RE	DM	RE	DM	RE	DM
Random Forest (15 estimadores)	0.95	0.71	0.96	0.74	0.94	0.66	0.95	0.76	0.95	0.71
Random Forest (30 estimadores)	0.95	0.71	0.96	0.77	0.94	0.61	0.96	0.82	0.95	.71
SVM (Kernel: rbf)	0.93	0.68	0.93	0.75	0.93	0.56	0.93	0.80	0.93	0.68
Bayes	0.83	0.66	0.91	0.74	0.74	0.49	0.93	0.83	0.83	0.66
Regresión Logística	0.83	0.67	0.85	0.73	0.80	0.55	0.86	0.79	0.83	0.67
Red neuronal (F. de activación: relu)	0.88	0.66	0.91	0.69	0.85	0.57	0.92	0.74	0.88	0.66

Tabla 5-28: Resultados métricas de desempeño métodos de identificación de DM y RE con partición hold out

K-fold

Este método de validación cruzada dividió aleatoriamente el conjunto de datos en K-subconjuntos, en este caso el valor de K varió de la siguiente manera: 10, 15, 20, 30, 39, 49, 60 y 70. De acuerdo a las particiones hechas según el valor de K, se tomó uno de los subconjuntos el cual utilizó como datos de prueba y el resto (K-1) como datos de entrenamiento, este método se repitió durante K iteraciones.

Por su parte la evaluación del desempeño de los clasificadores se realizó por cada iteración, es decir que según las iteraciones dadas por K se contaría con diferentes resultados de las métricas de evaluación, por lo que la estimación el desempeño de los clasificadores se realizó a través de la media de las métricas que se obtuvieron con cada una de las iteraciones. Cabe resaltar que este método se probó dado que reduce significativamente el sesgo del desempeño de los clasificadores ya que utiliza la mayoría de los datos para el ajuste.

Sin embargo, una de las desventajas de este método es que suele no ser estable por lo que se optó por repetir el proceso anteriormente descrito 10 veces por cada valor de K, y sacar la media de las medias de los resultados de las métricas de desempeño de los clasificadores, estos se pueden observar en las tablas **5-29** y **5-30**.

	Exactitud	Precisión	Sensibilidad	Especificidad	Score
	DM	DM	DM	DM	DM
8-Fold Random Forest (15 estimadores)	0.74	0.76	0.71	0.78	0.74
8-Fold Random Forest (30 estimadores)	0.74	0.76	0.70	0.78	0.74
10-Fold SVM (Kernel: rbf)	0.71	0.71	0.72	0.70	0.71
2-Fold Bayes	0.69	0.73	0.62	0.77	0.69
2-Fold Regresión Logística	0.70	0.71	0.66	0.72	0.70
2-Fold Red neuronal (F. de activación: relu)	0.69	0.69	0.68	0.69	0.69

Tabla 5-29: Resultados métricas de desempeño métodos de identificación de DM con partición k-fold

En la tabla **5-29** se observa que nuevamente el RF obtuvo los mejores resultados, con una cantidad de arboles de 15 y el valor de K igual a 8. Cabe mencionar que este clasificador obtuvo la mejor media con una desviación estándar pequeña para todas sus métricas (Exactitud: 0.74 ± 0.016 , Precisión: 0.76 ± 0.017 , Sensibilidad: 0.71 ± 0.019 , Especificidad: 0.78 ± 0.016 , Score: 0.74 ± 0.016); por consiguiente, se reafirma que el resultado es estable.

	Exactitud	Precisión	Sensibilidad	Especificidad	Score
	RE	RE	RE	RE	RE
20-Fold Random Forest (15 estimadores)	0.96	0.96	0.95	0.96	0.96
2-Fold Random Forest (30 estimadores)	0.96	0.96	0.95	0.96	0.96
10-Fold SVM (Kernel: rbf)	0.91	0.90	0.91	0.92	0.91
2-Fold Bayes	0.83	0.91	0.73	0.93	0.83
2-Fold Regresión Logística	0.82	0.84	0.79	0.85	0.82
25-Fold Red neuronal (F. de activación: relu)	0.88	0.91	0.85	0.91	0.88

Tabla 5-30: Resultados métricas de desempeño métodos de identificación de RE con partición k-fold

En la tabla **5-30** se observa que nuevamente el RF obtuvo los mejores resultados, con una cantidad de arboles de 30 y el valor de K igual a 2. Cabe mencionar que este clasificador obtuvo la mejor media con una desviación estándar pequeña para todas sus métricas (Exactitud: 0.96 ± 0.001 , Precisión: 0.96 ± 0.002 , Sensibilidad: 0.95 ± 0.001 , Especificidad: 0.95

± 0.002 , Score: 0.96 ± 0.001); por consiguiente, se reafirma que el resultado es estable.

Leave one out cross validation (LOOCV)

Conocida como la validación cruzada de dejar uno afuera, porque del conjunto de datos total se eligió un solo dato para realizar la prueba y el resto para el entrenamiento; esto se repitió por cada uno de los datos del conjunto total por lo que es computacionalmente costoso. Este se probó sobre todo para identificación de DM porque no se cuenta con una cantidad de datos considerable grande, contrario a lo que ocurre con el caso de identificación de RE, es por esto que desde un comienzo ya se pronosticaba que este método de partición no iba a ser el elegido para la detección de RE.

Por su parte la evaluación de desempeño de los clasificadores se realizó por cada iteración y al tener todos los resultados de estas, se dispuso a obtener la media, estos se muestran en la tabla 5-31. Cuyos resultados arrojaron que al igual que en los métodos de partición anteriores el clasificador RF obtuvo el mejor desempeño, en este caso con 15 árboles. Cabe resaltar que se espera que este método de partición no sea el elegido para la identificación de RE dado que tuvo un alto costo computacional, un mayor tiempo de ejecución y los resultados son similares a los que se obtuvieron con hold out y kfold.

	Exactitud		Precisión		Sensibilidad		Especificidad		Score	
	RE	DM	RE	DM	RE	DM	RE	DM	RE	DM
Random Forest (15 estimadores)	0.96	0.74	0.96	0.77	0.95	0.69	0.96	0.79	0.99	0.99
Random Forest (30 estimadores)	0.96	0.74	0.96	0.77	0.95	0.68	0.96	0.79	0.99	0.99
SVM (Kernel: rbf)	0.87	0.68	0.86	0.69	0.88	0.65	0.85	0.71	0.87	0.68
Bayes	0.83	0.68	0.91	0.71	0.72	0.60	0.93	0.76	0.82	0.68
Regresión Logística	0.86	0.68	0.84	0.69	0.79	0.65	0.85	0.71	0.82	0.68
Red neuronal (F. de activación: relu)	0.85	0.68	0.83	0.70	0.80	0.65	0.83	0.72	0.81	0.68

Tabla 5-31: Resultados métricas de desempeño métodos de identificación de DM y RE con partición leave one out

Evaluación de resultados

Para que sea más clara la evaluación de los resultados presentados anteriormente, se dividirá en dos secciones: la primera para identificación de DM y la segunda para identificación de RE.

Para identificación de DM:

Por cada método de partición anteriormente descrito se eligió un clasificador con los mejores resultados, es decir, el clasificador que obtuvo mejor desempeño; para todos los métodos el clasificador que se eligió fue el Random Forest pero con diferentes números de árboles. Para el método hold out fue de 30 árboles pero para k-fold y leave one out fue de 15. Los resultados de las métricas de desempeño de estos se observan en la tabla 5-32.

	Exactitud	Precisión	Sensibilidad	Especificidad	Score
	DM	DM	DM	DM	DM
(HOLD OUT) Random Forest (30 estimadores)	0.71	0.77	0.61	0.82	0.71
(K-FOLD) 8-Fold Random Forest (15 estimadores)	0.74	0.76	0.71	0.78	0.74
(Leave One Out) Random Forest (15 estimadores)	0.74	0.77	0.69	0.79	0.99

Tabla 5-32: Resultados métricas de desempeño para los métodos de identificación de DM

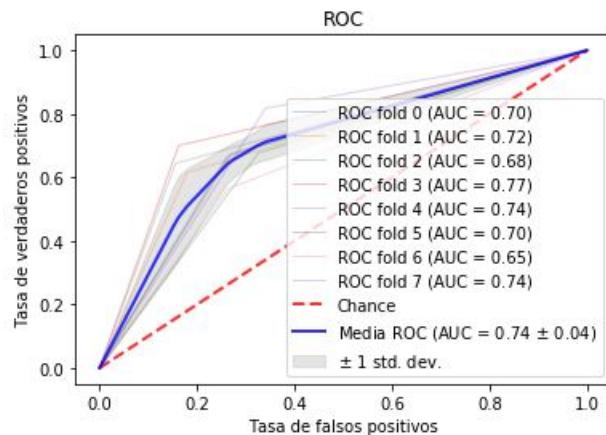


Figura 5-42: ROC del clasificador elegido para identificación de DM

Como se puede observar en la tabla 5-32, las métricas de los clasificadores elegidos son cercanas entre si, sin embargo, se optó por elegir el método de partición k-fold con una $k=8$ con el clasificador RF de 15 arboles, por otro lado, se puede observar en la figura 5-42 la gráfica de la ROC por cada uno de los 8-fold, y donde muestra que la media de el valor de área bajo la curva es de 0.74. Estos resultados fueron comparados con los que están presentes en la literatura como se visualiza en la tabla 5-33. Antes de analizar los datos presentes en dicha tabla, cabe resaltar la complejidad de la clasificación efectuada en este trabajo de grado, dado que los datos tienen una alta variabilidad como se mencionó en la subsección 5.1, pues las imágenes endoscópicas entregadas por la base de datos EAD2019 fueron tomadas con diferentes tipos de endoscopios, en diferentes partes del cuerpo, en diferentes personas y en diferentes establecimientos.

	Exactitud	Precisión	Sensibilidad	Especificidad	AUC
	DM	DM	DM	DM	DM
Método de identificación elegido	0.74	0.76	0.71	0.78	0.74
[Oh et al., 2007]	0.905	0.89	0.88	0.92	-
[Kaçmaz et al., 2018]	0.76	-	0.80	0.72	0.76

Tabla 5-33: Comparación de los resultados de las métricas de desempeño para los métodos de identificación elegido de DM contra resultados expuestos en la literatura

Aunque es evidente que las métricas reportadas en la literatura tienen un mejor desempeño que las arrojadas por el clasificador que se eligió, es importante tener en cuenta que tanto

el estudio realizado por [Oh et al., 2007] y [Kaçmaz et al., 2018] no contaron con tanta variabilidad en los datos dado que solo usaron imágenes colonoscópicas y sumado a esto, en el caso de [Oh et al., 2007] las imágenes con las que realizaron su estudio provenían de dos videos endoscopicos, es decir, solo de dos pacientes.

Sin embargo, se encontró un articulo que trabajó con la misma BD (EAD 2019), es decir, contó con la misma variabilidad en los datos; este estudio detectó los siete artefactos proporcionados por esta, y cuya media de precisión para la detección de DM fue de 0.71 [Oksuz et al., 2019]. Es decir, que la precisión obtenida con el clasificador elegido en este trabajo es ligeramente mayor que la obtenida en el trabajo de [Oksuz et al., 2019].

Para identificación de RE:

Por cada método de partición anteriormente descrito se eligió un clasificador con los mejores resultados, para todos los métodos el clasificador que se eligió como en el caso anterior fue el Random Forest pero con diferentes números de árboles. Para el método hold out y k-fold fue de 30 árboles pero para leave one out fue de 30. Los resultados de las métricas de desempeño de estos se observa en el tabla **5-34**.

	Exactitud	Precisión	Sensibilidad	Especificidad	AUC
	RE	RE	RE	RE	RE
(HOLD OUT) Random Forest (30 estimadores)	0.95	0.96	0.94	0.96	0.95
(K-FOLD) 2-Fold Random Forest (30 estimadores)	0.96	0.96	0.95	0.96	0.96
(LEAVE ONE OUT) Random Forest (15 estimadores)	0.96	0.96	0.95	0.96	0.99

Tabla 5-34: Resultados métricas de desempeño métodos de identificación de RE

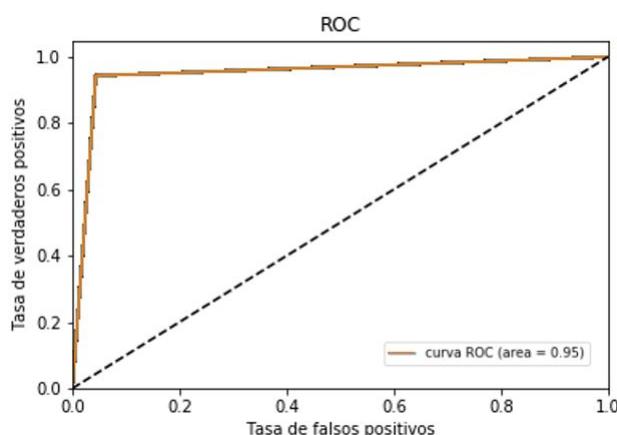


Figura 5-43: ROC del clasificador elegido para identificación de RE

Como se puede observar en la tabla **5-34**, las métricas de los clasificadores elegidos son cercanas entre sí, sin embargo, se optó por elegir el método de partición hold out con el

clasificador RF de 30 arboles que arrojó un AUC de 0.95 como se ve en la figura 5-43. Un dato adicional del porque se optó por escoger este método de partición es por que arroja resultados estables y tiene muy bajo costo computacional. Estos resultados fueron comparados con los que están presentes en la literatura como se puede observar en la tabla 5-35.

	Exactitud	Precisión	Sensibilidad	Especificidad	AUC
	RE	RE	RE	RE	RE
Método de identificación elegido	0.95	0.96	0.94	0.96	0.95
[Oh et al., 2007]	0.99	0.92	0.96	0.99	-
[Arnold et al., 2009]	98.3	81.3	75.3	99.2	-
[Arnold et al., 2011]	0.91	0.96	0.92	0.90	-

Tabla 5-35: Comparación de los resultados de las métricas de desempeño para los métodos de identificación elegido de RE contra resultados expuestos en la literatura

En la tabla 5-35 se puede observar que el método de clasificación que se eligió para la identificación de reflejos especulares dio un resultado comparable con los encontrados en la literatura, sin embargo, es importante es importante mencionar que ninguno de los estudios presentes en la tabla contaron con tanta variabilidad en los datos dado que solo usaron imágenes colonoscópicas.

Sin embargo, en el estudio de [Oksuz et al., 2019] que como se mencionó anteriormente utilizó las imágenes proporcionadas por la BD EAD 2019, obtuvo una media de la precisión de 0.56, es decir, que el método de clasificación elegido en este trabajo obtuvo una mejor precisión.

Evaluación de la pertinencia de unir características

La unión de características busca mejorar las métricas de desempeño de un clasificador, es decir que la idea es que las características que van a ser combinadas sean diferenciables, pero al mismo tiempo esa diferencia no debe representar un costo negativo en el desempeño del método de clasificación [Ellis, 2000]. Dicho lo anterior, tras obtener las características más informativas de cada clase se hizo la prueba para evaluar el desempeño del clasificador a partir de la unión de características; esta unión se hizo teniendo en cuenta la correlación, pues como se dijo inicialmente se busca que las características sean diferenciables.

A continuación se describe el proceso de combinación de características para DM y después para RE con el fin de evaluar dos hipótesis, de acuerdo a la descripción anterior:

1. El método de identificación no mostrará un mejor desempeño al unir características con alta correlación.

2. El método de identificación mostrará un mejor desempeño al unir características con baja correlación.

- **Unión de características para DM:** la combinación de características se hace concatenando una característica con otra, para ello, como se puede observar en la tabla **5-36**, se buscaron las características con mayor y menor correlación entre Entropía y SODMS (características de mayor separabilidad seleccionadas en la sección 5.5.2). Para el caso de la **Entropía** la característica con mayor correlación fue la Disimilitud y la de menor fue la media del coeficiente de detalle horizontal de la transformada Wavelet (para mayor información ver sección 4.2), en la tabla se encuentra como *media_LH*; para **SODMS** la característica con mayor correlación fue la Homogeneidad y la de menor fue la media de las altas frecuencias de la transformada Wavalet, en la tabla se puede encontrar como *media_altas*. Los datos de correlación de todas las características de DM se encuentran en la carpeta: **anexos/tablas/ (Correlacion-caracteristicasDM.xlsx**.

	Entropía	SODMS	media_LH	media_altas	Homogeneidad	Disimilitud
Entropía	1.00	0.70	0.01	0.04	0.69	0.71
SODMS	0.70	1.00	0.02	0.03	0.90	0.79
media_LH	0.01	0.02	1.00	0.92	0.01	0.01
media_altas	0.04	0.03	0.92	1.00	0.02	0.00
Homogeneidad	0.69	0.90	0.01	0.02	1.00	0.86
Disimilitud	0.71	0.79	0.01	0.00	0.86	1.00

Tabla 5-36: Correlación entre características de DM

Los resultados tras la unión de características de alta y baja correlación se muestran en la tabla **5-37**, como se observa, para la identificación de DM en imágenes de entrenamiento no se obtuvo un mejor desempeño cuando se unieron características de baja correlación, es decir que **media_LH** y **media_altas** no aportaron información relevante que permitiera mejorar el desempeño del clasificador. Sin embargo, en esta primer evaluación, el desempeño aumentó al unir características con alta correlación (**Entropía + Disimilitud / SODMS + Homogeneidad**), por lo tanto se hace una segunda evaluación con las imágenes de prueba entregadas por la BD, es decir con información completamente nueva para el método de identificación, para así evaluar la estabilidad del clasificador al unir dichas características, los resultados pueden ser observados en la tabla **5-38**.

	Exactitud	Precisión	Sensibilidad	Especificidad	Score
Método de identificación elegido (sin unión)	0.74	0.76	0.71	0.78	0.74
Unión de características con baja correlación	0.74	0.76	0.71	0.77	0.74
Unión de características con alta correlación	0.87	0.85	0.88	0.86	0.74

Tabla 5-37: Resultados métricas de desempeño en imágenes de entrenamiento con la unión de características DM

De acuerdo con los resultados de la tabla 5-38 se observa que al evaluar el método de identificación con datos nuevos (imágenes de prueba), la unión de características tanto de alta como de baja correlación reducen significativamente el desempeño del método, probablemente porque se agrega sesgo en la información al unir características con un alto índice de correlación. Sin embargo, la variación entre la exactitud del método de identificación seleccionado sin unión de características (K-fold - ver tabla 5-32) con imágenes de entrenamiento e imágenes de prueba es de 0.06. En consecuencia se demuestra que en este caso la unión de características no resultó pertinente, pues no aumenta el desempeño del método sino que disminuye en parámetros como exactitud, precisión y sensibilidad, además también se logra evidenciar la estabilidad del método de identificación seleccionado inicialmente.

	Exactitud	Precisión	Sensibilidad	Especificidad	Score
Método de identificación elegido (sin unión)	0.68	0.68	0.64	0.82	0.53
Unión de características con baja correlación	0.57	0.56	0.64	0.50	0.57
Unión de características con alta correlación	0.56	0.53	0.90	0.22	0.56

Tabla 5-38: Resultados métricas de desempeño con imágenes de prueba DM

- **Unión de características para RE:** la combinación de características se hace concatenando una característica con otra, para ello se buscaron las características con mayor y menor correlación entre Contraste, Desviación de intensidades y Brillo (características de mayor separabilidad para RE seleccionadas en la sección 5.5.2), como se puede observar en la tabla 5-39. Para el caso del **contraste** la característica con mayor correlación fue la Disimilitud y la de menor fue *curtosis del histograma*, para la **Desviación de intensidades** la característica con mayor correlación fue la *Varianza de intensidades* y la de menor fue *ASM (Segundo Momento Angular)*, por último para el **Brillo** la característica con mayor correlación fue la *Asimetría del histograma* y la de menor fue *Entropía*. Los datos de correlación de todas las características de RE se encuentran en la carpeta: **anexos/tablas/CorrelacioncaracteristicasRE.xlsx**.

	Contraste	Desviación_intensidades	Brillo	Curtosis_Hist	ASM	Entropía	Varianza_intensidades	Disimilitud	Asimetría_Hist
Contraste	1.00	0.42	0.72	0.11	0.10	0.24	0.34	0.95	0.38
Desviación_intensidades	0.42	1.00	0.25	0.04	0.02	0.30	0.96	0.45	0.15
Brillo	0.72	0.25	1.00	0.19	0.11	0.09	0.19	0.70	0.42
Curtosis_Hist	0.11	0.04	0.19	1.00	0.08	0.21	0.02	0.07	0.51
ASM	0.10	0.02	0.11	0.08	1.00	0.77	0.10	0.10	0.09
Entropía	0.24	0.30	0.09	0.21	0.77	1.00	0.14	0.41	0.03
Varianza_intensidades	0.34	0.96	0.19	0.02	0.10	0.14	1.00	0.34	0.10
Disimilitud	0.95	0.45	0.70	0.07	0.20	0.41	0.34	1.00	0.40
Asimetría_Hist	0.38	0.15	0.42	0.51	0.09	0.03	0.10	0.40	1.00

Tabla 5-39: Correlación entre características de RE

Los resultados en el desempeño del método de identificación de RE en imágenes de entrenamiento cuando se unen características de alta y baja correlación se muestran

en la tabla **5-40**. Como se observa, no se obtuvo un mejor desempeño al unir características de baja correlación, es decir que la varianza de intensidades, la disimilitud y la asimetría del histograma no aportaron información relevante que permitiera mejorar el desempeño del clasificador. En este caso también es importante evaluar la respuesta del clasificador ante imágenes de test dadas por la BD (datos completamente nuevos para el método), para así confirmar la elección del método de identificación y evaluar la estabilidad del mismo, los resultados pueden ser observados en la tabla **5-41**.

	Exactitud	Precisión	Sensibilidad	Especificidad	Score
Método de identificación elegido (sin union)	0.95	0.96	0.94	0.96	0.95
Unión de características con baja correlación	0.89	0.89	0.90	0.89	0.89
Unión de características con alta correlación	0.92	0.93	0.90	0.93	0.92

Tabla 5-40: Resultados métricas de desempeño en imágenes de entrenamiento con la unión de características RE

Los datos en la tabla **5-41** demuestran que aún con los datos nuevos el método de identificación sin unir características elegido anteriormente (Hold-out-ver tabla **5-34**) es el que mejor desempeño tiene.

	Exactitud	Precisión	Sensibilidad	Especificidad	Score
Método de identificación elegido (sin unión)	0.94	0.98	0.90	0.98	0.94
Unión de características con baja correlación	0.90	0.97	0.83	0.97	0.90
Unión de características con alta correlación	0.89	0.96	0.81	0.97	0.89

Tabla 5-41: Resultados métricas de desempeño con imágenes de prueba RE

De acuerdo a lo anterior, la unión de características no demostró ser pertinente, pues el método de clasificación redujo sus métricas de desempeño tanto en RE como en DM, por lo tanto para la identificación de DM se continuaron usando las características **Entropía** y **SODMS** y para RE **Contraste, Desviación de intensidades y Brillo**.

5.7. Método de corrección de RE

Como se mencionó anteriormente la presencia de RE en imágenes endoscópicas puede ser una fuente de error para muchos algoritmos de diagnóstico, puesto que pueden generar regiones falsas, afectar las características de textura, así como, causar errores en segmentación, reconocimiento de objetos, reconstrucción de formas, etc [Gao et al., 2017]. Todos estos posibles errores se convierten en un obstáculo para el posterior desarrollo o uso de una herramienta de apoyo al diagnóstico, por lo que se torna importante detectar RE y corregirlos.

Dicha corrección se hace para que apoye el desarrollo de algoritmos DAO, minimizando la incidencia de los errores mencionados en tareas como: detección de marcos no informativos, segmentación y detección de instrumentos quirúrgicos, reconstrucción de estructuras 3D o corrección de la desalineación del canal de color [Münzer et al., 2018]. Por tal motivo en este trabajo, se decidió evaluar el desempeño de algunos métodos de corrección existentes de la siguiente manera:

Montaje experimental

Para hacer una evaluación de los métodos de corrección lo primero que se hizo fue seleccionar las imágenes endoscópicas sin presencia de RE o DM, se obtuvieron **207 imágenes**. Posteriormente se simularon los reflejos especulares, para esto se tuvieron en cuenta ciertos aspectos que los caracterizan como lo son:

- Los RE son regiones fuertemente resaltadas en las imágenes que aparecen como un resplandor blanco o un resplandor de color claro en las imágenes, es decir, cuentan con muchos valores de intensidades muy altas [Marcinczak and Grigat, 2013].
- Varían de forma y tamaño [Akbari et al., 2018].
- Están distribuidos al azar en superficies de tejido rojas difusas o rugosas (pliegues) [Khanian et al., 2018].
- Tienen un brillo alto y una baja saturación [Oh et al., 2007].

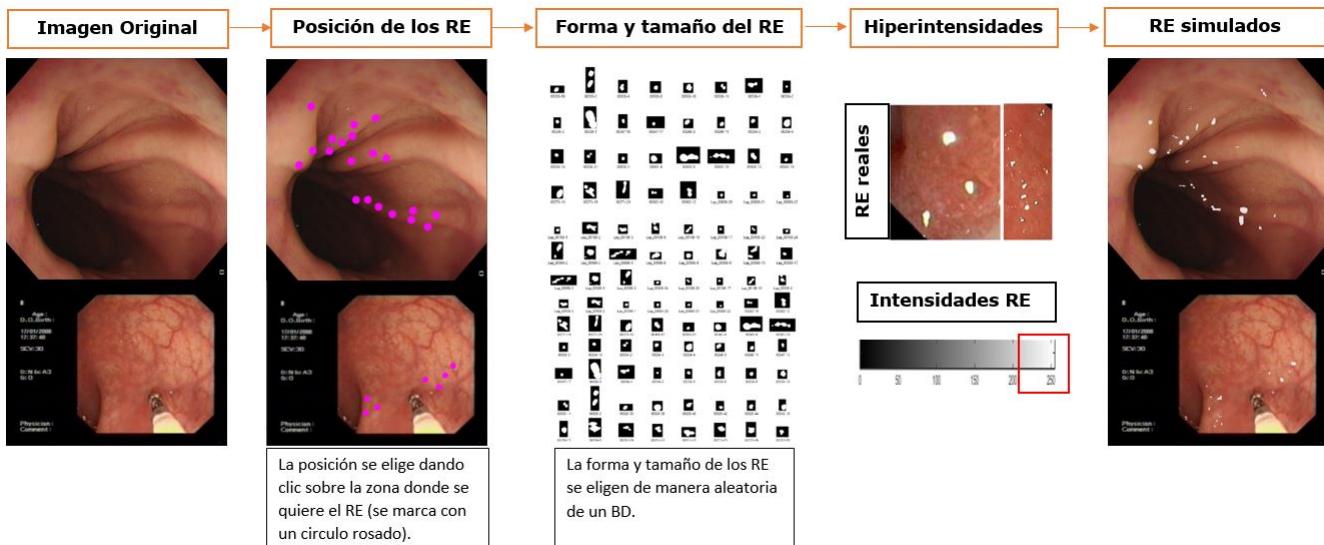


Figura 5-44: Método para simular RE

Teniendo en cuenta lo anterior se procedió a simular los reflejos, para lo cual se tomó imagen por imagen de las seleccionadas (sin presencia de RE o DM) a la que se le eligió la cantidad y posición de los RE, como se observa en la figura 5-44. La forma y el tamaño de los RE se eligen de manera aleatoria de una BD que contiene bbox de máscaras de RE reales, los cuales fueron tomados de las imágenes proporcionadas por la BD EAD 2019; dichos bbox de reflejos se agregan a la imagen como una intensidad alta, en tonos cercanos al blanco. Posteriormente, a la imagen con el artefacto simulado se le aplicaron técnicas de corrección como el suavizado y distintos métodos de inpainting, los cuales serán explicados a continuación.

Por ultimo, para evaluar cada técnica de corrección se emplearon las métricas de similitud (SSIM, MSE, y PSNR) explicadas en la sección 4.3. Con dichas métricas se procedió a elegir el método de corrección con mejor desempeño, es decir, el método que con la corrección dejara la imagen con artefacto (simulados) mas similar a la original (sin RE), en términos de las métricas, el método con un SSIM mas cercano a uno, un MSE más cercano a cero y con el PSNR más alto. A continuación se observa la media de los resultados de dichas métricas por cada uno de los métodos de corrección, sin embargo, todos los resultados se encuentran en la carpeta: **anexos/tablas/metodos_correccion.xlsx**.

Técnicas de corrección

- **Suavizado:** esta técnica consiste en aplicar un filtro de mediana cuya función principal es transformar los valores de intensidad de los RE, los cuales son muy distintos al resto de la vecindad, a valores más próximos de sus vecinos, para ello el filtro reemplaza el valor gris de un punto por la mediana de los niveles de gris de una cierta vecindad [Gonzalez and Woods, 2014], como se ve en la figura 5-45.

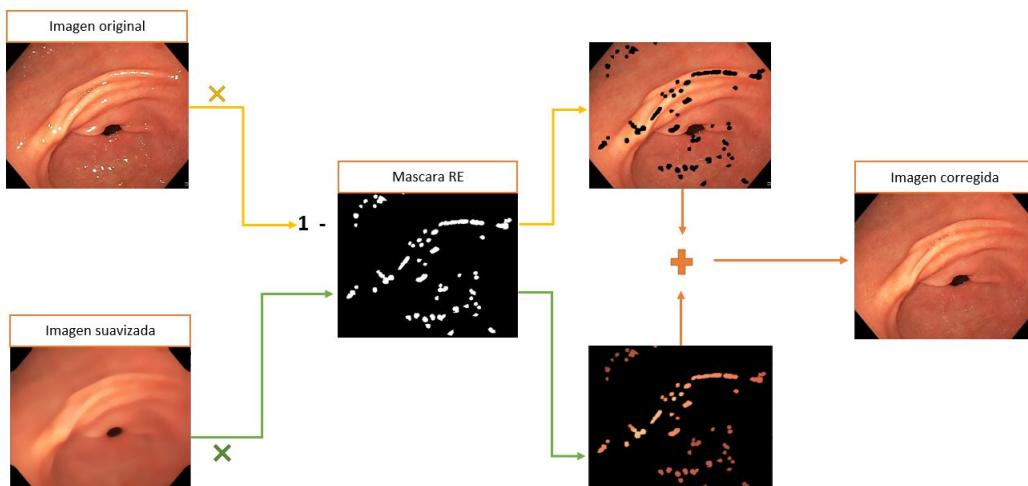


Figura 5-45: Técnica de suavizado

La elección de la máscara que lograba la mejor corrección de los RE se hizo comparando cada resultado con la imagen original mediante las métricas de similitud inicialmente mencionados. Como se puede observar en la tabla 5-42 la máscara con un valor de vecindad de 20 fue la que arrojó una mejor corrección, pues su MSE es el menor y el PSNR es el mayor.

	SSIM	MSE	PSNR
Con mascara de 20	0.996 ± 0.002	0.528 ± 0.370	51.644 ± 2.533
Con mascara de 25	0.996 ± 0.001	0.534 ± 0.344	51.567 ± 2.554
Con mascara de 30	0.996 ± 0.002	0.556 ± 0.386	51.426 ± 2.554

Tabla 5-42: Resultados métricas de similitud método corrección por suavizado

- **Inpainting_1:** la técnica de inpainting se entiende como el proceso de reconstrucción o llenado automático de partes dañadas o perdidas de una imagen o video, mediante la extracción de información presente en regiones que no estén dañadas [Damelin and Hoang, 2018]. La primer técnica de inpainting que se probó es de la herramienta *scikit-image*, la cual toma la imagen dañada y la máscara (RE segmentados) para corregirla mediante ecuaciones biarmónicas.

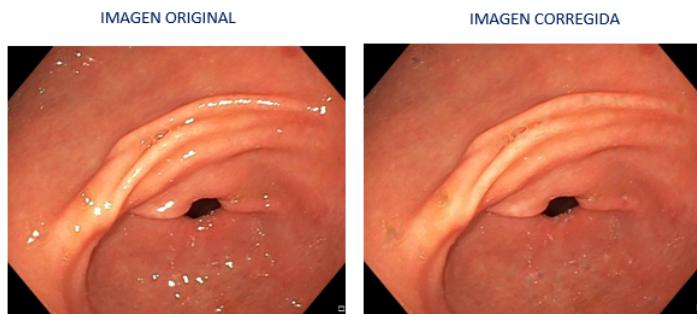


Figura 5-46: Corrección de RE mediate *Inpainting_1*

En la figura 5-46 se puede observar la media de los resultados de la corrección tras aplicar el método de inpainting 1 en las imágenes con RE simulados. Esta técnica también fue evaluada mediante la métricas de similitud y como se observa en la tabla 5-43 su desempeño es menor a la técnica de suavizado.

	SSIM	MSE	PSNR
Inpainting_1	0.995 ± 0.003	5.575 ± 36.349	46.294 ± 4.394

Tabla 5-43: Resultados métricas de similitud método corrección por Inpainting_1

- **Inpainting_2:** esta técnica de inpainting es de la librería de OpenCV y tal como lo anterior corrige una región específica (RE) teniendo en cuenta la información disponible de su vecindad mediante un método basado en las ecuaciones de Navier-Stokes.

Este enfoque usa la idea de la dinámica clásica de los fluidos para propagar las líneas de los isótopos continuamente desde el exterior hasta la región que debe ser corregida. Por lo tanto, la intensidad de la imagen se toma como una función de flujo para un flujo incompresible bidimensional [Au and Takei, 2001], como se muestra a continuación:

Dinámica de fluidos	Procesamiento de imágenes
función de flujo Ψ	Intensidad de la imagen I
velocidad del fluido $v = \nabla^\perp \Psi$	Dirección del isófoto $\nabla^\perp I$
Vorticidad $w = \Delta \Psi$	suavidad $w = \Delta I$
Viscosidad ν	Difusión anisotrópica ν

Tabla 5-44: Aplicación de la ecuación de Navier-Stokes en el método de Inpainting

De acuerdo a la tabla 5-44 la función de flujo en términos de procesamiento de imágenes queda como se muestra a continuación:

$$\omega_t + v \cdot \nabla w = \nu \nabla \cdot (g(|\Delta \omega|) \Delta \omega) \quad (5-34)$$

La letra g en la ecuación 5-34 hace referencia al término de difusión anisotrópico o el término de difusión de preservación de bordes. En general éste termino es mayor para áreas suavizadas y cercano a cero para áreas con bordes [Au and Takei, 2001].

	SSIM	MSE	PSNR
Con radio de 1	0.996 ± 0.002	0.744 ± 0.449	50.158 ± 2.655
Con radio de 3	0.996 ± 0.002	0.636 ± 0.443	50.865 ± 2.625
Con radio de 10	0.996 ± 0.002	0.578 ± 0.407	51.260 ± 2.574

Tabla 5-45: Resultados métricas de similitud método corrección por Inpainting_2

Para la aplicación de ésta técnica se modificó el radio de la vecindad con la que se hizo la corrección de cada región donde hay RE. Como se observa en el tabla 5-45, el radio de **10** fue el que mejor valores de MSE y PSNR obtuvo, pues el SSIM es el mismo para los diferentes valores de radio probados.

- **Inpainting_3:** la tercera técnica de inpainting también es de la librería OpenCV, pero en este caso se aplica el método desarrollado por Alexandru Telea, el cual se basa en el *Método de Marcha Rápida o FMM por sus siglas en inglés*.

El FMM resuelve los problemas de los valores de bordes en la Ecuación Eikonal, en dónde:

$$|\nabla T| = 1 \text{ para el inpaintg completo y con } T = 0 \text{ en el límite de la región} \quad (5-35)$$

La solución T de la ecuación 5-35 es el mapa de distancia del conjunto de píxeles al límite de la región. Los conjuntos de niveles de T son exactamente los límites sucesivos que se necesitan para llevar a cabo el inpainting. Dicho lo anterior el FFM garantiza que los píxeles (límite de la región) siempre se procesen de orden creciente de su distancia al límite T , es decir que siempre se llenen primero los píxeles más cercanos al área de la imagen conocida [Telea, 2004].

	SSIM	MSE	PSNR
Con radio de 1	0.996 ± 0.002	0.714 ± 0.450	50.359 ± 2.658
Con radio de 3	0.996 ± 0.002	0.555 ± 0.400	51.448 ± 2.568
Con radio de 10	0.996 ± 0.002	0.490 ± 0.358	51.962 ± 2.493

Tabla 5-46: Resultados métricas de similitud método corrección por Inpaiting_3

Para la aplicación de esta técnica se hicieron 3 pruebas, modificando en cada una el radio de la vecindad que se iba a tomar para hacer la corrección. Como se observa en la tabla 5-46 el radio de **10** fue el que mejores resultados arrojó.

Evaluación de resultados

Dado que por cada método descrito anteriormente se eligió el método con los mejores resultados se cuenta con cuatro opciones, donde para el suavizado se escogió una máscara de 20 y para los métodos de inpaiting 2 y 3 un radio de 10. Los resultados de estos se pueden observar en las gráficas de la figura 5-47.

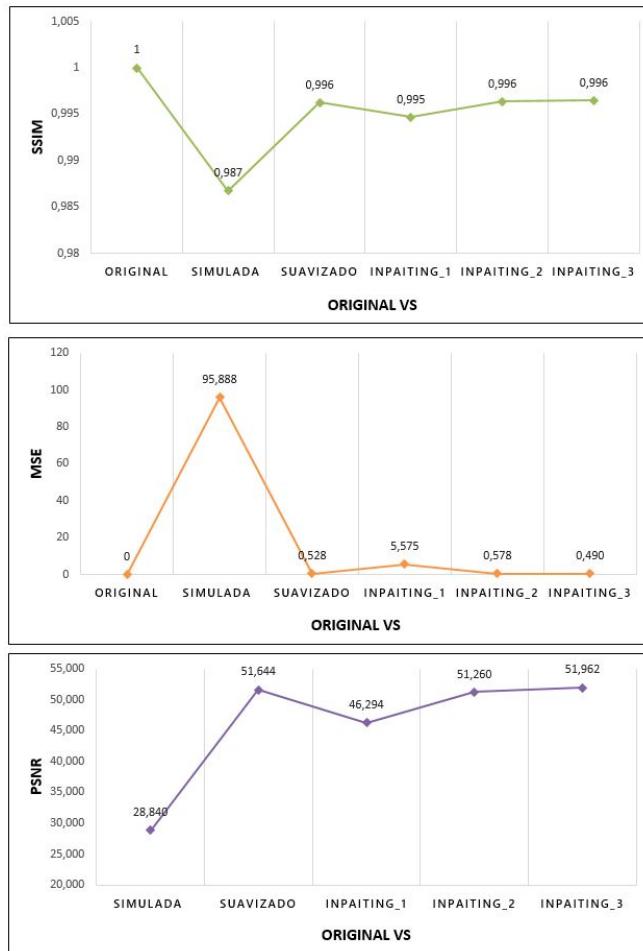


Figura 5-47: Medidas de evaluación

Como se observa en la figura 5-47, mediante los cuatro métodos de corrección mencionados anteriormente se logró que la imagen con artefactos se asemejara a la original, pues por ejemplo en la métrica MSE la imagen original comparada con la que tiene presencia de artefactos dio un resultado de 95,8 y después de aplicar la corrección ese valor disminuyó considerablemente, acercándose más a un valor de cero lo que da cuando la imagen original se compara con ella misma. El método elegido fue el de **inpaiting 3**, dado que es el que mayor valor de PSNR y menor MSE obtuvo.

5.8. Interfaz de presentación de resultados

Para la presentación de resultados se realizó una Interfaz Gráfica de Usuario (GUI), donde para su elaboración se tuvo en cuenta que el usuario al que va dirigido el método propuesto en este trabajo es para desarrolladores de posibles herramientas de apoyo al diagnóstico. Teniendo cuenta lo anterior, se plantearon algunos requerimientos con los que la interfaz

debía cumplir, los cuales son:

1. Debe permitir el ingreso de imágenes endoscópicas de cualquier tipo de formato.
2. Debe publicar el tipo de artefacto que se identificó o si no se detectó ninguno.
3. Debe mostrar la ubicación en la cual se detectó el artefacto.
4. Debe corregir la imagen automáticamente pero debe dar la opción de descargarla, si se identificaron RE.
5. Debe dar opción de eliminar la imagen, si se identificó DM.

Por cada uno de los requerimientos expuestos anteriormente se planteó el cómo y qué elementos dentro de la interfaz resultan necesarios para dar cumplimiento a dicha necesidad. Para satisfacer el **primer** requerimiento de permitir al usuario la selección e ingreso de imágenes en la GUI, partiendo de su localización en ficheros dentro del computador, encontramos como opción el ofrecer una ventana de explorador de carpetas accionada mediante un evento de *click* de botón en la interfaz. Asimismo, se planteó que nuestro visualizador debía presentar un espacio para el despliegue de la imagen que se escogió. Por otro lado, para el **segundo y tercero** requerimiento, se encontró que la interfaz debía contar con un botón que permitiera al usuario ejecutar el método propuesto y también debía contar con un espacio donde se indique qué tipo de detección hizo el método, asimismo, debía mostrar la imagen con la ubicación de los RE y/o DM en forma de bbox, logrando dejar en claro el bbox a que tipo de artefacto pertenece.

Por último, para el **cuarto y quinto** requerimiento, se determinó que la GUI debía contener un panel de opciones dentro del cual se incluyera la opción de guardar la imagen original, la imagen corregida o dado el caso eliminar la imagen de la carpeta. Teniendo en cuenta lo anterior se elaboró la interfaz como se ve en la figura 5-48.

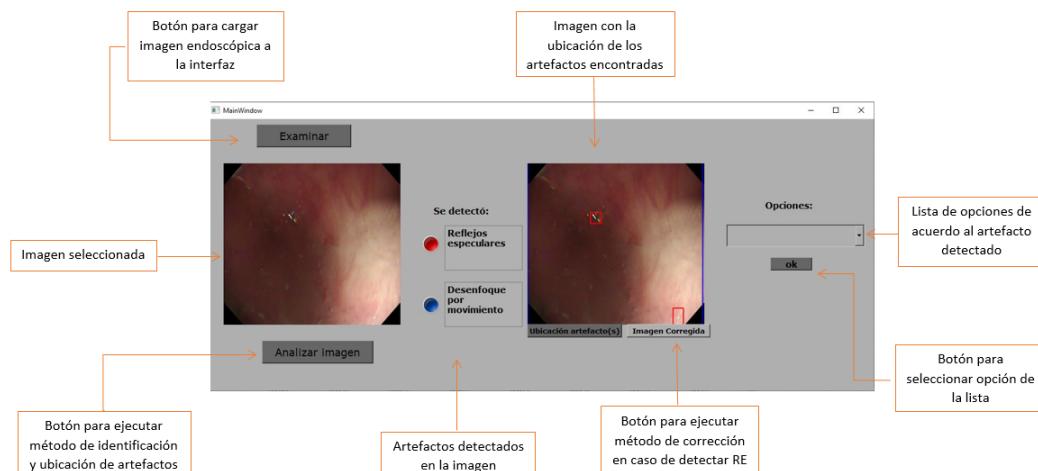


Figura 5-48: Interfaz de presentación de resultados

Cabe resaltar que, para usar la interfaz se deben seguir estos pasos:

1. Dar clic en **Examinar**, esto abrirá una ventana con el explorador de archivos para poder elegir la imagen endoscópica que se quiere analizar.
2. Elegir la imagen que se quiere analizar.
3. Dar clic en **Analizar imagen**, esto ejecutará todo el método de tratamiento de artefactos y dará 4 posibles resultados.

Posibles resultados:

1. Si el método detectó **reflejos especulares** en la imagen que se analizó, la interfaz gráfica dará las opciones que se observan en la figura 5-49.

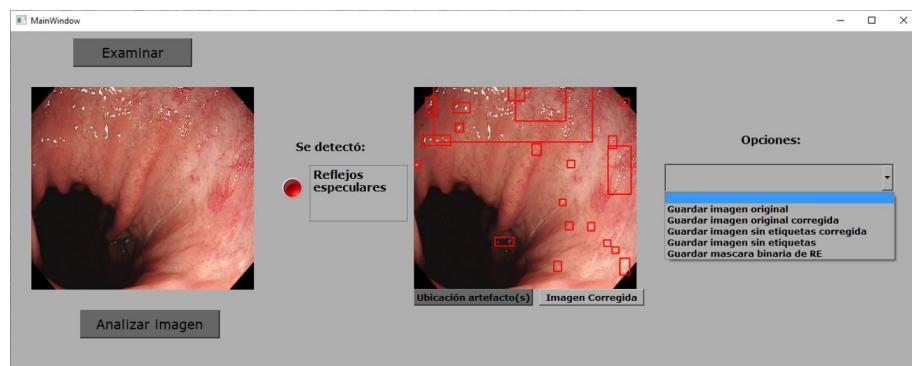


Figura 5-49: Interfaz cuando se detecta RE

Como se puede observar en la figura 5-49, la interfaz muestra la imagen que se escogió (original), el artefacto que se detectó (RE), la imagen con la ubicación (bbox rojo: RE) del artefacto, y las opciones de: guardar la imagen original, la imagen corregida, la imagen sin etiquetas y la máscara binaria de RE. Cabe resaltar que, las dos últimas opciones se agregaron a la interfaz como un extra, dado que el usuario puede llegar a necesitar la imagen sin etiquetas (rotulado que contienen algunas imágenes endoscópicas con información del paciente o el tipo de endoscopia) o necesitar las máscaras por la ubicación de los reflejos especulares.

2. Si el método detectó **desenfoque por movimiento** en la imagen que se analizó, la interfaz gráfica dará las opciones que se observan en la figura 5-50.

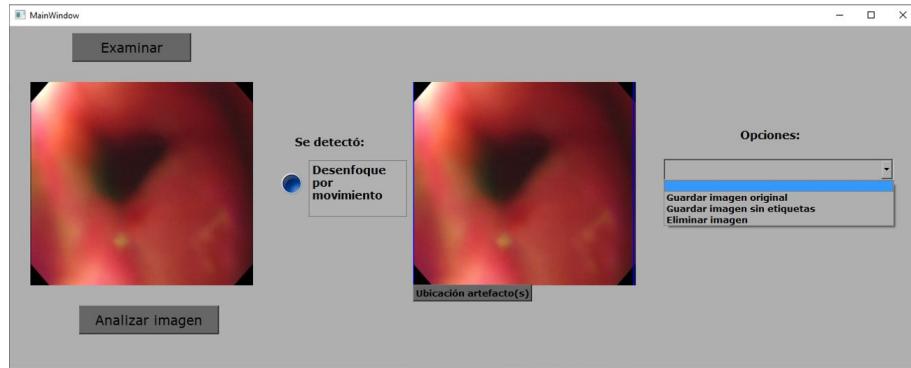


Figura 5-50: Interfaz cuando se detecta DM

Como se puede observar en la figura 5-50, al igual que con el resultado anterior la interfaz muestra la imagen que se escogió (original), el artefacto que se detectó (DM) y la imagen con la ubicación (bbox azul: DM) del artefacto; sin embargo, esta no cuenta con la opción de guardar la imagen corregida pero si cuenta con las siguientes opciones: guardar la imagen original, la imagen sin etiquetas y la opción de eliminar la imagen de la carpeta en la que se encuentre.

3. Si el método detectó **reflejos especulares** y **desenfoque por movimiento** en la imagen que se analizó, la interfaz gráfica dará las opciones que se observan en la figura 5-51.

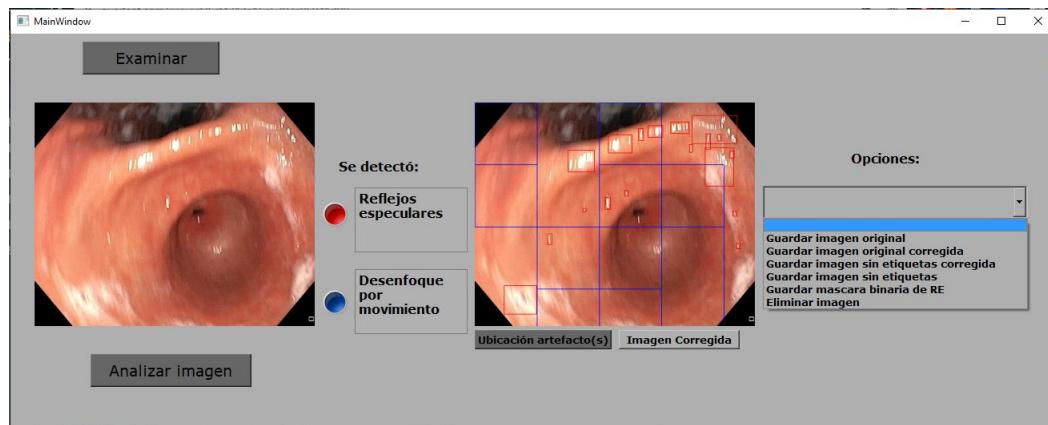


Figura 5-51: Interfaz cuando se detecta RE y DM

Como se puede observar en la figura 5-51, la interfaz muestra la combinación de las opciones presentadas en el ítem 1 y 2.

4. Si el método no detectó **ningún artefacto**, es decir, no detectó RE ni DM en la imagen que se analizó, la interfaz gráfica dará las opciones que se observan en la figura 5-52.

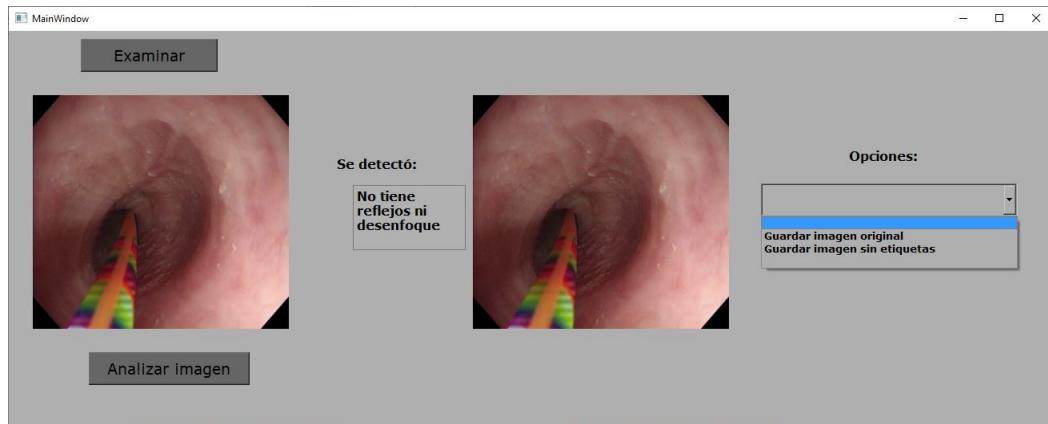


Figura 5-52: Interfaz cuando no se detecta RE ni DM

Como se puede observar en la figura 5-52, la interfaz muestra menos opciones a comparación de las que muestra si el resultado es alguno de los anteriores ítems, si no se detecto ningún artefacto el usuario solo contará con dos opciones: Guardar la imagen original y la imagen sin etiquetas.

Finalmente cabe resaltar que los códigos del método completo se encuentran en la **carpeta: anexos/Codigos** y los diagramas de flujo en el capítulo 8 (Anexos).

6 Resultados finales y discusión

Este capítulo se presenta como el resultado final del cumplimiento de los objetivos del trabajo de grado, asumiendo que en la metodología se cubrieron variedad de estrategias para cada una de las etapas metodológicas propuestas y se seleccionaron las mejores para conformar el método completo de tratamiento de artefactos para reflejos especulares y desenfoque por movimiento en imágenes endoscópicas.

1. Evaluar la capacidad de separabilidad de las características encontradas para identificación de artefactos:

Las características extraídas tanto para DM como para RE fueron evaluadas de manera cualitativa y cuantitativa, sin embargo, la primer evaluación no arrojó resultados que permitieran determinar con certeza las características de mayor separabilidad entre una clase y otra, lo que sí permitió evidenciar fue la necesidad de implementar dos clasificadores binarios, pues visualmente la mayor separabilidad se encontraba entre una de las clases vs el resto. Por otro lado, mediante la segunda evaluación (cuantitativa) se identificó claramente qué características separaban una clase del resto; en el caso de DM fueron: **SODMS** y **Entropía** y para RE fueron **brillo, contraste y desviación de intensidades**.

Igualmente se hicieron dos evaluaciones cuantitativas (t-student y prueba de rango Wilcoxon), para ratificar que las características que se mencionaron anteriormente si fueran separables. Efectivamente dichas características rechazaron las hipótesis planteadas por las evaluaciones, es decir, las medias y las medianas de las distribuciones no son iguales, por lo tanto las características son separables para una clase vs el resto (otros), los resultados de la evaluación de la separabilidad para las características seleccionadas se puede observar en la tabla 6-1.

Clases	DM						RE					
	SODMS			Entropía			Brillo			Contraste		
Clases	0	1	2	0	1	2	0	1	2	0	1	2
Percentil (p25-p35)	0.75-0.87	0.64-0.78	0.60-0.77	0.39-0.78	0.55-1	0.82-1.18	0.0043-0.0046	0.003-0.011	0.0043-0.0046	79-242	482-1295	87-294
Clases comparadas	0 vs 1	0 vs 2	0 vs 1	0 vs 2	1 vs 0	1 vs 2	1 vs 0	1 vs 2	1 vs 0	1 vs 2	1 vs 0	1 vs 2
T-student (valor de p)	6.57×10^{-23}	7.56×10^{-28}	1.25×10^{-19}	3.58×10^{-50}	2.673×10^{-293}	2.4006×10^{-305}	0	0	0	0	9.8757×10^{-232}	
Prueba Wilcoxon (valor de p)	7.05×10^{-20}	5.11×10^{-22}	2.40×10^{-12}	1.69×10^{-33}	2.7112×10^{-272}	3.4115×10^{-282}	0	0	0	0	1.6146×10^{-216}	

Tabla 6-1: Resultados de la evaluación de separabilidad de las características seleccionadas

2. Implementar un método de ubicación e identificación del tipo de artefacto:

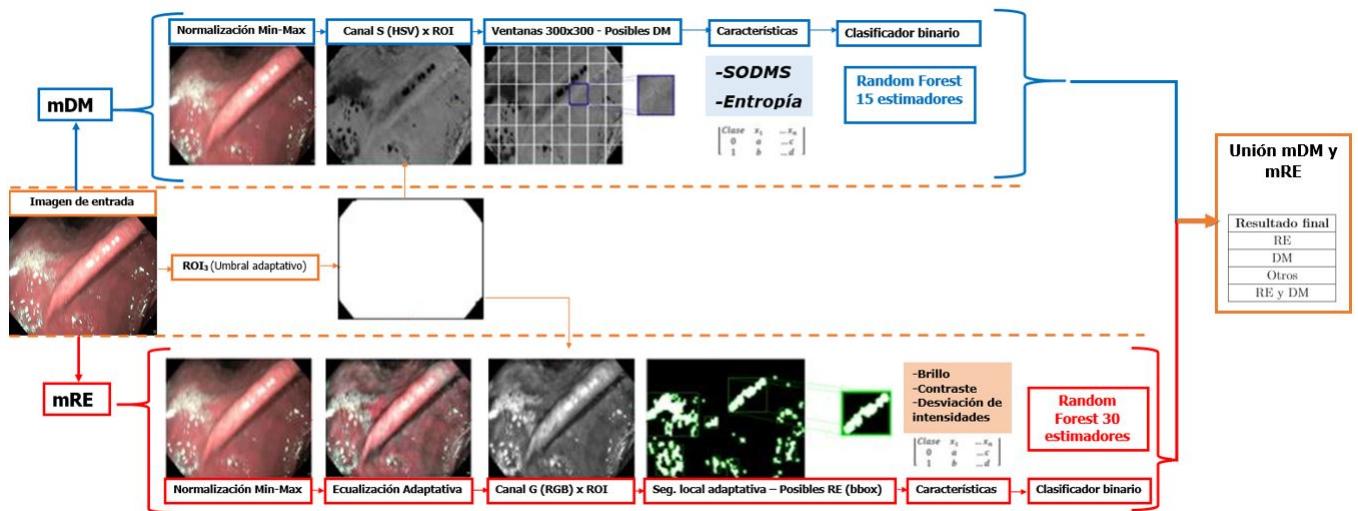


Figura 6-1: Método de ubicación e identificación del tipo de artefacto (DM y/o RE)

En esta sección se muestran los mejores resultados de cada etapa metodológica mostrada en la sección 5, que finalmente permitieron identificar y ubicar artefactos en las imágenes endoscópicas. Como se evidencia en la figura 6-1, el método propuesto consiste en dos metodologías una para identificación de RE (mRE) y otra para DM (mDM), las cuales al final se unen para conformar todo el método de ubicación e identificación de artefactos. Estas inician con la aplicación de una normalización Min-Max que según se evidencia en la subsección 5.2.1, mostró un mejor desempeño tanto en la disminución de la variabilidad de los rangos de intensidad como en las métricas de similitud entre histogramas propuestas (correlación, distancia Bhattacharyya y euclíadiana), es decir, fue la técnica con la que mejor se resaltó las regiones con artefactos del fondo. Posteriormente, para la mRE se procede a aplicar sobre la imagen normalizada una ecualización adaptativa, dado que aumentó el contraste de las imágenes y mejoró el desempeño respecto a las métricas de similitud (ver subsección 5.2.2), esto no pasó con DM por ende esta metodología no la contiene.

Para finalizar la etapa de preprocessamiento, se cambia el espacio de color al canal G (RGB) y S (HSV) para RE y DM, respectivamente (ver subsección 5.2.4). Seguido a esto, para las dos metodologías se multiplica el canal del espacio de color elegido por la ROI obtenida por el método de umbralización adaptativa, que según se evidencia en la sección 5.3 obtuvo el mejor desempeño en las métricas de similitud propuestas para comparar con los ground truth (DICE, SSIM y MSE). Posteriormente, para la mRE se

procede a aplicar sobre la región de interés encontrada en el canal G la umbralización local adaptativa, puesto que esta presenta un buen desempeño en las métricas de similitud, es decir, que esta técnica logró segmentar la mayoría de RE (verdaderos) según se evidencia en la sección 5.4. Por su parte, para la ubicación de DM no se realiza segmentación sino que se emplea una metodología basada en ventaneo (ventanas de 300x300 píxeles), es decir, la imagen se divide en ventanas.

Por último, para la mRE se extraen las características seleccionadas (**brillo, contraste y desviación de intensidades**) de los posibles candidatos a RE, obtenidos en forma de bbox (ver subsección 5.4.2), luego, se procede a emplear un *Random forest* con 15 estimadores dado que según la sección 5.6 obtuvo el mejor resultado en las métricas de desempeño y son comparables con las encontradas en la literatura; con este se identifica si el bbox de entrada al clasificador pertenece a la clase 0 (Otros) o a la clase 1 (RE). Por otro lado, para la mDM se extraen las características (**SODMS y Entropía**) de las ventanas, las cuales entran al clasificador *Random forest* con 15 estimadores, dado que obtuvo el mejor resultado en las métricas de desempeño y es comparable con los resultados encontrados en el estudio de [Oksuz et al., 2019] que usó de la misma BD; con este clasificador se identifica si la ventana de entrada pertenece a la clase 0 (DM) o a la clase 1 (Otros). Finalmente, el método identificará entre las tres clases como se observa en la tabla **6-2**.

Resultados clasificador mRE	Resultados clasificación mDM	Resultado final
1	1	RE
0	0	DM
0	1	Otros
1	0	RE y DM

Tabla 6-2: Resultados unión de clasificadores binarios

3. Desarrollar un algoritmo para la corrección o eliminación de imágenes

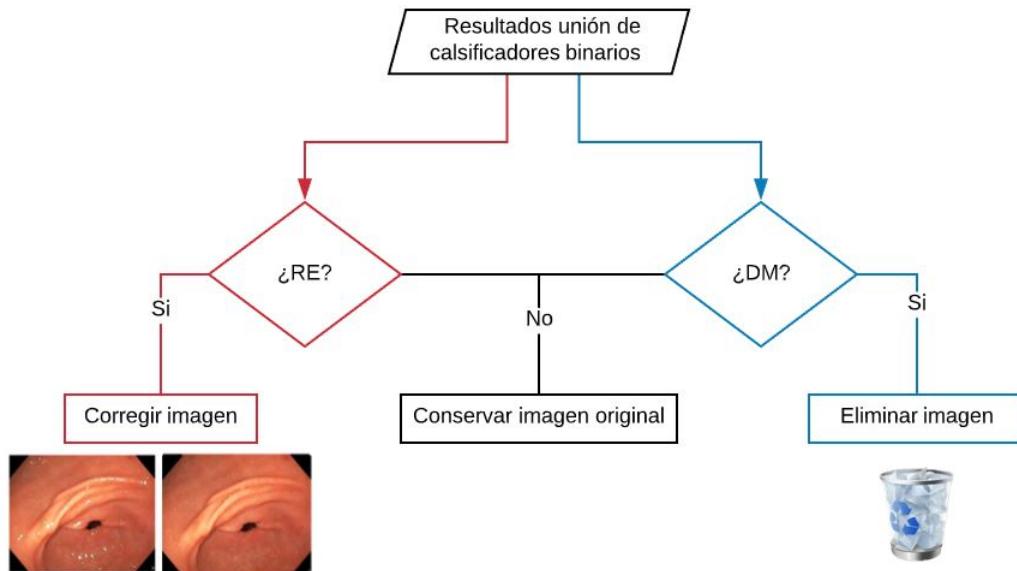


Figura 6-2: Algoritmo para la corrección o eliminación de imágenes endoscópicas

Antes que nada es necesario tener presente que, el algoritmo de corrección es únicamente para las imágenes endoscópicas que contengan RE y el de eliminación para las que contengan DM como se ve en la figura 6-2. Para el de corrección se evaluaron diferentes métodos como se explicó en la sección 5.7, donde el método de inpaiting 3 con un radio de 10 (basado en la marcha rápida de A. Telea) fue el que obtuvo el mejor desempeño (SSIM: 0.99, MSE: 0.49, PSNR: 51.96); por su parte el de eliminación es un algoritmo permite eliminar la imagen desenfocada (completa) de la carpeta en la que se encuentre.

4. Evaluar el rendimiento del método propuesto de identificación y corrección de artefactos

Durante el desarrollo de este trabajo de grado se evaluó cada fase metodológica, tal como se observa en la tabla 6-3, donde el método de corrección para reflejos especulares se evaluó con las siguientes métricas de similitud, SSIM: 0.99, MSE: 0.49 y PSNR: 51.96 (ver proceso de evaluación en la sección 5.7). Sin embargo, la evaluación del rendimiento del método completo de identificación para cada artefacto se hizo con 371 imágenes de prueba, es decir, aquellas que no estuvieron involucradas en el desarrollo del método. Dicha evaluación arrojó una precisión de 0.82, por otro lado la evaluación

de la ubicación, se hizo comparando las anotaciones dadas por la BD en imágenes de prueba con las del resultado del método propuesto lo que dio para RE un valor de MSE de 2842 ± 7377 y para DM de 6191 ± 8188 .

Cabe resaltar que todos los resultados de la evaluación de la ubicación de artefactos se encuentran en la **carpeta: anexos/tablas/evaluacionUbicacionArtefactos.xlsx**

Fase metodológica	Evaluación	
	DM	RE
Normalización min-max	<ul style="list-style-type: none"> -Rangos de intensidad: $0-255 \pm 0$ -Correlación: 0.43 ± 0.29 -Distancia Battacharyya: 0.56 ± 0.17 - Distancia Euclidiana: 340106 ± 334514 	<ul style="list-style-type: none"> -Rangos de intensidad: $0-255 \pm 0$ -Correlación: 0.26 ± 0.35 -Distancia Battacharyya: 0.66 ± 0.19 - Distancia Euclidiana: 120507 ± 99050
Ecuallización Adaptativa	Sin contiene la fase de ecualización	<ul style="list-style-type: none"> -Contraste: 9.90 ± 0.86 -Correlación: 0.16 ± 0.31 -Distancia Battacharyya: 0.66 ± 0.19 - Distancia Euclidiana: 115939 ± 59441
Espacios de color	S (HSV)	G y B (RGB)
ROI umbral adaptativo	<ul style="list-style-type: none"> -DICE: 0.97 ± 0.03 -SSIM: 0.97 ± 0.05 -MSE: 752 ± 1566 	<ul style="list-style-type: none"> -DICE: 0.97 ± 0.03 -SSIM: 0.97 ± 0.05 -MSE: 752 ± 1566
Segmentación	-Ventanas: 300x300	<ul style="list-style-type: none"> -Local Adaptativa -DICE: 0.69 ± 0.19 -MSE: 20959 ± 8797
Método de identificación	<ul style="list-style-type: none"> -Exactitud: 0.74 -Precisión: 0.76 -Sensibilidad: 0.71 -Especificidad: 0.78 -AUC: 0.74 	<ul style="list-style-type: none"> -Exactitud: 0.95 -Precisión: 0.96 -Sensibilidad: 0.94 -Especificidad: 0.96 -AUC: 0.95
Método de corrección	No contiene la fase de corrección, se hace la eliminación de la imagen completa	<ul style="list-style-type: none"> -SSIM: 0.99 ± 0.002 -MSE: 0.49 ± 0.35 -PSNR: 51.9 ± 2.49

Tabla 6-3: Resultados por fase metodológica del método

Teniendo en cuenta lo anterior, el método propuesto en el presente trabajo de grado logró detectar de manera automática RE y DM en imágenes endoscópicas, por ende se espera que la corrección o eliminación de imágenes con dichos artefactos disminuyan su efecto negativo en el desarrollo de futuras herramientas DAO.

7 Conclusiones y recomendaciones

A lo largo del desarrollo del trabajo de grado y el análisis de resultados se dieron diferentes tipos de conclusiones que son las que serán descritas a continuación:

- Tras los resultados obtenidos en la etapa de normalización se demostró que es necesario aplicar una metodología diferente para identificar cada tipo de artefacto (RE y DM).
- De acuerdo a las evaluaciones cuantitativas y cualitativas de características, se obtuvo que las más discriminantes para DM fueron *SODMS* y *Entropía* y para RE fueron *brillo*, *contraste* y *desviación de intensidades*.
- Fue necesario implementar dos clasificadores binarios complementarios para que identificaran RE vs el resto y DM vs el resto, cuya precisión fue de 96 % y 76 %, respectivamente.
- Dado lo anterior, el objetivo de la tesis se cumplió al realizar un método automático para la detección de dos artefactos (DM y RE) en imágenes endoscópicas y su corrección o eliminación a través de un algoritmo de análisis de imágenes.
- El método propuesto fue diseñado como herramienta de apoyo para el desarrollo de algoritmos DAO y aunque cumple con el objetivo propuesto, este aun debe pasar por pruebas de validación de software para poder llevarlo a escenarios de producción.
- Para que las máscaras binarias de posibles RE no contengan otros artefactos similares como saturación o burbujas, se recomienda tener una base lo suficientemente grande de máscaras de dichos artefactos para que así las características de forma y tamaño sean tenidas en cuenta en el proceso de identificación de reflejos.
- Es importante no sólo identificar dos tipos de artefactos sino el resto que se presenta frecuentemente en las imágenes de endoscopia y afecta de igual manera las herramientas de apoyo al diagnóstico.
- Para que el método desarrollado sea más completo, se recomienda que este pueda recibir más de una imagen, es decir, un dataset completo de frames o un vídeo de endoscopia.

8 Anexos

8.1. Diagrama de flujo de la normalización

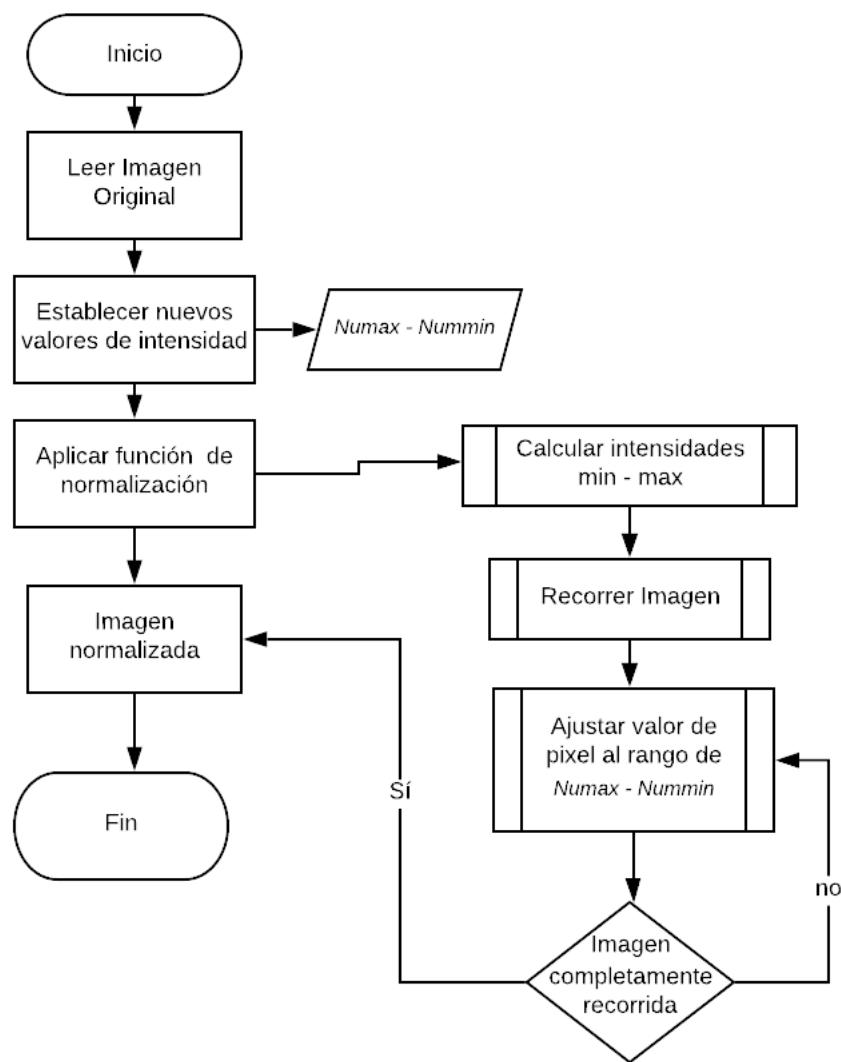


Figura 8-1: Diagrama de flujo de la normalización

8.2. Diagrama de flujo de la ecualización

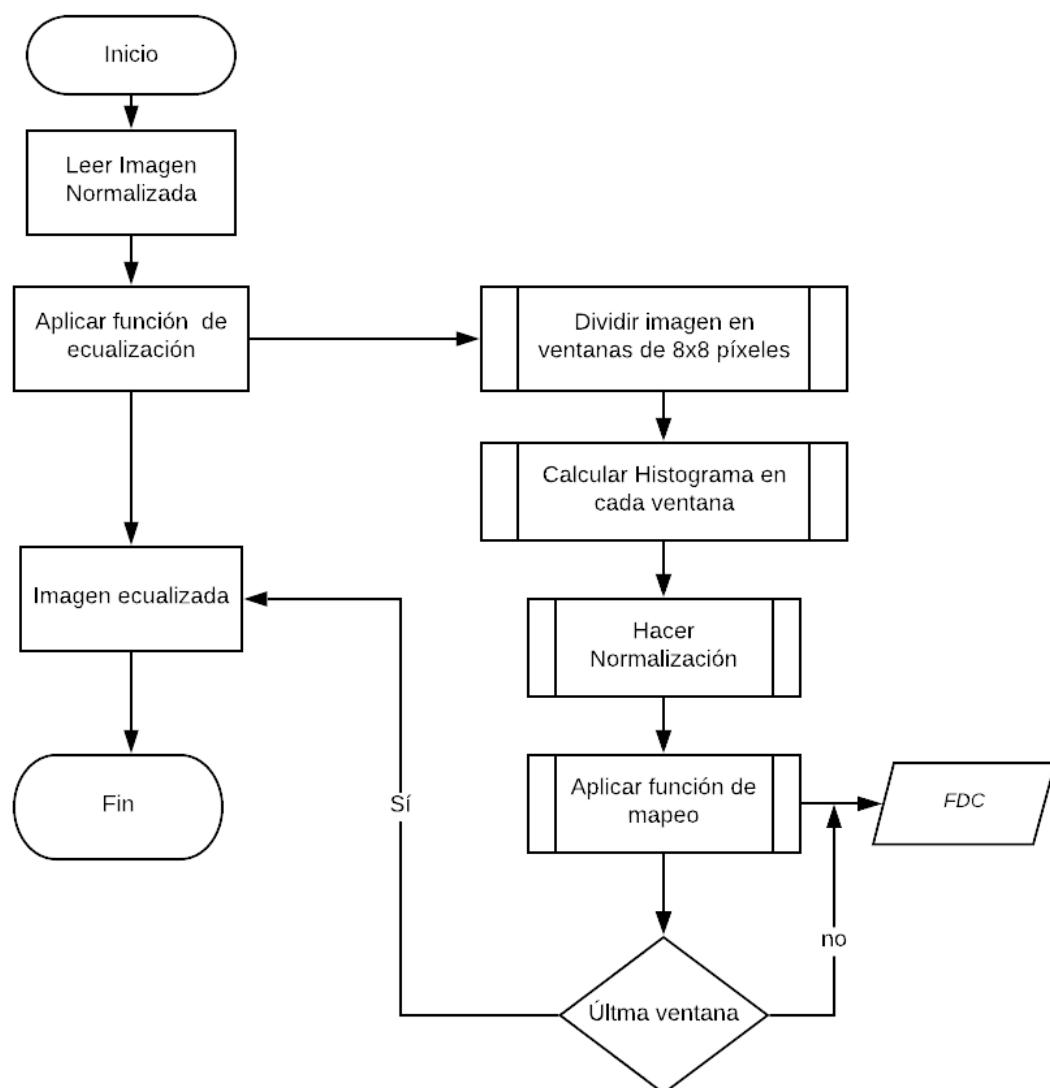


Figura 8-2: Diagrama de flujo de ecualización

8.3. Diagrama de flujo de la evaluación de las fases del preprocesamiento

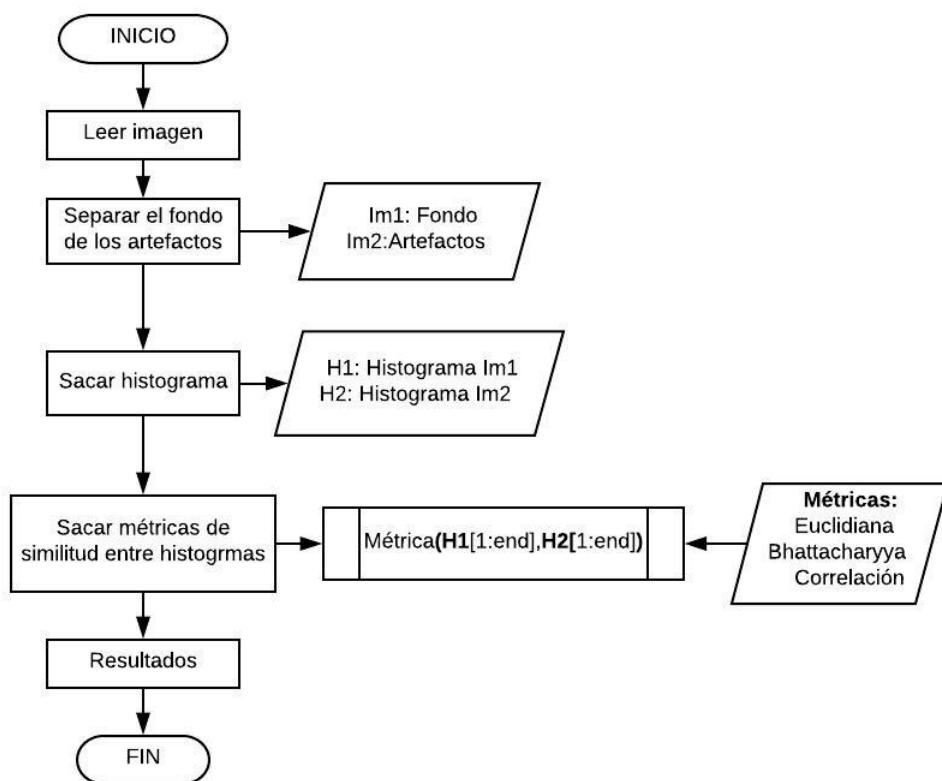


Figura 8-3: Diagrama de flujo de la evaluación del preprocesamiento

8.4. Diagrama de flujo sobre la región de interés

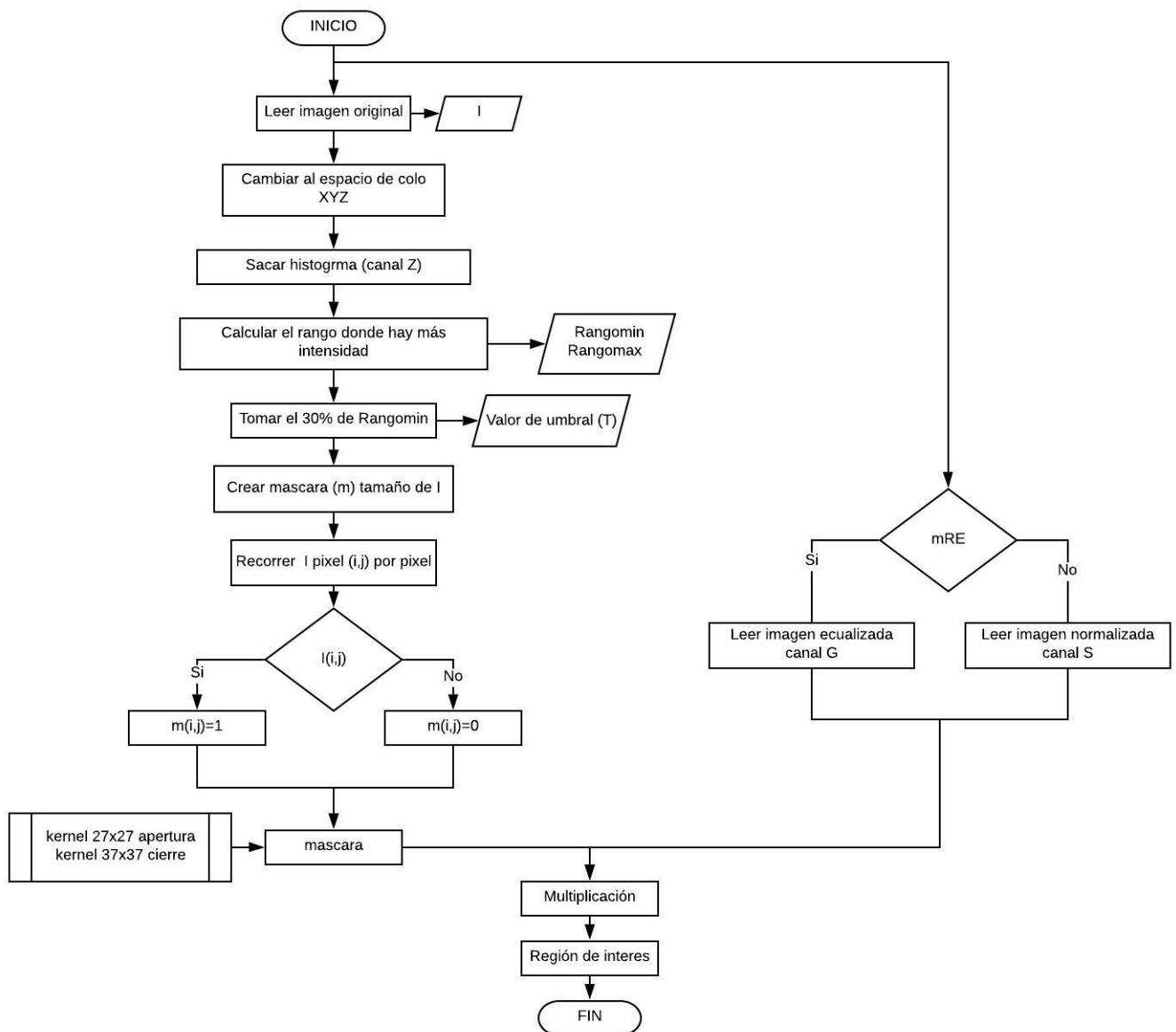


Figura 8-4: Diagrama de flujo de la región de interés (ROI)

8.5. Diagrama de flujo sobre la segmentación y extracción de posibles artefactos

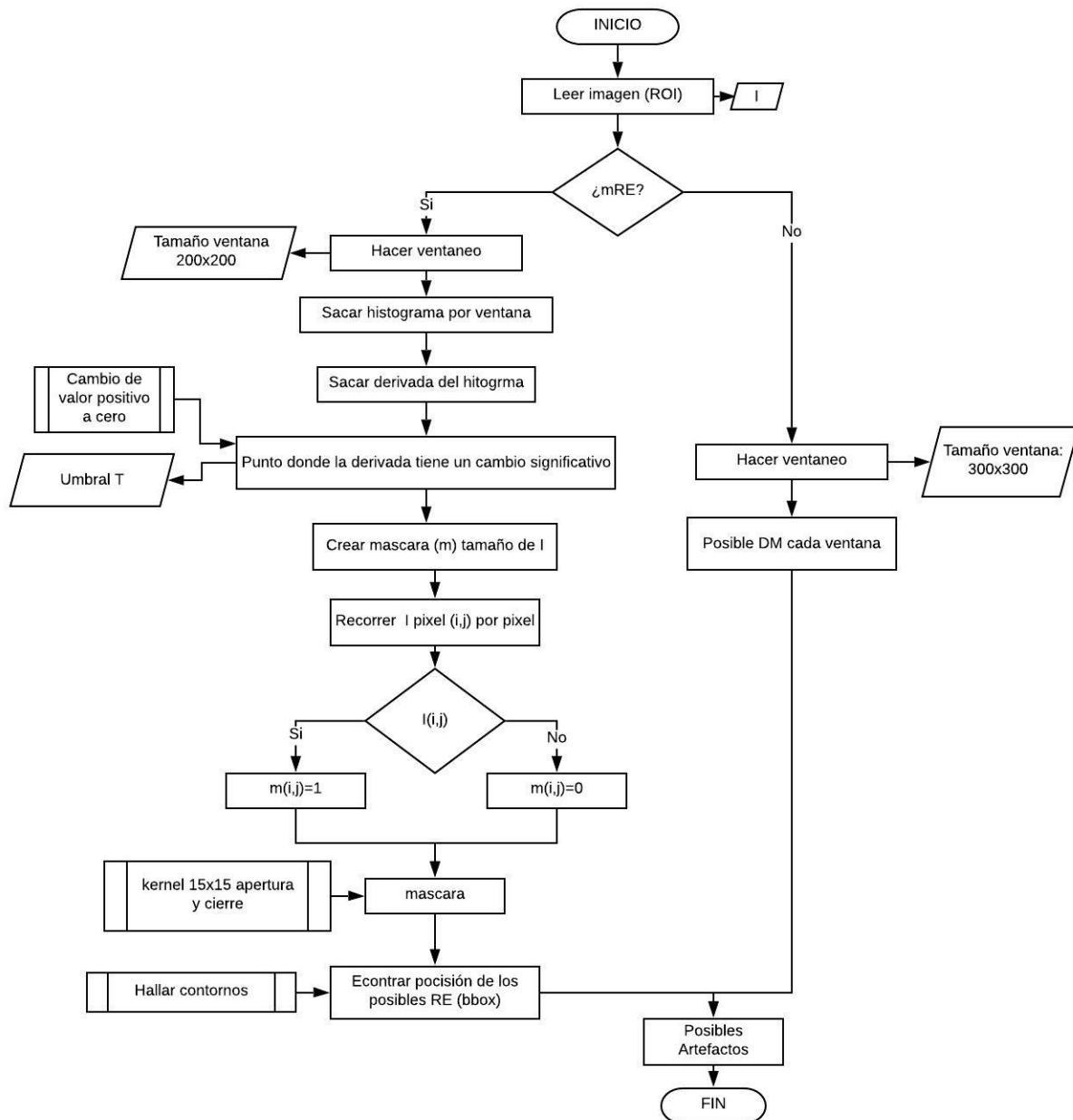


Figura 8-5: Diagrama de flujo de la segmentación y extracción de candidatos

8.6. Diagrama de flujo de la extracción de características

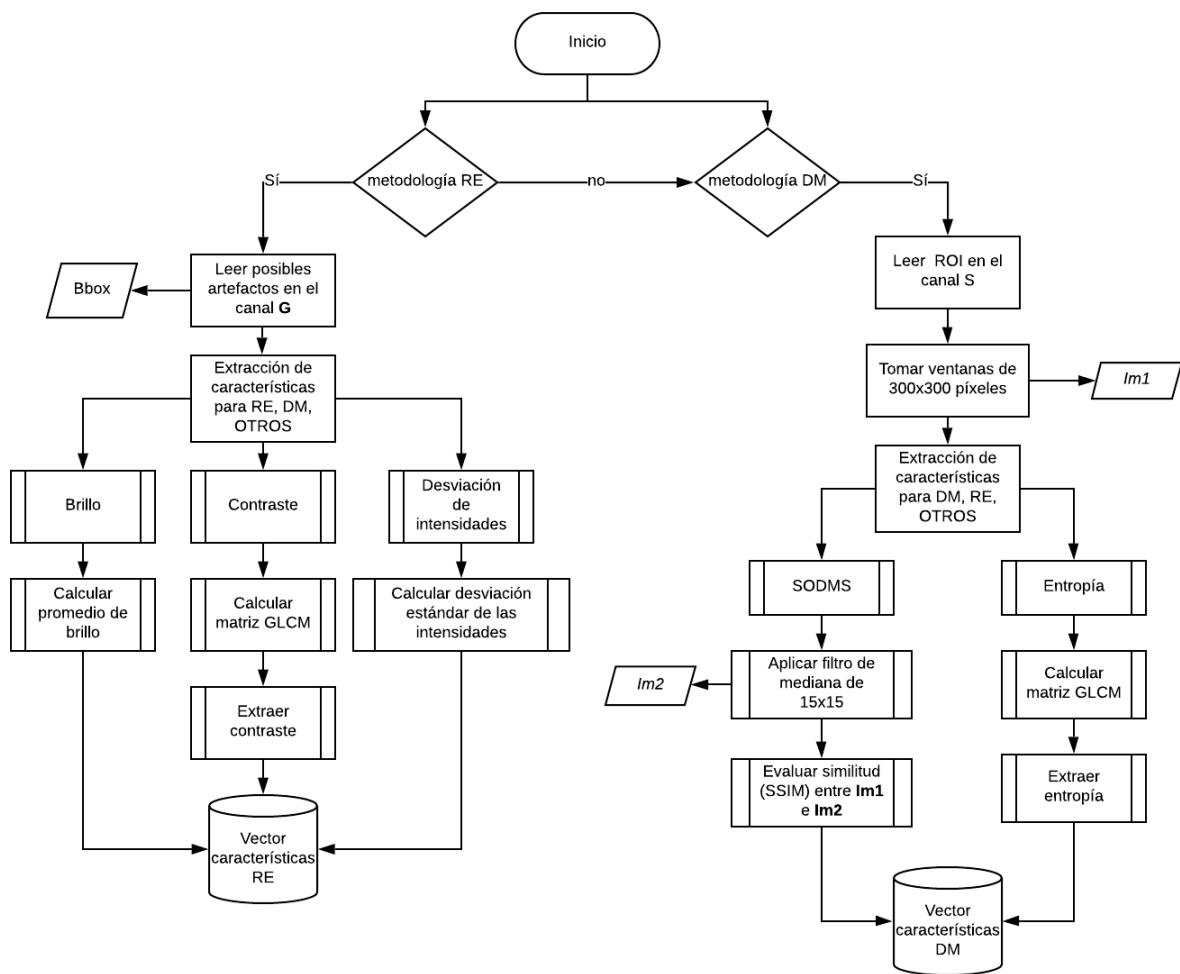


Figura 8-6: Diagrama de flujo de la extracción de características

8.7. Diagrama de flujo de la evaluación de separabilidad de características

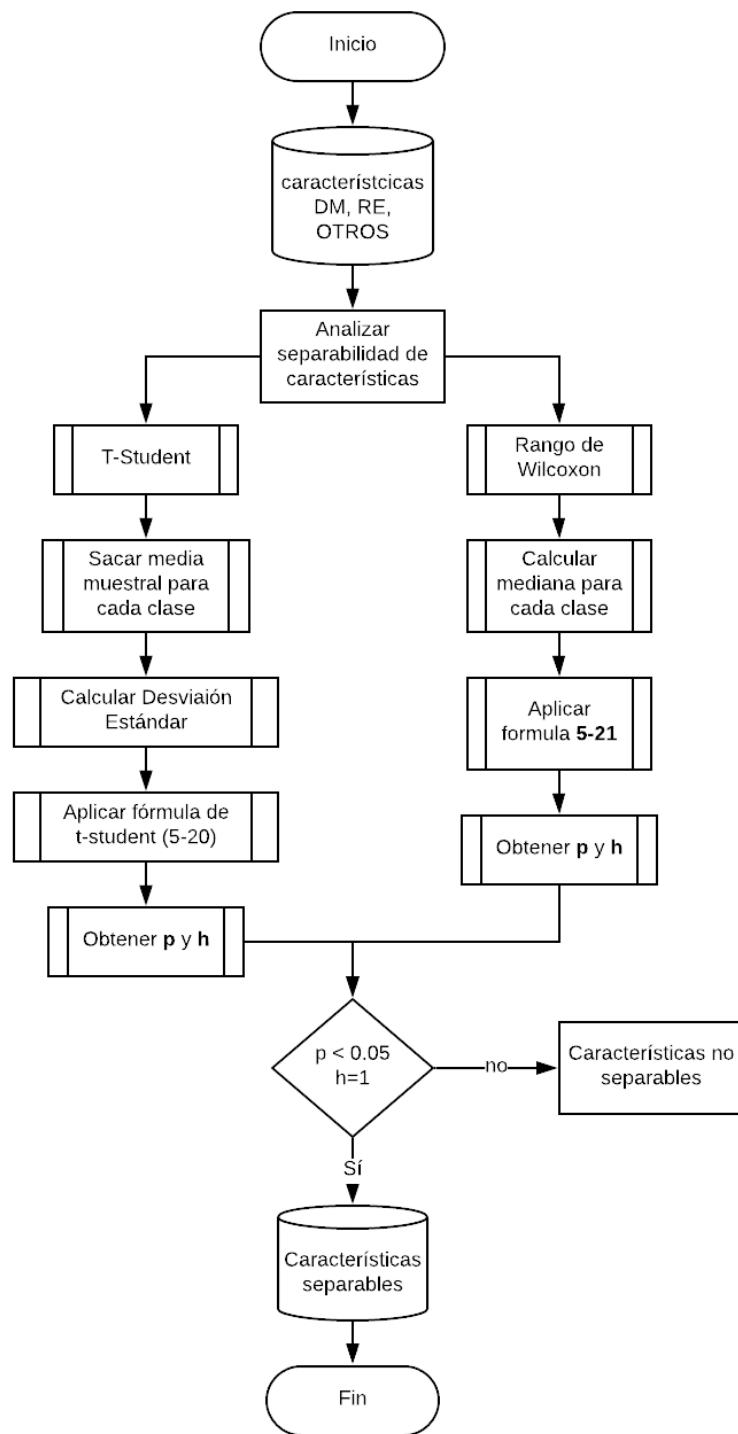


Figura 8-7: Diagrama de flujo de la evaluación de separabilidad de características

8.8. Diagrama de flujo del método de identificación

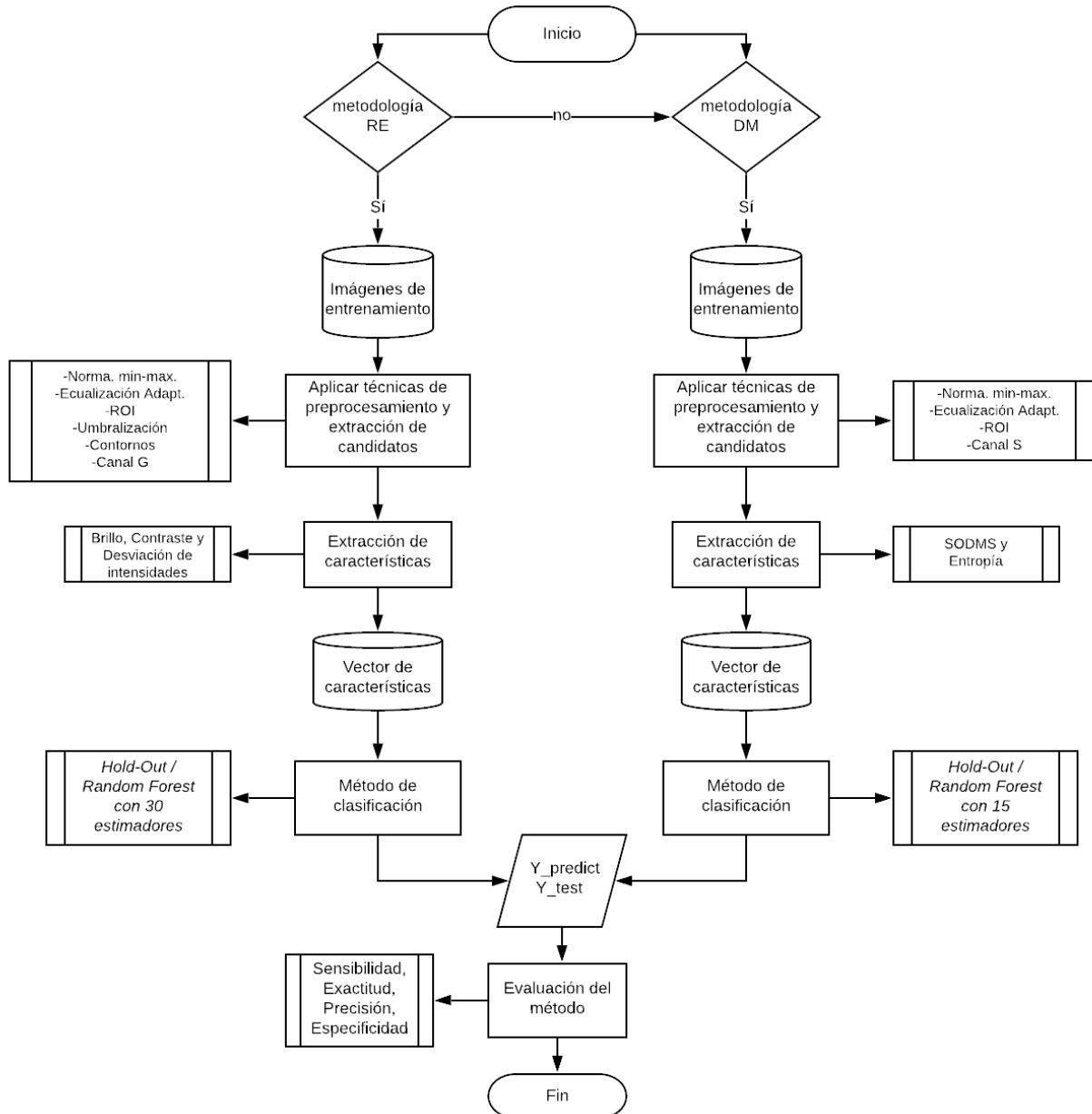


Figura 8-8: Diagrama de flujo método de identificación

8.9. Diagrama de flujo del método de corrección

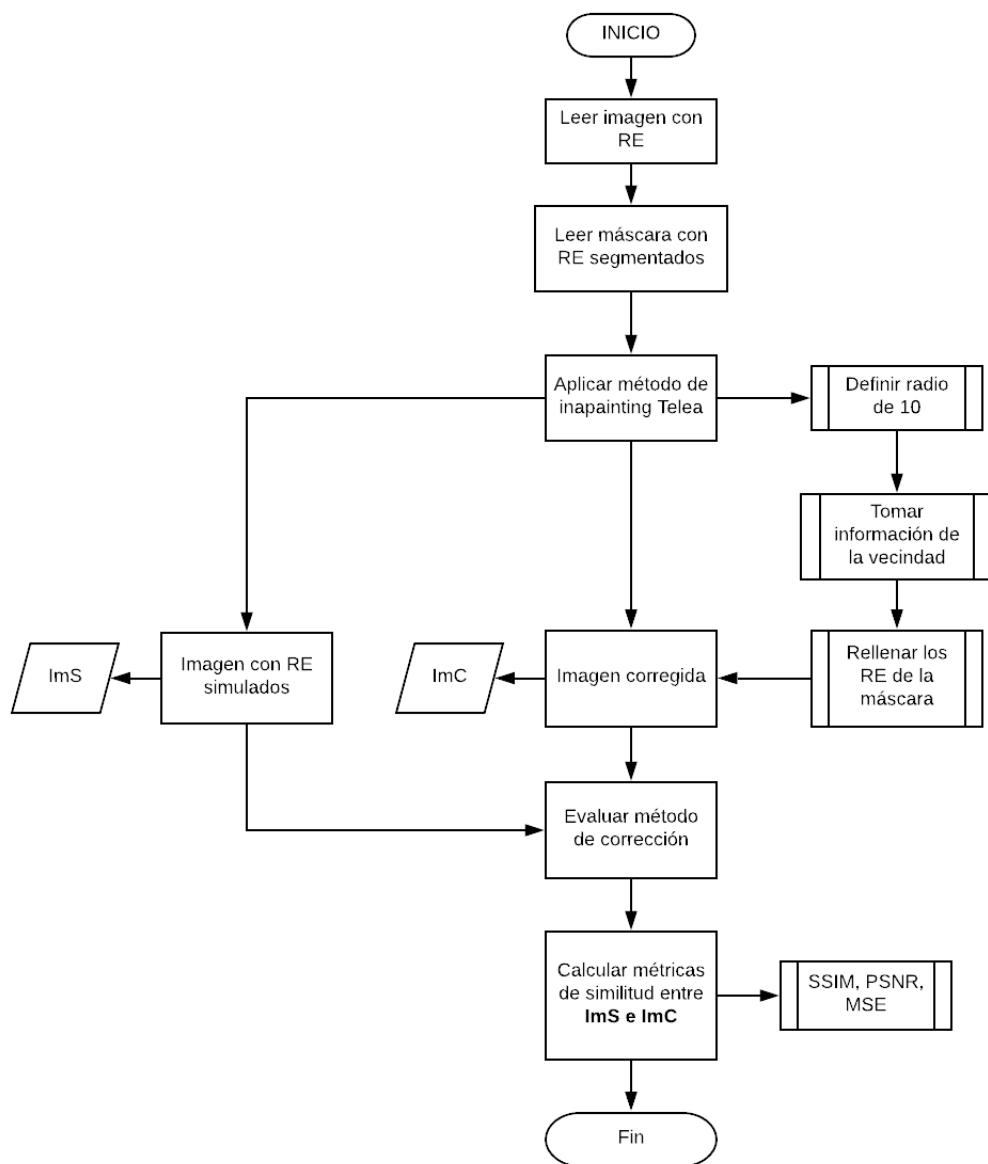


Figura 8-9: Diagrama de flujo del método de corrección

8.10. Diagrama de flujo de la interfaz

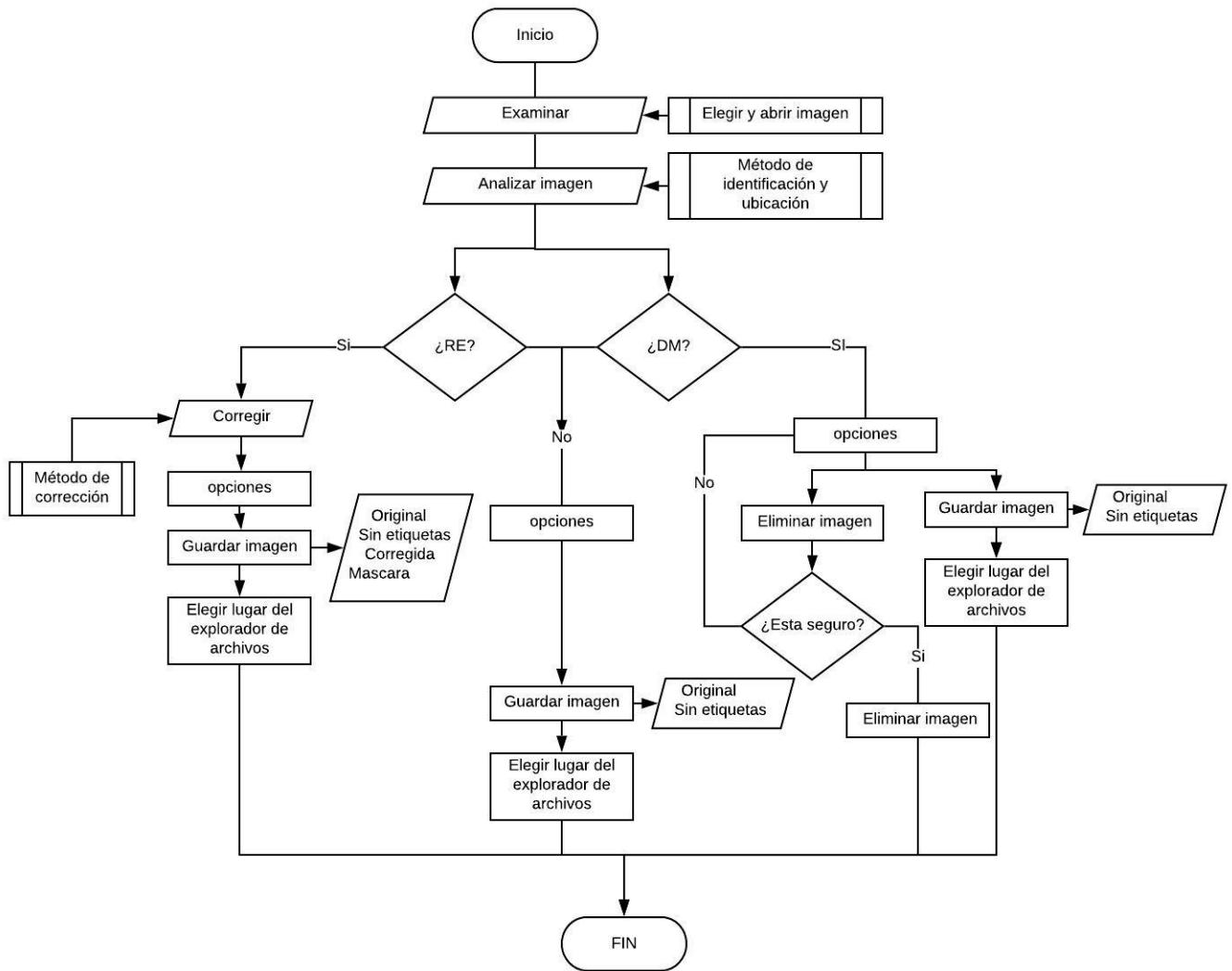


Figura 8-10: Diagrama de flujo de la GUI

Bibliografía

[Abel et al., 2014] Abel, E. W., Zhuo, Y., Ross, P. D., and White, P. S. (2014). Automatic glare removal in endoscopic imaging. *Surgical endoscopy*, 28(2):584–591.

[Ahsan et al., 2011] Ahsan, M., Babaei, L., Gholamrezaei, A., and Emami, M. H. (2011). Simethicone for the preparation before esophagogastroduodenoscopy. *Diagnostic and therapeutic endoscopy*, 2011.

[Akbari et al., 2018] Akbari, M., Mohrekesh, M., Najariani, K., Karimi, N., Samavi, S., and Soroushmehr, S. R. (2018). Adaptive specular reflection detection and inpainting in colonoscopy video frames. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 3134–3138. IEEE.

[Ali et al., 2019] Ali, Sharib and, F., Bailey, A., Braden, B., East, J., Lu, X., and Rittscher, J. (2019). A deep learning framework for quality assessment and restoration in video endoscopy. *arXiv preprint arXiv:1904.07073*.

[Ali and Zhou, 2019] Ali, S. and Zhou, F. (2019). Ead2019: Multi-class artefact detection in video endoscopy.

[American Cancer Society, 2015] American Cancer Society (2015). ¿cuándo se usa la endoscopia?

[Armin et al., 2016] Armin, M. A., Chetty, G., Jurgen, F., De Visser, H., Dumas, C., Fazlollahi, A., Grimpens, F., and Salvado, O. (2016). Uninformative frame detection in colonoscopy through motion, edge and color features. pages 153–162.

[Arnold et al., 2011] Arnold, M., Ameling, S., Ghosh, A., and Lacey, G. (2011). Quality improvement of endoscopy videos. In *Proceedings of the 8th IASTED International Conference on Biomedical Engineering, Innsbruck, Austria. ACTA Press*, page 72.

[Arnold et al., 2009] Arnold, M., Ghosh, A., Lacey, G., Patchett, S., and Mulcahy, H. (2009). Indistinct frame detection in colonoscopy videos. In *2009 13th International Machine Vision and Image Processing Conference*, pages 47–52. IEEE.

[Atasoy et al., 2010] Atasoy, S., Mateus, D., Lallemand, J., Meining, A., Yang, G.-Z., and Navab, N. (2010). Endoscopic video manifolds. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 437–445. Springer.

- [Au and Takei, 2001] Au, W. and Takei, R. (2001). Image inpainting with the navier-stokes equations. *Available at Final Report APMA*, 930.
- [Betancourt, 2005] Betancourt, G. A. (2005). Las máquinas de soporte vectorial (svms). *Scientia et technica*, 1(27).
- [Bigot and Alvarado, 2009] Bigot, M. A. S. and Alvarado, A. A. (2009). Diagnóstico de acalasia. *Revista Médica de Costa Rica y Centroamérica*, 66(588):177–179.
- [Briñez de León et al., 2012] Briñez de León, J. C., Restrepo Martínez, A., López Giraldo, F. E., et al. (2012). Métricas de similitud aplicadas para análisis de imágenes de fotoelasticidad. *Dyna; Vol. 80, núm. 179 (2013); 42-50 DYNA; Vol. 80, núm. 179 (2013); 42-50 2346-2183 0012-7353*.
- [Cheng et al., 2001] Cheng, H.-D., Jiang, X. H., Sun, Y., and Wang, J. (2001). Color image segmentation: advances and prospects. *Pattern recognition*, 34(12):2259–2281.
- [Damelin and Hoang, 2018] Damelin, S. and Hoang, N. (2018). On surface completion and image inpainting by biharmonic functions: Numerical aspects. *International Journal of Mathematics and Mathematical Sciences*, 2018.
- [de Salud Pública, 2016] de Salud Pública, M. (2016). Protocolos de atención para gastroenterología. Santo Domingo.
- [de vídeo endoscopia Gastrointestinal de El Salvador,] de vídeo endoscopia Gastrointestinal de El Salvador, A. Gastritis aguda.
- [Dey, 2018] Dey, S. (2018). *Hands-on Image Processing in Python*.
- [Ellis, 2000] Ellis, D. P. (2000). Improved recognition by combining different features and different systems.
- [Gallego Ojea, 2019] Gallego Ojea, J. (2019). Resonancia magnética (rm).
- [Gao et al., 2017] Gao, Y., Yang, J., Ma, S., Ai, D., Lin, T., Tang, S., and Wang, Y. (2017). Dynamic searching and classification for highlight removal on endoscopic image. *Procedia Computer Science*, 107:762–767.
- [Gaudó Perales and Revnic,] Gaudó Perales, C., a. e. C. F. and Revnic, A. N. Acalasia: A propósito de un caso clínico.
- [Gonzalez and Woods, 2014] Gonzalez, R. C. and Woods, R. E. (2014). *Digital image processing*. Dorling Kindersley.
- [Gross, 2014] Gross, S. (2014). *Automated Classification of Colon Polyps in Narrow Band Imaging Colonoscopy*. Sierke Verlag.

- [Grávalos and González Flores, 2017] Grávalos, C. and González Flores, E. (2017). Cáncer gástrico.
- [Gupta and chandel, 2013] Gupta, G. and chandel, r. (2013). Image filtering algorithms and techniques: A review.
- [Hansen and Carlos, 2015] Hansen, J. T. and Carlos, A. (2015). *Netter, cuaderno de anatomía para colorear*. Elsevier España SLU.
- [Haralick et al., 1973] Haralick, R. M., Shanmugam, K., et al. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621.
- [HENAO et al., 2005] HENAO, R., HURTADO, J. E., and CASTELLANOS, G. (2005). Selección de hiperparámetros en máquinas de soporte vectorial utilizando adaptación de matriz de covarianza. *Scientia et technica*, 1(27).
- [Hore and Ziou, 2010] Hore, A. and Ziou, D. (2010). Image quality metrics: Psnr vs. ssim. In *2010 20th International Conference on Pattern Recognition*, pages 2366–2369.
- [Hsu et al., 2003] Hsu, C.-W., Chang, C.-C., Lin, C.-J., et al. (2003). A practical guide to support vector classification.
- [Info, 2018] Info, R. (2018). Tac-abdomen y pelvis.
- [Iris et al.,] Iris, Palak, and Zander. Machine learning.
- [Jain et al., 2011] Jain, A., Kanjalkar, P., and Kulkarni, J. (2011). Estimation of image focus measure and restoration by wavelet. In *2011 4th International Conference on Intelligent Networks and Intelligent Systems*, pages 73–76. IEEE.
- [Jarrett et al., 2009] Jarrett, K., Kavukcuoglu, K., Ranzato, M., and LeCun, Y. (2009). What is the best multi-stage architecture for object recognition? In *2009 IEEE 12th International Conference on Computer Vision*, pages 2146–2153.
- [JungHwan Oh, 2004] JungHwan Oh, Sae Hwang, W. T. P. C. d. G. J. W. (2004). Blurry-frame detection and shot segmentation in colonoscopy videos. In *Storage and Retrieval Methods and Applications for Multimedia 2004*, volume 5307, pages 531–543. International Society for Optics and Photonics.
- [Kabbai et al., 2016] Kabbai, L., Sghaier, A., Douik, A., and Machhout, M. (2016). Fpga implementation of filtered image using 2d gaussian filter. *International Journal of Advanced Computer Science and Applications*, 7(7):514–520.
- [Kaçmaz et al., 2018] Kaçmaz, R. N., Yılmaz, B., Dündar, M. S., and Doğan, S. (2018). Motion artifact detection in colonoscopy images. *The EuroBiotech Journal*, 2(3):171–175.

- [Karapetyan and Sarukhanyan, 2013] Karapetyan, G. and Sarukhanyan, H. (2013). Automatic detection and concealment of specular reflections for endoscopic images. In *Ninth International Conference on Computer Science and Information Technologies Revised Selected Papers*, pages 1–8. IEEE.
- [Khanian et al., 2018] Khanian, M., Boroujerdi, A. S., and Breuß, M. (2018). Photometric stereo for strong specular highlights. *Computational Visual Media*, 4(1):83–102.
- [Kim et al., 2016] Kim, G. H., Cho, Y. K., Cha, J. M., Lee, S.-Y., and Chung, I.-K. (2016). Efforts to increase image quality during endoscopy: the role of pronase. *World journal of gastrointestinal endoscopy*, 8(5):267.
- [Kundu et al., 2018] Kundu, A. K., Fattah, S. A., and Rizve, M. N. (2018). An automatic bleeding frame and region detection scheme for wireless capsule endoscopy videos based on interplane intensity variation profile in normalized rgb color space. *Journal of healthcare engineering*, 2018.
- [Latarjet et al., 2019] Latarjet, M., Ruiz, A., and Pro, E. (2019). *Anatomía humana*. Editorial Médica Panamericana, Buenos Aires, 5 edition.
- [Lee et al., 2017] Lee, J. Y., Gong, E. J., Chung, E. J., Park, H. W., Bae, S. E., Kim, E. H., Kim, J., Do, Y. S., Kim, T. H., Chang, H.-S., et al. (2017). The characteristics and prognosis of diffuse-type early gastric cancer diagnosed during health check-ups. *Gut and liver*, 11(6):807.
- [Li and Tam, 1998] Li, C. and Tam, P. K.-S. (1998). An iterative algorithm for minimum cross entropy thresholding. *Pattern recognition letters*, 19(8):771–776.
- [Li et al., 2018] Li, Z., Chen, Q., and Koltun, V. (2018). Interactive image segmentation with latent diversity. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 577–585.
- [Liedlgruber and Uhl, 2011] Liedlgruber, M. and Uhl, A. (2011). Computer-aided decision support systems for endoscopy in the gastrointestinal tract: a review. *IEEE reviews in biomedical engineering*, 4:73–88.
- [Lincoln University, 2014] Lincoln University (2014). Normal, binomial, poisson distributions. Technical report, Lincoln University, New Zealand.
- [Liu et al., 2014] Liu, Y., Lian, J., Bartolacci, M. R., and Zeng, Q.-A. (2014). Density-based penalty parameter optimization on c-svm. *The Scientific World Journal*, 2014.
- [Lizarzábal García, 2016] Lizarzábal García, M. (2016). Lo que necesitas saber de los pólipos.

- [Löfstedt et al., 2019] Löfstedt, T., Brynolfsson, P., Asklund, T., Nyholm, T., and Garpebring, A. (2019). Gray-level invariant haralick texture features. *PloS one*, 14(2):e0212110.
- [Luis Gerardo de la Fraga, 2011] Luis Gerardo de la Fraga (2011). Procesamiento digital de imágenes. Cinvestav.
- [Luna-Sánchez et al., 2011] Luna-Sánchez, S., Machuca, S. M., and Díaz, M. C. (2011). Esofagitis eosinofílica. *SEMERGEN-Medicina de Familia*, 37(6):303–306.
- [López et al., 2017] López, J. A., de Vries, E., Espinosa, M. T. R., Giana Henríquez, M., Valencia, Y. M., Pardo, C., Puerto, D. N., Rodríguez, C. H., Rojas, M. P., Samudio, M. L., Serrano, C., and Wiesner, C. (2017). Análisis de la situación del cáncer en colombia 2015. In Mendoza, G. M. H., editor, *Instituto Nacional de Cancerología ESE*, volume 1.
- [Macías-Rodríguez et al., 2014] Macías-Rodríguez, R. U., García-Flores, O. R., Ruiz-Margáin, A., and Barreto-Zúniga, R. (2014). Evaluación de los pólipos gástricos en endoscopia. *Endoscopia*, 26(2):62–69.
- [Marcinczak and Grigat, 2013] Marcinczak, J. M. and Grigat, R.-R. (2013). Closed contour specular reflection segmentation in laparoscopic images. *Journal of Biomedical Imaging*, 2013:18.
- [Marques, 2011] Marques, O. (2011). *Practical image and video processing using MATLAB*. J. Wiley Sons, 1 edition.
- [MathWorks,] MathWorks. Edge detection methods for finding object boundaries in images.
- [Medina-Merino and Ñique-Chacón, 2017] Medina-Merino, R. F. and Ñique-Chacón, C. I. (2017). Bosques aleatorios como extensión de los árboles de clasificación con los programas r y python. *Interfases*, 0(010):165–189.
- [Mery, 2004] Mery, D. (2004). Visión por computador. *Santiago de Chile. Universidad Católica de Chile*.
- [Meslouhi et al., 2011] Meslouhi, O., Kardouchi, M., Allali, H., Gadi, T., and Benkaddour, Y. (2011). Automatic detection and inpainting of specular reflections for colposcopic images. *Open Computer Science*, 1(3):341–354.
- [Mishra, 2018] Mishra, M. (2018). Machine learning classification.
- [Moccia et al., 2018] Moccia, S., Vanone, G. O., De Momi, E., Laborai, A., Guastini, L., Peretti, G., and Mattos, L. S. (2018). Learning-based classification of informative laryngoscopic frames. *Computer methods and programs in biomedicine*, 158:21–30.
- [Molina Cabello et al., 2018] Molina Cabello, M. Á. et al. (2018). Segmentación y detección de objetos en imágenes y vídeo mediante inteligencia computacional.

- [Morgand and Tamaazousti, 2014] Morgand, A. and Tamaazousti, M. (2014). Generic and real-time detection of specular reflections in images. In *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*, volume 1, pages 274–282. IEEE.
- [Münzer et al., 2018] Münzer, B., Schoeffmann, K., and Böszörmenyi, L. (2018). Content-based processing and analysis of endoscopic images and videos: A survey. *Multimedia Tools and Applications*, 77(1):1323–1362.
- [Murra Saca, 2018] Murra Saca, J. (2018). Gastritis.
- [Nwankpa et al., 2018] Nwankpa, C., Ijomah, W., Gachagan, A., and Marshall, S. (2018). Activation functions: Comparison of trends in practice and research for deep learning. *arXiv preprint arXiv:1811.03378*.
- [of Edinburgh, 2010] of Edinburgh, T. U. (2010). Spatial domain methods.
- [Oh et al., 2007] Oh, J., Hwang, S., Lee, J., Tavanapong, W., Wong, J., and de Groen, P. C. (2007). Informative frame classification for endoscopy video. *Medical Image Analysis*, 11(2):110–127.
- [Oksuz et al., 2019] Oksuz, I., Clough, J. R., King, A. P., and Schnabel, J. A. (2019). Artefact detection in video endoscopy using retinanet and focal loss function.
- [OpenCV, 2015] OpenCV (2015). Canny edge detection.
- [Ortiz et al., 2017] Ortiz, R. G. P., Segura, B. A. B., Acuña, M. L. S., Martín, D. M. A., and Wilches, R. O. (2017). Procedimientos endoscópicos y terapias antitrombóticas, una visión actual. *Revista Colombiana de Gastroenterología*, 32(2):150–159.
- [Patro and Sahu, 2015] Patro, S. and Sahu, K. K. (2015). Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462*.
- [Pladellorrens Mallofré, 2010] Pladellorrens Mallofré, J. (2010). Determinación de propiedades superficiales del papel utilizando el análisis de la textura del patrón de speckle. *Óptica pura y aplicada*, 43(1):43–48.
- [Queiroz and Ren, 2019] Queiroz, F. and Ren, T. I. (2019). Endoscopy image restoration: A study of the kernel estimation from specular highlights. *Digital Signal Processing*, 88:53–65.
- [R. Fisher and Wolfart., 2003] R. Fisher, S. Perkins, A. W. and Wolfart., E. (2003). Gaussian smoothing.
- [Ramos, 2016] Ramos, F. A. C. (2016). *Clasificación automática de Tweets utilizando K-NN y K-Means como algoritmos de clasificación automática, aplicando TF-IDF y TF-RFL para las ponderaciones*. PhD thesis, Pontificia Universidad Católica de Valparaíso.

- [Rebaza, 2007] Rebaza, J. V. (2007). Detección de bordes mediante el algoritmo de canny. *Escuela Académico Profesional di Informática. Universidad Nacional de Trujillo.*
- [Ríos-Díaz et al., 2009] Ríos-Díaz, J., Martínez-Payá, J. J., and del Baño Aledo, M. E. (2009). El análisis textural mediante las matrices de co-ocurrencia (gclm) sobre imagen ecográfica del tendón rotuliano es de utilidad para la detección cambios histológicos tras un entrenamiento con plataforma de vibración.(textural analysis by means of a grey level co-occurrence matrix method on patellar tendon ultrasonography is useful for the detection of histological changes after whole-body vibration training). *CCD. Cultura_Ciencia_Deporte.* -- doi: 10.12800/ccd, 4(11):91–102.
- [Rodríguez García et al., 2017] Rodríguez García, J., Peñaloza-Quintero, R. E., and Amaya-Lara, J. L. (2017). Estimación de la carga global de enfermedad en colombia 2012: nuevos aspectos metodológicos. *Revista de Salud Pública*, 19:235–240.
- [Ruiz, 2018] Ruiz, J. Z. (2018). Comparativa y análisis de algoritmos de aprendizaje automático para la predicción del tipo predominante de cubierta arbórea. *Maestría, Universidad Complutense NSE de Madrid.*
- [Saint-Pierre et al., 2011] Saint-Pierre, C.-A., Boisvert, J., Grimard, G., and Cheriet, F. (2011). Detection and correction of specular reflections for automatic surgical tool segmentation in thoracoscopic images. *Machine Vision and Applications*, 22(1):171–180.
- [Sartori et al., 2015] Sartori, P., Rozowykniat, M., Siviero, L., Barba, G., Peña, A., Mayol, N., Acosta, D., Castro, J., and Ortiz, A. (2015). Artefactos y artificios frecuentes en tomografía computada y resonancia magnética. *Revista argentina de radiología*, 79(4):192–204.
- [Shamir et al., 2019] Shamir, R. R., Duchin, Y., Kim, J., Sapiro, G., and Harel, N. (2019). Continuous dice coefficient: a method for evaluating probabilistic segmentations. *arXiv preprint arXiv:1906.11031*.
- [Shih and Rong-Chi Chang, 2005] Shih, T. K. and Rong-Chi Chang (2005). Digital inpainting - survey and multilayer image inpainting algorithms. In *Third International Conference on Information Technology and Applications (ICITA'05)*, volume 1, pages 15–24 vol.1.
- [Singh and Patel, 2017] Singh, B. and Patel, S. (2017). Efficient medical image enhancement using clahe enhancement and wavelet fusion. *International Journal of Computer Applications*, 167(5):0975–8887.
- [Society, 2019] Society, A. C. (2019). Endoscopia superior.

- [Soille, 2002] Soille, P. (2002). On morphological operators based on rank filters. *Pattern recognition*, 35(2):527–535.
- [Stehle, 2006] Stehle, T. H. (2006). Specular reflection removal in endoscopic images. In *Proceedings of the 10th international student conference on electrical engineering*. Citeseer.
- [Sánchez-González and Soto, 2017] Sánchez-González, A. and Soto, B. G.-Z. (2017). Colonoscopy image pre-processing for the development of computer-aided diagnostic tools. In *Surgical Robotics*. IntechOpen.
- [Telea, 2004] Telea, A. (2004). An image inpainting technique based on the fast marching method. *Journal of graphics tools*, 9(1):23–34.
- [Thakur and Datar, 2014] Thakur, M. and Datar, S. (2014). Image restoration based on deconvolution by richardson lucy algorithm. *algorithms*, 8:9.
- [Tortora and Derrickson, 2018] Tortora, G. J. and Derrickson, B. (2018). *Principios de anatomía y fisiología*. Médica panamericana, Ciudad de México, 15 edition.
- [Triana et al., 2017] Triana, J. J., Aristizábal-Mayor, J. D., Plata, M. C., Medina, M., Baquero, L., Gil-Tamayo, S., Leguizamón, A. M., Leonardi, F., Castañeda-Cardona, C., and Rosselli, D. (2017). Carga de enfermedad en años de vida ajustados por discapacidad del cáncer gástrico en colombia. *Revista Colombiana de Gastroenterología*, 32(4):326–331.
- [Tsai et al., 2008] Tsai, D.-Y., Lee, Y., and Matsuyama, E. (2008). Information entropy measure for evaluation of image quality. *Journal of digital imaging*, 21(3):338–347.
- [Universidad De Jaén, 2017] Universidad De Jaén (2017). Reducción del ruido en una imagen digital. Technical report.
- [Urueña et al., 2011] Urueña, W. A., Osorio, J. A. C., and Vargas, J. A. M. (2011). Técnicas alternativas para la conversión de imágenes a color a escala de grises en el tratamiento digital de imágenes. *Scientia et technica*, 1(47):207–212.
- [U.S. National Library of Medicine, 2018] U.S. National Library of Medicine (2018). Enfermedades digestivas.
- [Vakil, 2019a] Vakil, N. (2019a). Gastritis.
- [Vakil, 2019b] Vakil, N. (2019b). Úlcera gastroduodenal.
- [Valdovinos Rosas et al., 2006] Valdovinos Rosas, R. M. et al. (2006). *Técnicas de submuestreo, Toma de decisiones y Análisis de diversidad en aprendizaje supervisado con Sistemas Múltiples de Clasificación*. Universitat Jaume I.

- [Vargas et al., 2012] Vargas, J., Conde, B., Paccapelo, V., and Zingaretti, L. (2012). Máquinas de soporte vectorial: Metodología y aplicación en r.
- [Vera, 2017] Vera, M. (2017). Endoscopia en la enfermedad de crohn.
- [Wang et al., 2019] Wang, X., Liang, X., Zheng, J., and Zhou, H. (2019). Fast detection and segmentation of partial image blur based on discrete walsh–hadamard transform. *Signal Processing: Image Communication*, 70:47–56.
- [Watson et al., 2010] Watson, C., Kirkcaldie, M., and Paxinos, G. (2010). *The brain: an introduction to functional neuroanatomy*. Academic Press.
- [Yao et al., 2013] Yao, K., Nagahama, T., Matsui, T., and Iwashita, A. (2013). Detection and characterization of early gastric cancer for curative endoscopic submucosal dissection. *Digestive Endoscopy*, 25:44–54.
- [Yoo et al., 2016] Yoo, I. K., Jeen, Y. T., Kang, S. H., Lee, J. H., Kim, S. H., Lee, J. M., Choi, H. S., Kim, E. S., Keum, B., Chun, H. J., et al. (2016). Improving of bowel cleansing effect for polyethylene glycol with ascorbic acid using simethicone: A randomized controlled trial. *Medicine*, 95(28).
- [Zhang and Brainard, 2004] Zhang, X. and Brainard, D. H. (2004). Estimation of saturated pixel values in digital color imaging. *JOSA A*, 21(12):2301–2310.
- [Zhang, 2012] Zhang, Z. (2012). K-means algorithm cluster analysis in data mining. Technical report, Zijun Zhang.
- [Zou et al., 2004] Zou, K. H., Warfield, S. K., Bharatha, A., Tempany, C. M., Kaus, M. R., Haker, S. J., Wells III, W. M., Jolesz, F. A., and Kikinis, R. (2004). Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. *Academic radiology*, 11(2):178–189.