

CONFORMADOR DE IMAGEN ULTRASÓNICA BASADO EN REDES NEURONALES

Óscar Martínez-Graullera^{1*}
Carolina Bertoncini¹
Luis Elvira Segura¹
Montserrat Parrilla Romero¹
Alberto Ibáñez Rodríguez¹

¹Instituto de Tecnología Físicas y de la Información (CSIC)
C/Serrano 144, Madrid 2006, España

RESUMEN

El presente trabajo analiza el potencial de las redes neuronales profundas para configurar un conformador de haz para imagen ultrasónica capaz de reducir las aberraciones generadas por los lóbulos secundarios. Trabajando sobre un conjunto de señales simuladas como elemento de aprendizaje, los resultados obtenidos muestran que en un proceso de imagen de apertura sintética éste es un objetivo alcanzable y que además es posible mejorar la resolución lateral.

ABSTRACT

The current investigation delves into the capabilities of deep neural networks for configuring an ultrasonic beamformer with the ability to mitigate secondary lobe structures within images. Through experimentation with a collection of simulated signals, the results obtained underscore the feasibility of achieving this objective within a synthetic aperture imaging framework. Notably, data structure, improvements in lateral resolution are also demonstrated to be attainable.

Palabras Clave— Imagen ultrasónica, apertura sintética, virtual source, high frequency ultrasound.

1. INTRODUCCIÓN

Uno de los objetivos clásicos del procesado del señal en ultrasonidos es limitar los efectos de los lóbulos secundarios generados por el patrón de difracción de la apertura. Sobre una imagen, los lóbulos secundarios generan pérdida de contraste, de resolución lateral, contribuyen al patrón de

speckles y en general producen una pérdida de calidad en la imagen que condiciona su interpretación.

Un sistema basado en arrays puede hacer uso de su potencial como filtro espacial para reducir el impacto de los lóbulos secundarios. Las soluciones más clásicas, en base a filtros convencionales, lo hacen a costa de una pérdida de la resolución lateral, lo que desvirtua la eficiencia de la apertura. Otras soluciones, basadas en filtros adaptativos, como el conformador de mínima varianza, aunque muestran cierta capacidad para resolver la imagen manteniendo resolución y rango dinámico, con excelentes resultados en un espacio simulado, suelen presentar una significativa pérdida de rendimiento en aplicaciones reales. Estas soluciones sin embargo muestran que analizando la información capturada por el array es posible encontrar una función capaz de discriminar situaciones donde el lóbulo secundario oculta la presencia de un reflector.

En un proceso de estas características tanto la cantidad de la información como la calidad de la misma es un tema relevante. Operar con un sistema phased-array de N -elementos, donde las señales están parcialmente focalizadas por el proceso de emisión significa trabajar con un conjunto muestral de N señales. Con el mismo array, emplear un sistema de apertura sintética sobre el Full Matrix Capture (FMC) implica tener $N \times N$ señales con un mayor nivel de información a costa de reducir la relación señal a ruido. Este conjunto organiza su información sobre un esquema de frecuencias espaciales, conocido como coarray, con distintos niveles de redundancia. En el caso de un array lineal de N elementos esto supone un subconjunto de $2N-1$ señales independientes [1].

Tomando un array de N elementos, donde de forma secuencial vamos disparando los i -elementos del mismo, y donde para cada disparo se reciben las señales en los j -

*Autor de contacto: primer.oscar.martinez@csic.es

Copyright: ©2023 Óscar Martínez Graullera et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

elementos ($i, j = 1, 2, \dots, N$). Podemos decir que las muestras que se corresponden por su tiempo de vuelo con el eco de un posible reflector en el punto x vienen dadas por el conjunto:

$$FMC(\vec{x}): \left\{ s_{ij}(\tau), \tau = \left| \frac{\vec{x} - \vec{x}_i}{c} \right| + \left| \frac{\vec{x} - \vec{x}_j}{c} \right|, \forall i, j = [1, N] \right\} \quad [1]$$

Dado que cada $s_{ij}(t)$ integra toda la señal de eco correspondiente a una elipse cuyos focos están en los elementos i (emisor) y j (receptor), esta señal se puede formular como:

$$s_{ij}(\tau) = m_{ij}(\vec{x}) + n_{ij}(\tau) \approx m(\vec{x}) + n_{ij}(\tau) \quad [2]$$

Donde $m_{ij}(x)$ hace referencia al eco de nuestro interés y $n_{ij}(t)$ es el eco del resto de la elipse que interfiriere en nuestra medida. Siendo x es el punto en que todas las elipses intersectan. Por conveniencia vamos a considerar ese valor de reflectividad $m_{ij}(x)$ constante para toda $s_{ij}(t)$. Entoces la reflectividad en el punto de interés puede estimarse a partir de la suma de $FMC(x)$ como:

$$\hat{m}(\vec{x}) = \left| m(\vec{x}) + \sum_{i=1}^N \sum_{j=i}^N \frac{n_{ij}(\tau)}{N^2} \right| = |m(\vec{x}) + n(\vec{x})| \quad [3]$$

Donde esta estima es un sesgo de continua introducido en todas las muestras junto con una combinación del ruido eléctrico y ruido acústico que incluye la perturbación generada por los lóbulos secundarios. Ordenando la información la matriz de emisión recepción, se observa que los lóbulos secundarios se presentan como una oscilación periódica cuya frecuencia depende de la distancia angular del foco respecto al origen de la perturbación. En este sentido, el conformador convencional actúa como un filtro paso bajo que integra toda información contenida en $FM(x)$ generando una estima robusta pero poco precisa. Obviamente, en una situación real, donde pueden intervenir múltiples fuentes, esta estimación de la reflectividad es aún menos eficiente.

En la figura 1 se representan a modo de ejemplo las 8×8 matrices de emisión/recepción de un array de 64 elementos en una rejilla espacial de 8×8 puntos alrededor de un reflector. Mientras que el ruido se representa como tal, las posiciones alrededor del reflector muestran un patrón oscilatorio de baja frecuencia que se desplaza en diagonal. Cuando este patrón está compuesto por un número impar de semiciclos se genera un lóbulo secundario.

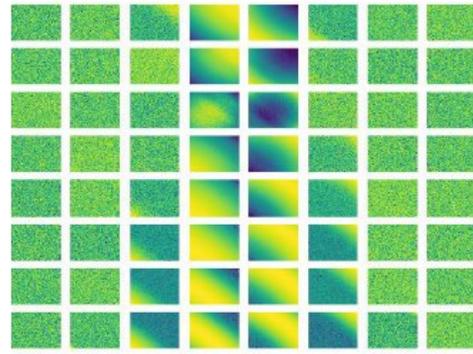


Figura 1. Las 8×8 matrices de emisión/recepción que corresponden a una rejilla espacial de 8×8 puntos.

Bajo esta interpretación, y tomando como referencia un modelo de pulso-eco basado en la respuesta espacial al impulso, simplificando el transductor como un elemento puntual [2], es posible diseñar un modelo de aprendizaje para redes neuronales que aprenda a identificar este hecho y reduzca su impacto en la imagen.

2. ESTRUCTURA DE LA RED NEURONAL

Dada la naturaleza de las redes neuronales un problema de estas características se puede plantear de diversas formas. Una opción es definir un problema de detección donde se genere un mapa de probabilidad de la presencia de blancos que pueda ser mapeado sobre la imagen convencional como una capa extra de información. Otra opción, la seguida en este trabajo, es generar una imagen con el potencial de ser sustitutiva de la imagen ultrasónica convencional donde se elimina la presencia de lóbulos secundarios [3,4,5].

2.1. La red neuronal

Nuestro diseño parte del análisis del contenido espectral del coarray para estimar la reflectividad de cada punto de la imagen. El modelo está basado en una red de 8 capas que al modelar una función regresiva se unen mediante funciones de activación lineales. La primera capa toma como entrada los componentes de más baja frecuencia del coarray y genera un total de 128 combinaciones como salida, procesando los componentes reales e imaginarios por separado se obtiene el envolvente para pasar a las siguientes capas. Estas, están formadas por 16 nodos cada una. El sistema se ha implementado con Tensorflow, empleando Python y Keras para diseñar la red e implementar las métricas de error [6]. En total, la red necesita estimar 5.889 coeficientes.

2.3. Modelo de referencia

La fase de aprendizaje se desarrolla en base a simulaciones. Usando un modelo sencillo de un array de 64 elementos puntuales, se han definido de forma aleatoria 1001 regiones de interés (o imágenes) de 64x64 puntos, donde hemos posicionado hasta un máximo de 12 reflectores distribuidos aleatoriamente y con una reflectividad aleatoria entre 0 y -30dBs. Con objeto de mejorar la casuística, garantizando siempre el correcto muestreo de la imagen, el espacio comprendido en la misma también se determina de forma aleatoria pudiendo parte de los reflectores quedar fuera de la imagen, aunque no así los lóbulos que generan.

Sobre este modelo se simulan las señales ultrasónicas correspondiente al Full Matrix Capture (FMC) para un nivel de SNR de 20dB y se determina para cada punto de la imagen las 4096 señales organizadas en matrices de emisión/recepción prefocalizadas. Los valores de las matrices se recombinan sobre el vector del coarray y se obtiene su FFT para determinar su distribución espectral. Dado que los lóbulos secundarios se encuentran en la parte más baja del espectro se valora únicamente el cuarto más bajo del rango de frecuencias. Para el caso que estamos considerando, esto son 16 coeficientes, el primero de los cuales constituye la estima del conformador DAS convencional.

Como modelo al que la solución debe converger, se diseña a partir del mismo patrón de ruido una nueva imagen donde los reflectores son modelados como campanas gaussianas generadas tomando como origen de referencia el centro del array en base a una distribución cilíndrica compatible con la naturaleza de la imagen ultrasónica (modulada axial y angularmente). El modelo mantiene la referencia de la resolución axial pero permite reducir la resolución lateral dentro de unos márgenes limitados.

3. RESULTADOS

Las figuras 2, 3 y 4 muestran un ejemplo asociado al modelo simulado. La figura 2 es la imagen ultrasónica obtenida a partir de un proceso de conformación DAS (delay-and-sum) convencional. La figura 3 es la imagen de referencia idealizada sobre la que se pretende converger. La figura 4 es el resultado de aplicar la red neuronal entrenada sobre el conjunto de datos simulados que genera la imagen 1. El resultado muestra que es posible reducir los lóbulos secundarios, aumentar la resolución lateral y mejorar el contraste de la imagen.

Para validar estos resultados con un ejemplo real, se analizaron datos tomados con una array de 64 elementos y 2.6MHz en un phantom CIRS 040GSE.

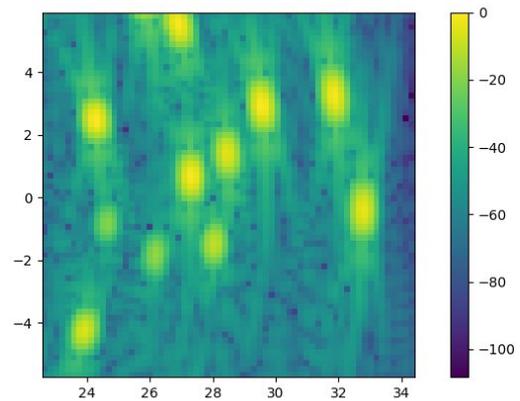


Figura 2. Imagen DAS convencional obtenida con datos simulados.

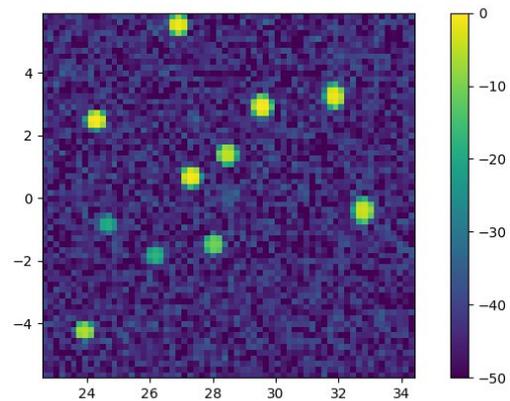


Figura 3. Imagen de referencia sobre la que se quiere converger.

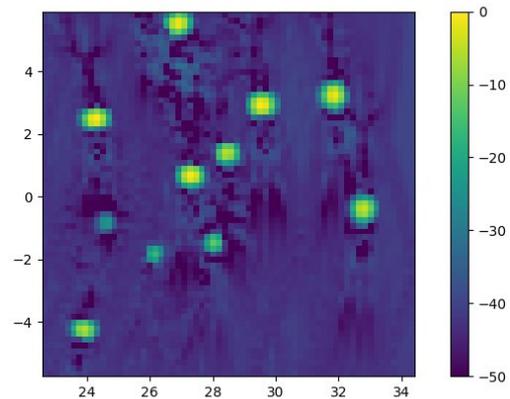


Figura 4. Imagen generada por la red neuronal.

En general todas las imágenes muestran que la red neuronal genera un mejor resultado que el DAS convencional. Incluso en un caso real, nuestra red elimina los lóbulos secundarios mejorando tanto el contraste como la resolución lateral. El resultado no afecta a la resolución axial. Un punto importante es que speckle del “tejido” mantiene los elementos principales y su reducción es un mero adelgazamiento debido al aumento de la resolución.

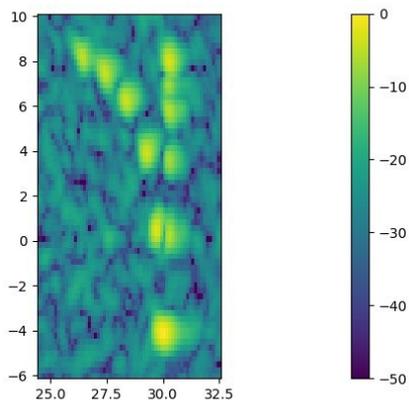


Figura 5. Imagen real obtenida con un conformador DAS.

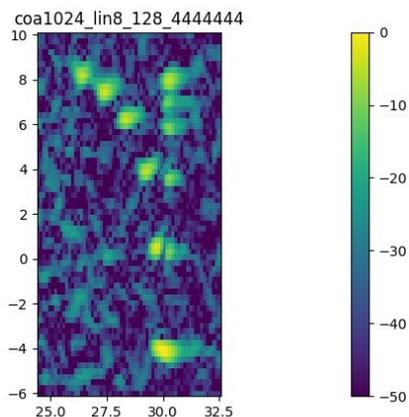


Figura 6. Imagen real obtenida con la red neuronal entrenada.

4. CONCLUSIONES

Los resultados muestran que es posible diseñar un conformador basado en redes neuronales. No obstante, aunque se obtiene una importante mejora en la calidad de la imagen, en lo que respecta a la resolución lateral y a la reducción de los lóbulos secundarios, no se ha valorado que exista un incremento del rango dinámico. De hecho, en los ejemplos simulados se ha detectado la pérdida de algunos blancos con reflectividad inferior a -30dBs. Este es un aspecto que está bajo estudio ya que nos interesa aumentar la

certidumbre en los niveles más bajos del rango dinámico teórico, que para esta apertura está en -35dB.

Desde el punto de vista de su implementación práctica, la red muestra una estructura razonablemente simple que debido a la naturaleza lineal de las funciones de activación es posible implementar como una sola operación.

5. AGRADECIMIENTOS

Los resultados de esta publicación son parte de los proyectos de I+D+i:

PID2019-111392RB-100/ MCIN/AEI/10.13039/501100011033/
PID2022-138013OB-I00/ MCIN/AEI/10.13039/501100011033/
FEDER, UE

ISABEL- 202150E058.

12. REFERENCIAS

- [1] Gordon S. Kino. Acoustic Waves: Devices, Imaging, and Analog Signal Processing. Prentice-Hall, 1987
- [2] Carlos Julián Martín Arguedas, “Técnicas de apertura sintética para la generación de imagen ultrasónica”, Tesis doctoral, Universidad de Alcalá, Escuela politécnica superior, Madrid, 2010.
- [3] D. Hyun, L. Brickson, Kevin T. Looby y Jeremy J. Dahl, “Beamforming and speckle reduction using neural networks”, IEEE Trans. on UFFC, vol. 66, núm. 5, May 2019.
- [4] Ben Luitjen, Regev Cohen, Frederik J. de Bruijn, Harold A. W. Schmeitz, Massimo Misch, Yonina C. Eldar and Ruud J. G. van Sloun, “Adaptive Ultrasound Beamforming Using Deep Learning, IEEE Trans on Medical Imaging, vol. 39, No. 12, December 2020.
- [5] Carolina Bertoncini Fernández. Diseño de un conformador de haz basado en Deep Learning para la mejora de la imagen ultrasónica. Trabajo fin de grado. Universidad Rey Juan Carlos, Madrid, 2022
- [6] Aurelien Geron. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. Ed. O'REILLY 2019