

Metodología basada en técnicas de procesamiento de imágenes para la Segmentación Automática de Microcalcificaciones en mamografías

Katheryn Bambague Ruíz^{1[0009-0003-4954-8544]}, Ronny Guevara^{1[0000-0002-1536-5948]}, Adolfo Giles²

¹ICIC, Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), Av. San Andrés 800, Bahía Blanca, Argentina, 8000

²Universidad Nacional del Sur, Departamento de ingeniería eléctricas y de computadoras, Av. San Andrés 800, Bahía Blanca, Argentina, 8000
katherymbr@gmail.com, ronny.guevara@uns.edu.ar, fito87@hotmail.com

Resumen. Las imágenes médicas son fundamentales en el campo de la salud, proporcionando datos esenciales para diagnósticos precisos y evaluación de tratamientos. Para maximizar su utilidad, se requiere la aplicación de técnicas y procesos de mejora de características para obtener resultados eficientes. En este contexto, el preprocesamiento de imágenes juega un papel crucial en la segmentación de microcalcificaciones en mamografías, resaltando características claves y detectando áreas con alta probabilidad de anomalías que podrían derivar en cáncer de mama. La detección de microcalcificaciones sigue siendo un desafío clínico, debido a su tamaño que oscila entre 0.1 y 1 mm de diámetro, densidad similar a otros tejidos circundantes ocasionando superposición, formas o estructuras variadas y la presencia de ruidos o artefactos en las mamografías que pueden interferir con su detección. Este estudio aborda los desafíos mencionados anteriormente y presenta una metodología robusta, fácil de manejar y de libre acceso para la segmentación automática de microcalcificaciones. Se utilizan diversos algoritmos de mejora de características y técnicas de procesamiento de imágenes en cascada para la segmentación automática. Posteriormente, los resultados obtenidos se comparan con una máscara de referencia marcada por un especialista, utilizando como métrica de evaluación el coeficiente Sørensen-Dice. Este enfoque permite evaluar la eficiencia de la metodología propuesta, obteniendo un coeficiente Dice de 0.68 (68% de coincidencia entre los conjuntos), lo que resalta la eficacia de la metodología propuesta.

Palabras clave: Imágenes médicas, Microcalcificaciones, Diagnósticos precisos

1 Introducción

Las calcificaciones mamarias son depósitos de sales de calcio en la mama, que se muestran con mayor densidad en las mamografías. De acuerdo con Arancibia et al., las calcificaciones mamarias son hallazgos comunes en la mamografía y su frecuencia aumenta con la edad de la paciente [1]. El cáncer de mama es el resultado del crecimiento anormal de células en el tejido del seno, comúnmente conocido como tumor. Un tumor no significa cáncer, estos pueden clasificarse en benignos (no cancerosos), premalignos (precancerosos) o malignos (cancerosos) [2].

La diferencia de densidad entre las calcificaciones benignas y malignas está dada principalmente por los distintos compuestos químicos que prevalecen en cada una [3]. Las calcificaciones benignas suelen ser más grandes, presentan un aspecto característico [15] y no requieren magnificación, están conformadas principalmente por oxalato de calcio, a diferencia de la calcificación maligna que están conformadas de fosfato de calcio y deben ser estudiadas con placas magnificadas [1].

2 K. Bambague and R. Guevara

En la actualidad, el desarrollo tecnológico permite crear aplicativos capaces de diagnosticar enfermedades rápida y verazmente [4]. Éstos requieren un conjunto de datos (imágenes) para su entrenamiento donde el preprocesamiento juega un rol fundamental. Sin embargo, en su adquisición existen factores físicos difíciles de controlar, tales como, la iluminación, la reflexión y el ruido aleatorio [5], los cuales pueden alterar el diagnóstico del paciente. Gran parte de los inconvenientes en el diagnóstico por imágenes se presentan en la etapa de preprocesamiento. Se ha demostrado que el preprocesamiento de imágenes es eficiente en imágenes dermatoscópicas [6], imágenes cardíacas [7], imágenes de retinografía a color [8], o imágenes de electroforesis [9].

Las técnicas de preprocesamiento desempeñan un papel vital para lograr el estado del arte en cualquier conjunto de datos. Coronel en su investigación manifiesta que la metodología cascada es la más tradicional y la base para muchas otras nuevas metodologías [10]. El preprocesamiento de las imágenes está dado por alineamiento, filtro gaussiano, filtro de agudización y ecualización. Además, es fundamental considerar la eficiencia y velocidad del algoritmo en el momento de su implementación, ya que esto habilita su uso en investigaciones que requieren análisis en tiempo real o en aplicaciones móviles [11].

Diversas investigaciones han incursionado en imágenes de mamografías, Álvarez et al., presenta una metodología para el preprocesamiento de mamografías de tres pasos: reducción del área de trabajo, reducción de ruido y realce de contraste. Los resultados en este estudio son: para la reducción del ruido la mejor técnica es la transformada wavelet (WT), en la etapa de realce de contraste los mejores resultados se obtuvieron con una combinación de técnicas de filtrado espacial y ecualización del histograma. [5]. Sandino en su trabajo de investigación, realiza el preprocesamiento de las imágenes, a fin de restringir el área de interés reduciendo el fondo de la imagen y el rótulo que indica el tipo de proyección en que fue realizada la mamografía. Para alcanzar el objetivo, binariza la imagen mediante un umbral, obteniendo como resultado un trabajo donde no se requiere de un conocimiento previo de la localización de las agrupaciones de microcalcificaciones, ni etapas de entrenamiento del algoritmo [12].

En este artículo, el objetivo es proponer una metodología en cascada para el procesamiento de imágenes de mamografías utilizando técnicas de mejora de características, con el fin de identificar zonas donde puedan estar presentes posibles microcalcificaciones. La metodología se distingue por su enfoque integral, donde no solo se basa en la forma y tamaño de las microcalcificaciones, sino que también agrega un conjunto de características adicionales como la forma, la textura y la densidad. Este enfoque robustece la fase de detección, permitiendo la detección más precisa de microcalcificaciones.

Adicionalmente, la integración de técnicas de umbralización adaptativa favorece aún más la detección de anomalías, ajustando el umbral de segmentación a las particularidades de la mamografía.

2 Metodología

La metodología empleada en este estudio se encuentra descrita en el diagrama de flujo representado en la Figura 1. Este describe los pasos involucrados en el proceso de segmentación automática de microcalcificaciones en mamografías. Los números 3.1, 3.2, 3.3, 4, 4.1 y 4.2 y 4.3 corresponden a las secciones del artículo donde se detallan los resultados obtenidos.

Segmentación Automática de Microcalcificaciones 3

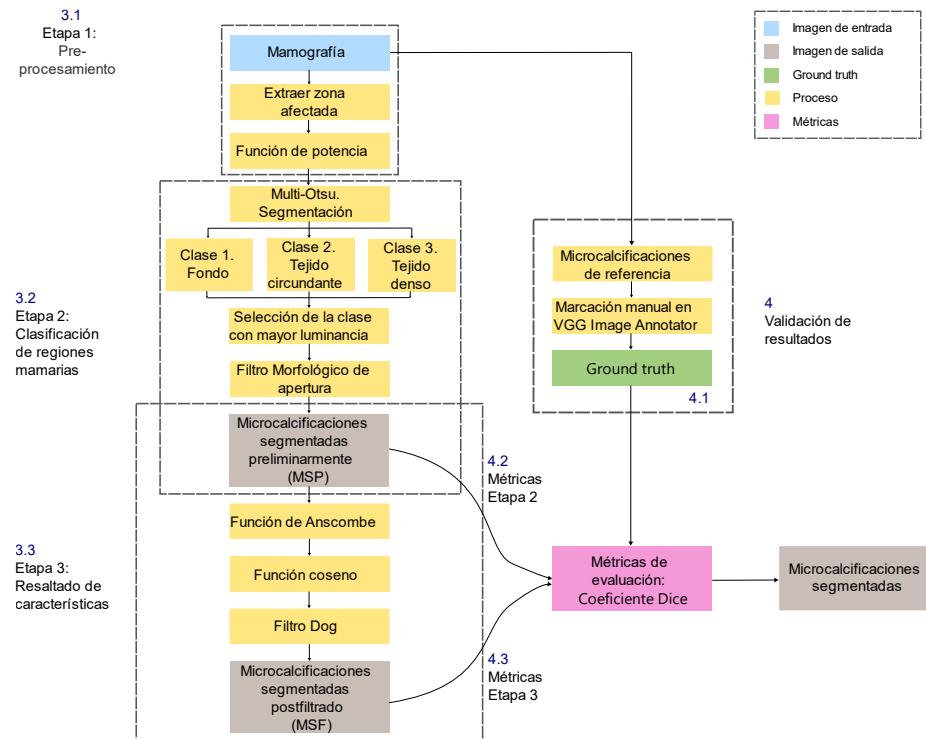


Fig.1. Descripción esquemática de la metodología propuesta para segmentación automática de microcalcificaciones en mamografías.

2.1 Base de datos

Se utilizó la base de datos mini-Mías de mamografías en esta investigación, disponible en <http://peipa.essex.ac.uk/info/mias.htm>. Esta base de datos consta de 322 imágenes de mamografías en blanco y negro, con tamaños de 1024×1024 píxeles, clasificadas por características del tejido de fondo, tipo de anormalidad presente y gravedad de esta. Se seleccionó este conjunto de datos por su accesibilidad y su capacidad para proporcionar una amplia variedad de casos que abarcan diferentes tipos de tejido mamario y diversas anormalidades. Este dataset ofrece una muestra completa que refleja la complejidad de la detección de microcalcificaciones en mamografías.

2.2 Etapa 1. Preprocesamiento

Fase inicial del enfoque metodológico, implementa una secuencia de procedimientos en cascada destinados a mejorar la imagen para la segmentación de artefactos en la mamografía. La secuencia busca resaltar estructuras mamarias y mejorar la calidad de la imagen sin comprometer la retención de detalles críticos. En la etapa 1, la imagen de entrada es una mamografía con microcalcificaciones, posteriormente se procede a identificar y delimitar la región de la imagen que contiene la patología de interés, seguido de la normalización de la imagen en el rango entre 0 y 1. Posteriormente, se aplica una

4 K. Bambague and R. Guevara

función de potencia, como se describe en la ecuación 1, para mejorar la calidad de la imagen y eliminar el ruido presente.

En la ecuación 1, $I_{potencia}$ es la imagen resultante, η el valor de la potencia e Img Es el recorte de la zona afectada con microcalcificaciones en la mamografía. Este filtro se basa en el análisis de las luminancias de los píxeles, realzando las características relevantes y atenuando las irrelevantes. El procesamiento de la imagen en esta etapa permite obtener una representación más clara y precisa de la zona afectada, lo que facilita el análisis posterior.

$$I_{potencia} = (\text{Img})^\eta \quad (1)$$

2.3 Etapa 2. Clasificación de Regiones mamarias

En esta etapa de la metodología se detectan las áreas de la imagen que exhiben similitudes con las microcalcificaciones. Para este fin, se recurre al algoritmo de Otsu, reconocido como un método de umbralización automática ampliamente utilizado en el procesamiento de imágenes. En el contexto específico de las mamografías, se emplea la técnica Multi-Otsu para clasificar la imagen en tres clases distintas (fondo, tejido mamario y microcalcificaciones), lo que implica una umbralización multinivel [16]. La expresión matemática del método Multi-Otsu se presenta en la ecuación 2.

$$\sigma^2 = \sum_{i=1}^n W_i * \sigma_i^2 + \dots W_n * \sigma_n^2 \quad (2)$$

Para tres clases: $I_{Multi-Otsu} = \sigma^2 = W_1 * \sigma_1^2 + W_2 * \sigma_2^2 + W_3 * \sigma_3^2$

La ecuación 2 presenta el método de umbralización Multi-Otsu, una extensión del método de Otsu diseñada para determinar múltiples umbrales de manera eficaz. En esta ecuación, σ^2 representa la varianza total de la imagen, W_i , W_n son los pesos asociados a las clases i hasta n, respectivamente, y σ_i^2 representa la varianza intraclase de las clases i hasta n. A diferencia del método convencional de Otsu, que divide la imagen en dos clases (fondo y primer plano), el método Multi-Otsu divide la imagen en múltiples clases. Dada una imagen en escala de grises y de 8 bits, el algoritmo Multi-Otsu busca segmentar la imagen en tres clases, utilizando los dos umbrales encontrados. Posteriormente se selecciona la región donde se encuentran las calcificaciones y se refinan los bordes usando operaciones morfológicas de filtrado. La imagen resultante de esta etapa se denomina MSP (microcalcificaciones segmentadas preliminarmente), como se describe en la metodología (ver figura 1).

2.4 Etapa 3. Resultado de características

Esta fase de la metodología se centra en resaltar las características de la región donde se encuentran calcificaciones, las cuales fueron previamente seleccionadas en la etapa 2. Con este propósito, se aplican tres funciones secuenciales: la Función de Anscombe (ver Ecuación 3), la Función Coseno (ver Ecuación 4) y el Filtro DOG o diferencia de Gaussianas (ver Ecuación 5). La primera transformación de esta etapa se aplica a la imagen MSP, resultante de la etapa 2. La transformación aplicada es una variante de la función de Anscombe. Esta función busca reducir la heterocedasticidad, estabilizar la varianza, normalizar la distribución, realzar los bordes y mejorar el contraste. La ecuación 3 proporciona la representación matemática de la función de Anscombe.

$$I_{Anscombe} = \frac{(I-A)}{B} + C \times \sqrt{\frac{(I-D)}{E}} \quad (3)$$

En la ecuación 3, la variable I corresponde a la imagen MSP, obtenida previamente en la etapa 2, mientras que $I_{Anscombe}$ representa la imagen transformada. Los parámetros de

Segmentación Automática de Microcalcificaciones 5

la transformación: A, B, C, D y E, permite ajustar el valor central de la imagen transformada, controlar la amplitud, regular la curvatura, determinar el centro de la transformación y controlar el ancho de la misma.

Posteriormente, después de aplicar la transformación de Anscombe, siguiendo la secuencia de la metodología propuesta en la Figura 1, se lleva a cabo una transformación con la función coseno a la imagen resultante del paso anterior, como se muestra en la Ecuación 4. Esto se realiza con el propósito de suavizar las transiciones de intensidad y resaltar los bordes de la imagen, lo que facilitará el proceso de detección.

$$I_{\text{coseno}} = 1 - \cos\left(\frac{\pi}{2} \times \frac{I}{255}\right) \quad (4)$$

En la ecuación 4, I_{coseno} representa la imagen transformada, mientras que I representa la imagen obtenida en la transformación anterior. El valor 255.0 representa la intensidad máxima en la imagen original.

En la fase final de esta metodología, se emplea un filtro DOG, conocido por realizar estructuras de distintos tamaños en una imagen, tales como bordes y texturas. Este filtro implica la convolución de la imagen con dos filtros Gaussianos, según se ilustra en la ecuación 5.

$$I_{\text{DOG}(x,y)} = k1 \times G(x, y, \sigma1) - k2 \times G(x, y, \sigma2) \quad (5)$$

El filtro de diferencias Gaussianas o comúnmente denominado DoG, se expresa mediante la ecuación 5, donde, $I_{\text{DOG}(x,y)}$ representa la salida del filtro en la posición (x, y) de la imagen. $k1$ y $k2$ son constantes de ponderación. $G(x, y, \sigma)$ representa la función Gaussiana en la posición (x, y) con desviación estándar σ . Mientras que $\sigma1$ y $\sigma2$ son las desviaciones estándar de las dos funciones Gaussianas utilizadas en el filtro. La imagen resultante de esta etapa final se denomina MSF (microcalcificaciones segmentadas post filtrado), como se menciona en la metodología (ver figura 1).

2.5 Validación de resultados

Para evaluar la precisión de la segmentación automática de microcalcificaciones obtenida en la etapa 2 (imagen MSP) y etapa 3 (imagen MSF) de la metodología, se utiliza el coeficiente Sørensen-Dice (Dice Similarity Coefficient, DSC) para la validación. Este coeficiente es una medida de similitud estadística empleada para comparar dos muestras y es ampliamente utilizado en el procesamiento de imágenes médicas para evaluar la superposición entre dos segmentaciones [13][14].

El coeficiente de Dice se define matemáticamente como:

$$DSC = \frac{2 \times |A \cap B|}{|A| + |B|} \quad (6)$$

En la ecuación 6, la letra A representa el conjunto de píxeles de la imagen segmentada. B representa el conjunto de píxeles de la imagen ground truth. $|A \cap B|$ es el número de píxeles comunes entre A y B. $|A|$ y $|B|$ son el número de píxeles en los conjuntos A y B, respectivamente. El valor del coeficiente Dice varía entre 0 y 1, donde 0 indica que no hay superposición entre los dos conjuntos, y 1 indica una superposición perfecta entre ellos.

3 Resultados

En esta sección se exponen los hallazgos obtenidos tras aplicar la metodología propuesta para la segmentación automática de microcalcificaciones en mamografías, tal como se

muestra en la Figura 1. Los resultados se estructuran por etapas de acuerdo con la propuesta metodológica, seguidos de una evaluación de su efectividad en la etapa 2 y 3. Por último, se calculan las métricas para evaluar el desempeño de la metodología.

3.1 Etapa 1: Preprocesamiento

El preprocesamiento de las imágenes mamográficas ha demostrado una mejora sustancial en la calidad de las mismas. Se observa una reducción notable del ruido y un incremento en la definición de los bordes de las estructuras mamarias. Se destaca especialmente el uso de la función potencia, que permite oscurecer el tejido mamario y resaltar las posibles microcalcificaciones. En la Figura 2(a), se presenta la mamografía con microcalcificaciones, mientras que en la Figura 2(b) se muestra el recorte de la zona afectada. El resultado de la aplicación del filtro de potencia a la figura 2(b), se ilustra en la Figura 2(c). Además, se presenta el histograma de luminancias de la mamografía inicial en la Figura 2(d), seguido por el histograma de luminancias de la zona afectada en la Figura 2(e), y finalmente, la Figura 2(f) muestra el histograma de luminancias después de aplicar la función de potencia.

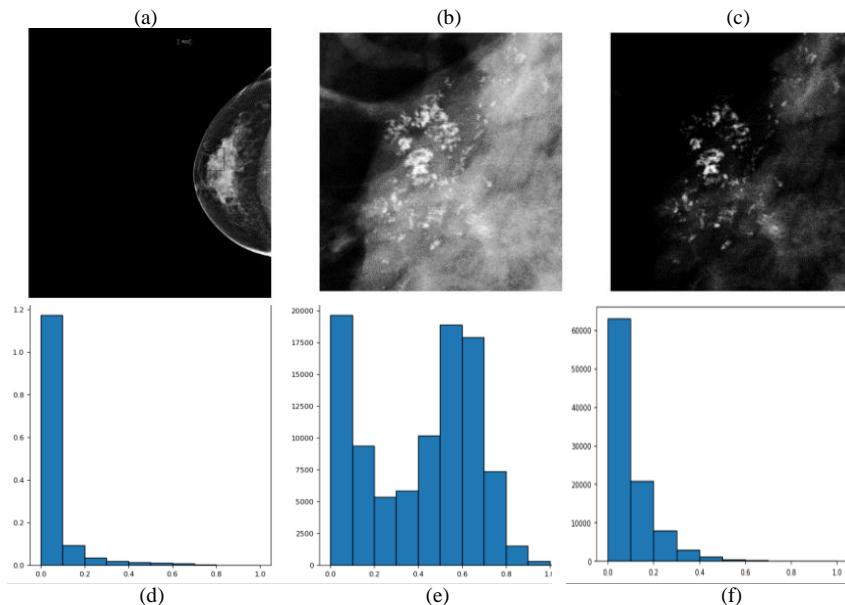


Fig. 2. Imágenes e Histogramas de luminancias de las diferentes fases en la etapa 1 de la metodología. 2(a). Mamografía con microcalcificaciones. 2(b) Recorte de la zona afectada. 2(c). Imagen después de aplicar la función de potencia. 2(d,e,f) Histogramas de luminancia de cada una de las imágenes.

El histograma de luminancias 2(d) revela una distribución predominante de píxeles en la imagen original, que se concentran mayormente en el extremo oscuro del rango. Estos valores representan el fondo de la imagen, que abarca la mayor parte de la mamografía. Por consiguiente, resulta crucial enfocarse únicamente en la región afectada, como se muestra en 2(b), donde la mayoría de los píxeles se encuentran dentro del rango de tonos más claros (histograma de luminancias 2(e)), correspondiente a la zona donde suelen ubicarse las microcalcificaciones. En la imagen 2(c), se puede observar que después de

aplicar el filtro de potencia de la ecuación 1, los valores de los píxeles se ajustan de acuerdo con la función.

Al aplicar la ecuación 1, con $n=4$ en la mamografía inicial (imagen 2(a)), se evidencia que los píxeles con valores de gris más bajos se tornan aún más oscuros, mientras que aquellos con valores más altos se vuelven aún más brillantes. Este fenómeno conlleva a una reducción del ruido, facilita un mayor contraste entre las diferentes regiones de la imagen y destaca áreas cruciales para el diagnóstico de calcificaciones. Además, esta técnica mejora la visualización de detalles importantes en la mamografía, lo cual puede ser fundamental para una detección y caracterización eficiente de posibles anomalías mamarias.

3.2 Etapa 2: Clasificación de Regiones mamarias.

El desafío en esta etapa consistió en separar las microcalcificaciones del tejido mamario circundante. Se logra clasificar las regiones empleando la biblioteca scikit-image de Python. Para ello, se importó la función threshold_multiotsu del módulo skimage.filters. Con este algoritmo, y aplicando la ecuación 2 para el caso particular de tres clases, se consiguió separar las regiones mamarias.

La primera clase representa el fondo de la imagen, la segunda incluye tejidos grasos y tejido mamario circundante, mientras que la tercera corresponde al tejido mamario denso, área con alta probabilidad de contener tumores, lesiones o posibles microcalcificaciones, que requieren una evaluación más detallada.

La imagen obtenida como resultado de la etapa 1 (Figura 2(c)), se utiliza como imagen de entrada en la etapa 2. En esta etapa, se aplica el algoritmo Multi-Otsu para segmentarla en las tres clases mencionadas anteriormente. El color azul identifica la clase 1, el verde la clase 2 y el rojo la clase 3. La clase 3 (tejido denso) es nuestra zona de interés, ya que en estas áreas pueden haber depósitos de calcio o microcalcificaciones. En la figura 3, se observan visualizaciones de procedimientos correspondientes de la etapa 2 de la metodología.

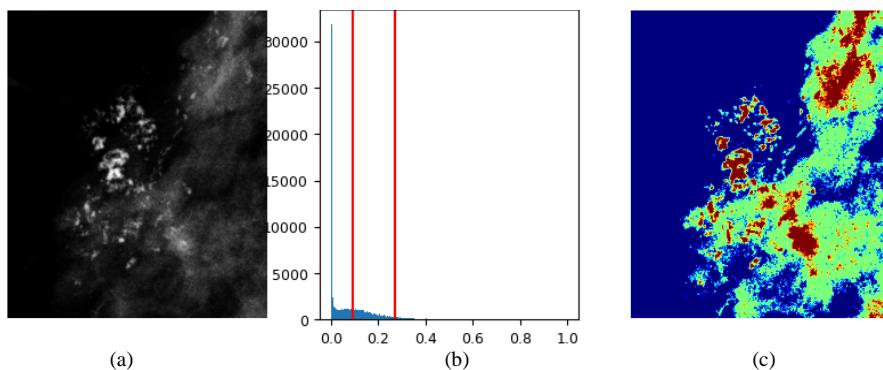


Fig. 3. Visualización de procedimientos en la etapa 2 de la metodología. 3(a). Imagen de entrada a la etapa 2. 3(b) Histograma de luminancias de la imagen de entrada y definición de las tres clases por el método Multi-otsu). 3(c). Visualización de las tres clases en la mamografía. Fondo (clase 1- color azul), tejido graso o circundante (clase 2 - color verde), tejido denso (clase 3 - color rojo).

Para profundizar en el análisis de la imagen y obtener información específica de las tres clases segmentadas con el método Multi-Otsu, se realizó un estudio cuantitativo de la distribución espacial y las propiedades de luminancia de cada clase. La Figura 4 presenta el número de píxeles pertenecientes a cada clase, lo que permite visualizar la distribución

8 K. Bambague and R. Guevara

espacial de las mismas a lo largo de la imagen. Esta información es fundamental para comprender la representatividad de cada clase en la imagen y su relación con las estructuras anatómicas presentes.

La figura 4, muestra un gráfico de barras que representa la distribución de píxeles en tres clases después de aplicar el método Multi-Otsu. El eje X describe las tres clases identificadas por el método, mientras que, el eje Y representa la cantidad de píxeles que pertenecen a cada clase.

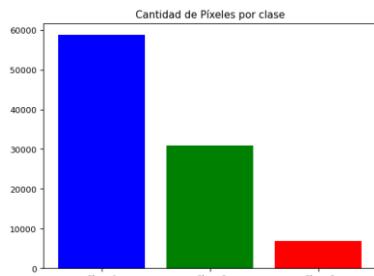


Fig. 4. Distribución de píxeles en tres clases tras la aplicación del método Multi-Otsu

El método Multi-Otsu es un algoritmo de segmentación de imágenes que utiliza una umbralización automática para dividir la imagen en diferentes regiones o clases. El algoritmo proporciona dos valores de umbral, con los cuales se forman las tres clases correspondientes. Dado que se trabaja con imágenes normalizadas en el rango de 0 a 1, las clases se definen de la siguiente manera, como se muestra en la tabla 1.

Tabla 1. Características de las clases segmentadas con Multi-Otsu en una mamografía con microcalcificaciones.

Características	Multi-Otsu	Umbral multi-Otsu	Porcentaje píxeles (%)	Valor medio (\bar{X})	Desviación estándar (s)
Clase 1	Fondo	$P \leq 0.0917$	60.90	0.020	0.028
Clase 2	Tejido circundante	$0.091 < P \leq 0.271$	32.00	0.159	0.047
Clase 3	Tejido denso	$P > 0.271$	7.08	0.383	0.120

En la Tabla 1, se evidencia que la mayoría de los píxeles (P) corresponden al fondo de la imagen, con valores de luminancia iguales o inferiores a 0.0917. Por otro lado, en la región de interés donde se buscan posibles microcalcificaciones, se observa una cantidad significativamente menor de píxeles, los cuales presentan valores de luminancia superiores a 0.271. Los valores porcentuales muestran la significancia de cada clase en la imagen, facilitando la comprensión de la abundancia relativa de las diferentes regiones. La Clase 1 predomina en la imagen con un 60.90%, seguida de la Clase 2 (Tejido circundante) con un 32.00%, y la Clase 3 (Tejido denso) es la menos presente con un 7.08%. Finalmente, en esta etapa se logró clasificar la mamografía en tres clases utilizando el método Multi-Otsu, como se muestra en la Figura 5(a). Sin embargo, dado que el objetivo principal es detectar calcificaciones, hay que centrarse únicamente en dos clases: el fondo (clase 1) y las calcificaciones (clase 3), como se ilustra en la Figura 5(b). La clase 2 se descarta. Después de obtener la máscara binaria mediante el método Multi-Otsu, se procedió a aplicar una operación morfológica. Considerando el reducido tamaño de las microcalcificaciones, se utilizó un filtro morfológico de apertura con un kernel de

2x2 para suavizar los bordes, eliminar el ruido y eliminar las pequeñas áreas que fueron falsamente detectadas como microcalcificaciones. Esta operación permite preservar las regiones más grandes y significativas que representan las verdaderas microcalcificaciones o estructuras de interés en la imagen, como se ilustra en la Figura 5(c). La imagen obtenida como resultado de la etapa 2, se denomina: microcalcificaciones segmentadas preliminarmente (imagen MSP).

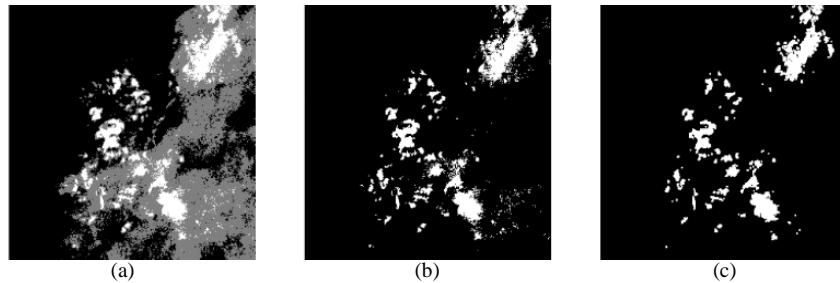


Fig. 5. Visualización del proceso de segmentación usando Multi-Otsu y postprocesamiento usando filtro morfológico de apertura. 5(a). Segmentación de la zona afectada en tres clases .5(b). Máscara binaria seleccionando 2 clases del Multi-Otsu (clase 1- fondo y clase 3-calcificaciones). 5(c).Imagen después de aplicar filtro de apertura (MSP).

3.3 Etapa 3. Resaltado de características

El sistema visual humano (HVS) proporciona un enfoque para evaluar la calidad de las imágenes utilizando el ojo humano como referencia. Esta sección presenta los resultados de las operaciones morfológicas y filtrados aplicados a la imagen MSP obtenida en la etapa 2, abordando atributos como brillo, contraste, textura y orientación. La Figura 6 muestra el proceso de separación del tejido denso en la mamografía y las diversas etapas de mejora. La imagen resultante de la etapa 3 se denomina microcalcificaciones segmentadas postfiltrado (imagen MSF).

En la Figura 6 se muestra la secuencia visual de las mejoras realizadas en el área de interés, junto con los valores que describen la calidad de la imagen en la experiencia final del usuario. A continuación, se detallan los parámetros utilizados en cada proceso de esta etapa.

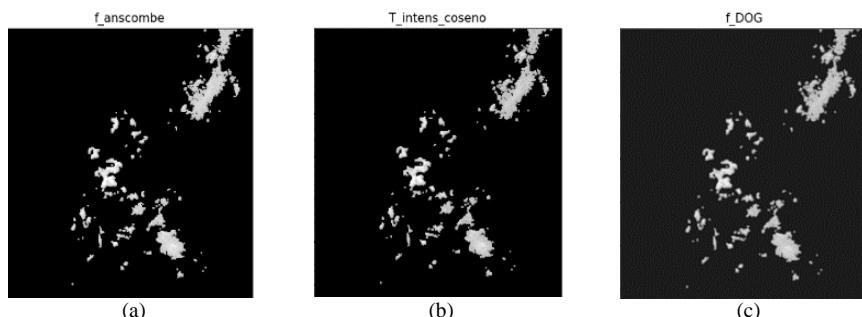


Fig. 6. Resultados de operaciones morfológicas aplicadas a la imagen MSP en la etapa 3 de la metodología. 6(a) Filtro de Anscombe. 6(b) Transformación de intensidad por coseno. 6(c)Filtro DOG.

La estabilización estadística de varianza en la imagen se logra mediante la aplicación de una variante de la función de Anscombe, definida por la ecuación 3, con parámetros $C = 2$, $D = 3$, $E = 8$. La calidad de esta estabilización se evalúa mediante la relación señal - ruido, obteniendo un valor de 27.68. Este valor indica que la señal de información en la imagen es 27.68 veces más fuerte que el ruido, lo cual sugiere una calidad visual óptima. Para suavizar el contraste de la imagen, se utiliza la transformación ajustada en la función coseno descrita en la ecuación 4. La calidad del ajuste del contraste se evalúa utilizando la fórmula de Michelson, con un valor de 0.57 indica que la transformación aplicada mejora notablemente la percepción de los bordes y detalles en la imagen.

Finalmente, para resaltar las diversas estructuras presentes en la imagen, se aplica un filtro DoG descrito en la ecuación 5, con parámetros $k1 = 1.2$ y $k2 = 0.6$. La mejora en la visualización de la imagen se cuantifica mediante el MSE, obteniendo un valor de 1.67. Este resultado demuestra una clara diferencia entre la imagen resultante de la intensidad de coseno y la imagen resultante del filtro DoG, conservando características esenciales, eliminando ruido y resaltando bordes.

4 Validación de resultados

Para evaluar con precisión la segmentación automática de microcalcificaciones propuesta en la metodología, se realizó una segmentación manual y semántica a cargo de un experto en el campo para establecer el Ground Truth. A continuación, se detalla el proceso de obtención del Ground Truth y se presenta la métrica del coeficiente Sørensen-Dice, que compara las imágenes segmentadas de las etapas 2 y 3 (imagen MSP e imagen MSF) con este estándar de referencia.

4.1 Obtención del Ground truth

Para validar los resultados obtenidos en los pasos anteriores, se realizó una validación utilizando la misma mamografía con microcalcificaciones.

En esta fase, la segmentación fue llevada a cabo manualmente por personal experto en el tema, utilizando la herramienta de código abierto VGG Image Annotator (VIA), como se muestra en la figura 7(a). La máscara generada a partir de esta segmentación manual, representada en la figura 7(c), es decir el Ground truth, se utilizó como referencia para validar la imagen obtenida en la etapa 2, numeral 3.2 (imagen MSP) y la imagen obtenida en la etapa 3, numeral 3.3 (imagen MSF).

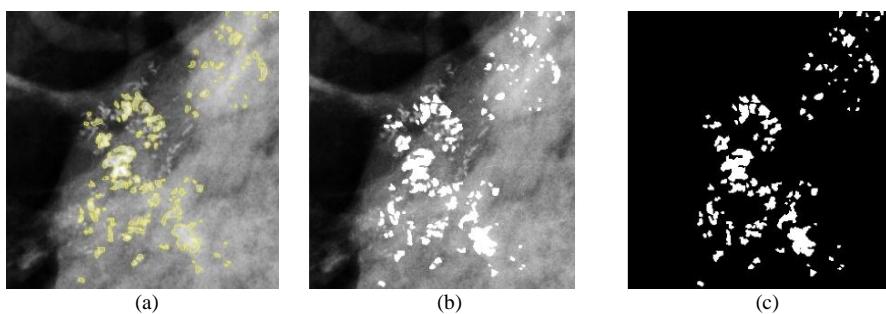


Fig. 7. Obtención del Ground truth. 7(a). Imagen segmentada manualmente por personal experto usando VGG Image Annotator (VIA). 7(b). Imagen segmentada. 7(c). Máscara de la imagen segmentada manualmente (Ground truth).

4.2 Métricas etapa 2

En esta etapa, se llevó a cabo una evaluación comparativa entre la imagen obtenida en la etapa 2 de la metodología (Imagen MSP) y el Ground truth obtenido previamente. Las dos imágenes para comparar se muestran en la figura 8: en la subfigura 8(a) se muestra el Ground truth y en la subfigura 8(b) se presenta la imagen obtenida en la etapa 2. Para este propósito, como se mencionó anteriormente, se utilizaron como métrica el coeficiente Sørensen-Dice.

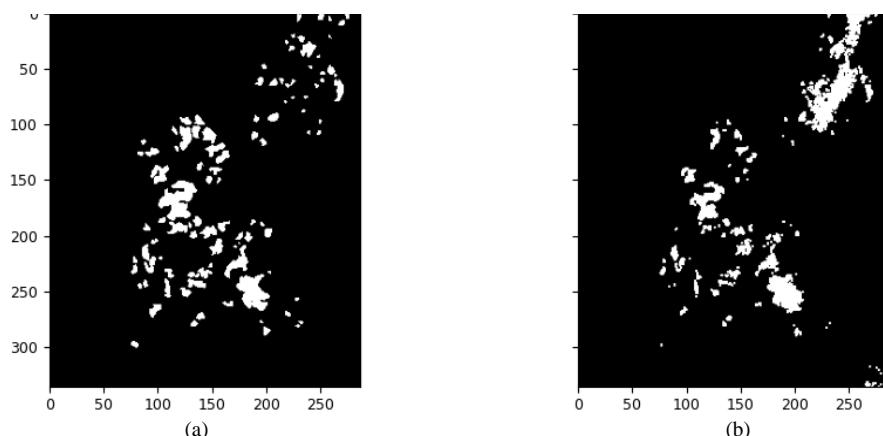


Fig. 8. Evaluación de la segmentación obtenida en la etapa 2 de la metodología usando el coeficiente Sørensen-Dice. 8(a) Ground truth. 8(b) Imagen obtenida en la etapa 2 (Imagen MSP).

A continuación, en la tabla 2, se presenta la métrica correspondiente al coeficiente Sørensen-Dice entre las dos máscaras mostradas en la figura 8.

Tabla 2. Métricas de clasificación de regiones mamarias: Comparación entre Ground Truth e imagen MSP (imagen obtenida en la etapa 2 de la Metodología)

Métricas	Segm manual	Segm Imagen MSP	Descripción
DICE	1	0.60	Indica que hay una superposición considerable entre dos conjuntos de datos o segmentaciones. Indica que el 60% de los elementos de los dos conjuntos coinciden entre sí. Es decir, hay una similitud significativa entre las dos segmentaciones evaluadas, pero aún presentan diferencias en algunos puntos.

4.3 Métricas etapa 3

En la fase anterior, se evaluó la segmentación de la imagen MSP obtenida en la etapa 2 con el Ground truth. Ahora, nos centraremos en validar la imagen MSF obtenida como producto final en la etapa 3 de la metodología (ver Fig. 6(c)). Dado que la imagen MSF (microcalcificaciones segmentadas postfiltrado) representa el último paso de la metodología propuesta para la segmentación automática de microcalcificaciones, es crucial validar este resultado final con el Ground truth obtenido previamente.

Las microcalcificaciones detectadas con la metodología propuesta fueron validadas comparándolas con una imagen de referencia como se muestra en la figura 9. Se encontró

una alta concordancia entre las microcalcificaciones identificadas y las presentes en la imagen de referencia o Ground Truth, encontrando un coeficiente Sørensen-Dice de 0.68 como se expone en la tabla 3.

La métrica presentada en la tabla 3 resalta la relevancia de las operaciones y filtros aplicados secuencialmente en la etapa 3 de la metodología propuesta. Se evidencia que, tras la aplicación sucesiva de la Función de Anscombe, la función Coseno y el filtro DOG en la mamografía para resaltar las características de las microcalcificaciones, el coeficiente Dice mejora en un 10%.

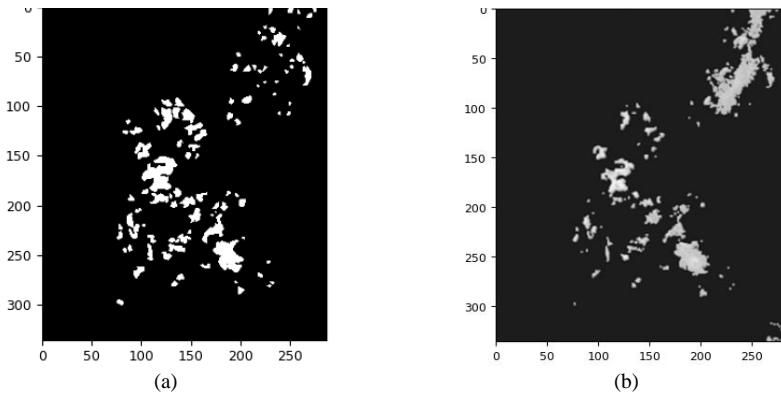


Fig. 9. Evaluación de la segmentación obtenida en la etapa 3 de la metodología. 9(a) Ground Truth. 9(b) Imagen MSF obtenida en la etapa 3

Tabla 3. Métricas de clasificación de regiones mamarias: Comparación entre Ground Truth e imagen MSF (imagen obtenida en la etapa 3 de la Metodología)

Métricas	Segm manual	Segm Imagen MSF	Descripción
DICE	0	0.68	Significa que aproximadamente el 68% de los elementos de los dos conjuntos coinciden entre sí. Las dos segmentaciones evaluadas tienen una superposición significativa pero no perfecta; hay áreas donde difieren, pero en gran parte se superponen.

En la figura 10 también se aprecia la alta concordancia entre las microcalcificaciones identificadas con la metodología propuesta (imagen MSF - obtenida en la etapa 3) y aquellas detectadas manualmente por el experto.

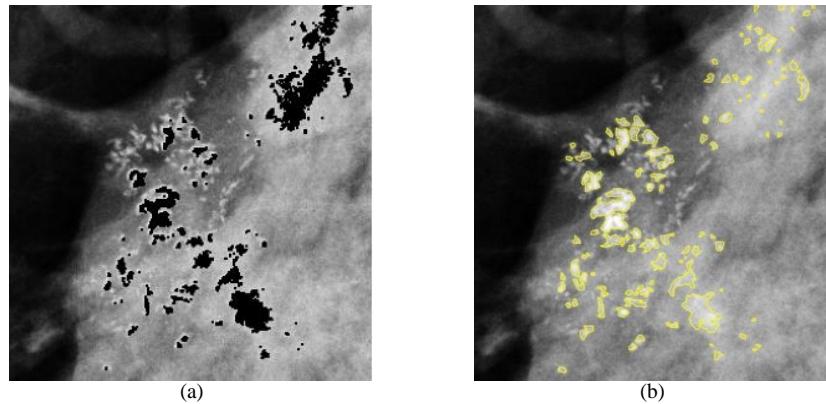


Fig. 10. Segmentación de microcalcificaciones en mamografías. 10(a). Segmentación automática obtenida con la metodología propuesta. 10(b). Segmentación manual realizada por un profesional (Ground Truth)

5 Conclusiones

Esta investigación presenta una metodología en cascada para la detección automatizada de microcalcificaciones en mamografías. A través de un enfoque en tres etapas, que abarca desde el preprocesamiento hasta el resaltado de características, se ha desarrollado un proceso integral que mejora significativamente la detección de estas anomalías. La validación de los resultados obtenidos en las etapas 2 y 3, utilizando el coeficiente Sørensen-Dice como métrica de evaluación, revela una buena precisión y fiabilidad en la detección de microcalcificaciones, considerando la complejidad de la imagen trabajada en este caso específico. Los resultados obtenidos en la validación de la segmentación correspondiente a la etapa 3, es decir, las microcalcificaciones segmentadas postfiltrado (imagen MSF), con un coeficiente Dice de 0.68, indican que esta es la mejor segmentación (68% de los elementos de los dos conjuntos coinciden entre sí). Esto valida que la mejor segmentación se obtiene al aplicar las tres etapas en cascada propuestas en la metodología inicial. En contraste, si se limita la metodología solo hasta la etapa 2, se obtiene un coeficiente Dice menor, de 0.60, lo que representa una reducción del 8% en la segmentación de microcalcificaciones. Además, como es sabido este tipo de metodologías livianas para la segmentación rápida y accesible pueden adaptarse fácilmente a diversos procesos donde se analicen imágenes y convertirse en una herramienta de apoyo, con el fin de brindar mayor velocidad a la hora de dar un diagnóstico médico.

El análisis cuantitativo de las clases obtenidas mediante el método Multi-Otsu en la etapa 2, detallado en la tabla 1, numeral 3.2, arroja información crucial sobre la distribución espacial y las propiedades de luminancia de las microcalcificaciones en mamografías. Se observa que, en imágenes con este tipo de distribución de luminancias, los píxeles con valores superiores a 0.27 representan un área con alta probabilidad de contener estas anomalías mamarias.

La aplicación de operaciones morfológicas y filtros en la etapa 3 a la imagen MSP (microcalcificaciones segmentadas posfiltrado), obtenida previamente en la etapa 2 del proceso metodológico, ha evidenciado mejoras significativas en la calidad y la percepción visual de las mamografías. La estabilización estadística de la varianza mediante la función de Anscombe, la suavización del contraste a través de la transformación de coseno y el resultado de estructuras mediante el filtro DoG han sido elementos clave para realzar la visualización de microcalcificaciones y otras características relevantes. Estos hallazgos respaldan la eficacia y utilidad de nuestra metodología en la detección automatizada de calcificaciones en imágenes mamográficas, proporcionando herramientas valiosas, de bajo costo y fácil manejo para un diagnóstico temprano y eficiente.

Referencias

1. Arancibia P., Teresa E., López A., Díaz M., & Sáez C.: Calcificaciones mamarias: descripción y clasificación según la 5.^o edición BI-RADS. Revista Chilena de Radiología, 22(2), 80-91 (2016). DOI: 10.1016/j.rchira.2016.06.004
2. Rodríguez F.: Pronóstico de cáncer de mama benigno y maligno: comparación de nueve métodos de clasificación usando R. Universidad Nacional de Colombia. (2019).
3. Arancibia P., Taub J., de Grazia J., Díaz M., López A., Sáez C., et al.: Microcalcificaciones mamarias: revisión de los descriptores y categorías BI-RADS. Rev Chil Obstet Ginecol., 78, 383-394 (2013).
4. Guevara R., Delrieux C.: Aplicación de redes neuronales densas y convolucionales para detección de COVID_19 en imágenes de rayos X. Revista conectividad, 4(2) (2023).
5. Álvarez D., Guevara M., & Holguín G.: Preprocesamiento de imágenes aplicadas a mamografías digitales. Scientia et Technica, 31 (2006).
6. Catillo A., Gallegos F., Rosales A., & Ramos R.: Preprocesamiento de imágenes dermatoscopicas para extracción de características. Research in Computing Science, 114, 59-70 (2016).
7. Vera M., Huérano Y., Contreras J., Valbuena O., et al.: Técnicas de preprocesamiento de imágenes cardíacas: fundamentos y alcance. Revista Latinoamericana de Hipertensión, 11, 60-66 (2016).
8. Pérez M., Pedraza J., & Tovar S.: Estudio comparativo de técnicas de preprocesamiento en imágenes de retina. Asociación Mexicana de Mecatrónica A.C, 11, 10-19 (2022).
9. Mora M.: Pre-procesamiento de imágenes de electroforesis bidimensional en gel aplicado al análisis proteómico. Instituto Tecnológico Metropolitano (2017).
10. Coronel B.: Impacto del preprocesamiento de imágenes en la efectividad de la verificación facial empleando visión computacional. Universidad ESAN (2020).
11. Adouani, A., Wiem, M., & Zied, L.: Comparison of Haar-like, HOG and LBP approaches for face detection in video sequences. International Multi-Conference on Systems, Signals & Devices, 266-271, Istanbul: IEEE (2019).
12. Sandino, A.: Detección de microcalcificaciones mamarias agrupadas. Universidad Distrital Francisco José de Caldas (2017).
13. Carass A., Roy S., Gherman A., Reinhold JC., Jesson A., Arbel T., et al.: Evaluating White Matter Lesion Segmentations with Refined Sørensen-Dice Analysis. Sci Rep, 10(1), (2020). DOI: 10.1038/s41598-020-64803-w
14. Wang R., Lei T., Cui R., Zhang B., & Meng H.: Medical image segmentation using deep learning: A survey. IET Image Processing, 16(5), 1243-1267 (2022). DOI: 10.1049/ipt2.12419
15. Ali M., Czene K., Hall P., & Humphreys K.: Microcalcification on mammography: approaches to interpretation and biopsy. British Journal of Radiology. (2016)
16. Liao P., Chen T., & Chung P.: A fast algorithm for multilevel thresholding. J. Inf. Sci. Eng. Journal of Information Science and Engineering, 17(5), 713-727 (2001)