

# Redes neuronales convolucionales: un modelo de *Deep Learning* en imágenes diagnósticas. Revisión de tema

Convolutional neural networks, a model for deep learning in diagnostic imaging. A topic review

Federico Lubinus Badillo<sup>1</sup>  
César Andrés Rueda Hernández<sup>2</sup>

Boris Marconi Narváez<sup>3</sup>

Yhary Estefanía Arias Trillos<sup>4</sup>

DOI: <https://doi.org/10.53903/01212095.161>



## Palabras clave (DeCS)

Inteligencia artificial  
Aprendizaje profundo  
Radiología

## Key words (MeSH)

Artificial intelligence  
Deep learning  
Radiology

## Resumen

Los avances en la inteligencia artificial han repercutido en varios espacios de la vida cotidiana, así como en la medicina. En vista de la rápida aplicación del aprendizaje profundo —conocido como *Deep Learning*— en los datos biomédicos, las imágenes radiológicas han comenzado a adoptar esta técnica. En lo que respecta, se espera que el aprendizaje profundo tenga un efecto en el proceso de adquisición e interpretación de imágenes, así como en la toma de decisiones. Esta revisión ofrece en primer lugar una descripción general del funcionamiento de las redes neuronales convolucionales, los conceptos básicos de estas, y las perspectivas actuales en la aplicación médica centrada en imágenes diagnósticas.

## Summary

Advances in artificial intelligence have impacted several areas of everyday life, as well as in the area of medicine. Due to the rapid application of deep learning in biomedical data, radiological and nuclear imaging has begun to adopt this technique. Deep learning is expected to have an effect on the process of image acquisition and interpretation, as well as on decision making. This review first provides an overview of the basic concepts and operation of convolutional neural networks, as well as current insights into the medical application focused on diagnostic imaging.

**Abreviaciones:** inteligencia artificial (IA); *Deep Learning* (DL); redes neuronales convolucionales (CNN); competencia de reconocimiento visual a gran escala de ImageNet (ILSVRC); tomografía axial computarizada (TAC); capa de reducción (*max pooling*); capas densas (*fully connected*).

## Introducción

El concepto de inteligencia artificial (IA) se ha venido desarrollando desde la década de los cincuenta, aunque las limitaciones tecnológicas en los primeros años significaron un bajo rendimiento comparado con los seres humanos. Actualmente, con la rápida progresión del diseño de algoritmos, el crecimiento de conjuntos de datos digitales y el desarrollo de una capacidad computacional, la IA tiene la posibilidad de superar a los humanos en muchas tareas. En consecuencia, su exploración potencial en los últimos diez años se ha incrementado en el campo médico, en especial en el diagnóstico por imágenes.

El *aprendizaje profundo* conocido como el *Deep Learning* (DL), es considerado por muchos como una parte integral de la cuarta revolución (1). Investiga el uso de redes neuronales artificiales con un algoritmo inspirado en la estructura y función del cerebro humano, mediante el reconocimiento o categorización de imágenes jerarquizadas de los datos distribuidos en múltiples capas compuestas de módulos simples

y no lineales, realizando así la transformación de los datos para su discriminación. Las redes neuronales convolucionales (CNN, del inglés *Convolutional Neural Networks*) son el modelo de DL de mayor importancia en la actualidad (2).

Las CNN tienen sus orígenes en el neocognitrón propuesto por Fukushima et al. (3), cuya idea se basaba en la biología del reconocimiento de la corteza visual primaria del ser humano investigada por Hubel y Wiesel (4). Estas CNN están constituidas por múltiples neuronas artificiales (figura 1). El incremento en la popularidad de las CNN fue secundario a la victoria obtenida en la competencia de reconocimiento visual a gran escala de ImageNet (ILSVRC) en 2012 (5). En este evento, Krizhevsky y Hinton (5) desarrollaron una CNN llamada AlexNet, que superó otras técnicas de aprendizaje automático de la competencia. Hoy en día se considera esa CNN como una de las principales influenciadoras en el análisis de imágenes (6,7). Desde entonces, se han desarrollado muchos modelos de reconocimiento y clasificación de imágenes basados en este modelo de DL.

<sup>1</sup>Radiólogo, Director Académico del programa de Radiología e Imágenes Diagnósticas UNAB. *Staff* de Radiólogos Especializados Clínica FOSCAL. Bucaramanga, Colombia.

<sup>2</sup>Médico general, Universidad Autónoma de Bucaramanga. Bucaramanga, Colombia.

<sup>3</sup>Residente de Radiología e Imágenes Diagnósticas, Universidad Autónoma de Bucaramanga. Bucaramanga, Colombia.

<sup>4</sup>Ingeniera de Sistemas, Universidad Autónoma de Bucaramanga. Bucaramanga, Colombia.

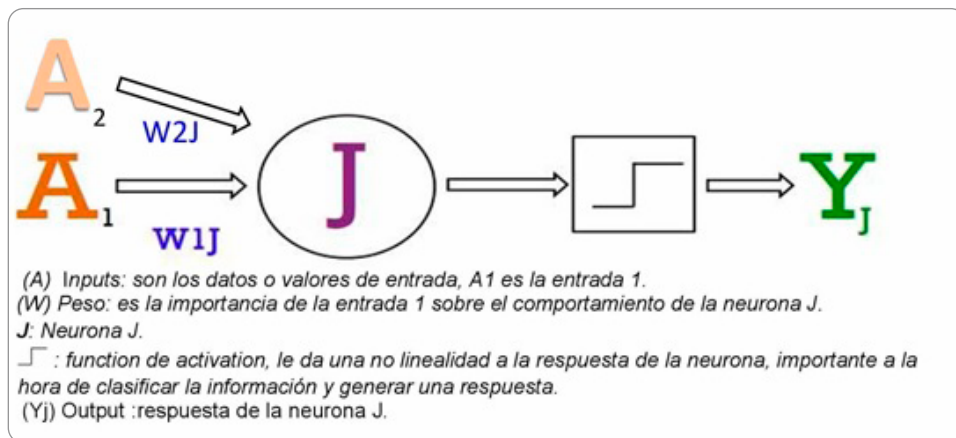


Figura 1. Esquema general de una neurona artificial.

Fuente: elaboración propia.

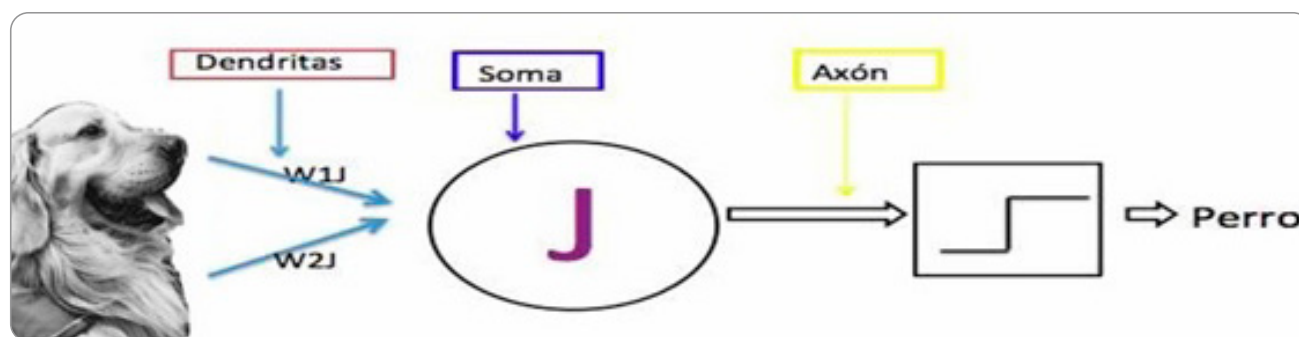


Figura 2. Neurona artificial. La base de las redes neuronales artificiales está inspirada en la biología de las neuronas y sus conexiones. Se observa lo que se podría llamar “neurona artificial” y cómo esta trata de seguir la misma arquitectura de una neurona biológica.

Fuente: elaboración propia.

Inicialmente, el desarrollo del DL fue lento debido a que los datos médicos no se encontraban bien estructurados ni etiquetados; sin embargo, el número de informes en la aplicación de este modelo en datos clínicos ha aumentado rápidamente en los últimos años (8). Su aplicación actual incluye la segmentación y clasificación de procesos patológicos. En múltiples campos se han desarrollado estudios de investigación médica mediante CNN, como en la detección de retinopatía diabética (9), para la clasificación de lesiones de piel (10) o para la detección de metástasis en nódulos linfáticos (11). En radiología se han publicado infinidad de estudios con aplicaciones de las CNN (6,12) que van desde la detección de nódulos pulmonares mediante radiografías de tórax (13) o de la tomografía axial computarizada (TAC) (14), clasificación de nódulos pulmonares (15), detección y clasificación de masas en mamografía (16,17) y la aplicación en cardiología e imágenes cardíacas (18).

El objetivo de este artículo es aclarar el orden de los conceptos básicos de las CNN, de tal forma que permita una aproximación al tema y pueda ser una guía para los que deseen profundizar al respecto. Se realizó una revisión de la bibliografía, aportando las definiciones más relevantes y la aplicación de la IA en la actividad del médico radiólogo y nuclear para en un futuro cercano apoyar las decisiones clínicas.

## 1. Función de una red neuronal convolucional

Las CNN aplicadas a las imágenes diagnósticas buscan actuar de manera muy similar a la corteza visual primaria del cerebro humano. Los

seres humanos clasifican y diferencian perfectamente entre múltiples objetos, dada la capacidad de distinguir diversas características, como el color, los bordes, las curvas, las sombras; es decir, se enfocan en todo aquello que les permite distinguir o clasificar cada objeto de manera individual. Por esa razón, muchos de estos sistemas intentan imitar el funcionamiento de las neuronas de la corteza visual (4,19) (figura 2).

Las CNN funcionan tanto en imágenes 2D como en 3D y están típicamente constituidas por tres capas: capa de convolución, capa de reducción (*max pooling*) y capas densas (*fully connected*) (figura 3), de las cuales, las dos primeras llevan a cabo una fase de extracción de características, mientras que las últimas (*fully connected*) son las encargadas de realizar la fase de clasificación (20).

Las primeras capas le permiten a la CNN ejecutar una fase de extracción de características, que actúa de manera similar al cerebro humano cuando busca ese tipo de características que definen un objeto. En la fase de clasificación o la fase de regresión que se realiza en las capas densas se relacionan todas las características ya extraídas y se obtiene directamente una clasificación (20). Por ejemplo, se crea una CNN y se entrena para detectar perros. La primera fase (convolución y *max pooling*) extrae las principales características de un perro —el pelaje, el color, la forma que tiene el perro—. La segunda fase (*fully connected*) utiliza todas las características extraídas y puede decir si lo que hay en la imagen es un perro o no. Las CNN establecen relaciones entre los diferentes píxeles de una imagen, buscando relaciones en ella, y eso les permite tener un contexto mucho más general con coherencia espacial.

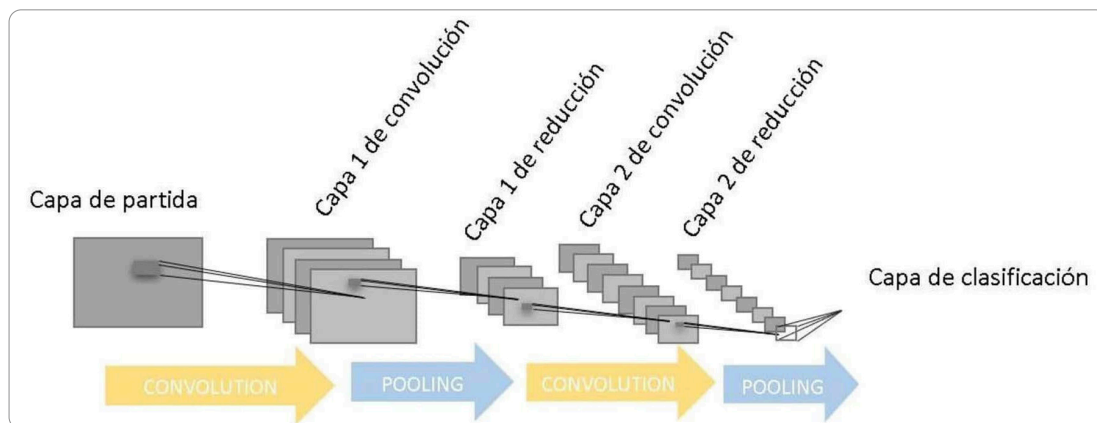


Figura 3. Arquitectura de una red neuronal convolucional. Capa de partida (imagen de entrada, o *input*).

Fuente: elaboración propia.

## 2. Conceptos claves de las redes neuronales

### 2.1. Píxeles y redes neuronales

**Pixel:** Del inglés *picture element*, “elemento de imagen”, es un único punto o un pequeño cuadro que contiene color, puede ser blanco, negro o matices de gris en el caso de las imágenes diagnósticas, y forma parte de una imagen digital (figura 4) (20).

**Kernel:** Conocido como *feature detector*, *filter* o “detector de propiedades”, es una matriz de números que se utiliza para enfocar, desenfocar, grabar, detectar bordes y características de la imagen (20).

**Preprocesamiento:** Una imagen digital es una matriz almacenada en el computador donde, para el caso de imagen en escala de grises (imágenes radiológicas), cada uno de los píxeles está representado con un valor que va de 0 (tonalidades oscuras) hasta 255 (tonalidades claras o brillantes). La normalización de los datos hace referencia a la división del valor del píxel en 255. Esta transformación asigna valores de píxeles entre 0 y 1 (21), lo que permite disminuir la carga computacional y generar una mejor red neuronal con mayor agilidad en el aprendizaje, que se verá reflejado de manera positiva en las predicciones de las imágenes (figura 5).

## 3. Funcionamiento de la CNN

### 3.1. Capas de convolución

Lo que hace la convolución es que a través de un filtro (*kernel*) que se aplica a la imagen permite extraer ciertas características. Entonces, si se tiene para una imagen un filtro que permite detectar bordes, lo que se verá en la convolución es una imagen como la figura 6. La idea general de la convolución es aplicar un patrón a una imagen, para extraer ciertas características o patrones dentro de esa imagen (7).

Para entender la operación matemática (figura 7) se dibujó una imagen sencilla de  $6 \times 6$  píxeles, del lado derecho se ve un *kernel* o filtro que tiene una orientación vertical con un tamaño de  $3 \times 3$  (figura

7). Generalmente, el filtro tiene un tamaño menor que la imagen. La idea de la convolución es que ese filtro se vaya desplazando por la imagen y a medida que lo hace realiza una serie de operaciones, generando de esta manera un mapa con las características de mayor peso o importancia.

Es un proceso iterativo, en la primera iteración lo que hace la convolución es multiplicar punto a punto los coeficientes del filtro por la porción de la imagen que está debajo de ese filtro (píxeles correspondientes de la porción de la imagen) y después se suman los valores para generar el píxel correspondiente en la imagen de salida. Al lado derecho se ve la imagen de salida y este primer píxel que resulta de multiplicar los coeficientes del filtro por los valores correspondientes de la imagen. En la segunda iteración se realiza exactamente la misma operación con la única diferencia de que el filtro desplaza una posición a la derecha y acá se obtiene el siguiente píxel de salida, y así, sucesivamente. Cuando llega al extremo derecho, el *kernel* se desplaza una posición hacia abajo y vuelve al lado izquierdo, y repite las mismas operaciones hasta que el *kernel* hace un barrido de toda la imagen, obteniendo así la imagen resultante para posteriormente pasar a la capa de reducción.

### 3.2. Capa max pooling o capa de reducción

La capa de reducción es la segunda de la CNN y es la encargada de reducir el flujo de datos obtenidos de las capas de convolución, extrayendo la información más importante o las características más importantes de la imagen convolucionada (22) (figura 8); dicha capa reduce la carga computacional que va a tener la CNN, ayudando a que la red entrene y segmente mejor cuáles son las características más importantes, para quitar el ruido de la imagen.

En el caso de que se exceda el *max pooling* no solo se quitará ruido, sino también podría eliminar características que son importantes, por lo que no se debe abusar de esta capa. En últimas, el fin de estas capas es reducir el tamaño de la imagen conservando las principales características (mayores activaciones) de la imagen convolucionada.

La imagen resultante después de la convolución y la reducción va a ser más pequeña, con sus principales características, una imagen más limpia en el contexto computacional comparada con la imagen de entrada. En este caso, se partió con una imagen de  $4 \times 4$  y se obtuvo una imagen de  $2 \times 2$  del lado derecho (22,23).

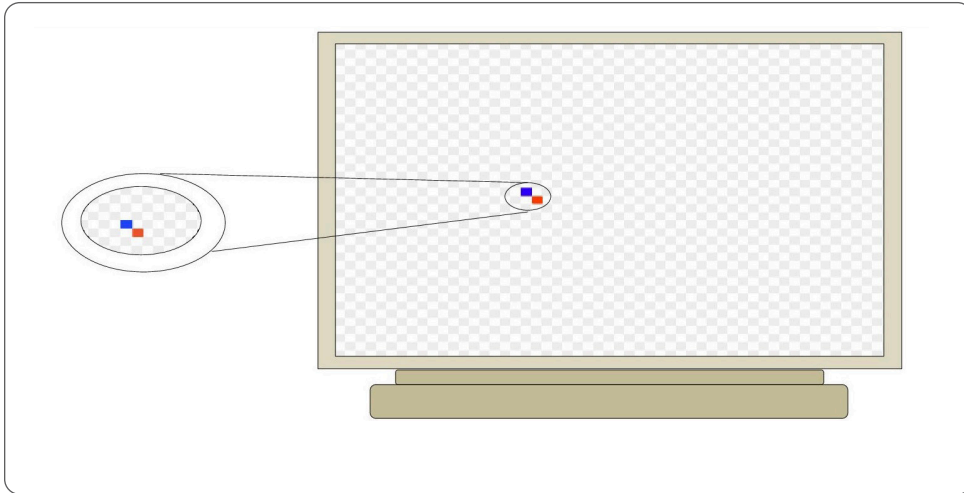


Figura 4. Esquema de un pixel.  
Fuente: elaboración propia.

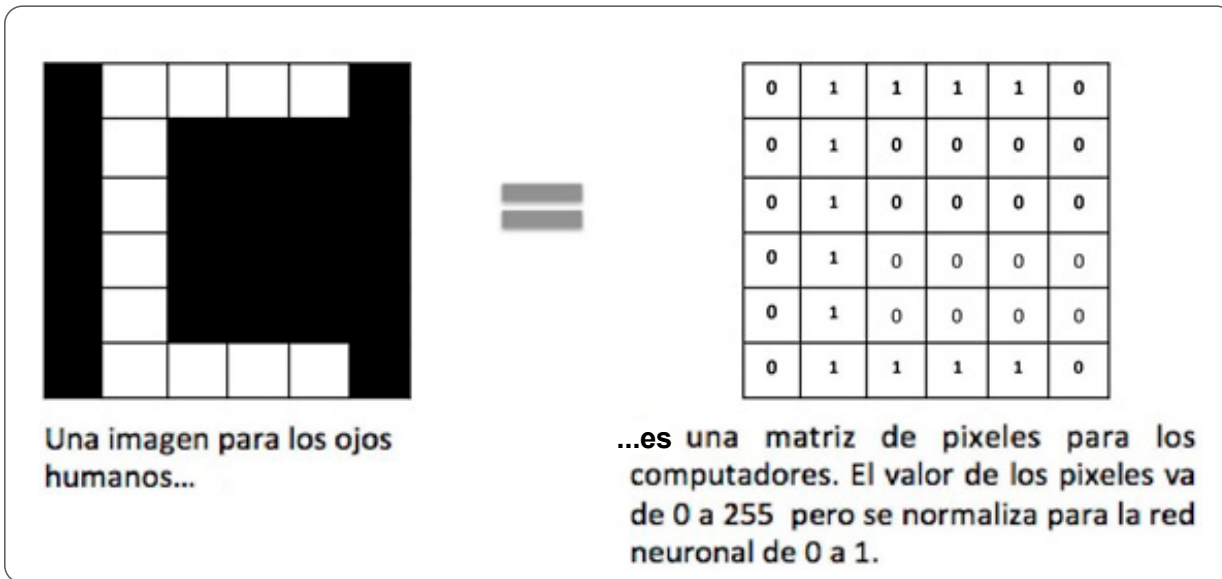


Figura 5. Preprocesamiento.  
Fuente: elaboración propia.



Figura 6. Convolución  
Fuente: elaboración propia.

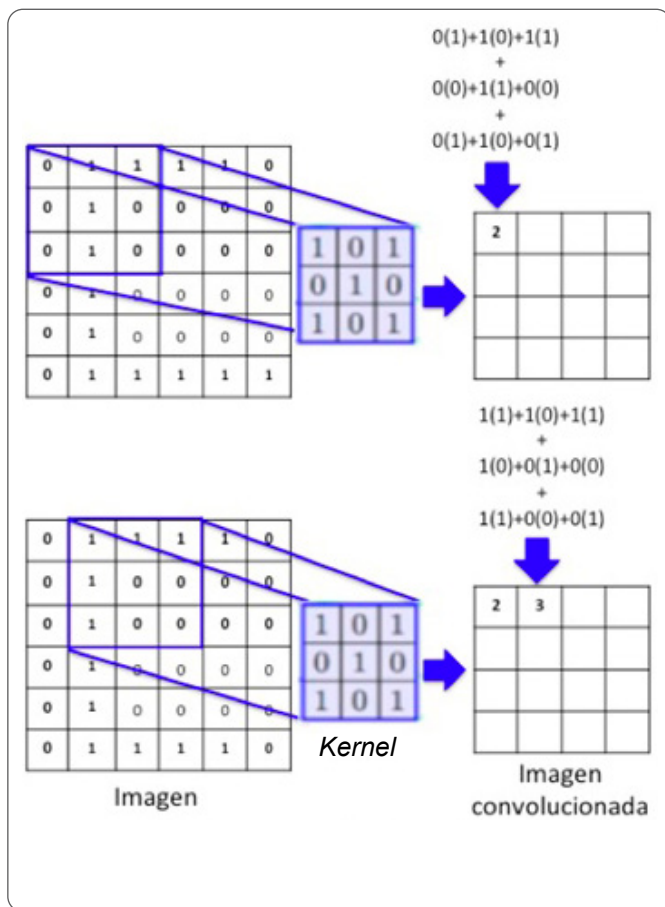


Figura 7. Función del *kernel*: la imagen resultante será más pequeña con características que priman respecto a la imagen de entrada. En este caso, se partió con una imagen de 6x6 y se obtuvo una imagen de 4x4 del lado derecho; se repite la misma operación hasta que el *kernel* haga una barrido de toda la imagen.  
Fuente: elaboración propia.

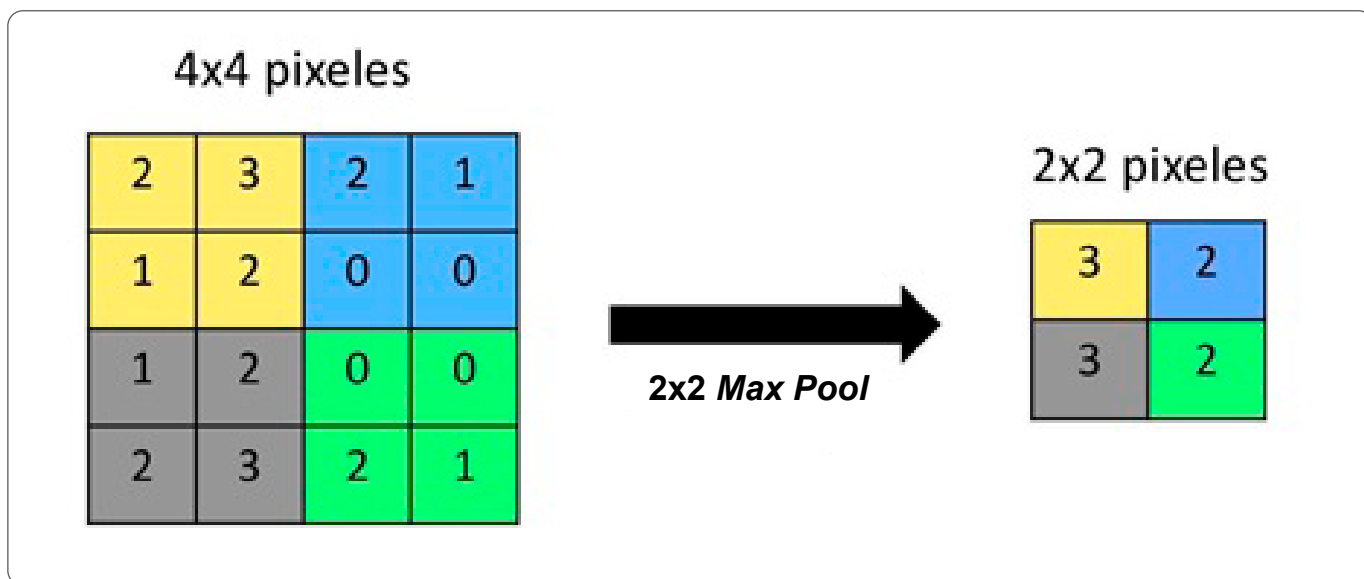


Figura 8. Imagen convolucionada del ejemplo anterior con un tamaño de *pool* o filtro. Su tamaño puede variar en una u otra red neuronal. En este ejemplo se impone un tamaño de 2x2, lo que quiere decir que se recorre cada uno de los 4x4 pixeles de la imagen convolucionada de izquierda-derecha, arriba-abajo, pero de 2x2 (2 de alto por 2 de ancho = 4 pixeles) y se irá preservando en la salida el valor más alto de entre esos 4 pixeles, por eso el término "Max". En este caso, la imagen resultante se reduce a la mitad para obtener una imagen de 2x2 pixeles. De esta manera, también se disminuye la cantidad de neuronas necesarias para continuar el almacenamiento de la información más importante para detectar las características deseadas.  
Fuente: elaboración propia.

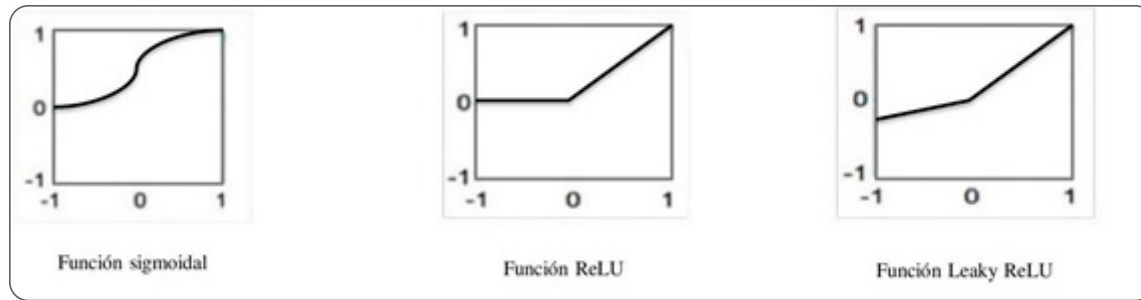


Figura 9. Representación de la función de activación.  
Fuente: elaboración propia.

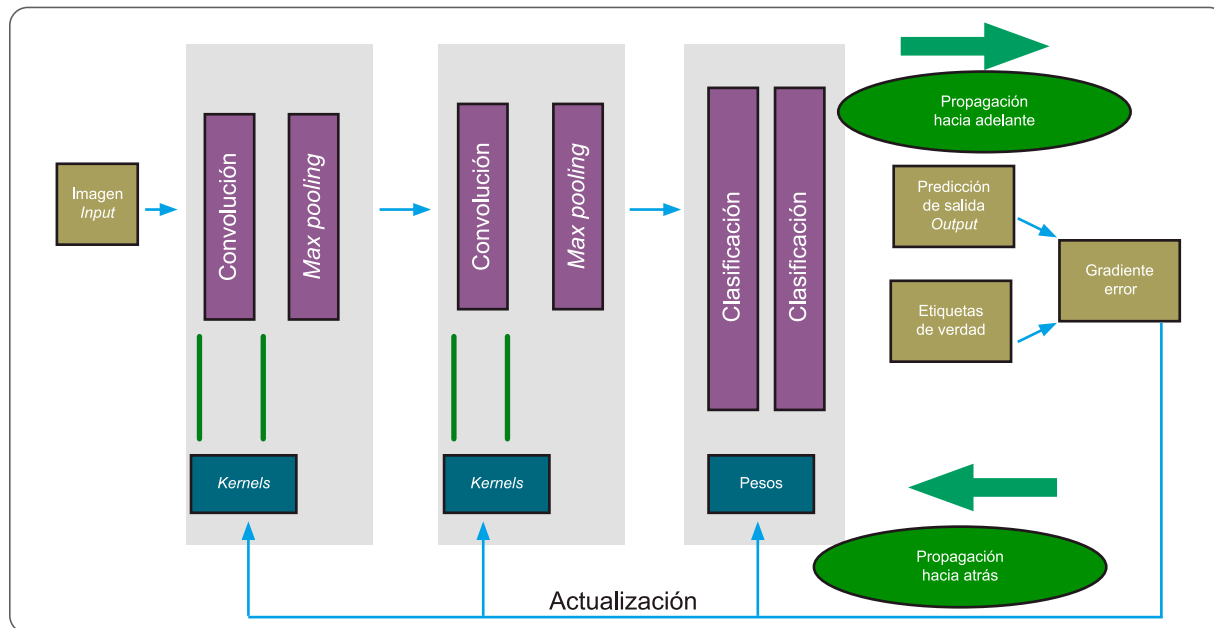


Figura 10. Esquema de entrenamiento de una red neuronal convolucional.  
Fuente: elaboración propia.

### 3.3. Capas densas o capas de clasificación (red neuronal multicapa o fully connected):

Estas capas calculan la puntuación de cada clase de características extraídas de las capas de convolución y del *max pooling*, relacionando todas las características o todos los datos que se han extraído previamente al entrenamiento de la red neuronal para obtener una clasificación final (24). No existen reglas en cuanto al número de capas de neuronas tipo *fully connected* que se deben utilizar en este bloque final de la CNN; sin embargo, en la literatura se describe el uso de 2 a 4 capas; tal es el caso de LeNet (25), VGG Net (26) y AlexNet que es la más influyente (5).

Cada neurona de la capa *fully connected* contiene una función no lineal y la elección de esta dependerá del tipo de tarea que va a realizar la CNN. Sin embargo, debido a que las capas de clasificación son tan pesadas computacionalmente, durante los últimos años se han propuesto otros enfoques. Estos incluyen la capa de agrupación promedio global (*the global average pooling layer*) y la capa de agrupación promedio (*the average pooling layer*), que ayudan a reducir significativamente la carga computacional (27).

#### 3.3.1. Función de activación

Es importante añadir que las funciones de activación se encuentran presentes en las capas de convolución y en las capas *fully connected*. Es decir, estas funciones de activación hacen parte de la estructura de cada neurona artificial que conforma las capas mencionadas, y son estructuralmente muy semejantes a la neurona artificial esquematizada en la figura 1; gracias a esta función de activación, la neurona —como unidad funcional— clasifica los datos recopilados para obtener una respuesta final (27).

Existen varios tipos de funciones de activación. Por un lado, se encuentra la función sigmoideal empleada en la clasificación binaria, como se observa en la figura 9 varía entre 0 y 1, donde 0 es una respuesta negativa y 1 es una respuesta positiva, dada en términos de probabilidad (28). El funcionamiento sigmoideal se entiende de la siguiente manera: se tiene un modelo CNN para identificar perros, al ingresar la imagen de un perro a la red neuronal esta arroja un resultado de 0,9, muy cercana a 1, por lo tanto, tiene una probabilidad de que la imagen ingresada sea un perro; caso contrario, si ingresa la imagen de un gato, la red

neuronal arroja un resultado de 0,2, lo cual está cerca al cero, lo que significaría que no es un perro.

En el caso de que se tenga una clasificación multiclase, donde hay más de una posible salida, se creó la función *soft max*, la cual permite obtener una probabilidad normalizada y es una extensión de la función sigmoideal. Retomando el ejemplo anterior, si se desea obtener la raza del perro, se entrena un modelo de red neuronal para diferenciar cuatro posibles clases de perros (labrador, terrier, bulldog y pincher); lo que arroja la red neuronal será una probabilidad por cada una de esas razas de perros y la suma de todas esas posibilidades será 1 (29). Por consiguiente, la imagen de un terrier en una red neuronal ya entrenada podría dar como respuesta: terrier 0,9, labrador 0,05, bulldog 0,03, pincher 0,02, la sumatoria de la respuesta es igual a 1. Si la red neural arroja el valor más alto, en este caso 0,9, existe una probabilidad que el perro sea de raza terrier.

Por otro lado, se ha demostrado la alta eficiencia y eficacia que tiene la función lineal rectificadora o función ReLU (por sus siglas en inglés de *Rectifier Linear Unit*), para desarrollar redes mucho más profundas (figura 9). Esta es capaz de eliminar el desvanecimiento de gradiente (el cual se refiere al error de aprendizaje en el cual la red deja de aprender) (5).

Existen varias versiones mejoradas de ReLU con alto desempeño para hacer una CNN. La más conocida es la *leaky ReLU*, que añade una pequeña pendiente en esa zona negativa, para impedir que se lleven todas las activaciones negativas a cero y permitir que la red siga aprendiendo, en menor medida, pero sin pausar el entrenamiento; de esta manera, la CNN diariamente podrá aprender y se hará más poderosa con el tiempo. Estas son las dos funciones de activación más usadas actualmente en las CNN (30).

## 4. Entrenamiento de una CNN

El entrenamiento de una red es un proceso de búsqueda de *kernel* en capas de convolución y pesos en las capas de clasificación que minimizan las diferencias entre las predicciones de salida, las cuales son las respuestas de la CNN que se está entrenando, y las etiquetas de verdad (*truth labels*), que hacen referencia a las respuestas correctas con las que se está entrenando la CNN. En el caso de radiología, corresponde a los diagnósticos correctos de las imágenes con las que se entrena la CNN. El aprendizaje de una CNN es un proceso iterativo de ir y venir de las capas de las neuronas. La *propagación hacia adelante* se denomina *forward propagation*, y hace alusión a los procesos matemáticos que se realizan en la CNN al paso de información por cada capa de la red neuronal para generar una respuesta o predicción de salida; posteriormente esta se compara con la etiqueta de verdad y a través de procesos matemáticos genera lo que se conoce como *gradiente de error*, el cual se podría decir que corresponde matemáticamente al ajuste que se hace en los *kernel* de las capas de convolución y pesos de las capas de clasificación. A este proceso retrógrado y de retroalimentación que hace ajustes en cada capa de la red neuronal se le conoce como *propagación hacia atrás* (*back propagation*). De esta manera, si la respuesta de la red neuronal es equivocada, se hará la respectiva corrección en la CNN (29) (figura 10).

Las *forward propagation* y *back propagation* se llevan a cabo con todos los datos, mediante sucesivas iteraciones a través de la base de datos, de esta manera, si se tuvieran 10.000 datos, se tendrían que realizar múltiples iteraciones para cubrir todos esos datos de entrenamiento.

Una vez se realizan las modificaciones correspondientes y se han generado las iteraciones suficientes para cubrir todas las bases de datos, culminará una época (*epochs*) (20). Normalmente, estos entrenamientos se realizan durante varias épocas, quiere decir que no ven solo una vez los datos de la base de datos, sino que los ven las veces que sea necesario.

## 5. Aplicaciones del DL en radiología

### 5.1. Detección de lesiones

Diversos estudios han evaluado el uso de las CNN en la identificación de lesiones en la radiografía (RX) de tórax con resultados muy significativos. Es así como, para detección de tuberculosis el área bajo la curva (ABC) que se encontró fue de 0,99 (31); en derrame pleural, 0,96; en edema pulmonar, 0,87; para consolidaciones, 0,85; cardiomegalia, 0,88 y neumotórax, 0,86, (32,33). También se han descrito modelos de CNN basados en ecografías mamarias para detección temprana de cáncer de seno con ABC entre 0,79-0,87 (34,35).

Se han propuesto modelos de CNN en TAC para la detección de hallazgos críticos en TAC de cráneo sin medio de contraste. Un estudio realizado demostró un modelo para la detección de hemorragia, efecto de masa o hidrocefalia obteniendo una sensibilidad del 90 %, especificidad del 85 % y un ABC de 0,91 (36). Por otro lado, Nakao et al., desarrollaron un método para detección de aneurismas en RM con una sensibilidad del 70 % (37).

### 5.2. Segmentación de lesiones

La segmentación precisa es la clave de una planeación efectiva de la radioterapia en el cáncer de cabeza y cuello. Ibragimov y Xing desarrollaron un modelo de CNN para una segmentación rápida y consistente de estas estructuras (38). También es el caso de Men et al., quienes propusieron una segmentación automática usando DL en la planeación de radioterapia del cáncer rectal, demostrando alta eficacia en cuanto a precisión y rapidez (39).

### 5.3. Pronóstico de pacientes

Los avances en la computación, inteligencia artificial y especialmente en el análisis de imágenes médicas han permitido extraer datos cuantitativos que permiten un mejor enfoque oncológico. Este concepto fue descrito en 2012 como *radiomics* (40,41), y hace referencia a la extracción de grandes cantidades de características cuantitativas de las diferentes imágenes médicas correlacionándolas entre sí para generar un diagnóstico, pronóstico y tratamiento óptimo del cáncer (42). Es el caso de un modelo de DL basado en *radiomics* de RM para predicción de supervivencia en el glioblastoma multiforme (43) y de una CNN basada en TAC de tórax para predecir mortalidad en paciente con enfermedad pulmonar obstructiva crónica en fumadores crónicos (44).

## Conclusión

Las decisiones médicas se toman con base en la interpretación integral de los datos relevantes del paciente, como los signos, síntomas, pruebas de laboratorio e imágenes diagnósticas. El DL permite la extracción automática de las características discriminativas de datos de alta dimensión, por lo tanto, supone un gran impacto en el campo médico y en particular en el análisis cuantitativo (figura 11).

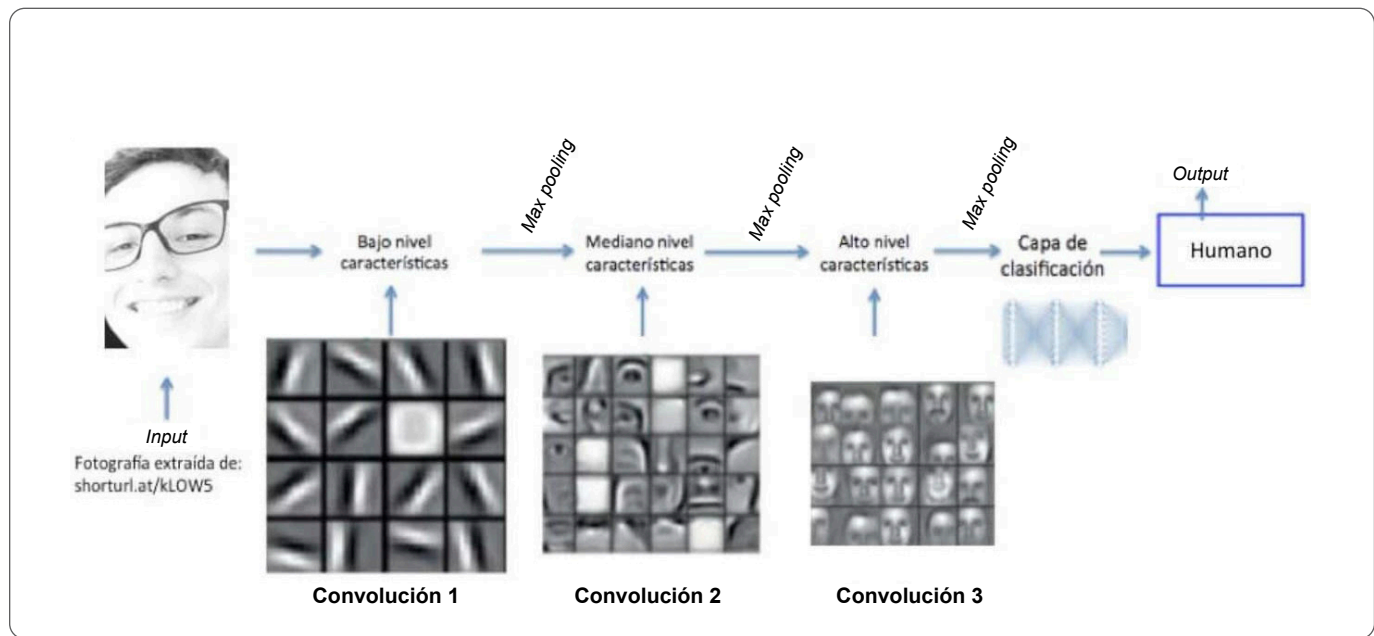


Figura 11. Esquema general y resumido del funcionamiento de un modelo CNN con tres capas convolucionales, tres capas *max pooling* y tres capas de clasificación para identificar rostros. Convolución 1: Al pasar por la primera capa convolucional, la red busca relaciones simples en la imagen, como líneas o sombras; son filtros muy simples. Convolución 2: Una vez pasa a la segunda capa convolucional, se tiene a la entrada una imagen ya filtrada por los filtros de la primera capa convolucional y reducida, debido a que pasó por la capa de *max pooling*, conservando sus características; por lo tanto, lo que produce es la abstracción en el sentido de que el filtrado de la segunda capa será sobre la salida del primer filtrado, buscando relaciones más complejas y concretas; en esta segunda capa, por ejemplo, se buscan curvas típicas del rostro, la forma que pueden tener, la forma ovalada de los ojos, la forma de la nariz y la forma de las cejas. Convolución 3: Tras la tercera capa convolucional, de nuevo se obtienen las salidas, tras aplicarles los dos primeros filtrados, una vez más aumenta la complejidad y especificidad de los filtros. Los filtros ahora buscan aspectos tan complejos como la relación de la forma y ubicación de los ojos y otras estructuras con respecto a las demás, como la nariz, labios, bigote, cejas, etc., características del rostro que lo diferencian de los demás rostros. Es decir, a medida que se avanza hacia capas más profundas, se aumenta el nivel de abstracción del contenido de la imagen al que se muestran sensibles, tal y como se observa en las imágenes.

Fuente: elaboración propia.

Hay escepticismo sobre su precisión y los desafíos que enfrenta, así como el miedo de que reemplace a los médicos radiólogos o nucleares; sin embargo, la IA permite una práctica más rápida y reproducible. La participación adecuada de la comunidad médica permite asegurar un desarrollo tecnológico óptimo para mejorar la calidad de vida laboral y un mejor servicio de salud.

## Referencias

- Hricak H. New horizons lecture: beyond imaging—radiology of tomorrow. *Radiology*. 2018;286(3):764-75.
- Soffer S, Ben-Cohen A, Shimon O, Amitai M, Greenspan H, Klang E. Convolutional neural networks for radiologic images: A Radiologist's Guide Shelly. *Radiology*. 2019;290(3):590-606.
- Fukushima K, Miyake S. Neocognitron: a new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognit*. 1982;15(6):455-69.
- Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*. 1962;160(1):106-54.
- Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks. *NIPS'12*. 2012;1:1097-105.
- Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60-88.
- Chartrand G, Cheng PM, Vorontsov E, et al. Deep learning: a primer for radiologists. *RadioGraphics*. 2017;37(7):2113-31.
- Wang F, Casalino L, Khullar D. Deep learning in medicine. Promise, progress, and challenges. *JAMA Intern Med*. 2019;179(3):293-4.
- Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402-10.
- Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115-8.
- Baback EB, Mitko V, Paul J D, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA*. 2017;318(22):2199-210.
- Dreyer KJ, Geis JR. When machines think: radiology's next frontier. *Radiology*. 2017;285(3):713-8.
- Shiraishi J, Li Q, Suzuki K, et al. Computer-aided diagnostic scheme for the detection of lung nodules on chest radiographs: Localized search method based on anatomical classification. *Medical Physics*. 2006;33:2642-53.
- Arimura H, Katsuragawa S, Suzuki K, et al. Computerized scheme for automated detection of lung nodules in low-dose computed tomography images for lung cancer screening I. *Acad Radiol*. 2004;11(6):617-29.
- Shah SK, McNitt-Gray MF, Rogers SR, et al. Computer aided characterization of the solitary pulmonary nodule using volumetric and contrast enhancement features I. *Acad Radiol*. 2005;12(10):1310-9.
- Delogu P, Evelina FM, Kasae P, et al. Characterization of mammographic masses using a gradient-based segmentation algorithm and a neural classifier. *Comp Biol Med*. 2007;37(10):1479-91.
- Giger ML. Update on the potential of computer-aided diagnosis for breast cancer. *Future Oncol*. 2010;6(1):1-4.
- Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial intelligence in cardiology. *J Am Coll Cardiol*. 2018;71(23):2668-79.



19. Cadieu CF, Hong H, Yamins DLK, et al. Deep neural networks rival the representation of primate it cortex for core visual object recognition. *PLoS Comp Biol.* 2014;10.
20. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521:436-444.
21. Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge. *Int J Comput Vis.* 2015;115:211-52.
22. Graham B. Fractional max-pooling. Cornell University Library arXiv. 2015.
23. Lee CY, Gallagher PW, Tu Z. Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree. Cornell University Library arXiv. 2015.
24. Yasaka K, Akai H, Kunimatsu A, et al. Deep learning with convolutional neural network in radiology. *Japan J Radiol.* 2018;36(4):257-72.
25. Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *IEEE.* 1998;86:2278-324.
26. Simonyan K, Zisserman A. Deep convolutional networks for large-scale image recognition. Cornell University Library arXiv.2014.
27. Alom M Z, Taha T M, Yakopcic C, et al. A state-of-the-art survey on deep learning theory and architectures. *Electronics.* 2019;8(3).
28. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput.* 1997;9(8):1735-80.
29. Yamashita R, Nishio M, Do RKG, et al. Convolutional neural networks: an overview and application in radiology. *Insights Imag.* 2018;9:611-29.
30. Yang L, Chunxiao F, Yong L, et al. Improving deep neural network with multiple parametric exponential linear units. Cornell University Library arXiv. 2017.
31. Lakhani P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology.* 2017;284(2):574-82.
32. Cicero M, Bilbily A, Colak E, et al. Training and validating a deep convolutional neural network for computer aided detection and classification of abnormalities on frontal chest radiographs. *Invest Radiol.* 2017;52(5):281-7
33. Aoyama M, Li Q, Katsuragawa S, et al. Computerized scheme for determination of the likelihood measure of malignancy for pulmonary nodules on low-dose CT images. *Medical Physics.* 2003;30(3):387-94.
34. Becker AS, Marcon M, Ghafoor S, et al. Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. *Invest Radiol.* 2017;52(7): 434-40.
35. Kooi T, Litjens G, van Ginneken B, et al. Large scale deep learning for computer aided detection of mammographic lesions. *Med Image Anal.* 2017;35:303-12.
36. Prevedello LM, Erdal BS, Ryu JL, et al. Automated critical test findings identification and online notification system using artificial intelligence in imaging. *Radiology.* 2017;285(3):923-31.
37. Nakao T, Hanaoka S, Nomura Y, et al. Deep neural network-based computer-assisted detection of cerebral aneurysms in MR angiography. *J Mag Reson Imag.* 2018;47(4):948-53.
38. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Medical Physics.* 2017;44(2):547-57.
39. Men K, Dai J, Li Y. Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks. *Medical Physics.* 2017;44(12):6377-89.
40. Kumar V, Gu Y, Basu S, et al. Radiomics: the process and the challenges. *Magn Reson Imaging.* 2012.
41. Lambin P, Ríos-Velázquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012;48(4):441-6.
42. Aerts HJ, Velázquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun.* 2014.
43. Lao J, Chen Y, Li ZC, et al. A deep learning-based radiomics model for prediction of survival in glioblastoma multiforme. *Sci Rep.* 2017;7(1).
44. González G, Ash SY, Vegas Sánchez-Ferrero G, et al. Disease staging and prognosis in smokers using deep learning in chest computed tomography. *Am J Respir Crit Care Med.* 2018;197(2):193-203.

## Correspondencia

César Andrés Rueda Gutiérrez  
 Universidad Autónoma de Bucaramanga  
 Av. 42 # 48-11  
 Bucaramanga, Colombia  
 crueda552@unab.edu.co

Recibido para evaluación: 24 de enero de 2021

Aceptado para publicación: 15 de junio de 2021