

Detección y Clasificación de Tejidos Anómalos en Mamografías Digitales Mediante Redes Neuronales Convolucionales

J. González-Bueno Puyal¹, B. Ríos Sánchez², C. Sánchez Ávila²

¹Universidad Politécnica de Madrid, España, juana.gonzalezbueno.puyal@alumnos.upm.es

²Grupo de Biometría, Bioseñales y Seguridad, Universidad Politécnica de Madrid, España, {brios,csa}@cedint.upm.es

Resumen

El cáncer de mama constituye un problema de salud global que supone más del 25% de los nuevos casos de cáncer en mujeres y en el que la detección precoz mediante la realización de mamografías juega un papel fundamental. Este trabajo presenta un sistema novedoso de detección y clasificación de anomalías en imágenes mamográficas mediante redes neuronales convolucionales (CNN). Se trata de un sistema ambicioso que permite distinguir entre cinco clases de mamografías: sin anomalías, con masas tumorales benignas, con masas tumorales malignas, con microcalcificaciones benignas o con microcalcificaciones malignas. Este trabajo evalúa no solo la precisión de las CNN aplicadas a este problema concreto, sino también la influencia de otros parámetros como la inclusión de una etapa de mejora de la calidad de la imagen o la resolución y cantidad de imágenes utilizadas para entrenar la red.

1. Introducción

El cáncer de mama es una de las enfermedades más comunes hoy en día entre mujeres, con 1.7 millones de casos nuevos al año en el mundo [1]. Dado que un diagnóstico precoz permite mejorar considerablemente el pronóstico de los pacientes, es cada vez más frecuente la realización periódica de mamografías para localizar posibles anomalías como masas tumorales o microcalcificaciones, pequeñas acumulaciones de calcio cuya presencia está relacionada con una alta actividad celular. Pese a no implicar necesariamente el desarrollo de la enfermedad, las microcalcificaciones son un buen marcador, ya que permiten su detección incluso cuando el tumor aún no es palpable.

Sin embargo, la correcta interpretación de estas imágenes es complicada, incluso para radiólogos experimentados, debido a la presencia de ruido, la similitud entre tejidos y la alta variabilidad y pequeño tamaño de las anomalías. Los sistemas de ayuda al diagnóstico asistido por ordenador (CAD) ayudan a mejorar la interpretación de las mamografías favoreciendo la reducción de costes y la mejora del diagnóstico. Estos sistemas están generalmente basados en técnicas de aprendizaje automático mediante las cuales es posible clasificar las imágenes según distintos criterios, como la naturaleza de la lesión presente o su grado de malignidad, y habitualmente se componen de los siguientes módulos: mejora de la calidad de la imagen, obtención de los parámetros de entrada al clasificador y clasificación.

La fase de mejora de la calidad de las imágenes tiene como finalidad el incremento del contraste y la eliminación del ruido de las mamografías. Para ello es frecuente utilizar métodos de realce de imágenes en el dominio espacial o de realce basado en características de la imagen. Los primeros se basan en la manipulación directa de los píxeles, siendo los más comunes los que se basan en una modificación del histograma como la ecualización [2]. Esta categoría incluye también los métodos de realce basado en regiones, que operan sobre cada píxel teniendo en cuenta el valor de sus vecinos. Estas técnicas son de gran utilidad para aumentar la nitidez de las mamografías y permiten detectar o realzar las propiedades de las masas tumorales [3]. Por otro lado, las técnicas basadas en características utilizan propiedades de la imagen generalmente extraídas mediante operaciones en el dominio de la frecuencia como la transformada Wavelet [4], permitiendo realzar tanto masas tumorales como microcalcificaciones.

A continuación, las imágenes son adaptadas a las necesidades del clasificador utilizado. Por un lado, si el clasificador utiliza ciertas características de la imagen, como factores de forma [5], medidas de energía de la textura de la imagen [6] o propiedades obtenidas mediante transformadas Wavelet [7][8], es necesario obtener dichas características. Es frecuente utilizar técnicas de segmentación para identificar y etiquetar regiones con características homogéneas previamente a la extracción de características [2][3]. Por otro lado, cuando el clasificador se alimenta directamente de imágenes, resulta recomendable introducir imágenes de un tamaño reducido con el objetivo de minimizar la carga computacional y, por tanto, el tiempo requerido. Para ello, las imágenes son subdivididas en imágenes mucho más pequeñas cuyo tamaño varía de un trabajo a otro (32x32 [9], 55x55 [10] o 64x64 [11] píxeles).

Finalmente, el módulo de clasificación es el encargado de decidir la clase a la que pertenece una imagen según el aprendizaje realizado en la etapa de entrenamiento. Entre los tipos de aprendizaje automático utilizados para la clasificación de mamografías podemos encontrar los árboles binarios [8], los árboles de regresión [12] o las máquinas de soporte vectorial [13]. Sin embargo, este tipo de técnicas presentan algunos inconvenientes en la clasificación de mamografías como la acumulación de errores o un elevado tiempo de entrenamiento. Numerosos trabajos optan por utilizar redes neuronales

artificiales (ANN) para clasificar mamografías [14] ya que presentan buenos resultados. No obstante, en sistemas alimentados directamente por imágenes, dichas redes son lentas debido al elevado número de entradas presente (tantas como píxeles en la imagen) y la alta interconectividad en su estructura interna. Ante este problema surgen las redes neuronales convolucionales (CNN), cuya estructura explota las propiedades locales de la imagen y reduce considerablemente el número de conexiones de la red, disminuyendo significativamente la carga computacional. Aunque su uso no está aún muy extendido, en los últimos años se han realizado prometedores estudios que aplican este tipo de redes a la clasificación de mamografías [9][11].

En este trabajo se presenta un sistema CAD de clasificación de mamografías mediante redes neuronales convolucionales que es capaz de distinguir entre cinco tipos distintos de mamografías en función de si la imagen presenta o no anomalías y en caso afirmativo su tipología (masas tumorales o microcalcificaciones) y su malignidad. Esto supone un gran reto frente a la mayoría de los sistemas actuales, que realizan una clasificación binaria de las imágenes, bien evaluando la tipología de las lesiones [6], bien identificando si se trata de una anomalía benigna o maligna [7]. Además, con el objetivo de evaluar la influencia de distintos parámetros, se han realizado diversas pruebas en las que la red ha sido entrenada utilizando distintas configuraciones en las que se varían el número de imágenes del conjunto de pruebas y su resolución y se han utilizado distintas técnicas para mejorar la calidad de las imágenes: ninguna, ecualización del histograma y transformada Wavelet.

2. Estructura del sistema

El sistema propuesto, siguiendo con la estructura habitual de este tipo de sistemas, consta de tres módulos: mejora de la calidad de la imagen, adaptación de las imágenes a los requisitos del clasificador y clasificación, tal y como se puede observar en la Figura 1.

2.1. Mejora de la calidad de la imagen

En este módulo se han implementado dos técnicas distintas con el fin de evaluar los resultados de cada una de ellas.

Por un lado se ha aplicado la ecualización del histograma

de las imágenes, un método sencillo pero muy efectivo que permite incrementar el contraste entre los distintos tejidos de la mama.

Por otro lado se ha realizado una descomposición mediante paquetes de wavelets que pretende minimizar la mayor cantidad de ruido posible sin alterar la información útil de la imagen, de forma similar a la realizada en [16]. En concreto, se aplica una descomposición espacial que traslada la imagen al dominio de la frecuencia donde se realiza una umbralización con el fin de eliminar las altas frecuencias, correspondientes a las componentes de ruido. A continuación se reconstruye la imagen a partir de las componentes de frecuencia filtradas.

2.2. Adaptación de las imágenes de entrada a los requisitos del clasificador

Dada la naturaleza del clasificador elegido, una CNN, que se alimenta directamente de imágenes, en este módulo las mamografías son divididas en sub-imágenes más pequeñas. Además, para que la red aprenda a diferenciar entre las distintas clases es necesario conocer la naturaleza de cada una de estas sub-imágenes en la etapa de entrenamiento.

Para ello, en primer lugar se aísla la región que contiene la mama y, en caso de contener algún tipo de anomalía, se separan el tejido sano y el anómalo. A continuación ambas regiones son divididas en sub-imágenes de 32x32 píxeles.

En este módulo también se ha implementado una segunda metodología de división inspirada en [11] con el fin de analizar cuál ofrece mejores resultados, y que difiere de la anterior en que una vez detectado el tejido anómalo, y de forma previa a la subdivisión, se hace un reescalado de la imagen de forma que el lado más pequeño de la región anómala ocupe 64 píxeles.

2.3. Clasificación

Este módulo consta de una red neuronal convolucional. Dicha red se compone de capas organizadas en dos etapas: la etapa de extracción de características y la etapa de clasificación [17].

La primera etapa está compuesta por una serie de capas convolucionales, encargadas de extraer parámetros locales de la imagen tales como bordes o texturas, además de

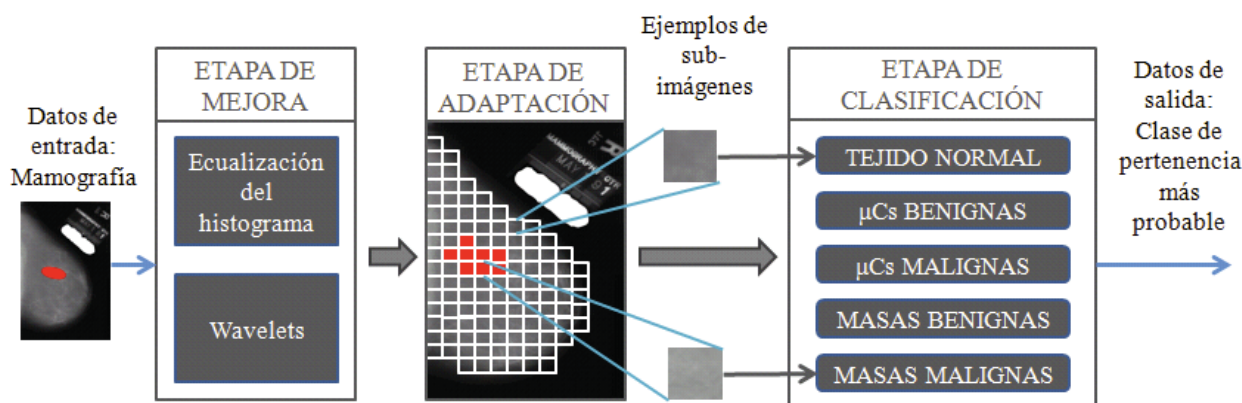


Figura 1. Esquema de la estructura del sistema, donde μCs es abreviatura de Microcalcificaciones

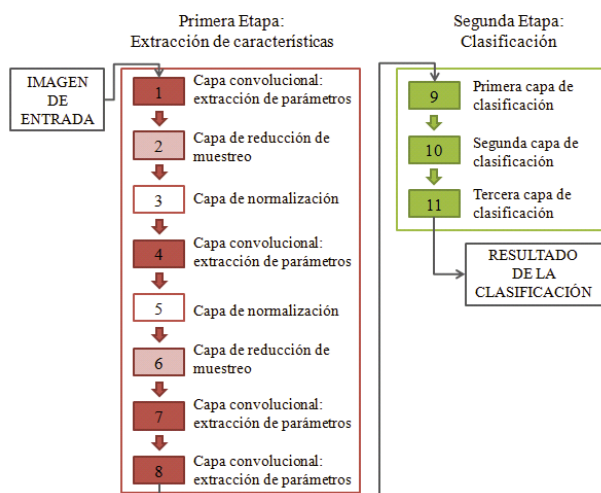


Figura 2. Configuración de las capas de la CNN.

capas de normalización y capas de reducción de muestreo (capas de fondo rojo, blanco y rosa respectivamente en la Figura 2).

La segunda etapa está compuesta por tres capas y clasifica la imagen de entrada en función de las características proporcionadas por la etapa anterior. La primera de dichas capas genera cinco salidas, a partir de las cuales la siguiente capa obtiene la probabilidad de la imagen analizada de pertenecer a cada una de las cinco clases posibles. La última capa genera el resultado de la clasificación, que corresponde a la clase más probable de pertenencia de la imagen analizada.

3. Pruebas

3.1. Base de Datos

Para entrenar y testear el sistema se ha utilizado la base de datos DDSM [15], en la que las anomalías han sido marcadas y etiquetadas por expertos. Dado el gran tamaño de esta base de datos se ha seleccionado un subconjunto de sus imágenes para realizar las pruebas. A la hora de diseñar este subconjunto se han tenido en cuenta diversos factores: variabilidad entre distintas mamografías, número de imágenes necesarias para el entrenamiento y que el número de imágenes en cada clase esté equilibrado para evitar la obtención de resultados sesgados.

Nombre de la prueba	Training set	Mejora	Adaptación
P. 1	1º	NO	1
P. 2	2º	NO	1
P. 3	1º	1	1
P. 4	2º	1	1
P. 5	1º	2	1
P. 6	2º	2	1
P. 7	1º	NO	2
P. 8	2º	NO	2
P. 9	1º	1	2
P. 10	2º	1	2
P. 11	1º	2	2
P. 12	2º	2	2

Tabla 1. Descripción de las pruebas realizadas

En concreto se han utilizado dos subconjuntos de imágenes: uno formado por 50 mamografías, 10 de cada tipo, denominado Primer training set y otro compuesto por 200 mamografías de cada tipo (Segundo training set).

3.2. Pruebas Realizadas

Con el objetivo de evaluar cada uno de los módulos se ha realizado un plan de pruebas con distintos entrenamientos de la red (ver Tabla 1) que cubre los siguientes aspectos:

- Comportamiento de la CNN funcionando de forma independiente al resto de los módulos.
- Influencia del subconjunto de imágenes seleccionado para el entrenamiento y la validación del sistema.
- Mejora aportada por los distintos métodos del módulo de mejora de la imagen. De aquí en adelante serán referidas como Mejora 1 la equalización del histograma y Mejora 2 la transformada Wavelet.
- Influencia del módulo de adaptación de imágenes a los requisitos del clasificador. De aquí en adelante se denominará Adaptación 1 al método sin reescalado y Adaptación 2 al método con reescalado.

3.3. Resultados

La Tabla 2 muestra los resultados obtenidos al validar el sistema con cada una de las pruebas realizadas.

Estos resultados muestran que al aumentar el número de mamografías utilizadas para el entrenamiento, la precisión global disminuye. Esto no concuerda con los resultados que cabría esperar al aumentar el training set y puede deberse a la variabilidad entre sujetos.

Los resultados muestran que al realizar la equalización del histograma la precisión mejora ligeramente en comparación con las pruebas en las que se ha omitido la etapa de mejora. La precisión aumenta de forma considerable en las clases de masas benignas y malignas, mientras que en las de microcalcificaciones disminuye o se mantiene estable. Se confirma por tanto que la equalización del histograma constituye una mejora en la clasificación de masas tumorales.

Prueba	Precisión (%)					Total
	CB	MB	CM	MM	Normal	
P. 1	77	56	71	72	82	72,1
P. 2	66	41	50	47	56	53,6
P. 3	62	73	73	74	90	74,8
P. 4	50	55	53	50	75	56,9
P. 5	65	50	66	61	47	56,6
P. 6	33	40	36	39	41	38,1
P. 7	25	27	60	67	56	45,7
P. 8	40	33	26	50	46	40,0
P. 9	60	50	83	75	58	62,9
P. 10	36	37	32	44	49	40,3
P. 11	40	33	60	67	88	57,1
P. 12	42	44	37	44	94	52,7

Tabla 2. Precisiones de cada clase y totales obtenidas para cada una de las pruebas realizadas (CB: Calcificaciones benignas, MB: Masas Benignas, CM: Calcificaciones Malignas, MM: Masas malignas)

Por otro lado, al realizar una comparación entre los resultados obtenidos al aplicar paquetes de wavelets para mejorar las imágenes y los obtenidos sin la etapa de mejora, se obtienen resultados contradictorios en función del tipo de adaptación utilizada. Mientras que con la Adaptación 1 se aprecia una disminución de la precisión, con la Adaptación 2 se obtiene un aumento considerable. Esto puede explicarse por el hecho de que al reescalar las imágenes en la Adaptación 2, la transformada Wavelet elimine menor cantidad de información útil de la mamografía.

Por último, comparando la influencia de las dos adaptaciones, se observa que la Adaptación 1 presenta mejores resultados. La Adaptación 2 conlleva la generación de una cantidad menor de sub-imágenes debido al cambio de tamaño aplicado, lo que implica que se entrena la red con un conjunto de imágenes menor, repercutiendo en las precisiones obtenidas.

4. Conclusiones

El presente trabajo muestra un sistema novedoso en el campo de la clasificación de mamografías puesto que, pese a los buenos resultados de las redes neuronales convolucionales en análisis de imágenes, apenas han sido utilizadas en este ámbito. Además, se trata de un sistema ambicioso, puesto que realiza una clasificación de las imágenes en 5 tipos distintos, dando con ello un paso más allá de la tradicional clasificación binaria. Dada la complejidad de la clasificación propuesta, los resultados obtenidos son muy prometedores, alcanzando una precisión del 75%.

Asimismo, se ha realizado un extenso estudio acerca de la influencia de diversos factores en la clasificación, tales como el número de imágenes del conjunto de pruebas y su resolución o las técnicas de mejora de la calidad de las imágenes.

De este modo, este trabajo abre nuevas vías de estudio hacia la obtención de un sistema CAD de clasificación de imágenes mamográficas preciso y útil que permita ayudar al facultativo en la toma de decisiones, y por tanto mejorar las estadísticas de prevención y detección del cáncer de mama.

Referencias

[1] Página web de la World Cancer Research Fund International. <http://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/breast-cancer-statistics> (Consultada: Agosto 2015).

[2] Bovis, K. y Singh, S., Detection of masses in mammograms using texture features. *Proceedings of the 15th International Conference on Pattern Recognition*, vol.2, 2000, pp.267-270.

[3] Dominguez, A.R. y Nandi, A.K. Detection of masses in mammograms via statistically based enhancement, multilevel-thresholding segmentation, and region selection, *Computerized Medical Imaging Graphics*, vol. 32, No 4, 2008, pp. 304-315.

[4] Dabour, W., Improved Wavelet Based Thresholding for Contrast Enhancement of Digital Mammograms, *Proc. of*

the 2008 International Conference on Computer Science and Software Engineering, vol 4, 2008, pp 948-951.

[5] Tralic D, Bozek J, Grgic S. Shape Analysis and Classification of Masses in Mammographic Images using Neural Networks. *18th International Conference on Systems, Signals and Image Processing (IWSSIP)*. Zagreb, 2011, pp 1-5.

[6] Dheebea J, Singh NA, y Selvi ST. Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach. *Journal of Biomedical Informatics*, vol 49, 2014, pp 45-52.

[7] Jasmine JSL, Govardhan A, Baskaran S. Microcalcification Detection in Digital Mammograms based on Wavelet Analysis and Neural Networks. *International Conference on Control, Automation, Communication and Energy Conservation*, Perundurai, 2009, pp 1-6 (ISBN:978-1-4244-4789-3).

[8] Sun Y, Babbs CF, Delp EJ. Normal mammogram classification based on regional analysis. *The 2002 45th Midwest Symposium on Circuits and Systems (MWSCAS-2002)*, West Lafayette, 2002, pp 375-8 (ISBN: 0-7803-7523-8).

[9] Rose DC, Arel I, Karnowski TP, Paquit VC. Applying Deep-Layered Clustering to Mammography Image Analytics. *Biomedical Science & Engineering Conference Annual (BSEC)*, Oak Ridge, 2010, pp 1-4 (ISBN: 978-1-4244-6713-6).

[10] Zhang W, Doi K, Giger ML, Wu Y, Nishikawa RM, Schmidt RA. Computerized detection of clustered microcalcifications in digital mammograms using a shift-invariant artificial neural network. *Medical Physics*, vol 21, No 4, 1994, pp 517-524.

[11] Agarwal V, Carson C. Using Deep Convolutional Neural Networks to Predict Semantic Features of Lesions in Mammograms. 2015.

[12] Vibha L, HarshaVardhan GM, K, Deepa SP et al. Classification of mammograms using decision trees. *10th International Database Engineering and Applications Symposium (IDEAS'06)*, Delhi, 2006, pp 263-6. (ISBN: 0-7695-2577-6).

[13] Cao A, Song Q, Yang X, Liu S. Mammographic mass detection by vicinal support vector machine. *IEEE International Joint Conference on Neural Networks*, vol 3, Singapore, 2004, pp 1953-1958 (ISBN: 0-7803-8359-1).

[14] Ayer T, Chen Q, Burnside ES. Artificial neural networks in mammography Interpretation and decision making. *Computational and Mathematical Methods in Medicine*. Georgia, 2013, pp 11.

[15] Heath M, Bowyer K, Kopans D, Moore R, Kegelmeyer P. The Digital Database for Screening Mammography. *Proceedings of the Fifth International Workshop on Digital Mammography*, 2000, pp 212-218 (ISBN 1-930524-00-5).

[16] Ausin, M. Detección automatizada de microcalcificaciones en mamografía digitalizada utilizando paquetes de wavelets. *Proyecto de Fin de Carrera*, Madrid.

[17] Página web del tutorial "UFLDL Tutorial" de la Universidad de Stanford. <http://ufldl.stanford.edu/tutorial/>. (Consultada: Septiembre 2015).