

This is a postprint version of the following published document:

Sakaltras, N., et al. (2020, noviembre). *Compensación de radiación dispersa en radiografía digital a través del aprendizaje automático: resultados preliminares*. En: CASEIB 2020: XXXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica, libro de actas, 25–27 Nov, 2020. Sociedad Española de Ingeniería Biomédica, 2020, pp. 73–76.

URL: http://caseib.es/2020/wp-content/uploads/2020/12/CASEIB2020_LibroActas.pdf



This work is licensed under a [Se distribuye bajo una licencia de Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Compensación de radiación dispersa en radiografía digital a través del aprendizaje automático: resultados preliminares

N. Sakaltras^{1,2}, F. A. Tovar¹, C. Martínez^{1,2}, C. F. Del Cerro^{1,2}, M. Desco^{1,2,3,4}, M. Abella^{1,2,4}

¹ Dept. Bioingeniería e Ingeniería Aeroespacial, Universidad Carlos III de Madrid, España

² Instituto de Investigación Sanitaria Gregorio Marañón, Madrid

³ Centro de investigación en red en salud mental (CIBERSAM), Madrid

⁴ Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, España

Resumen

La dispersión de los rayos X reduce significativamente la resolución de contraste de la imagen en radiografía digital de tórax. La estrategia convencional para la reducción de la radiación dispersa es el uso de rejillas antidifusoras que, aunque mejoran la calidad de la imagen, aumentan la dosis de radiación absorbida por el paciente y plantean problemas en técnicas no estándar. En este trabajo, proponemos un método de corrección de la radiación dispersa basado en técnicas de aprendizaje profundo, que adopta una red neuronal convolucional de arquitectura U-net con 4 bloques tanto en el codificador como en el decodificador. Debido a la falta de pares de adquisiciones reales con y sin rejilla antidifusoras, se realizaron simulaciones de Monte Carlo para generar los datos de entrenamiento. El presente estudio demuestra el potencial del método propuesto, con un error inferior al 5%.

1. Introducción

En radiología digital de tórax, la radiación dispersa reduce significativamente la calidad de la imagen en términos de disminución de la resolución de contraste, limitando seriamente el uso clínico de la imagen. El uso de rejillas antidifusoras [1], actual estándar en el entorno clínico, permite la eliminación de una parte significativa de los fotones dispersos que llegan al detector a costa de un aumento de dosis de entre 3 y 6 veces. Además, cada rejilla está optimizada para una distancia y posición específicas entre la fuente y el detector, lo que dificulta su uso para otras geometrías de adquisición, como en tomosíntesis [2] o en radiografía portátil.

La mayoría de los trabajos para la corrección de radiación dispersa se centran en la tomografía computarizada (TAC), donde se dispone de información 3D de la muestra. Enfoques basados en hardware como en los trabajos propuestos en [2-4] hacen uso de un *Beam-Stopper* (BS), una placa con agujeros de un material con un alto coeficiente de atenuación a los rayos X, que se coloca entre la fuente de rayos X y el paciente, estimándose el campo de radiación dispersa a partir de la región sombreada, que se sustraerá de los datos completos. Alternativamente, en los métodos basados en software como [5, 6], los autores proponen el uso de convoluciones con *kernels* de distinto tamaño y valor, basándose en la observación de que la radiación de dispersión se puede aproximar con una versión paso bajo de la radiación total (primaria + dispersa). La optimización del tamaño, forma y valor de los

kernels se realiza adquiriendo maniqués sencillos. Sin embargo, al no simular las estructuras complejas del cuerpo humano, la búsqueda de los *kernels* no es óptima, lo que impide la obtención de un mapa de radiación disperso perfecto. En radiografía plana, también se ha utilizado la idea del BS, basada en una corrección hardware, para la estimación de la radiación dispersa. En [4], se realizan dos adquisiciones, con y sin el BS, estimando el componente de radiación dispersa del objeto y restándolo de la imagen total. Otros trabajos utilizan una adquisición con un BS detrás del detector para adquirir la señal de dispersión [7], pero a costa de disminuir la señal a ruido de la imagen.

Los métodos software más comunes para la corrección de radiación dispersa en radiografía plana, al igual que en TAC, se basan en convolucionar la imagen original para obtener el mapa de radiación dispersa y substraerlo [8, 9]. Sin embargo, estos métodos no son capaces de eliminar completamente la radiación dispersa sin reducir la resolución de la imagen original. En los últimos años se ha sustituido la utilización de convoluciones por el uso de algoritmos de optimización basados en la norma L1 [10, 11] para preservar los bordes y mejorar el mapa de radiación dispersa obtenido. El más interesante es el propuesto por Kim et al. [12], en el que se consigue una mejora sustancial del contraste. Sin embargo, el esquema de corrección necesita de un parámetro que da mayor o menor peso según los valores de atenuación. En imagen de tórax, en los que el rango dinámico de la imagen es tan amplio, desde valores muy bajos en los pulmones a valores muy altos en la zona columna, no es posible obtener un parámetro óptimo para toda la imagen.

Los recientes avances en el campo de la inteligencia artificial y el desarrollo de las unidades de procesamiento gráfico (GPU) han proporcionado nuevas herramientas para abordar la corrección de la radiación dispersa. En TAC, varios autores [13, 14] utilizan redes neuronales convolucionales para estimar el componente de radiación dispersa y eliminarlo de las proyecciones originales. En [13], Maier et al. proponen el método *Deep Scatter Estimation* (DSE), que usa la arquitectura U-net para obtener el mapa de radiación dispersa y muestra resultados prometedores. En este trabajo, se presenta una aproximación más directa para la corrección de la radiación dispersa basada en el uso de la arquitectura U-net. En lugar

de estimar primero la radiación dispersa y restarla de la adquisición original, como en DSE [13], se obtiene directamente la imagen de radiación primaria. Los resultados preliminares muestran mejoras significativas con respecto DSE y el método que usa una optimización basada en la norma L1 de Kim et al. [12].

2. Métodos

2.1. Descripción de la red

La red utilizada es una modificación del modelo U-Net propuesto por Ronneberger et al. [15], que se compone de tres bloques principales: el codificador, el cuello de botella y el decodificador. El codificador está formado por cuatro bloques compuestos por dos capas convolucionales con un tamaño de *kernel* de 3×3, una capa de ReLU y una capa de 2×2 *max-pooling* (no se utiliza la normalización de cada conjunto ya que los datos de entrada tienen el mismo rango de valores [16]). Después de cada bloque, el número de filtros se duplica para compensar la pérdida de resolución espacial causada por las capas de *max-pooling* y aprender características más complejas. Después de la última capa de *max-pooling* del codificador, se forma el cuello de botella de la U-net con una doble capa convolucional. El decodificador está formado por cuatro bloques compuestos por una capa 2×2 de *up-sampling* concatenada con el correspondiente mapa de características de la parte codificadora y una doble capa de convolución seguida de una activación ReLU. La idea principal del decodificador es recuperar la resolución espacial original utilizando la información contextual extraída en la parte del codificador. La arquitectura U-net utilizada se muestra en la Figura 1.

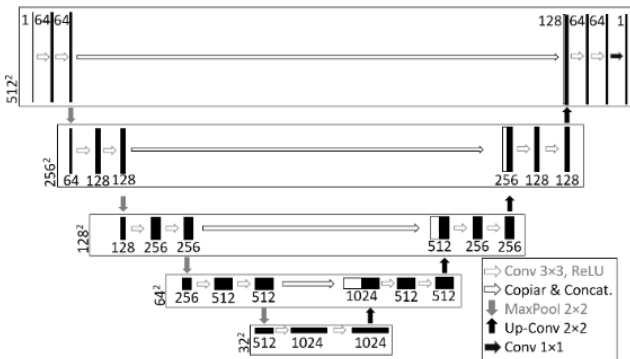


Figura 1. La arquitectura de U-net.

El número de filtros iniciales utilizados es 64, duplicándose en cada bloque subsiguiente. Para el entrenamiento de la red se usa un tamaño de lote (*batch-size*) de 8 imágenes, el optimizador Adam con una tasa de aprendizaje de 10^{-4} , y un número de épocas de 125. La función de coste utilizada durante el entrenamiento fue el error cuadrático medio (ECM).

2.2. Conjunto de datos y simulaciones

Las simulaciones se han hecho con MC-GPU v1.5b [17], basadas en la geometría estándar de adquisición de tórax, con un espectro de 120 kVp y un filtro de aluminio de 3 mm, obtenidos con la herramienta Spektr [18]. La distancia fuente-detector (DFD) se fijó en 1800 mm, con una separación entre paciente y el detector de 50 mm. El

tamaño del detector es de 512×512 píxeles con un tamaño de píxel de 0,1 mm×0,1 mm.

Las simulaciones se basan en cinco volúmenes de TAC de tórax descargados del *National Biomedical Imaging Archive* (NBIA). Estos volúmenes se convirtieron en unidades Hounsfield (HU) a mapas de densidad y se segmentaron en tres tejidos con valores de densidad de 0,001205 g/cm³ (aire), 1,060 g/cm³ (tejidos blandos) y 1,920 g/cm³ (hueso) según la base de datos del *National Institute of Standard and Technology* (NIST) [19]. Un ejemplo de esta segmentación se muestra en la Figura 2.

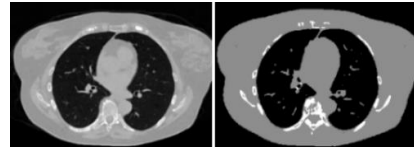


Figura 2. Corte axial de un paciente (izquierda) y segmentación del mismo corte con las correspondientes densidades asociadas (derecha).

Para el entrenamiento y evaluación de la red se generaron 82 proyecciones de cada volumen cubriendo un rango de ± 20 grados, en pasos de un grado, alrededor de las posiciones radiológicas antero-posterior (AP) y postero-anterior (PA). En total se generaron 410 proyecciones, 328 para el conjunto de entrenamiento y 82 para el conjunto de evaluación. Para cada proyección, se obtuvieron los componentes de radiación primaria, radiación dispersa y Rayleigh. A partir de estas componentes se crearon dos imágenes: total (radiación primaria + Rayleigh + radiación dispersa) y referencia (radiación primaria + Rayleigh). La radiación producida por el efecto Rayleigh se añadió a la radiación primaria para formar la imagen referencia, debido a que los ángulos de desviación son mínimos y consiguen pasar a través de las rejillas antidifusoras. En la Figura 3 se muestra un ejemplo de las imágenes de salida.



Figura 3. Imagen total (izquierda), imagen referencia (centro) y radiación dispersa (derecha).

La Figura 4 muestra la evolución del error en los conjuntos de entrenamiento y validación, que indica que la red está aprendiendo y que el número de muestras utilizadas son suficientes para el estudio preliminar llevado a cabo.

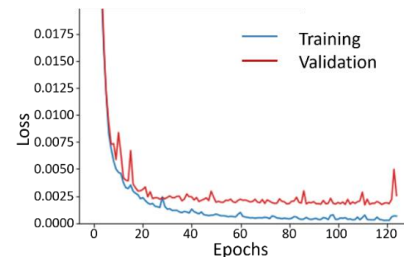


Figura 4. Errores de entrenamiento y validación de la red propuesta.

3. Evaluación

El método propuesto se comparó, usando la proyección AP central del volumen de evaluación, con dos métodos recientes: DSE [13], basado también en redes neuronales, y el método propuesto por Kim et al. [12], basado en la optimización de la norma L1. Para el método de post-procesamiento de Kim et al., usamos dos valores diferentes para el parámetro de peso, $A_{\text{Pulmones}} = 1$ and $A_{\text{Columna}} = 2.5$, para las regiones de los pulmones y la columna vertebral, respectivamente, resultando en dos imágenes corregidas diferentes. Como medidas cuantitativas se obtuvieron perfiles y la raíz del error cuadrático medio (RECM) entre las imágenes corregidas y la referencia en regiones de interés situadas en los pulmones y la columna vertebral (Figura 5).

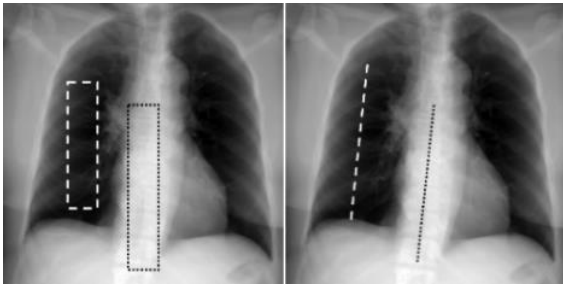


Figura 5. ROI para el cálculo del RECM (izquierda) y los perfiles elegidos para la evaluación (derecha).

4. Resultados

En la Figura 6, podemos ver un aumento de la resolución de contraste al utilizar el método propuesto, especialmente en la región de la columna vertebral, similar a la imagen referencia. También los bronquios se vuelven más visibles. De manera similar, el método DSE recupera el contraste en ambas regiones. El método de Kim et al. recupera el contraste de los pulmones y la columna vertebral, sin embargo, no es posible mejorar el contraste de ambas regiones en una misma imagen.

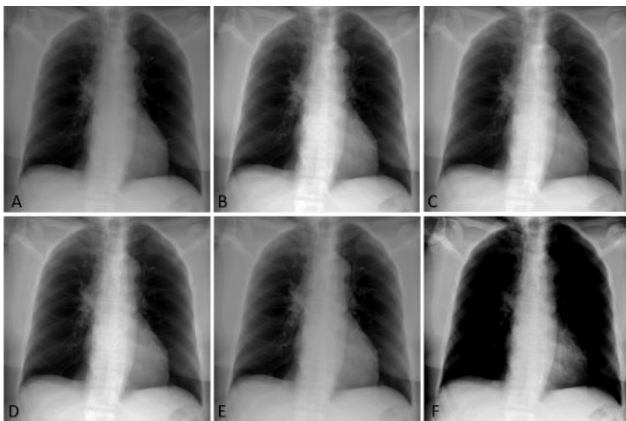


Figura 6. Imagen total (A), imagen referencia (B), imagen corregida con el método propuesto (C), imagen corregida con DSE (D), imagen corregida con el método de Kim et al. optimizado para los pulmones (E) e imagen corregida con el método de Kim et al. optimizado para la columna (F).

Estas diferencias pueden verse de forma más cuantitativa en los perfiles de que se muestran en la Figura 7.

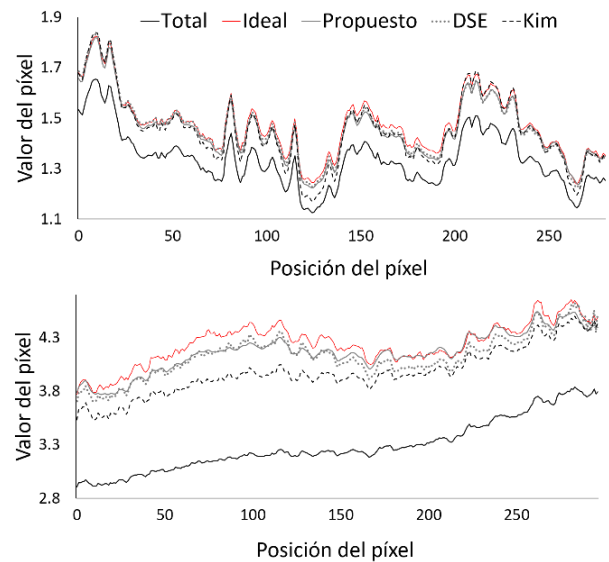


Figura 7. Perfiles trazados en los pulmones (arriba), y la columna vertebral (abajo).

La Tabla 1 muestra la RECM de todos los métodos en comparación con la imagen referencia, donde se puede ver que el método propuesto resulta en el menor error, especialmente en la región de los pulmones.

RECM	Método propuesto	DSE	Kim et al.
Pulmones	0.009	0.036	0.093
Columna	0.066	0.082	0.214

Tabla 1. Errores de los distintos métodos en comparación con la referencia.

5. Discusión y conclusiones

En este trabajo, se presenta un método de corrección de radiación dispersa en radiografía digital de tórax basado en la estimación de la radiación primaria a través de redes convolucionales con arquitectura U-net. Los resultados preliminares, en datos simulados, muestran una recuperación de la resolución de contraste similar al resultado de usar una rejilla antidisfusa. La arquitectura U-net utilizada no incorpora las normalizaciones por conjuntos de la original propuesta por Ronneberger et al. [15], eliminando un paso adicional que consume tiempo, debido a que los valores de nuestro conjunto de datos se encuentran en el mismo rango.

Comparando con el método de software propuesto por Kim et al., nuestro método logró una mejor corrección de la radiación dispersa sin necesidad de elegir una región específica, como los pulmones o la columna vertebral. El método propuesto mostró resultados similares que el método DSE, pero con la ventaja de la estimación directa de la radiación primaria, que permite un menor error de los valores obtenidos en la zona de los pulmones y la columna.

Las imágenes de radiografía plana de tórax tienen un tamaño superior a 3000×3000 píxeles con un tamaño de píxel en el detector de alrededor de 0.1 mm. Sin embargo, en este trabajo hemos utilizado imágenes de 512×512 píxeles, debido a las limitaciones actuales en la memoria

de la GPU. En futuros trabajos se hará uso de imágenes con el tamaño original, lo que podría lograrse mediante el uso de parches de la imagen para entrenar la red. Por otra parte, la red neuronal se ha entrenado con imágenes simuladas con un voltaje de fuente y una geometría de adquisición específicos. El trabajo futuro también incluye el entrenamiento con un conjunto de datos más grande y diverso, lo que podría conducir a la optimización del número de capas y filtros, reduciendo los tiempos de entrenamiento. Por último, actualmente la red se entrena con un ritmo de aprendizaje constante, siendo posible la utilización de una tasa de aprendizaje variable a lo largo de las épocas [20] en un futuro. Para reducir el error también se considerará el uso de funciones de pérdida más sofisticadas como *perceptual loss* [21] o la inclusión de arquitecturas basadas en *Generative Adversarial Networks* (GAN) [22].

6. Agradecimientos

Este trabajo ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades (Instituto de Salud Carlos III, proyecto DTS17/00122; Agencia Estatal de Investigación, proyecto DPI2016-79075-R-AEI/FEDER, UE), cofinanciado por Fondos de la Unión Europea (FEDER), “A way of making Europe”. Además, ha sido financiado por el Programa de apoyo a la realización de proyectos interdisciplinarios de I+D para jóvenes investigadores de la Universidad Carlos III de Madrid 2019-2020 en el marco del Convenio Plurianual Comunidad de Madrid- Universidad Carlos III de Madrid (proyecto DEEPCT-CM-UC3M). El CNIC está financiado por el Ministerio de Ciencia, Innovación y Universidades y la fundación PRO-CNIC y es un centro de excelencia Severo Ochoa (SEV-2015-0505).

7. Referencias

- Neitzel, U., *Grids or air gaps for scatter reduction in digital radiography: a model calculation*. Medical physics, 1992. **19**(2): p. 475-481.
- Dong, X., et al., *Low-dose and scatter-free cone-beam CT imaging using a stationary beam blocker in a single scan: phantom studies*. Computational and mathematical methods in medicine, 2013. **2013**.
- Maltz, J.S., et al., *Focused beam-stop array for the measurement of scatter in megavoltage portal and cone beam CT imaging*. Med Phys, 2008. **35**(6): p. 2452-62.
- Sakaltras, N., et al., *Scatter Correction for planar radiography with a Beam-hole array: preliminary results*, in *IEEE Nuclear Science Symposium & Medical Imaging Conference (NSS/MIC)*. 2019: Manchester.
- Star-Lack, J., et al., *Efficient scatter correction using asymmetric kernels*. 2009. **7258**: p. 72581Z.
- Sun, M. and J.M. Star-Lack, *Improved scatter correction using adaptive scatter kernel superposition*. Phys Med Biol, 2010. **55**(22): p. 6695-720.
- Lo, J.Y., et al., *Scatter compensation in digital chest radiography using the posterior beam stop technique*. Medical physics, 1994. **21**(3): p. 435-443.
- Kruger, D.G., et al., *A regional convolution kernel algorithm for scatter correction in dual-energy images: Comparison to single-kernel algorithms*. Medical physics, 1994. **21**(2): p. 175-184.
- Love, L.A. and R.A. Kruger, *Scatter estimation for a digital radiographic system using convolution filtering*. Med Phys, 1987. **14**(2): p. 178-85.
- Kim, K., et al., *A new software scheme for scatter correction based on a simple radiographic scattering model*. Medical & biological engineering & computing, 2019. **57**(2): p. 489-503.
- Meng, G., et al. *Efficient image dehazing with boundary constraint and contextual regularization*. in *Proceedings of the IEEE international conference on computer vision*. 2013.
- Kim, K.S., et al., *Improvement of radiographic visibility using an image restoration method based on a simple radiographic scattering model for x-ray nondestructive testing*. NDT & E International, 2018. **98**: p. 117-122.
- Maier, J., et al., *Deep scatter estimation (DSE): Accurate real-time scatter estimation for X-ray CT using a deep convolutional neural network*. Journal of Nondestructive Evaluation, 2018. **37**(3): p. 57.
- Nomura, Y., et al., *Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network*. Medical physics, 2019. **46**(7): p. 3142-3155.
- Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation*. in *International Conference on Medical image computing and computer-assisted intervention*. 2015. Springer.
- Ioffe, S. and C. Szegedy, *Batch normalization: Accelerating deep network training by reducing internal covariate shift*. arXiv preprint arXiv:1502.03167, 2015.
- Badal, A. and A. Badano, *Accelerating Monte Carlo simulations of photon transport in a voxelized geometry using a massively parallel graphics processing unit*. Medical physics, 2009. **36**(11): p. 4878-4880.
- Siewerdsen, J., et al., *Spektr: A computational tool for x-ray spectral analysis and imaging system optimization*. Medical physics, 2004. **31**(11): p. 3057-3067.
- Hubbell, J. and S. Seltzer, *NIST standard reference database 126*. Gaithersburg, MD: National Institute of Standards and Technology, 1996.
- Smith, L.N. *Cyclical learning rates for training neural networks*. in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 2017. IEEE.
- Johnson, J., A. Alahi, and L. Fei-Fei. *Perceptual losses for real-time style transfer and super-resolution*. in *European conference on computer vision*. 2016. Springer.
- Wolterink, J.M., et al., *Generative adversarial networks for noise reduction in low-dose CT*. IEEE transactions on medical imaging, 2017. **36**(12): p. 2536-2545.