

ARTÍCULO DE REVISIÓN

Revisión crítica de los Métodos de Supresión Ósea en Imágenes de Rayos X de Tórax

*Critical Review of Bone Suppression Methods
in Chest X-ray Images*

Eduardo Garea Llano

eduardo.garea@cneuro.edu.cu • <https://orcid.org/0000-0001-9101-1589>

CENTRO DE NEUROCIENCIAS DE CUBA

Ernesto Santiesteban Torres

ernestost@automatica.cujae.edu.cu

UNIVERSIDAD TECNOLÓGICA DE LA HABANA "JOSÉ ANTONIO ECHEVERRÍA", CUJAE, CUBA

Recibido: 2022-01-19 • Aceptado: 2022-03-06

RESUMEN

El examen de imágenes de radiografía de tórax es un método de evaluación del grado de efectividad de los protocolos aplicados a pacientes de COVID-19 en estado grave o crítico. Un elemento que influye significativamente en la efectividad de estos exámenes es la presencia en las imágenes de los huesos que interfieren en la correcta detección y evaluación de las lesiones provocadas por la enfermedad. El trabajo tiene por objetivo el estudio de los Métodos de Supresión Ósea (MSO) que han sido propuestos como paso para el pre-procesamiento de imágenes de radiografía de tórax. La metodología empleada se basó en la búsqueda, selección, revisión y análisis de los trabajos más actuales publicados en la temática. Se analizaron los trabajos más representativos. Se realizó un análisis del papel de la supresión ósea para mejorar la eficacia del diagnóstico. Se presentó una clasificación taxonómica de los métodos estudiados. Se realizó una propuesta de posible solución sobre el enfoque de aprendizaje no supervisado, en aras de mejorar el desempeño del diagnóstico tanto de los radiólogos como de sistemas automatizados.

PALABRAS CLAVE: Aprendizaje profundo, Imágenes CXR; Máquinas de aprendizaje, Supresión Ósea.

ABSTRACT

The examination of chest X-ray images is a method of evaluating the degree of effectiveness of the protocols applied to COVID-19 patients in serious or critical condition. An element that significantly influences the effectiveness of these examinations is the presence in the images of the bones that interfere with the correct detection and evaluation of the lesions caused by the disease. The aim of the work is the study of the Bone Suppression Methods (MSO) that have been proposed as a step for the pre-processing of chest radiography images. The methodology used was based on the search, selection, review and analysis of the most current works published on the subject. The most representative works were analyzed. An analysis of the role of bone suppression was performed to improve diagnostic efficiency. A taxonomic classification of the studied methods was presented. A possible solution proposal was made on the unsupervised learning approach, in order to improve the diagnostic performance of both radiologists and automated systems.

KEYWORDS: *Deep Learning; CXR Images; Machine Learning; Bone Suppression.*

INTRODUCCIÓN

Luego de casi dos años en la lucha contra la pandemia de la COVID-19 en Cuba, los protocolos y formas de actuación en el tratamiento de pacientes en estado grave o crítico han estado en constante revisión y perfeccionamiento. Un método de evaluación del grado de efectividad de los protocolos aplicados en estos casos ha sido el examen de imágenes de radiografía de tórax (CXR) para buscar indicadores visuales asociados con la evolución de la infección viral del SARS-CoV-2 (Ming-Yen *et al.*, 2020; Huang *et al.*, 2020).

Un elemento que influye significativamente en la efectividad de los exámenes realizados tanto de forma visual por especialistas como de forma automática por *software* es la presencia de los huesos que interfieren en la correcta detección y evaluación de las lesiones provocadas por las enfermedades pulmonares (Rajaraman *et al.*, 2021).

La supresión ósea es una técnica de pre-procesamiento en la radiografía de tórax (Baltruschat I. M., 2019) y es un paso importante que apoya a la segmentación pulmonar y la extracción de características. Las costillas y la clavícula pueden bloquear las anomalías pulmonares, lo que complica la fase de extracción de características de un sistema CAD (Diagnóstico Asistido por Computadora, por sus siglas en inglés). Por lo tanto, existe la necesidad de eliminar las estructuras esqueléticas, especialmente las costillas posteriores y las estructuras de la clavícula, para aumentar la visibilidad de la densidad del tejido blando (Qin *et al.*, 2018).

Una de las mayores dificultades para interpretar las radiografías de tórax se debe al hecho de que muchas estructuras anatómicas normales se muestran superpuestas sobre estructuras posiblemente anormales. Por esta razón, muchos esquemas de diagnóstico CAD en la radiografía de tórax pueden beneficiarse de la supresión de la mayor cantidad posible de estructuras normales. En el caso ideal, esto significaría que, después de procesar una radiografía de tórax, se obtiene una imagen que muestra solo las anomalías presentes en la radiografía original. En la práctica, sin embargo, el resultado será una imagen que contiene menos de las estructuras normales y, con suerte, respuestas más fuertes a las anomalías. Especialmente la supresión de las estructuras óseas que recubren los campos pulmonares, por ejemplo, clavículas, costillas y escápulas, es interesante porque en muchas tareas de detección conduciría a una reducción en el número de falsos positivos (Loog & van Ginneken, 2006). Por ejemplo, en (Nguyen & Dang, 2015) se mostró que la mayoría de las lesiones de cáncer de pulmón que se pasan por alto en las radiografías frontales de tórax se encuentran detrás de las costillas y que la inspección de una imagen de tejido blando puede mejorar el rendimiento de la detección de estas lesiones.

En este trabajo se presenta una revisión de los métodos de supresión ósea para imágenes CXR, presentamos una clasificación taxonómica de los métodos estudiados según el enfoque que utilizan. Continuamos con el análisis de las ventajas y desventajas de los métodos más representativos del estado del arte a partir de la clasificación realizada y realizamos una propuesta para la obtención de un método propio a desarrollar en futuras investigaciones. Finalmente damos las conclusiones del estudio.

METODOLOGÍA

Para el desarrollo de esta investigación se consultaron más de 30 artículos publicados que recogen los resultados de diferentes autores en los últimos 15 años referentes a la temática a partir de una investigación a fondo en las principales bases de datos como *IEEE Xplore*, *ACM Digital Library*, *ScienceDirect* y se seleccionaron los artículos más representativos. Seguidamente, se llevó a cabo una evaluación exhaustiva detallada y un análisis en profundidad de los mismos. Estos artículos reflejan, hasta cierto punto, las tendencias en la dirección de la investigación de supresión ósea en imágenes de rayos X, tomando como base los principios físicos y la naturaleza de los algoritmos sobre los cuales funcionan los dividimos en tres categorías.

En la figura 1 se presenta el esquema general de la clasificación taxonómica propuesta. En este sentido decidimos realizar tres grandes agrupaciones de los métodos revisados. En una primera a la que nombramos “Métodos instrumentales” agrupamos aquellos métodos que se basan en el uso de equipamiento comercial específicamente diseñado y fabricado para resolver esta tarea y que tienen como principio de funcionamiento la sustracción de energía dual.

En una segunda división nombrada “Métodos clásicos” agrupamos métodos que se basan en el uso de técnicas de procesamiento digital de imágenes (PDI), como son las técnicas de separación ciega de fuentes (BSS, por sus siglas en inglés *Blind Source Separation*) y transformaciones geométricas basadas en modelos matemáticos.

En la tercera agrupación nombrada “Métodos DL” agrupamos los métodos basados en el uso de técnicas de aprendizaje profundo (DL, por sus siglas en inglés *Deep Learning*). Durante el proceso de análisis de la literatura más reciente es evidente que con el auge de las técnicas de DL, estas son la tendencia más actual de los últimos cinco años no solo en el tema de la supresión ósea sino en el procesamiento de imágenes médicas en general. Cada subdivisión de esta agrupación se corresponde con el tipo de arquitectura de DL utilizada por los algoritmos para resolver el problema.



Figura 1. Clasificación taxonómica de los métodos de supresión ósea para imágenes CXR.

RESULTADOS Y DISCUSIÓN

Los MSO clasificados como “Métodos Instrumentales” utilizan dispositivos médicos con *hardware* desarrollado específicamente para enfrentar este problema. Con estos métodos se obtiene una supresión ósea de buena calidad, pero son caros debido al alto precio del *hardware* por lo que resulta muy costosa su generalización.

- **MSO por sustracción de energía dual (DES por sus siglas en inglés):** la sustracción de energía desde el punto de vista físico se basa en el hecho de que la atenuación de los materiales por rayos X difiere en dependencia de la energía de los fotones *roentgen*. A mayores energías de rayos X (aproximadamente > 100 kVp, expresada por el voltaje acelerador de los electrones en un tubo de rayos X), prevalece la llamada *Dispersión de Compton*, en este caso la absorción no es muy diferente para los huesos y los tejidos blandos que a energías más bajas, donde la absorción fotoeléctrica es el mecanismo más importante. La foto-absorción, sin embargo, es mucho más eficaz con átomos de alto número atómico (como el calcio) que, con hidrógeno, carbono y oxígeno, que son los átomos dominantes en los tejidos blandos. A pesar de que los tubos de rayos X producen espectros en lugar de fotones mono-energéticos, el uso de diferentes magnitudes kVp influirá significativamente en la absorción de rayos X. Existen dos enfoques fundamentales. Ambos enfoques se utilizan actualmente para la radiografía de sustracción de energía y producen dos radiografías, una a kVp más alto y otra a más bajo (Vock & Szucs-Farkas, 2009).
- **Radiografía DES de simple tiro:** utiliza una exposición única del paciente para producir una radiografía normal de tórax, una imagen de hueso y una imagen de tejido blando. El dispositivo utilizado contiene un filtro de cobre delgado intercalado entre dos placas

de fósforo. El haz de rayos X que llega a la primera placa produce una radiografía de tórax normal que contiene huesos y tejidos blandos. La segunda placa es alcanzada por un número menor de fotones roentgen, en su mayoría de alta energía, que han pasado la primera placa y el filtro. Debido a la alta energía y el bajo recuento de fotones, la imagen generada en la segunda placa tiene un bajo contraste óseo y también es más ruidosa que la imagen normal de la primera placa. La cantidad de ruido de las radiografías de un solo disparo se compensa parcialmente durante un proceso pos-procesamiento de las imágenes. Para generar una imagen ósea, la señal de la imagen de alta energía en la segunda placa se mejora hasta que la intensidad de los tejidos blandos alcanza la de la imagen de baja energía en la primera placa; la resta ponderada de estas imágenes da como resultado la cancelación de la señal de los tejidos blandos, dejando solo los huesos y las estructuras calcificadas visibles (imagen del hueso). Para generar la imagen de tejido blando, las señales de la imagen de alta energía en la segunda placa se ajustan para igualar la intensidad de los huesos en ambas placas; la resta ponderada de las imágenes se produce de nuevo en la imagen de tejido blando (Vock & Szucs-Farkas, 2009).

- **Radiografía DES de doble tiro:** las radiografías de alta y baja energía se generan mediante dos exposiciones del paciente. El sistema utiliza un detector digital con una capacidad de detección superior, en comparación con las placas de fósforo. El detector se expone consecutivamente dos veces a dos niveles de energía diferentes, típicamente a 120 y 60 kVp, generando dos radiografías y permitiendo su resta ponderada.

El uso del detector digital da como resultado una excelente calidad de imagen con alto contraste y baja presencia de ruido en la imagen. Adicionalmente, los espectros de energía de las dos imágenes se separan mejor que con el filtro de cobre en la técnica de disparo único, lo que da como resultado una muy buena sustracción de los huesos por ende mejoran la precisión del diagnóstico de varias enfermedades pulmonares (fibrosis quística, tuberculosis) en comparación con radiografías convencionales (Obmann *et al.*, 2017; Manji *et al.*, 2016), particularmente estos resultados son más visibles cuando los especialistas que realizan la evaluación no son radiólogos especializados en tórax.

El problema del método DES de doble tiro es que el corto intervalo de 150-200 ms entre las dos exposiciones sucesivas hace que las imágenes sean susceptibles a artefactos producidos por los movimientos cardíaco, respiratorio y muscular, por lo tanto, el pos-procesamiento se hace necesario para reducir estos artefactos (Vock & Szucs-Farkas, 2009). Por otro lado, el paciente se debe someter a una mayor dosis de radiación.

Para la aplicación de ambos métodos DES es necesario contar con el equipamiento especialmente desarrollado para ello por empresas especializadas que lo comercializan a altos precios en el mercado internacional. Además requieren de locales especialmente adaptados para su instalación y funcionamiento. Este hecho encarece sumamente su generalización y conlleva una inversión considerable de compra, soporte y mantenimiento. Estas condiciones bajo las características de la pandemia de COVID-19 hace inviable su uso pues el tiempo de desinfección entre pacientes para el equipo y la sala de estudio es

de 15 minutos aproximadamente, lo cual no es viable en condiciones de muy alta presión asistencial. Además, la exposición del personal médico hace que esta técnica no se recomienda como método de diagnóstico para pacientes con COVID-19.

MÉTODOS CLÁSICOS (NO SUPERVISADOS)

Este segundo grupo de métodos utiliza diversas técnicas de procesamiento digital de imágenes como manipulaciones de contraste y diferentes tipos de filtros en el dominio espacial o frecuencial, o el análisis estadístico de los componentes de la imagen. Fueron los métodos no instrumentales más utilizados en los primeros años de la década de los 2000, hasta que comenzaron a tomar protagonismo los métodos basados en aprendizaje profundo.

Según la revisión realizada, los métodos estudiados fundamentalmente abordan el problema de la obtención de la supresión ósea con algoritmos BSS y transformaciones geométricas. La diversidad de técnicas usadas en estos métodos, hace que difieran bastante entre sí. En este trabajo abordaremos los aspectos y resultados más generales de los métodos de esta categoría.

En Erdogan (2006) se presenta un enfoque de BSS. El autor presupone que la mezcla de observación se obtiene a través de un mapeo lineal desconocido sin memoria de fuentes de magnitud independientes y limitadas. Además se asume un pre-blanqueamiento inicial adaptativo del vector de observación original que lo transforma en un vector blanco con la misma dimensión que el vector fuente original.

Rasheed, *et al.* (2007), presentan un método basado en Análisis de Componentes Independientes (ICA por sus siglas en inglés). Los autores se basan en que la radiografía de tórax se puede considerar como una combinación lineal de varias fuentes, y que el ruido introducido durante el proceso de adquisición puede considerarse como de una fuente única. De manera similar, asumen que las costillas también constituyen una fuente separada que puede eliminarse para mejorar la radiografía y utilizan ICA para realizar la separación basada en la distribución de probabilidad de los componentes individuales.

El método ICA necesita al menos dos vectores de observación de entrada. Existen diferentes enfoques para obtener este par de imágenes, entre ellos se encuentra el propuesto por Nguyen & Dang (2015). Los autores utilizan la ecualización del histograma de la imagen original de CXR y un filtro para crear artificialmente las dos imágenes.

En el análisis de las imágenes biomédicas solo se trabaja con componentes no negativos, es decir, la intensidad de cada pixel será siempre mayor o igual que cero. En (Plumbley, 2003) implementan el método ICA, incorporándole esta restricción de no negatividad de las fuentes.

El problema de este primer grupo es que, como inicialmente solo se tiene una imagen de CXR, se hace necesario obtener dos imágenes diferentes a partir de la imagen inicial de CXR. Este hecho propicia la creación no controlada de artefactos que pueden ser después introducidos en la imagen resultante y conllevar a la que hacen muy complejo este enfoque.

En (Wang, *et al.*, 2006) se propone otro BSS basado en el análisis de los componentes no negativos poco correlacionados. Este método se basa en encontrar una función de correlación conjunta de múltiples señales para revelar y confirmar que las observaciones después de la

mezcla no negativa tendrían una correlación conjunta más alta que las fuentes desconocidas originales. En consecuencia, se propone un método de análisis de componentes menos correlacionados no negativos (nLCA) para diseñar la matriz de desmezcla minimizando la función de correlación conjunta entre las fuentes no negativas estimadas.

En (Wang *et al.*, 2009) se propone una optimización del algoritmo general de nLCA para el caso de múltiples fuentes que se desarrolla en base a un principio de maximización de volumen iterativo (IVM) y programación lineal.

En (Chan *et al.*, 2008) se presenta otro método BSS. El método propuesto, denominado por los autores como análisis convexo de mezclas de fuentes no negativas (CAMNS), es determinista y no requiere la suposición de independencia de la fuente y se basa en un supuesto especial llamado dominio local. Bajo el dominio local y varios supuestos se aplica el análisis convexo para establecer un nuevo criterio BSS, que establece que las señales de origen se pueden identificar al encontrar los puntos extremos de un conjunto poliédrico construido por observación. También se derivan métodos para cumplir con el criterio CAMNS, utilizando programación lineal (CAMNS-LP) o geometría *simplex* (CAMNS-G, geométrico).

Los autores compararon sus resultados con otros métodos clásicos como nLCA, nICA y el algoritmo de Erdogan sobre un mismo conjunto de datos. Concluyeron que su propuesta garantiza una separación perfecta bajo algunos supuestos que pueden ser realistas en muchas aplicaciones.

El problema aquí es que dada la gran cantidad de supuestos que asumen, el método es aplicable solo bajo aquellas condiciones, las que generalmente se basan en una buena calidad de las imágenes captadas. Esto es contraproducente cuando se trabaja con imágenes de pacientes COVID-19 que presentan una baja calidad respecto a su nitidez producto a las afectaciones provocadas por la enfermedad.

En (Von Berg *et al.*, 2016) se presenta un algoritmo MSO. Dado un contorno que describe un hueso, la región que ocupa en la imagen es **espacialmente transferida** para separar los componentes de gradiente de la imagen normal donde el contorno del hueso aparecerá como una línea recta. En este espacio se calcula una derivada parcial a través del contorno óseo, se suaviza a lo largo del contorno y luego se reintegra hasta la línea central. La idea del suavizado es eliminar todas las demás señales además del contorno óseo de la imagen. Al reintegrarse, solo se restaura la señal ósea y la sustracción de esta imagen ósea de la imagen original da como resultado una eliminación visual del hueso.

En (Liu *et al.*, 2019) se implementa un método que usa lo que ellos nombran “imágenes priori no paramétricas”. Para una radiografía de tórax en estudio, se seleccionan radiografías de tórax similares con las correspondientes imágenes de tejido blando y óseo DES de una base de datos de entrenamiento. Los campos correspondientes entre la radiografía de tórax observada y los ejemplos se resuelven mediante un algoritmo de coincidencia jerárquicamente denso. A continuación, se construyen *prioris* no paramétricos de tejidos blandos y componentes óseos muestreando parches de imágenes de las imágenes óseas y de tejidos blandos. Finalmente, estos *prioris* no paramétricos se integran en el modelo de descomposición, cuya

función energética se optimiza mediante un esquema de mínimos cuadrados re-ponderados iterativamente.

Los métodos anteriores para la tarea de supresión ósea se pueden categorizar como métodos no supervisados. La ventaja de los métodos no supervisados es que no necesitan de radiografías para el entrenamiento; sin embargo, requieren de la segmentación de los huesos y luego reconstruir imágenes libres de huesos. El rendimiento de los métodos no supervisados se ve muy afectado por la precisión de la segmentación lo que los hace muy dependientes de la calidad de las imágenes.

MÉTODOS DL (SUPERVISADOS)

Estos son métodos que pueden catalogarse como métodos supervisados ya que suprimen la estructura ósea en las radiografías de tórax mediante predicción de regresión.

Estos métodos utilizan diferentes arquitecturas de aprendizaje profundo para estimar la imagen de supresión ósea. Son generalmente entrenados usando pares de imágenes de CXR y DES, por lo tanto se necesita disponer de una base de datos con gran cantidad de estos tipos de imágenes para realizar el entrenamiento de la red neuronal y esta es su principal desventaja.

En (Suzuki *et al.*, 2006) se propone una técnica de procesamiento de imágenes para suprimir el contraste de costillas y clavículas en radiografías de tórax mediante una red neuronal artificial de entrenamiento masivo multi-resolución (MTANN). Un MTANN es un filtro altamente no lineal que puede entrenarse mediante el uso de radiografías de tórax de entrada y las imágenes DES correspondientes.

Los experimentos realizados mostraron que el método propuesto puede suprimir bien las costillas, no suprimieron muy bien los bordes de las costillas, las costillas cercanas a la pared pulmonar y las clavículas porque la orientación, el ancho, el contraste y la densidad de los huesos son diferentes de un lugar a otro, esto se debe a que la capacidad de una MTANN de resolución múltiple es limitada.

Para lidiar con las limitaciones de MTANN, en (Chen & Suzuki, 2014) se propone una mejora del método desarrollado en (Suzuki *et al.*, 2006), esta vez se implementa un modelo anatómico-específico múltiple de MTANNs (ASMTANN). Para ello entrenaron múltiples modelos MTANNs sobre regiones anatómicas específicas de la región pulmonar de interés, luego las imágenes de salida de los múltiples MTANN se fusionan para producir una imagen de huesos, a esta se le aplica un suavizado para reducir el ruido mientras se conservan los bordes y luego se le resta a la CXR original para obtener la imagen de tejido blando.

Las Redes Neuronales Convolucionales (CNN, por sus siglas en inglés) (Albawi *et al.*, 2017) son un tipo de redes neuronales artificiales. Este tipo de red se aplica en matrices bidimensionales, son muy efectivas para tareas de visión artificial, como en la clasificación y segmentación de imágenes, entre otras aplicaciones. Existen reportadas en la literatura formas muy variadas de arquitecturas CNN para resolver diferentes problemas de la visión por computadoras (Voulodimos *et al.*, 2018). Diferentes autores utilizan arquitecturas del tipo CNN para desarrollar algoritmos de supresión ósea con diferentes enfoques.

En (Gusarev *et al.*, 2017) se proponen dos arquitecturas que suprimen los huesos en las radiografías tratándolos como ruido. El primer modelo es una familia de autocodificadores convolucionales dividido en tres capas de codificación de la información con 16, 32 y 64 neuronas respectivamente; mientras que el segundo es una familia de las CNN simples.

En (Gozes & Greenspan, 2020) utilizan la red convolucional completamente conectada (FCN por sus siglas en inglés) *U-Net*. Para crear los pares de imágenes de entrenamiento, emplean radiografías simuladas de rayos X o radiografías reconstruidas digitalmente (DRR), derivadas de 664 tomografías computarizadas (TC). Utilizando la segmentación de estructuras óseas basada en la Unidad *Hounsfield* (HU) en el dominio de TC, se produce una DRR sintética de “rayos X óseos” en 2D que se utiliza para entrenar la red. Una vez que se extraen las estructuras óseas, la imagen original se mejora fusionando la radiografía de entrada original y la “radiografía ósea” sintetizada.

En (Matsubara *et al.*, 2020) se propone un filtro neuronal convolucional (CNF por sus siglas en inglés) para la supresión ósea basado en una CNN. CNF genera un valor para el componente óseo del píxel de destino ingresando valores de píxel en la vecindad del píxel de destino. Al procesar todas las posiciones en la imagen de entrada, se genera una imagen extraída del hueso. Finalmente, la imagen con supresión ósea se obtiene restando la imagen extraída del hueso de la imagen original de la radiografía de tórax.

En (Wang *et al.*, 2020) los autores combinan el método de fusión de características multiescala *U-Net*, la conexión densa *DenseNet* y el mecanismo multitarea para construir una red denominada MDU-Net. Luego presentan un mecanismo de codificación de máscara que puede obligar a la red a aprender las características de fondo. El aprendizaje por transferencia se introduce en última instancia para ayudar a la red a extraer suficientes funciones.

En (Sirazitdinov *et al.*, 2020) se evalúan varios modelos DL adaptados a la tarea de supresión ósea. Fueron implementadas varias arquitecturas DL de última generación

Este tipo de arquitecturas basadas en CNN han demostrado ser muy eficaces en la tarea de supresión ósea y comparativamente superiores a los métodos no supervisados clásicos, pero se puede observar, que la mayoría de los trabajos utilizan conjuntos de pares de imágenes DES disponibles a nivel internacional. Estos conjuntos presentan la dificultad de pertenecer a orígenes distintos, y las enfermedades presentes en ellas se corresponden con estos orígenes. De esta forma, los métodos pueden aprender a sustraer los huesos ante la presencia de esas enfermedades o de personas sanas bajo las condiciones con las que fueron entrenados. Esto puede llevar a una falta de generalización de los modelos cuando son utilizados ante imágenes de pacientes afectados por COVID-19 de las cuales no se cuentan con suficientes imágenes DES, por lo que el principal problema de estos modelos es que hasta el momento no ha existido un protocolo de evaluación para los modelos propuestos en el sentido de su aplicación a imágenes de pacientes COVID-19

En los estudios analizados, no se presentan resultados que utilizan imágenes que no pertenezcan a ninguna de las fuentes de procedencia de los conjuntos de imágenes usados en el entrenamiento de los modelos.

Las redes generativas antagónicas o *Generative Adversarial Network* en inglés (GAN) usan dos redes neuronales profundas (red generadora y red discriminadora) donde una de las redes

se enfoca en generar una salida y la otra red se encarga de discriminar las salidas incorrectas de la red generadora. Estas dos redes son adversarias, es decir, “juegan” un juego de suma cero donde lo que una red gana, la otra pierde.

En (Liang *et al.*, 2020) se implementan dos modelos diferentes de GANs como MSO, uno de ellos utiliza el método con imágenes en parejas y un modelo Pix2Pix y el otro usa imágenes desemparejadas y la arquitectura *Cycle-GAN*. Los autores concluyen que, si bien el entrenamiento con imágenes en parejas produce un rendimiento ligeramente mejor que el de imágenes no emparejadas cuando se mide con dos métricas objetivas de calidad de imagen, el índice de similitud estructural (SSIM) y la relación señal-ruido máxima (PSNR), el entrenamiento con imágenes no emparejadas demuestra una mejor capacidad de generalización en radiografías anteroposteriores (AP) invisibles que en el entrenamiento emparejado.

En (Li *et al.*, 2020) se propone un enfoque MSO de “grueso a fino” mediante el uso de *prioris* estructurales derivados de imágenes TC. Se utiliza la imagen de radiografía reconstruida digitalmente (DRR) que se calcula a partir de TC en forma de puente para conectar la imagen TC y CXR. Luego, se realiza la descomposición ósea aprovechando el modelo de descomposición ósea DRR aprendido de TCs no apareados y la adaptación de dominio entre CXR y DRR. Los autores utilizan un modelo generativo basado en *CycleGAN* (Zhu *et al.*, 2017) que se entrena entre los dominios de CXR y DRR. Después de la adaptación del dominio, las imágenes sin hueso se generan con *U-Net* (Ronneberger *et al.*, 2015), luego se retroalimentan a la red para obtener CXR con aumento de supresión ósea. También usan una red adversaria generativa de descomposición (DecGAN) (Li *et al.*, 2019) entrenada con DRR para incorporar el conocimiento previo de descomposición de CT en el espacio latente de CXR y separar los componentes óseos en el espacio latente.

El problema de estos enfoques es que utilizan las *Cycle-GAN* que son predicciones de “cómo sería si...” y las predicciones, que se consideran plausibles, pueden diferir en gran medida de la realidad básica. Las *CycleGAN* solo deben usarse con gran cuidado y calibración en dominios donde se deben tomar decisiones críticas en función de su salida. Esto se debe tomar muy en cuenta en aplicaciones médicas, y la supresión ósea es una de ellas. *CycleGAN* puede, por ejemplo, agregar tumores en imágenes médicas donde no existen, o eliminarlos de las que los tienen. En este sentido Cohen (*et al.*, 2018) presentaron evidencia experimental de que estas imágenes traducidas no deben utilizarse para interpretación (por ejemplo, por parte de médicos) sin las herramientas adecuadas para verificar el proceso de traducción. Los autores ilustraron este problema utilizando ejemplos de tumores que se agregan y eliminan de las imágenes de resonancia magnética.

DISCUSIÓN

Después del análisis de la literatura descrita podemos observar que la supresión ósea en imágenes de CXR es un problema que se ha abordado con una gran diversidad de enfoques. Los MSO por DES proporcionan buenos resultados y, aunque presentan desventajas, como el aumento de la dosis de radiación a los pacientes y la necesidad de hardware especializado muy caros y de complicado manejo en las condiciones de la pandemia de COVID-19, constituyen el patrón de referencia para desarrollar nuevos MSO.

Destaca como tendencia en las investigaciones actuales la preferencia de los investigadores por los MSO basados en DL como métodos supervisados, frente a los métodos clásicos o no supervisados, esto se evidencia en la creciente cantidad de publicaciones de MSO de esta categoría en los años más recientes, fenómeno que caracteriza a todo las ramas de la visión por computadoras.

Del análisis realizado pudimos observar que los métodos CAMNS y nLCA tienen muy buen rendimiento, pero como en la literatura analizada estos métodos no se evalúan con las mismas métricas, no es posible realizar una comparación precisa, aunque es importante resaltar que los métodos basados en nLCA tienen un menor costo computacional por lo que son mucho más rápidos que los MSO basados en CAMNS (Wang *et al.*, 2009). Comparar los MSO de esta categoría con los MSO basados en *Deep Learning* es complicado, ya que casi todas las métricas usadas son diferentes en ambos enfoques. De forma general, los enfoques basados en arquitecturas GANs tienen un rendimiento superior al resto de los enfoques analizados, aunque es necesario tomar en cuenta las deficiencias que fueron analizadas en los epígrafes anteriores

Atendiendo a esto, se propone como solución al tema implementar una arquitectura GAN usando el método *Pix2Pix*. Se elige este enfoque debido a la simplicidad de implementación de esta arquitectura y la flexibilidad de optimización que presenta a causa de las múltiples formas de implementar la red generadora y la red discriminadora. Del mismo modo, se sugiere para trabajos futuros implementar los métodos clásicos basados en nLCA y CAMNS, para su posterior comparación respecto a los MSO basados en DL. También se sugiere explorar el efecto que tienen estos MSO sobre los sistemas CAD. Como solución a largo plazo se sugiere, como una posible alternativa para lograr mayor eficacia en la obtención de imágenes sin huesos, implementar un método anatómico-específico con múltiples arquitecturas *CycleGAN*. La hipótesis de esta sugerencia es que la reducción de la región de interés combinado con un marco de trabajo de aprendizaje no supervisado puede aumentar la precisión de la imagen obtenida.

En la Tabla 1 se realiza una comparación del desempeño de algunos de los métodos de supresión ósea del estado del arte, en el caso de los MSO *Pix2Pix*, *CycleGAN* y *Piors* estructurales despareados, para la comparación se consideró la métrica AUC, la cual es el área bajo la curva ROC (acrónimo de *Receiver Operating Characteristic*, o Característica Operativa del Receptor) que se obtiene en tareas de clasificación de las imágenes CXR sobre la respuesta de sistemas CAD. La métrica se calcula a partir de la estimación de la clasificación de la imagen ante la presencia o no de lesiones características de las enfermedades en estudio (COVID-19, Neumonías, Cáncer) comparando las imágenes clasificadas sin aplicar previamente la supresión ósea con las imágenes clasificadas después de aplicar algún MSO. Para ello se cuantifican las imágenes correctamente clasificadas (verdaderos positivos y verdaderos negativos) y las imágenes incorrectamente clasificadas (falsos positivos y falsos negativos)

En los otros dos casos se utilizaron radiólogos profesionales, en el caso del MSO por DES, y no expertos, en el caso de MSO por Transformación al espacio ST. En los resultados se observa un mejor desempeño de los métodos basados en DL

En la figura 2 se presentan algunos ejemplos de los resultados obtenidos por varios de los “Métodos Clásicos” analizados. De (a - e) se muestran las imágenes de tejido blando y huesos

separadas mediante los diferentes MSO mientras que en (e) solo se muestra la CXR original y la imagen resultante. (a) CAMNS-LP (Chan *et al.*, 2008), (b) nLCA-ES (Wang *et al.*, 2009), (c) Algoritmo de Erdogan (Erdogan, 2006), (d) Priors de imagen no paramétricos (Liu *et al.*, 2019), (e) Transferencia espacial (Von Berg *et al.*, 2016).

Tabla 1. Comparación de algunos MSO respecto a al AUC.

Método	AUC
DES (Kodama, et al., 2018)	0,933
Transformación al espacio ST (von Berg et al.,2016)	0,655
Pix2Pix (Liang et al., 2020)	0,948 4 0,004
CycleGan (Liang et al., 2020)	0,9534 0,003
Priors Estructurales (Li et al., 2020)	0,824

En la figura 3 se presentan ejemplos de pares de imágenes CXR y la imagen resultante de diferentes MSO basados en DL donde: (a) ASMTAN (Chen & Suzuki, 2014), (b) CNN (Gozes & Greenspan, 2020), (c) CNN (Matsubara et al, 2020), (d) GAN (Liang *et al.*, 2020) y (e) Priors Estructurales (Li *et al.*, 2020).

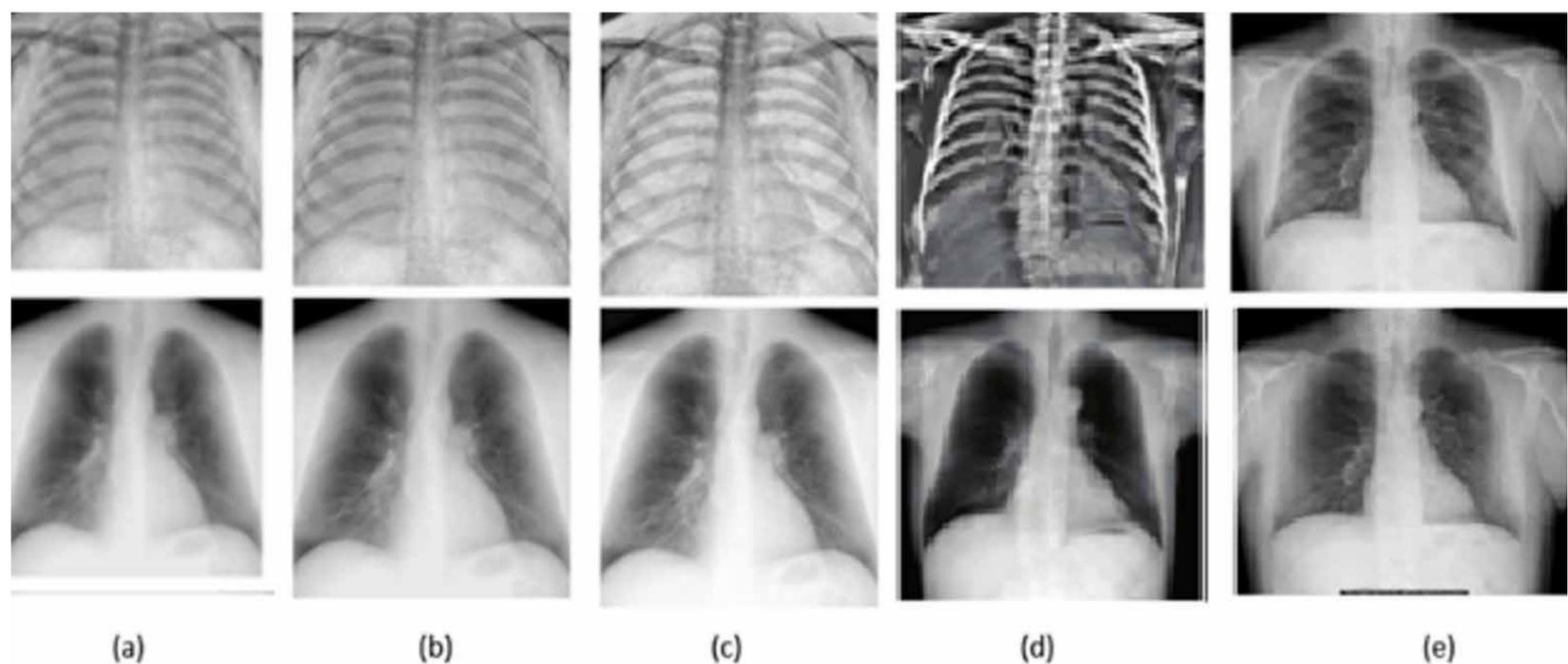


Figura 2. Comparación entre las imágenes resultantes de los “métodos clásicos”.

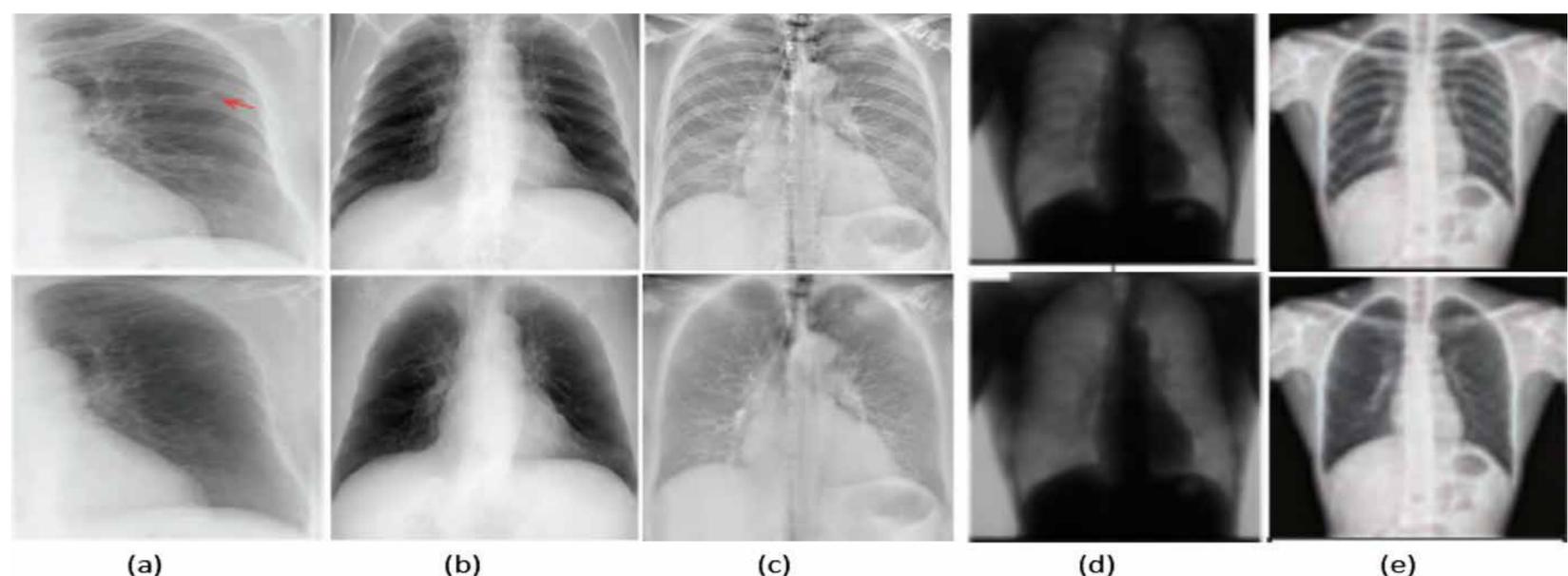


Figura 3. Pares de imágenes CXR e imágenes resultantes de diferentes MSO basados en DL.

CONCLUSIONES

En este trabajo se presentó un análisis del papel de la supresión ósea en las imágenes de CXR para mejorar la eficacia diagnóstica en la radiología, se presenta una clasificación taxonómica de los métodos estudiados según el tipo de enfoque que utilizan para obtener la imagen final. Se realiza el análisis de los artículos que contienen los métodos del estado del arte a partir de la clasificación taxonómica llegando a la conclusiones que permiten realizar una propuesta de cómo enfrentar el problema a partir del análisis de las ventajas y desventajas de los métodos estudiados, la cual se centra en dos variantes de solución, una a corto plazo basada en la adaptación de una arquitectura *Pix2Pix* la que requiere de un proceso de evaluación, implementación y adecuación a las características de las imágenes de pacientes con COVID-19. Una segunda propuesta de solución a más largo plazo que nos permitiría tener un método propio basado en el enfoque de aprendizaje no supervisado basado en un modelo anatómico-específico de múltiples *CycleGAN*, que permita mejorar el desempeño diagnóstico tanto de los radiólogos como de sistemas CAD.

REFERENCIAS

- Albawi. S., T. Mohammed A. and Al-Zawi S.(2017). Understanding of a convolutional neural network. 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.
- Baltruschat I. M. (2019). When Does Bone Suppression And Lung Field Segmentation Improve Chest X-Ray Disease Classification?. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 2019, pp. 1362-1366, doi: 10.1109/ISBI.2019.8759510.
- Chan, T.-H., Ma, W.-K., Chi, C.-Y., & Wang, Y. (2008). A convex analysis framework for blind separation of non-negative sources. *IEEE Transactions on Signal Processing*, 56(10), 5120-5134.
- Chen, S., & Suzuki, K. (2014). Bone suppression in chest radiographs by means of anatomically specific multiple massive-training ANNs combined with total variation minimization smoothing and consistency processing. *Computational Intelligence in Biomedical Imaging* (pp. 211-235): Springer.
- Cohen, J., Luck, M., & Honari, S. (2018). Distribution Matching Losses Can Hallucinate Features in Medical Image Translation. *ArXiv*, abs/1805.08841.
- Erdogan, A. T. (2006). A simple geometric blind source separation method for bounded magnitude sources. *IEEE Transactions on Signal Processing*, 54(2), 438-449.
- Gozes, O., & Greenspan, H. (2020). Bone Structures Extraction and Enhancement in Chest Radiographs via CNN Trained on Synthetic Data. Paper presented at the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI).
- Gusarev, M., Kuleev, R., Khan, A., Rivera, A. R., & Khattak, A. M. (2017). Deep learning models for bone suppression in chest radiographs. Paper presented at the 2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB).

- Huang C., Wang Y., Li X., Ren L., Zhao J., Hu Y., Zhang L., Fan G., Xu J., Gu X., Cheng Z., Yu T., Xia J., Wei Y., Wu W., Xie X., Yin W., Li H., Liu M., Xiao Y., Gao H., Guo L., Xie J., Wang G., Jiang R., Gao Z., Jin Q., Wang J., Cao B. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, china. *The Lancet*, 395(10223), 2020. 1
- Li, H., Han, H., Li, Z., Wang, L., Wu, Z., Lu, J., & Zhou, S. K. (2020). High-resolution chest x-ray bone suppression using unpaired CT structural priors. *IEEE Transactions on medical imaging*, 39(10), 3053-3063.
- Li, Z., Li, H., Han, H., Shi, G., Wang, J., & Zhou, S. K. (2019). Encoding CT anatomy knowledge for unpaired chest X-ray image decomposition. Paper presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention.
- Liang, J., Tang, Y.-X., Tang, Y.-B., Xiao, J., & Summers, R. M. (2020). Bone suppression on chest radiographs with adversarial learning. Paper presented at the Medical Imaging 2020: Computer-Aided Diagnosis.
- Liu, Y., Yang, W., She, G., Zhong, L., Yun, Z., Chen, Y., Feng, Q. (2019). Soft Tissue/Bone Decomposition of Conventional Chest Radiographs Using Nonparametric Image Priors. *Applied bionics and biomechanics*, 2019.
- Loog, M., & van Ginneken, B. (2006). Bony structure suppression in chest radiographs. Paper presented at the International Workshop on Computer Vision Approaches to Medical Image Analysis.
- Manji, F., Wang, J., Norman, G., Wang, Z., & Koff, D. (2016). Comparison of dual energy subtraction chest radiography and traditional chest X-rays in the detection of pulmonary nodules. *Quantitative imaging in medicine and surgery*, 6(1), 1.
- Matsubara, N., Teramoto, A., Saito, K., & Fujita, H. (2020). Bone suppression for chest X-ray image using a convolutional neural filter. *Physical and Engineering Sciences in Medicine*, 43(1), 97-108.
- Ming-Yen N., Lee Y. P. , Yang J., Yang F. , Li X. , Wang Hongx , Lui M.M. , Lo C. , Leung B. , Khong PL., Hui C. , Yuen K. , Kuo M.D. (2020) . Imaging profile of the COVID-19 infection: Radiologic findings and literature review. *Radiology: Cardiothoracic Imaging*, 2(1), 2020. 1
- Nguyen, H. X., & Dang, T. T. (2015). Ribs suppression in chest x-ray images by using ICA method. Paper presented at the 5th International Conference on Biomedical Engineering in Vietnam.
- Obmann, V. C., Christe, A., Ebner, L., Szucs-Farkas, Z., Ott, S. R., Yarram, S., & Stranzinger, E. (2017). Bone subtraction radiography in adult patients with cystic fibrosis. *Acta radiologica*, 58(8), 929-936.
- Plumbley, M. D. (2003). Algorithms for nonnegative independent component analysis. *IEEE Transactions on Neural Networks*, 14(3), 534-543.
- Qin, C., Yao, D., Shi, Y., & Song, Z. (2018). Computer-aided detection in chest radiography based on artificial intelligence: a survey. *Biomedical engineering online*, 17(1), 1-23.

- Rajaraman, S.; Zamzmi, G.; Folio, L.; Alderson, P.; Antani, S. Chest X-ray Bone Suppression for Improving Classification of Tuberculosis-Consistent Findings. *Diagnostics* 2021, 11, 840. <https://doi.org/10.3390/diagnostics11050>
- Rasheed, T., Ahmed, B., Khan, M. A., Bettayeb, M., Lee, S., & Kim, T.-S. (2007). Rib suppression in frontal chest radiographs: A blind source separation approach. Paper presented at the 2007 9th International Symposium on Signal Processing and Its Applications.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. Paper presented at the International Conference on Medical image computing and computer-assisted intervention.
- Sirazitdinov, I., Kubrak, K., Kiselev, S., Tolkachev, A., Kholiavchenko, M., & Ibragimov, B. (2020). Evaluation of Deep Learning Methods for Bone Suppression from Dual Energy Chest Radiography. Paper presented at the International Conference on Artificial Neural Networks.
- Suzuki, K., Abe, H., MacMahon, H., & Doi, K. (2006). Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN). *IEEE Transactions on medical imaging*, 25(4), 406-416.
- Vock, P., & Szucs-Farkas, Z. (2009). Dual energy subtraction: principles and clinical applications. *European journal of radiology*, 72(2), 231-237.
- Voulodimos A, Doulamis N., Doulamis A., Protopapadakis E (2018). Deep Learning for Computer Vision: A Brief Review, *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 7068349, 13 pages, 2018. <https://doi.org/10.1155/2018/7068349>
- Von Berg, J., Young, S., Carolus, H., Wolz, R., Saalbach, A., Hidalgo, A., Franquet, T. (2016). A novel bone suppression method that improves lung nodule detection. *International journal of computer assisted radiology and surgery*, 11(4), 641-655.
- Wang, F.-Y., Wang, Y., Chan, T.-H., & Chi, C.-Y. (2006). Blind separation of multichannel biomedical image patterns by non-negative least-correlated component analysis. Paper presented at the International Workshop on Pattern Recognition in Bioinformatics.
- Wang, F.-Y., Chi, C.-Y., Chan, T.-H., & Wang, Y. (2009). Nonnegative least-correlated component analysis for separation of dependent sources by volume maximization. *IEEE transactions on pattern analysis and machine intelligence*, 32(5), 875-888.
- Wang, W., Feng, H., Bu, Q., Cui, L., Xie, Y., Zhang, A., Chen, Z. (2020). MDU-Net: A Convolutional Network for Clavicle and Rib Segmentation from a Chest Radiograph. *Journal of Healthcare Engineering*, 2020.
- Zhu, J.-Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. Paper presented at the Proceedings of the IEEE international conference on computer vision.