

Sistema de segmentación y clasificación de lunares sospechosos de melanoma

Kalafatic E. Emiliano, González Budiño Joaquín
Ingeniería en Informática - FICH-UNL.

Resumen— En este artículo se presenta un enfoque alternativo para el desarrollo de un sistema de segmentación y clasificación de lunares sospechosos de melanoma, con la motivación de afianzar y utilizar las herramientas aprendidas en la asignatura, como así también para evaluar la implementación sobre datos reales. Para lograrlo, se investigó sobre los diferentes métodos, donde cada uno de estos ofrecen diferentes resultados y observan diferentes características del lunar. Entre ellos, este documento se basa en el método ABCD. Se obtuvieron imágenes de lunares de la base de datos proporcionada por “The International Skin Imaging Collaboration (ISIC)” [4], de las cuales se obtuvieron las imágenes con un notable marco de la lente. El desarrollo se llevó a cabo en el lenguaje Python, haciendo uso de la librería OpenCV y MedPy para muchos de los cálculos. Se desarrollaron finalmente dos métodos, clasificación por color por histograma y clasificación por bordes más el uso de la varianza de Hue (del modelo de color en HSV). Con las medidas de desempeño calculadas se determinó que el porcentaje medio de aciertos del primer método fue del 75%, mientras que para el segundo fue del 81%.

Palabras clave—melanoma, lunar, maligno, asimetría, colores, bordes, dermatología.

I. Introducción

Un melanoma maligno tiene la contrapartida de que, si es pronosticado como tal de forma tardía, será muy complicado su tratamiento. Además, la mayoría de los pacientes, concurren a un médico en las últimas etapas, cuando los melanomas ya evidencian características malignas como son el tamaño, color (más oscuro, variaciones, entre otros) y si probablemente han llegado al sangrado. Es importante, entonces, que ante las mínimas cualidades sospechosas detectadas en algún lunar, se recurra a un especialista para un análisis.

Aun así, posterior a la visita al especialista en piel, se requerirá también hacer un seguimiento del estado del melanoma, se haya o no identificado el lunar en cuestión como maligno durante la primera visita, ya que los lunares malignos tienen la característica de evolucionar y acrecentar sus características negativas con el paso del tiempo. Cuando se da este periodo de tiempo para observar la evolución de dicha lesión, el médico no tiene acceso al paciente hasta la siguiente revisión. Sin embargo, se puede utilizar una herramienta de control que detecte si la lesión ha tenido cambios significativos.

De esta manera, este trabajo tiene el objetivo de desarrollar un método que le sirva de apoyo tanto a una persona para analizar el estado de un melanoma y acudir a un médico en caso de síntomas sospechosos, o bien, que sirva de apoyo al análisis rápido de melanomas por parte de un profesional en la salud.

II. Antecedentes

REGLA ABCD

La regla ABCD es uno de los métodos clásicos, entre otros, para determinar si un melanoma es benigno, sospechoso o maligno. Su nombre hace referencia a las cuatro características que serán evaluadas:

- **Asimetría:** corresponde a la característica de los lunares de tener una forma o estructura amorfa, esparcida, no simétrica comparando rápidamente con un círculo o elipse como se da comúnmente. La asimetría se determina dividiendo la lesión por dos ejes perpendiculares ubicados de tal manera de producir el menor valor posible de asimetría. Si la lesión muestra propiedades de asimetría de

acuerdo a ambos ejes con respecto a color y diferencia estructural, el valor asignado para la asimetría es 2. Si existe asimetría en relación a un solo eje, el valor que se le asigna es 1. Si no existe asimetría en relación a ambos ejes, el valor es 0. Además, se suman otros descriptores más básicos como son la relación de aspecto y la extensión, que mide la razón entre el área del contorno y el área del rectángulo delimitador mínimo que encierra al lunar.

- **Borde:** estudia la regularidad del borde del melanoma, es decir, si dicho melanoma posee un borde liso o por el contrario presenta rugosidades, múltiples grietas, entradas y también existe un cambio brusco de color desde el centro hacia el borde, cuando esto debería ser gradual.
- **Diferencias estructurales u homogeneidad:** la homogeneidad mira si existen manchas dentro del melanoma. También se tiene en cuenta otras características como ser la red de pigmentos, áreas homogéneas o sin estructura, líneas, puntos y glóbulos. Las áreas homogéneas o sin estructuras deben ser más del 10% del total de la lesión. Las líneas y puntos son contados solo cuando más de dos son claramente visibles mientras que la presencia de un solo glóbulo es necesaria para tenerlo en cuenta.
- **Color:** un melanoma se considera sospechoso de ser maligno (con respecto a este descriptor) si posee un color general muy oscuro, casi negro. Otro aspecto es la variación de color dentro del lunar, respecto a una serie de colores considerados como indicadores malos. Los colores posibles son: blanco, rojo, marrón oscuro, marrón claro, azul grisáceo y negro.
- **Evolución:** el área de la lesión ha crecido en un corto periodo de tiempo o ha evolucionado rápidamente en alguna de sus otras características negativas.

PREPARACIÓN

La dermatoscopia consiste en la utilización de un instrumento óptico llamado dermatoscopio, que permite una observación detallada de las lesiones de la piel y así detectar estructuras y características. El dermatoscopio convencional es un microscopio manual con capacidades de ampliación de 10 o más, así como una fuente de luz convencional o polarizada. La utilización de polarización cruzada permite reducir los efectos de la reflexión y refracción de la luz, con el fin de poder apreciar estructuras anatómicas de la epidermis.

Los dermatoscopios actuales introducen ruidos parecidos a un ruido Gamma o Rayleigh, como se observa en el histograma de la Figura 1:(a). La presencia de este ruido es mínima, y puede incrementarse ante mala limpieza de la lente o presencia de polvo. Por el mismo motivo de mala limpieza de la lente, otros ruidos presentes pueden ser de tipo impulsivo o degradación por desenfoque gaussiano en una región de la imagen. Normalmente se omitirá estos últimos ya que los dermatoscopios modernos trabajan de forma remota mediante Bluetooth o Wireless para transmitir de forma segura una imagen a un computador.

Por último, los dermatoscopios introducen una estructura indeseable dentro de la imagen con la forma de la lente, en caso de que se haga mucho zoom. Esto se puede presenciar en la Figura 1: (b).

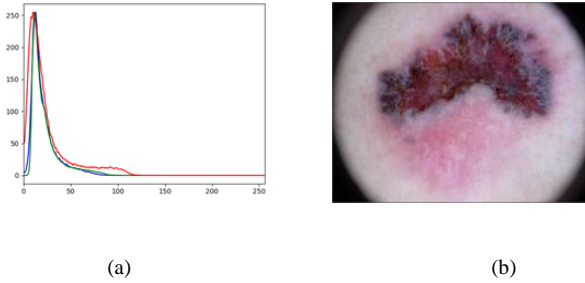


Figura 1: (a) Histograma de lunar tomado con un dermatoscopio. (b) Lunar tomado con un dermatoscopio

III. Extracción de características

SUPUESTOS

A modo de simplificar el trabajo realizado para centrar las tarea en la segmentación y detección de características malignas, se introdujeron algunos supuestos:

- El lunar se debe ubicar aproximadamente en el centro de la imagen.
- El lunar debe estar poco o nada afectado por la presencia de vello que afecte su estudio.
- Se presencia la forma entera del lunar, evitando estar recortado por los bordes de la imagen.
- Se presencia tanto el lunar (objeto de interés) como el fondo (la piel) para mejores resultados.
- La lente de la cámara siempre está limpia a la hora de capturar una imagen.
- Los bordes de la lente no deben inferir por completo en la imagen, es decir, el borde

VISIÓN GLOBAL

En la Figura 2, se presenta en un esquema el proceso completo realizado, desde el preprocesamiento, segmentación hasta el análisis de características de la regla ABCD anteriormente mencionada.

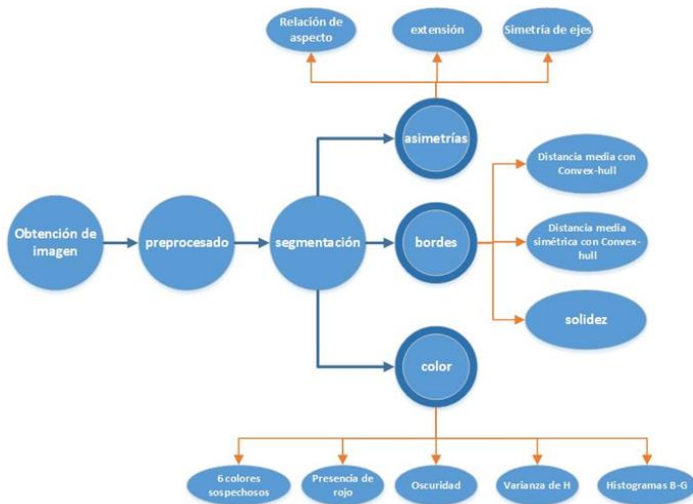


Figura 2: Esquema global de procesos

PREPROCESADO

Para eliminar la lente oscura en las esquinas, se procedió a crear una máscara circular binaria, tomando como diámetro del círculo el eje de tamaño mayor en la imagen (alto o ancho). Esto brinda una aproximación a la forma de la lente de la cámara para suprimir valores indeseados en las esquinas. Los resultados se pueden observar en la Figura 3.

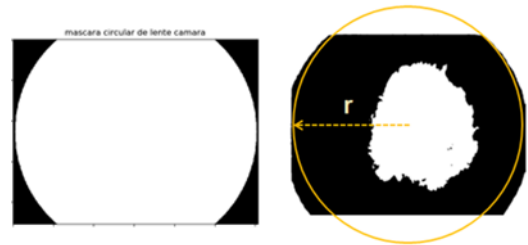


Figura 3: Eliminación del marco de la lente.

SEGMENTACIÓN

En la Figura 4 se puede observar el proceso de segmentación.



Figura 4: Pasos para la segmentación

Para lograr segmentar el lunar, con la imagen en escala de grises, primero se aplicó un filtrado de medias no locales para la eliminación de ruido ya que este método hace una ponderación teniendo en cuenta píxeles tanto locales como globales de la imagen [1][3], dando mejor resultado con menor pérdida de detalles que es lo que se precisa para el lunar (especialmente para los bordes).

Luego se aplicó un umbral óptimo de otsu para determinar una división precisa con los objetos más grandes.

Seguidamente se aplicó morfología para eliminar ruidos de fondo del lunar. Para ellos se hizo uso de dos aperturas, una con kernel elíptico y otra con kernel cruz (debido a la forma del lunar y vello), seguido de un cierre y dilatación con kernel elíptico para una recuperación del lunar.

Como el objeto de mayor área en el 99% de los casos es el lunar, se aplicó el método de "Componentes conectadas" etiquetando todos los objetos y obteniendo una máscara del de mayor área.

Finalmente se extrajo el lunar multiplicando la máscara por la imagen original. Los resultados se pueden observar en la Figura 5.

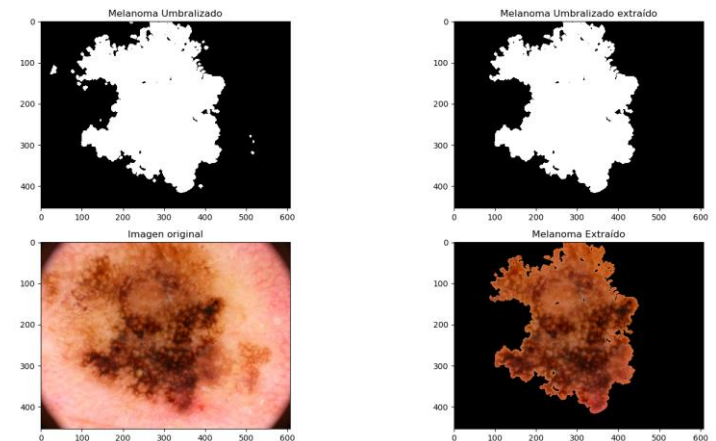


Figura 5: Obtención del objeto de mayor área y extracción del lunar de la imagen original

ASIMETRÍAS

Relación de Aspecto

Es la razón entre el ancho y la altura del contorno del objeto. Para calcularlo se hizo uso de la función *boundingRect()* de la librería OpenCV. No es un buen indicador, pero aporta al cálculo junto a la extensión.

Extensión

Se mide con la razón entre el área del contorno del lunar y el área del círculo delimitador mínimo que encierra al lunar. De esta manera es posible determinar la variación entre una forma simétrica regular de un lunar sano, respecto al lunar estudiado. Para calcularlo se hizo uso de la función *minEnclosingCircle()* de la librería OpenCV obteniendo un círculo delimitador.

Simetría respecto a ejes

A partir del centroide del lunar como centro de un sistema de referencia, se divide al lunar en 4 cuadrantes [1][2]. Luego se analiza en cada cuadrante la cantidad de píxeles del lunar en esa porción, respecto a la cantidad de píxeles totales del lunar completo. Las proporciones mediante esta evaluación deben ser cercanas a $\frac{1}{4}$ para considerar un lunar simétrico sano.

BORDES

Distancia media con convex-hull

Una de las medidas más sustanciales para el resultado fue encontrar la deformación de los bordes del lunar. Para esto se obtuvo un convex-hull del lunar (con la ayuda de la función *findContours* de la librería OpenCV) y se calculó la distancia media del contorno del lunar con la del convex-hull (ASD), utilizando la función *asd()* de la librería MedPy.

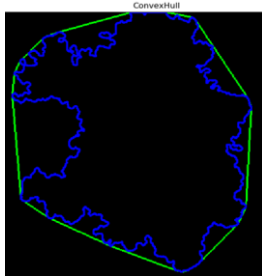


Figura 6: Convex-hull del lunar extraído.

Distancia media simétrica con convex-hull

La distancia media simétrica (ASSD) se calculó de manera similar a la anterior (utilizando la función *assd()*) pero esta vez se tiene en cuenta la simetría de la distancia. Fue de gran importancia para obtener una aproximación de la forma del lunar.

Distancia máxima entre contorno y convex-hull

Para calcular esta distancia se aplicó la distancia Hausdorff (función *hd()* de la librería MedPy). Fue útil también para obtener la deformación del contorno.

Solidez

La solidez es la razón entre el área del contorno y el área de su envoltura convexa. Útil para detectar múltiples cuencas, entradas o grietas en los bordes que evita que sea de carácter uniforme.

COLOR

Histogramas B-G

Mediante observación de los histogramas en muestras de melanoma benignos y malignos, se detectó un patrón de variación entre los planos azul y verde. Para melanomas benignos, el histograma de los planos B-G son muy parecidos entre sí y al calcular la media y varianza de ambos, daban valores muy similares. Con los melanomas malignos no ocurre así, la varianza es mayor y la medida de media de la distribución son un poco más distintas a la de los melanomas malignos. Por ello se consideró la distancia entre medias de las distribuciones verde y azul en el histograma.

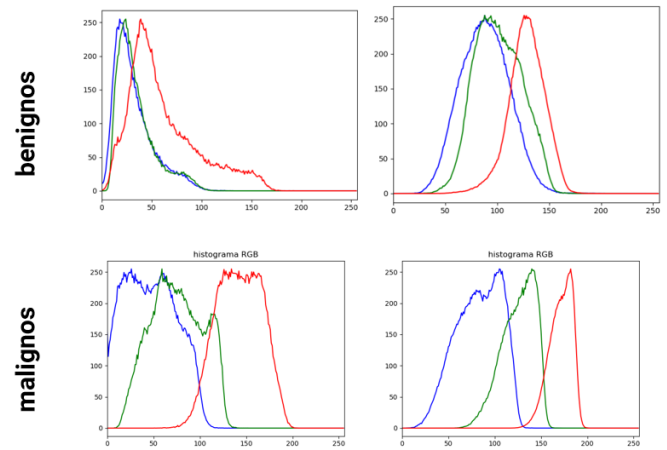


Figura 7: Histogramas de lunares benignos y malignos.

Varianza de H

Se utilizó el canal del Hue en el modelo de color HSV para detectar, de forma global, la variación de color dentro del lunar ya que, como se observa en la Figura 8, tiene valores más bajos donde los tiene la imagen y viceversa. Al mismo tiempo, se utilizó para determinar la presencia de mucho color rojo como indicador negativo con ayuda de la saturación. Para evitar errores respecto al tono de piel de la persona, se detectó cual es el color de piel predominante en el plano H, para medir variaciones respecto a ese valor, junto al valor global del melanoma [2].

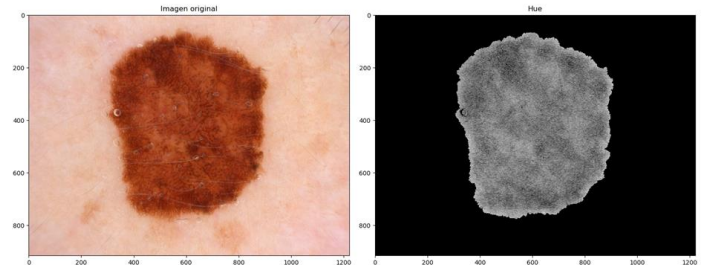


Figura 8: Comparación imagen RGB e imagen de Hue

Oscuridad

Se utilizó el modelo de color LUV, trabajando sobre el plano L-luminancia, que arrojó mejores resultados que utilizar escala de grises o el plano Value. Se calculó la media y varianza global de todo el lunar donde mediante múltiples experimentos se encontró un umbral de valor 120, para determinar que todo lunar con luminancia global por debajo de dicho valor, es considerado muy oscuro [3].

6 colores sospechosos

En la Tabla 1, se muestran los valores de interés. Si una lesión contiene dos o más de estos valores, es más probable que sea maligna [1]. Por ello, se crearon seis contadores, que se incrementan dependiendo de la cercanía del color de determinado píxel (distancia euclídea) con los valores de la Tabla. Esto se hace con todos los píxeles que pertenecen a la lesión y, por último, se divide por el área de la lesión para normalizar. Esto será problemático si más de uno de estos rangos supera el 10% de del total de píxeles recorridos, ya que los melanomas sanos tienden a tener solo un color que supere este porcentaje.

También, se propone en el presente estudio añadir la media y varianza de una nueva imagen obtenida como la distancia euclidiana entre el color de la imagen original para cada píxel y los tonos de interés, estos parámetros dan información de qué tan alejados están los colores de los tonos de interés en promedio y que tanto se alejan de ellos.

Color	Rojo	Verde	Azul
Blanco (B)	1	1	1
Rojo (R)	0.8	0.2	0.2
Café claro (CC)	0.6	0.4	0
Café oscuro (CO)	0.2	0	0
Gris azulado (GA)	0.2	0.6	0.6
Negro (N)	0	0	0

Tabla 1: Puntuación por color.

IV. Clasificación

A continuación, se mencionan los dos métodos de clasificación que mejores resultados dieron de los desarrollados utilizando las características mencionadas anteriormente.

COMPARACIÓN ASD vs ASSD vs STD H

Las medidas que mejores resultados diferenciales dieron fueron la distancia media y media simétrica entre contorno del lunar y su convex-hull (Tabla 2) y la varianza del Hue (H) (Tabla 3).

Archivo	Hausdorff	ASD	ASSD	Maligno	Benigno
ISIC_0000022	90,14	11,05	17,09 x		
ISIC_0000026	49,25	5,86	7,6 x		
ISIC_0000029	36,41	4,74	5,54 x		
ISIC_0000030	124,2	18,64	24,32 x		
ISIC_0000031	57,14	10,58	14,45 x		
ISIC_0000003	20,63	2,92	2,45		x
ISIC_0000012	6,32	0,9	1,04		x
ISIC_0000008	33,5	4,13	5,44		x
ISIC_0000010	63,26	4,23	9,08		x
ISIC_0000019	14,03	2,75	3,12		x

Tabla 2: distancias con contorno y convex-hull.

Archivo	Media H	std H	Maligno	Benigno
ISIC_0000026	57,13	79,78 x		
ISIC_0000029	19,06	46,43 x		
ISIC_0000030	157,21	32,56 x		
ISIC_0000031	79,03	84,76 x		
ISIC_0000169	41,81	73,33 x		
ISIC_0000003	11	5,99		x
ISIC_0000008	9,99	3,62		x
ISIC_0000010	11,83	20,1		x
ISIC_0000019	11,06	26,98		x
ISIC_0000032	8,84	3,86		x

Tabla 3: media y varianza de H.

Finalmente, bajo experimentación y con los resultados mencionados, se llegó al condicional de la Figura 8, utilizando umbrales de 4.7, 35 y 12 respectivamente.

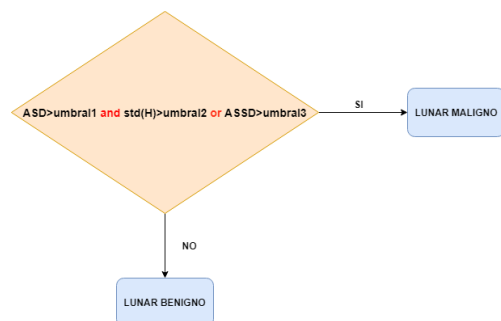


Figura 8: Condicional para clasificación

CLASIFICACIÓN POR COLOR

Tanto la oscuridad general del lunar, los histogramas de los planos B-G y la varianza del tono H cercano a los rojos, no dieron resultados muy precisos como se había estimado. Su tasa de acierto fue de 60%. Sin embargo, este resultado mejoro considerablemente al aplicar sobre el análisis la presencia de los seis colores malignos. De esta manera se

aplicó primero la comparación con la tabla de colores malignos, luego ante la presencia de 2 de ellos se identificó como sospechoso y más de 3 colores muy sospechoso. Luego de este test se procedió con el resto de las evaluaciones para dar más precisión a la clasificación momentánea. De esta manera, la tasa de acierto creció a un 75% de aciertos.

V. Resultados

Los resultados se pueden observar en la Tabla 4, siendo el mejor método el condicional con la distancia media y media simétrica entre contorno de lunar y su convex-hull más la varianza de H.

Método	Porcentaje medio de aciertos
ASD + ASSD + std H	81%
Tonalidad global + Gaussianas Color + 6 colores malignos	75%

Tabla 4: Resultados de métodos.

VI. Referencias

- [1] Cristian Felipe Ocampo (2011). "Herramienta Soporte al Diagnóstico del Melanoma usando Imágenes Dermatoscópicas". Tesis de Magister en Ingeniería de Automatización Industrial. Universidad Nacional de Colombia, Facultad de Ingeniería y Arquitectura
- [2] Roberto Mantas, Julio Prieto, José Antonio López, Francisco Varela (2018). "Clasificación de melanomas por imágenes". Universidad de Sevilla, Sevilla, España.
- [3] Miguel A. Castillo Martínez, Francisco J. Gallegos Funes, Alberto J. Rosales Silva. "Preprocesamiento de imágenes dermatoscópicas para extracción de características". Instituto Politécnico Nacional, Escuela Superior de Ingeniería Mecánica y Eléctrica, México.
- [4] The International Skin Imaging Collaboration (ISIC). Estados Unidos: Recuperado de: <https://www.isic-archive.com>.