

Asimetría de regiones digitales bajo el concepto de compacidad en el diagnóstico de cáncer de piel

Raul Santiago, Antonio Zamarrón y Angélica Ortiz-Gaucin

Instituto Tecnológico de León, León, Gto,
México
rsantiago66@gmail.com

Resumen El cáncer de piel tiene tasas de crecimiento del 400% a nivel mundial y en México este porcentaje es más alto. Entre todos los cánceres detectados en el mundo, uno de cada tres corresponde a la categoría de cáncer cutáneo. Entre todos los cánceres de piel, el más letal es el melanoma. Detectar en etapas tempranas y de forma masiva se ha convertido en un tema de interés mundial. La regla ABCD es el principal instrumento de que se sirven los especialistas, de manera cualitativa, en el diagnóstico temprano de este padecimiento. La A se refiere al grado de asimetría que presenta la lesión, B hace alusión al borde, C a las tonalidades presentes y D al tamaño. Los sistemas de diagnóstico asistido por computadora se han convertido en una opción para ayudar al diagnóstico masivo y en donde el principal componente es la descripción de estas propiedades de las lesiones. En el presente trabajo se hace uso de la geometría digital para genera una descripción de la primera característica, la asimetría, utilizando el concepto de compacidad para general un vector de características cuatro-dimensional. Esta propuesta se prueba con varias medidas que ilustran la utilidad del concepto en procesos de diagnóstico asistido mediante procesos de clasificación automática.

Palabras clave: Geometría digital, reconocimiento de patrones, análisis de forma, compacidad discreta.

1. Introducción

Uno de los cánceres más agresivos es el melanoma. El melanoma es una clase de cáncer de piel generado a partir de melanocitos atípico. Este tipo de cáncer se caracteriza por su gran capacidad de producir metástasis. La incidencia de este tipo de cáncer se está convirtiendo en un problema de alcances mundiales. La Organización Mundial de la Salud reporta que uno de cada tres pacientes diagnosticado con cáncer tiene algún tipo de cáncer de piel [1]. Del total de casos de cáncer de piel, el cuatro por ciento es melanoma. Acorde a las estadísticas del Instituto Mexicano de Cancerología, en México la incidencia ha aumentado en 500% [2].

El enorme crecimiento en la incidencia de cáncer de piel y la alta agresividad del melanoma ha despertado el interés de la comunidad científica por el desarrollo de sistemas automáticos de detección temprana de melanoma [3, 4]. Hasta el 2009 se habían escrito más de 400 estudios sobre descripción para el diagnóstico a través de

sistemas asistido por computadora [4]. La caracterización de una lesión sospechosa de ser cáncer de piel se realiza principalmente por la regla ABCD, donde A se refiere a la Asimetría de la lesión, B a la irregularidad del Borde, C a la diferencia de tonalidad en la lesión y D es la dimensión de la lesión.

Para el caso de la asimetría, ésta se evalúa a partir de un punto dentro de la región. Una vez localizado en punto se divide la región de la lesión por uno o más ejes. Las subregiones obtenidas a partir de los ejes generados se someten a diversos métodos de evaluación. Entre los métodos más generalizados esta la correlación de puntos entre las subregiones obtenidas por un traslape de estas subregiones [3, 5, 6].

El presente trabajo se propone un proceso para evaluar la asimetría de las lesiones de piel utilizando el concepto de compacidad. La compacidad hace alusión al grado de dispersión que un objeto o figura plana muestra. Existen más de 20 propuestas para medir este concepto en el espacio continuo. No obstante, estas propuestas presentan inconvenientes cuando son aplicadas sobre imágenes digitales, que hoy en día son la naturaleza donde se capturan las imágenes de las lesiones en piel. Es por ello que se aplican medidas de compacidad diseñadas específicamente para el espacio digital.

Con las medidas de compacidad en el espacio digital, se propone un vector de características 4 dimensional que es presentado a un proceso de inferencia automático; en nuestro caso hemos elegido el algoritmo KNN, por su sencillez y alto grado de eficacia. Los vectores de características de cada medida son comparados y es medida la eficacia de cada descripción, así como de la propuesta.

EL resto del documento se organiza como sigue: en la sección dos se describe el concepto de compacidad y las medidas de compacidad para el espacio digital utilizadas en el proceso comparativo. En la sección tres se muestra la propuesta de descripción. En la sección cuatro se presentan los resultados obtenidos y finalmente se presentan algunas conclusiones.

2. Compacidad

El concepto de compacidad hace referencia al grado de densidad-porosidad de un objeto o al grado disperso-compacto de la forma de este objeto [7]. En el caso de la forma, el concepto se aplica tanto en el plano como en el espacio. Cuantificar el concepto ha recaído sobre la teoría isoperimétrica desde hace varias décadas [8] y su relación base se ha aplicado como medida de compacidad de las formas [8-12], que está dada por:

$$C = \frac{p^2}{A} \quad (1)$$

No obstante su amplio uso, se detectaron efectos no deseables cuando la medida es aplicada sobre formas contenidas en imágenes digitales [13-15, 14]. La medida, en ámbito de las imágenes digitales, es afectada de manera importante por la resolución y la rotación de la forma bajo análisis. Esto es, la forma de un objeto contenida en una imagen digital a diferente resolución produce diferentes valores de compacidad, bajo la medida isoperimétrica. Esta inconsistencia se incrementa cuando se agrega una rotación a la forma.

Con el incremento y uso de imágenes digitales en diversas aplicaciones, el generar una medida adecuada para la compacidad se ha convertido, en si mismo, en un problema de investigación. Hasta ahora se encuentran reportadas 12 propuestas de medir el concepto de compacidad sobre formas en imágenes digitales [16, 17]. De todas ellas, 4 propuestas se han tomado en función de su sencillez de implementación y que realizan la evaluación de la compacidad aplicando el enfoque de similitud entre la forma bajo estudio y una forma referencia; mismo enfoque que presenta la medida clásica de compacidad.

La primera propuesta que describimos fue diseñada por Bribiesca y se apega al comportamiento que tiene la medida isoperimétrica en espacio continuo. La medida recibió el nombre de Compacidad Discreta Normalizada [18] y ha mostrado su utilidad para evaluar la evolución de tumores cancerígenos; Braumman et al. utilizan como único parámetro los valores de la C_{DN} en tumores cervico-uterinos [19]. La C_{DN} emplea una relación entre representaciones lineales de áreas. Esta relación compara el área de la región digital de la forma bajo estudio contra una región digital cuadrada con igual área a través de su compacidad discreta C_D . Los valores de la C_{DN} están contenidos en el rango 0-1; mientras que las formas compactas producen valores cercanos a 1 y las más dispersas tenderán al valor de 0. Esta relación se expresa como:

$$C_{DN} = \frac{C_D - C_{Dmin}}{C_{Dmax} - C_{Dmin}}. \quad (2)$$

donde

$$C_{D=\frac{4n-P}{2}} \quad (3)$$

donde n es el número de pixel que contiene la región digital y P es el perímetro de la región. C_{Dmin} es la compacidad discreta mínima de una región digital con n pixel y está dada por:

$$C_{Dmin} = n - 1. \quad (4)$$

y C_{Dmax} es la compacidad discreta máxima para una región compuesta por n pixels y se expresa como

$$C_{Dmax} = \frac{4n - 4\sqrt{n}}{2}. \quad (5)$$

Otra propuesta que incluimos en este trabajo es la medida de circularidad de Haralick [20]. Esta medida determina la compacidad de una región utilizando el conjunto de distancias entre en centroide de la región y las celda en su borde. Una vez que es determinado el conjunto de distancias, se obtiene una valor de compacidad mediante la media de las distancias entre su desviación estándar. No obstante el mayor proceso de cómputo de la medida, la medida de circularidad de Haralick se muestra sensible a los cambios en la morfología de región. El primer paso para determinar la medida es calcular el centroide de la región digital. Shapiro define el centroide como la localización promedio de los elementos de una región digital [21], denotado como (\bar{r}, \bar{c}) , donde:

$$\bar{r} = \sum_{(r,c) \in R} r, \quad (6)$$

$$\bar{c} = \sum_{(r,c) \in R} c. \quad (7)$$

Donde r y c son las posiciones en renglón y columna de las celdas de la región digital. A pesar de que los renglones y columnas son números enteros, el centroíde frecuentemente resulta en una posición en el espacio continuo, por lo tanto, se ubicará en la celda más próxima a su valor de renglón y de columna.

A partir de este punto de la región digital se calculan distancias euclidianas hacia cada celda en el borde de la región. La medida de circularidad de Haralick queda entonces expresada como:

$$C = \frac{\mu_R}{\sigma_R}. \quad (8)$$

donde

$$\mu_R = \frac{1}{K} \sum_{k=0}^{K-1} \|(r_k, c_k) - (\bar{r}, \bar{c})\|, \quad (9)$$

y

$$\sigma_R = \left(\frac{1}{K} \sum_{k=0}^{K-1} (\|(r_k, c_k) - (\bar{r}, \bar{c})\|)^2 \right)^{1/2}. \quad (10)$$

La última medida de compacidad que se emplea en este trabajo es la propuesta por Peura [25]. Esta medida, al igual que la CDN, establece una relación entre dos regiones. No obstante, la relación compara la región digital bajo estudio contra un círculo en el espacio continuo, en vez de una región digital cuadrada como en el caso de la CDN. La relación de circularidad de Iivarién y Peura se expresa como una relación entre el perímetro de la región digital y un círculo con igual área. La medida de circularidad de Peura está dada por:

$$C = \frac{2\sqrt{A\pi}}{P_D} \quad (11)$$

Donde A es el número de celdas de la región digital y P_D es el perímetro de la región digital.

En la Tabla 1 se muestran dos regiones digitales compuestas por celdas cuadradas y los valores obtenidos por cada una de las medidas de compacidad probadas.

3. Descripción de la asimetría por compacidad

La asimetría de las neoplasias en piel es uno de los parámetros que contribuye de manera importante, bajo la regla ABCD, para el diagnóstico de melanoma maligno. El proceso general para evaluar la asimetría de una lesión en piel consiste en: determinar un par de ejes ortogonales que dividan la forma de la lesión en cuatro secciones y que tengan como origen el centro de masa; se evalúan las diferencias entre las regiones contenidas en cada cuadrante y se determina un grado de asimetría [3, 22].

Ng y Cheung utilizan un la distancia de simetría [6, 23] para establecer el grado de asimetría de la lesión. La distancia de simetría se genera a partir de medir la distancia entre un conjunto de pares de puntos entre la región a evaluar y una figura simétrica de referencia. La medida se aproxima a cero si la región guarda una similitud con la

Tabla 1. Valores obtenidos por las diversas propuestas para medir la compacidad de regiones digitales; (a) forma dispersa; (b) forma compacta.

Región digital	Clásica	Discreta Norm	Haralick	Peura
 <p>(a)</p>	40.01	0.9599	5.44	1.784
 <p>(b)</p>	20.49	0.9999	84.35	1.276

forma simétrica de referencia. Conforme se requiere una mejor evaluación de la asimetría de la región es necesario evaluar más puntos. No obstante para un conjunto relativamente pequeño, de alrededor de cientos de puntos, el proceso eleva significativamente el trabajo de cómputo. La solución propuesta por Ng y Cheung es limitar el número de puntos haciendo un muestreo a intervalos angulares regulares, con el fin de reducir el proceso de cálculo.

Por otra parte, Bono et al. [22] describen la asimetría de una lesión como el grado de similitud entre la lesión y el círculo. El grado de similitud se obtiene mediante (1). Propuestas más sofisticadas se han aplicado en la medición de la asimetría. Tal es el caso propuesto en [5] donde emplean un proceso topológico ubicando el centro de masa de la región y un conjunto de distancias aplicadas a la función de tamaño [24]

La propuesta en [5] trata las regiones digitales como regiones continuas e interpola los resultados. Sin embargo, las regiones digitales y los parámetros obtenidos dependen fuertemente de la resolución de la imagen digital. Esto es: a diferente resolución, diferentes valores.

Santiago et al. [16] utilizan el concepto de centroide, propuesto por Shapiro [21], como punto donde los ejes tengan su origen. A partir de las posiciones de renglón y columna se traza el par de ejes que dividen la región en dos pares. Un par es por el eje vertical y otro par por el eje horizontal

Se realiza una medición usando la CDN y se obtiene un vector 4 dimensional que sirve como la firma de asimetría de cada lesión.

El proceso se ilustra con las regiones de la Figura 3, donde son evaluadas mediante el proceso de asimetría propuesto por Santiago et al. Los valores de compacidad entre regiones deberán tener menor discrepancia si son simétricas y lo contrario para regiones asimétricas.











Región	(a)	(b)	(c)	(d)
				
1	(a)	(b)	(c)	(d)
				
2	(a)	(b)	(c)	(d)

Fig. 1. División de las regiones de la Tabla 1 en cuatro sub-regiones a partir de dos ejes que cruzan por el centro de masa de la región.

En el presente trabajo se realiza un proceso comparativo entre la medida de compacidad discreta normalizada de Bribiesca frente a otras cuatro propuestas de medidas de compacidad [17]. Ahora un vector de características se forma por cada medida de compacidad.

Tabla 2. Descripción de la asimetría de las lesiones en piel representadas en la Figura 3 utilizando diferentes medidas de compacidad. Las celdas resaltadas muestran las diferencias en valores de compacidad más notorias.

Región 1	1(a)	1(b)	1(c)	1(d)
Clásica	36.081818	36.623077	35.135135	32.100209
CDN	0.94725531	0.95079233	0.95175802	0.95821481
C Haralick	4.0552708	2.6715716	3.3698502	3.1490382
C Peura	1.694491	1.7071531	1.672114	1.5982658
Region 2	2(a)	2(b)	2(c)	2(d)
Clásica	22.760536	23.142857	23.29306	23.036649
CDN	0.99118411	0.99055613	0.99034196	0.99088374
C Haralick	4.190369	4.0694204	4.0059821	4.173777
C Peura	1.3458179	1.3570741	1.3614708	1.3539565

La Tabla 2 muestra en cada renglón un vector de características de compacidad para cada región de la Figura 3, obtenidos de cada una de las sub-regiones, ilustradas en la Figura 3. Se espera que una región simétrica obtenga valores entre sus subregiones (a) y (b) o (c) y (d) que produzcan diferencia mínimas entre ella, lo contrario sería para regiones con asimetría. En el ejemplo de la Figura 3 y con los datos de la Tabla 2 se observa que para la región 1 de la Figura 3 siempre existen notorias diferencias entre las subregiones. No obstante que todas las medidas utilizadas en esta propuesta muestran asimetría en la región 1, la medida *C* de Haralick produce un mejor vector de características para la evaluación de la asimetría; esto al marcar diferencias en los dos pares de regiones divididas por los ejes resultantes.





















<i>Benignas</i>				
 1	 2	 3	 5	 6
<i>B1</i>	<i>B2</i>	<i>B3</i>	<i>B4</i>	<i>B5</i>
 7	 8	 9	 10	 11
<i>B6</i>	<i>B7</i>	<i>B8</i>	<i>B9</i>	<i>B10</i>
<i>Malignas</i>				
 1	 2	 6	 7	 8
<i>M1</i>	<i>M2</i>	<i>M3</i>	<i>M4</i>	<i>M5</i>
 9	 10	 11	 12	 15
<i>M6</i>	<i>M7</i>	<i>M8</i>	<i>M9</i>	<i>M10</i>

Fig. 2. Regiones de entrenamiento (regiones B1-B5 y M1-M5) y de prueba (regiones B6-B10 y M6-M10) para el algoritmo de clasificación KNN.

4. Pruebas experimentales

Para probar la capacidad de evaluar la asimetría bajo el concepto de la compacidad, la descripción de asimetría propuesta se emplea en un proceso de clasificación automática. El proceso de clasificación utiliza el algoritmo estadístico, no-paramétrico y supervisado KNN, con K igual a uno. En nuestro experimento se utilizan las descripciones de las cinco primeras lesiones benignas y las cinco primeras lesiones malignas de la Figura 2. Con este conjunto se configura el conjunto de entrenamiento del clasificador, las cuales ya han sido diagnosticadas previamente. Posteriormente se someten las siguientes las restantes 10 regiones de la Figura 2 para obtener la una matriz de confusión que valora la eficacia de la propuesta.

El experimento se realiza utilizando el algoritmo KNN con una K=1, donde se someten vecotes de características 4-dimensional. Se realiza un proceso de clasificación por cada una de las medidas utilizadas. Las matrices de confusión

mostradas en las Tablas 3, 4, 5 y 6 contienen los resultados de clasificación para cada medida.

Tabla 3. Matriz de confusión para la clasificación de la muestra utilizando la medida clásica de compacidad.

Clásica	Benignas	Malignas	%
Benignas	3	2	60
Malignas	0	5	100

Tabla 4. Resultado de la clasificación por KNN empleando la medida C de circularidad de Peura.

Peura	Benignas	Malignas	%
Benignas	4	1	80
Malignas	2	3	60

Tabla 5. Eficacia de la medida de circularidad de Haralick para el conjunto muestra.

Benignas	2	3	40
Malignas	1	4	80

Tabla 6. Matriz de confusión resultado del proceso de clasificación aplicando la medida de compacidad discreta normalizada.

Discreta Norm.	Benignas	Malignas	%
Benignas	4	1	80
Malignas	1	4	80

5. Conclusiones

La propuesta de evaluación de la asimetría utilizando las medidas de compacidad diseñadas para el espacio digital tienen las características de ser de bajo gasto computacional y simples de implementar. Los resultados muestran que aplicando un proceso de división mediante dos ejes arbitrarios y evaluando las compacidades de cada región puede describirse la asimetría de una región.

El trabajo presenta una aplicación donde la evaluación de esta propiedad resulta ser indispensable, como es el caso de la detección temprana de melanomas. Además, se muestra lo robusto que resulta ser medida de compacidad discreta normalizada.

Aunque los resultados no pueden ser concluyentes para inferir que la descripción de la asimetría propuesta es un indicador para el diagnóstico automático del melanoma humano, si se establece una metodología que puede ser extendida a evaluar las otras propiedades geométricas de las lesiones en piel.

Referencias

1. A. Alfaro, L. Castrejón y M. Rodríguez Ortíz: Cáncer de piel. Estudio epidemiológico a 10 años de derechohabientes del ISSSTE en Nuevo León. *Dermatología Rev Mex*, 54(6), 321-325 (2010)
2. R.M. Gutierrez Vidrio y N. Cortés Lozano: Confrontando al melanoma en el siglo XXI. *Med Cutan Iber Lat Am*, 35(1), pp. 3-13 (2007)
3. I. Maglogiannis, C.N. Doukas: Overview of Advanced Computer Vision Systems for Skin Lesions. *Characterization, Information Technology in Biomedicine, Trans on.*, 13(5), pp. 721-733 (2009)
4. S.M. Rajpara, A.P. Botello, J. Townend y A.D. Ormerod: Systematic review of dermoscopy and digital dermoscopy/ artificial intelligence for the diagnosis of melanoma. *British Journal of Dermatology*, 161, pp. 591-604 (2009)
5. M. d'Amico, M. Ferri, and I. Stanganelli, I.: Qualitative asymmetry measure for melanoma detection. In: *ISBI2004: Proceedings of the IEEE International Symposium on Biomedical Images (Arlington, VA)* (2004)
6. V. Ng and D.W.L. Cheung: Measuring asymmetries of skin lesion. *Computational Cybernetics and Simulation, IEEE International Conference on Systems, Man, and Cybernetics Conference Proceedings, Orlando, Florida, USA, 12-15 October 1997*, v. 5, p. 4211-4216 (1997)
7. Real Academia Española, *Diccionario de la lengua española*, Ed. 23 (2010)
8. R.O. Duda and P.E. Hart: *Pattern Classification and Scene Analysis*. Wiley-Interscience, New York, USA (1973)
9. K.R. Castleman: *Digital Image Processing*. Prentice Hall (1995)
10. L.F. Costa and R.M. Cesar: *Shape Analysis and Classification: Theory and Practice*, CRC Press, Boca Raton, Florida, USA (2009)
11. R.C. Gonzalez and R.E. Woods: *Digital Image Processing*. Prentice Hall, Upper Saddle, New Jersey, USA (2007)
12. A.K. Jain, *Fundamentals of Digital Image Processing*. Prentice Hall, USA (1989)
13. A. Rosenfeld: Compact figures in digital pictures. *Systems, Man and Cybernetics, Trans on.*, 4, pp. 221-223 (1974)
14. T.K. Lee and Ela Claridge: Predictive power of irregular border shapes for malignant melanomas. *Skin Research and Technology*, 11, pp. 1-8 (2005)
15. M.J. Bottema: Circularity of objects in images. In: *International Conference on Acoustic, Speech and Signal Processing, ICASSP 2000, Istanbul*, pp 2247-2250 (2000)
16. Santiago Montero, R. and Ernesto Bribiesca: State of the art of compactness and circularity measures. *International of Mathematical Forum*, 4(27), pp. 1305-1335 (2008)
17. N. Ritter and James Cooper: New Resolution Independent Measures of Circularity. *J Math Imaging Vis*, 35, pp. 117-127 (2009)
18. E. Bribiesca: Measuring 2D Shape Compactness using the Contact Perimeter. *Computer and Mathematics with Applications*, 33(11), pp. 1-9 (1997)
19. U.D. Braumann et al.: Three-dimensional reconstruction and quantification of cervical carcinoma invasión fronts from histological serial section. *Trans on Medical Imaging*, 24(10), pp. 1286-1307 (2005)
20. R.M. Haralick: A measure for circularity of digital figures. *Systems, Man and Cybernetics, Trans on.*, SMC- 4(4), pp. 334-336 (1974)
21. L.G. Shapiro and G.C. Stockman: *Computer Visión*. Prentice Hall, Upper Saddle River, New Jersey, USA (2001)
22. A. Bono et al.: The ABCD System of Melanoma Detection, A spectrophotometric Analysis of the Asymmetry, Border, Color, and Dimension. *Cancer*, 85(1), pp. 72-77 (1998)
23. H. Zabrodsky, Peleg, S. and D. Avnir: Continuous symmetry measure. *J. Am. Chem. Soc.*, 114, pp. 7843-7851 (1992)

24. Frosini P.: Connections between Size Functions and Critical Points. *Math. Meth. Appl. Sci.* 19, pp. 555–569 (1996)
25. M. Peura and J. Iivarinen: Efficiency of simple shape descriptors. In: C. Arcelli, P. Cordella and G Sanniti de Baja, editors, *Advances in Visual Form Analysis*, pp. 443–451, World Scientific, Singapore (1997)