

Método híbrido de reconstrucción de resonancia magnética mediante representaciones neuronales implícitas

B. Ramírez-Martínez¹, A. Castillo-López³, L. Cordero-Grande^{1,2}

¹ Tecnología de Imágenes Biomédicas, ETSI Telecomunicación, Universidad Politécnica de Madrid (UPM)

² CIBER-BBN, ISCIII

³ Depto. Matemática Aplicada a las TIC, Information Processing and Telecommunications Center (IPTC), ETSI Telecomunicación, UPM

Resumen

Este trabajo explora un enfoque no supervisado de reconstrucción de imagen de resonancia magnética (MRI) basado en representaciones neuronales implícitas (INRs). Las INRs representan la imagen como una función continua de sus coordenadas espaciales, cuya estructura responde a una red neuronal de tipo perceptrón multicapa (MLP). En particular se estudia el rendimiento de una red de representación sinusoidal (SIREN) en reconstrucción MRI. Como contribución principal, se propone un esquema híbrido que integra la INR como regularizador explícito sobre un método clásico de reconstrucción basado en codificación de sensibilidad (SENSE). La propuesta se evalúa sobre datos sintéticos, realizando un estudio de hiperparámetros, y posteriormente sobre datos reales del dataset fastMRI, usando métricas de error cuantitativo y perceptual. Los resultados muestran que INR+SENSE mejora las reconstrucciones frente al método clásico, alcanzando valores competitivos respecto a los métodos supervisados en el public leaderboard de fastMRI.

1. Introducción

La MRI es una técnica de imagen médica que permite obtener imágenes tanto anatómicas como funcionales de alta resolución sin emplear radiación ionizante. Sin embargo, su principal limitación es el elevado tiempo necesario para adquirir imágenes de calidad diagnóstica. Submuestrear el espacio k permite acelerar la adquisición pero introduce artefactos de solapamiento espacial en la imagen reconstruida. Los métodos de imagen en paralelo como SENSE [1] usan múltiples bobinas detectoras y sus mapas de sensibilidad para deshacer este solapamiento. No obstante, su rendimiento empeora en casos de elevado submuestreo o ante errores en los mapas de sensibilidad, lo que limita en la práctica la aceleración alcanzable en la adquisición.

En los últimos años, la reconstrucción de MRI ha incorporado métodos de aprendizaje automático generalmente basados en enfoques supervisados. En ellos se entrenan redes neuronales, típicamente convolucionales, con grandes conjuntos de datos donde cada entrada tiene asociada una salida correcta. Los enfoques supervisados han demostrado resultados sobresalientes, pero presentan limitaciones como la necesidad de grandes volúmenes de datos completamente muestreados o imágenes *ground truth* (GT), difíciles o directamente imposibles de obtener en la práctica clínica. Otras limitaciones son el riesgo de sobreajuste y la falta de generalización fuera

del conjunto de entrenamiento [2].

En este contexto, los métodos no supervisados emergen como una alternativa prometedora. Particularmente, las INRs permiten modelar la imagen como una función continua a partir de los datos adquiridos en el espacio k , sin necesidad de entrenamiento previo. Las propuestas basadas en INRs para reconstrucción de MRI son muy recientes. En 2024 se publica el método NeRP [3] que emplea una red MLP con activaciones sinusoidales SIREN [4] e introduce una imagen previa del mismo sujeto como *prior*. Por otra parte, IMJENSE [5] propone combinar la reconstrucción INR con la estimación simultánea de los mapas de sensibilidad mediante ajuste polinómico, alcanzando mejoras sobre métodos clásicos en escenarios de alto submuestreo.

A diferencia de los anteriores, en nuestro trabajo prescindimos de imágenes *prior*, lo que aumenta su aplicabilidad clínica y asumimos mapas de sensibilidad dados, centrándonos exclusivamente en el problema de reconstrucción. Como contribución novedosa, estudiamos la fusión de reconstrucción mediante INRs con el método clásico SENSE.

2. Métodos

La reconstrucción de MRI puede formularse matemáticamente como un problema inverso donde se estima la imagen $\hat{\mathbf{x}}$ a partir de las observaciones \mathbf{y} , adquiridas en el espacio k . La estimación se plantea como la minimización del error cuadrático medio:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2, \quad (1)$$

donde \mathbf{x} representa la imagen a estimar y \mathbf{A} el operador físico de adquisición, el cual incluye los mapas de sensibilidad, la transformada de Fourier y el submuestreo aplicado. Cuando la matriz \mathbf{A} es conocida y bien condicionada, la reconstrucción puede realizarse de la siguiente manera:

$$\hat{\mathbf{x}} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{y} \quad (2)$$

Donde \mathbf{A}^H representa la transpuesta conjugada de \mathbf{A} . Sin embargo, en escenarios con submuestreo, este problema se vuelve mal condicionado y requiere técnicas adicionales para mejorar los resultados.

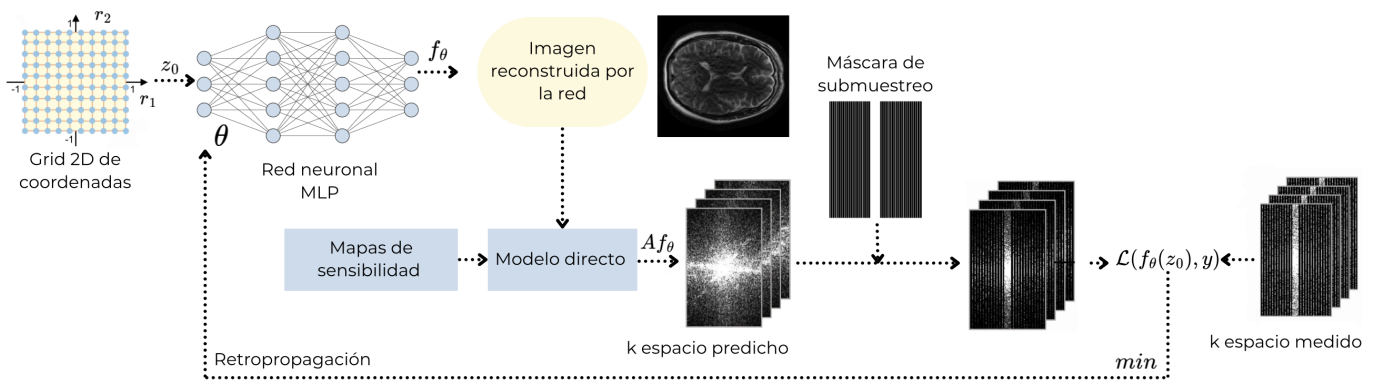


Figura 1. Diagrama de funcionamiento del método de reconstrucción basado en INRs.

2.1. INRs

Las INRs se basan en una red neuronal MLP completamente conectada que representa la imagen como una función continua de sus coordenadas espaciales tal que $\mathbf{x} = f_{\theta}(\mathbf{r})$ (véase la Figura 1). En imagen 2D, la entrada de la red, denotada por \mathbf{r} , corresponde al vector de coordenadas espaciales en el plano de la imagen. Este vector se toma de un *grid* bidimensional $\mathbf{r} = (r_1, r_2)$ que indexa los píxeles de la imagen.

A la entrada de la red se usa codificación por características de Fourier, la cual proyecta las coordenadas de entrada a un espacio de mayor dimensión mediante una matriz aleatoria B , cuyas entradas se muestrean de una distribución gaussiana $\mathcal{N}(0, \gamma)$, aplicando funciones seno y coseno:

$$\mathbf{z}_0 = [\sin(2\pi B\mathbf{r}), \cos(2\pi B\mathbf{r})]. \quad (3)$$

A continuación, la salida \mathbf{z}_0 se procesa mediante una secuencia de capas ocultas SIREN, que utilizan funciones de activación sinusoidales. Cada capa i aplica la transformación:

$$\mathbf{z}_{i+1} = \phi_i(\mathbf{z}_i), \quad \text{con } \phi_i(z) = \sin(W_i z + b_i) \quad (4)$$

donde W_i y b_i son los pesos y sesgos de la capa i -ésima. Finalmente, la salida de la red se obtiene mediante una capa lineal. A su vez, se emplea una inicialización específica de los pesos para garantizar estabilidad en el entrenamiento.

Los hiperparámetros clave de la red son:

- Depth (D): número de capas ocultas del MLP, que influye en su capacidad de modelado no lineal.
- Width (W) y Embedding size (E): correspondientes respectivamente al número de neuronas por capa y a la dimensión de la primera capa de la red. Determinan la complejidad y capacidad representativa de la red.
- Modelscale (γ): parámetro clave que controla la dispersión espectral o frecuencias espaciales representables por la red a partir de la codificación de entrada.

2.2. Reconstrucción mediante INRs

Usando INRs, la reconstrucción se plantea como la búsqueda de los parámetros de la red $\hat{\theta}$ que minimicen el error cuadrático medio entre las medidas adquiridas en el *espacio* k , \mathbf{y} , y

el ajuste realizado por el modelo, $\hat{\mathbf{y}}_{\theta} = \mathbf{A}f_{\theta}(\mathbf{r})$, con \mathbf{A} el operador de adquisición (ver Figura 1). Formalmente:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \|\mathbf{A}f_{\theta}(\mathbf{r}) - \mathbf{y}\|_2^2. \quad (5)$$

2.3. Fusión INR-SENSE

Además del uso directo de la INR como procedimiento para resolver el problema de reconstrucción propuesto en métodos previos, en nuestro trabajo se estudia la integración de la INR como regularizador explícito de Tikhonov dentro de esquema de reconstrucción SENSE:

$$(\hat{\mathbf{x}}, \hat{\theta}) = \underset{\mathbf{x}, \theta}{\operatorname{argmin}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \|\mathbf{x} - f_{\theta}(\mathbf{r})\|_2^2. \quad (6)$$

Esta combinación pretende aprovechar tanto la fidelidad radiológica del método SENSE como la flexibilidad representativa de las INRs.

2.4. Conjunto de datos

Los datos reales utilizados provienen del proyecto NYU fastMRI Initiative¹, que proporciona un entorno estandarizado para evaluar algoritmos de reconstrucción MRI a partir de datos submuestreados [6]. El dataset utilizado en este trabajo es el conjunto *Multi-Coil Brain*, que incluye 6.970 estudios completos obtenidos con distintos escáneres 1.5T y 3T y múltiples secuencias: T1, T1 PRE/POST, T2 y FLAIR.

Los datos se organizan en bloques de entrenamiento, validación y test. En este trabajo se utilizaron 10 volúmenes del bloque de validación, con factor de aceleración $R = 4$ (lo cual corresponde a reducir a 1/4 el tiempo de adquisición). El patrón de submuestreo es cartesiano, con adquisición completa en el centro del *espacio* k y muestreo uniforme en los extremos. Las imágenes de referencia fueron generadas mediante *Root-Sum-of-Squares* (RSoS). Nuestro trabajo usa la tabla de resultados de fastMRI como marco comparativo para situar el rendimiento del modelo INR en relación a los métodos supervisados del *leaderboard*. Al tratarse de un reto cerrado, no fue posible seguir los canales oficiales de validación. Por ello, se usaron los resultados de reconstrucción

¹<https://fastmri.med.nyu.edu>

zero-filled (ZF) como referencia con la que establecer una comparativa aproximada con los resultados del *leaderboard*.

2.5. Implementación y validación

El ajuste de la red se realiza en PyTorch y la configuración experimental se gestiona mediante archivos YAML. En todos los casos se empleó el optimizador ADAM y una estrategia de *learning rate* adaptativo que hace uso de la dinámica de la función de pérdida. Cada imagen fue procesada de forma independiente, entrenando la red desde cero. Las ejecuciones se realizaron remotamente en un servidor de cómputo equipado con una GPU NVIDIA RTX 4090 (24 GB VRAM).

Como paso inicial, se realizó un estudio sistemático de los hiperparámetros que definen la arquitectura INR, con el objetivo de identificar la configuración óptima para un factor de aceleración $R = 4$. Para ello, se utilizó una simulación de imagen MRI paralela obtenida de un tutorial sobre SENSE². La calidad de las reconstrucciones se evaluó según la métrica PSNR, que mide el error en decibelios respecto a la referencia, y la métrica SSIM, que cuantifica la similitud estructural.

3. Resultados

3.1. Datos sintéticos

En la Tabla 1 se expone la configuración óptima hallada para datos sintéticos así como el rendimiento obtenido en comparación con el de la configuración inicial propuesta en [3]. El ajuste del parámetro γ resultó determinante ya que disminuir su valor se tradujo en mejoras de hasta 4 dB en PSNR.

Seguidamente se evaluó el rendimiento del modelo para distintos factores de aceleración, $R = 1$ a 5. Como se observa en la Figura 2, INR superó consistentemente a SENSE, especialmente en condiciones más exigentes donde la ganancia relativa fue mayor.

3.2. Datos reales

Una vez ajustados los hiperparámetros del modelo para datos reales sobre un caso independiente, se aplicó el modelo a 10 volúmenes reales y se comparó la reconstrucción ZF, el método clásico SENSE, el ajuste directo mediante INR y el enfoque híbrido INR+SENSE propuesto.

Los valores de las métricas en reconstrucción se condensan en la Tabla 2. En primer lugar, INR como modelo independiente muestra un $\Delta\text{PSNR} = 3.83$ dB y $\Delta\text{SSIM} = 0.026$ frente a ZF, lo que se traduce en una ligera mejora respecto a SENSE (0.31 dB en PSNR). Sin embargo, la eficacia del método depende de la secuencia empleada, en volúmenes con contraste FLAIR no se obtuvo mejora. En contraste, la combinación INR+SENSE alcanza una media de 35.61 dB en PSNR y 0.903 en SSIM, lo que supone una diferencia media de $\Delta\text{PSNR} = 5.21$ dB y $\Delta\text{SSIM} = 0.021$ sobre ZF, y una ganancia de casi 2 dB frente al método clásico.

Este comportamiento se explica porque los errores cometidos

Configuración	E	W	D	γ	PSNR	SSIM
Inicial	96	96	8	10.666	26.12	0.7062
Óptima	64	48	6	0.5	32.56	0.8392

Tabla 1: Configuraciones inicial y óptima.

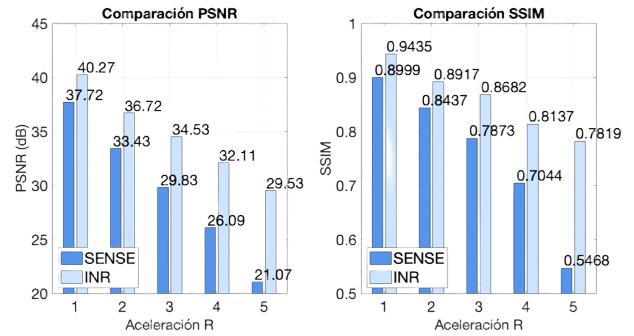


Figura 2. PSNR y SSIM SENSE frente a INR con distintas R .

por INR y SENSE son de naturaleza distinta y, al integrarlos en un esquema híbrido se compensan parcialmente. Así, los volúmenes reconstruidos con INR eliminan artefactos de solapamiento característicos de SENSE, mientras que SENSE preserva mejor el contraste radiológico (ver Figura 3).

Además, al comparar los resultados con los métodos presentes en el *leaderboard* del reto fastMRI, la propuesta INR+SENSE se posiciona por encima de 5 soluciones supervisadas entre las cuales están *MONAI U-Net* o *IR_FRestormerF72* y por detrás de arquitecturas más avanzadas como *AIRS-Net*³. Nuestro enfoque es no supervisado y se ajusta de forma independiente para cada imagen. Ello favorece la generalidad pero limita el rendimiento frente a redes supervisadas optimizadas para el *benchmark*. No obstante, los resultados evidencian el potencial de los métodos no supervisados y, sobre todo, de enfoques híbridos como el presentado.

4. Discusión

Los resultados obtenidos indican que el modelo INR ofrece una alternativa no supervisada capaz de competir con métodos supervisados en reconstrucción de MRI. Una de sus principales fortalezas es que mejora a SENSE en escenarios de mayor aceleración, como se observa en la Figura 2. En datos reales, la INR aplicada de forma independiente supera ligeramente a SENSE (0.3 dB en PSNR). Los errores cometidos por la INR son complementarios a los de SENSE, lo que motiva su uso como regularizador explícito en un esquema de Tikhonov, donde la combinación INR+SENSE logra una mejora de casi 2 dB sobre SENSE y alcanza un rendimiento competitivo frente a alternativas supervisadas en el *leaderboard* de fastMRI. En las soluciones INR y INR+SENSE observamos que la SSIM desciende muy ligeramente respecto a SENSE, mientras que se produce una mejora de la PSNR. Dado que la SSIM se calcula promediando similitudes en ventanas locales, puede verse influida por estructuras

²<https://github.com/mchiew/SENSE-tutorial.git>

³<https://web.archive.org/web/20230324102125/https://fastmri.org/leaderboards>

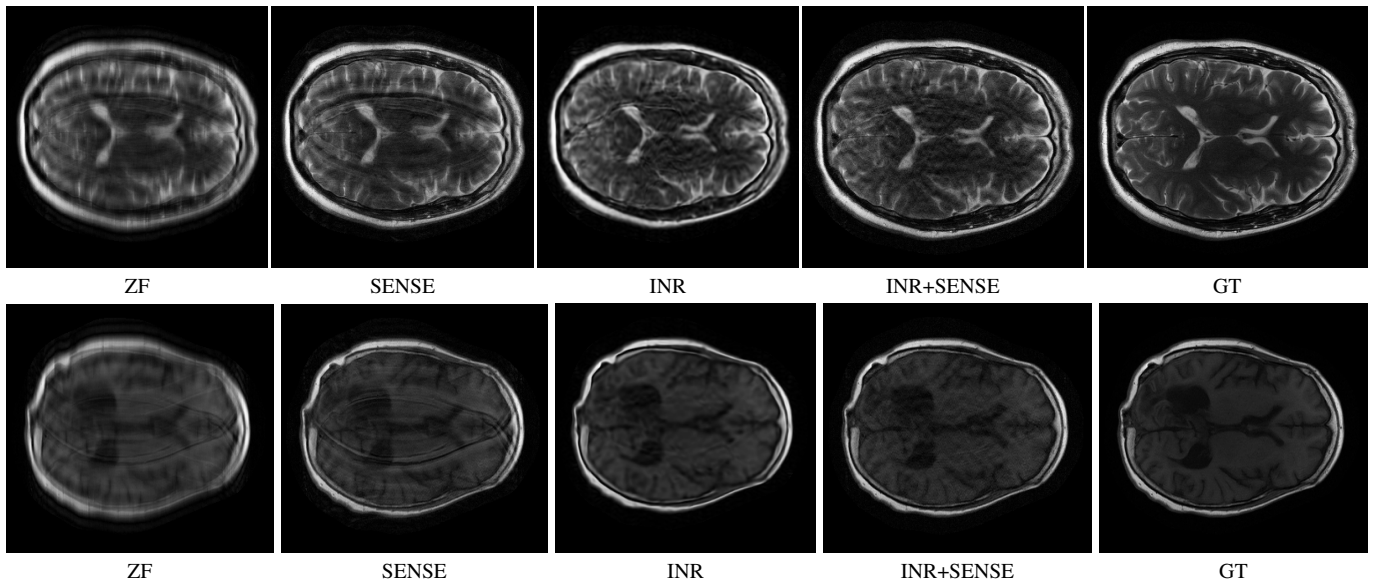


Figura 3. Resultados ejemplo sobre imágenes pesadas por T2 (panel superior) y T1 (panel inferior) junto a referencia GT.

Métrica	ZF	SENSE	INR	INR+SENSE
PSNR	30.40 ± 1.3	33.83 ± 2.1	34.23 ± 2.6	35.61 ± 2.7
SSIM	0.882 ± 0.02	0.912 ± 0.01	0.908 ± 0.03	0.903 ± 0.02

Tabla 2: Resultados medios en 10 volúmenes cerebrales reales de fastMRI ($R=4$).

de alto contraste pero limitado interés radiológico. Por ello se usa PSNR como métrica principal.

De cara al desarrollo futuro se plantean dos líneas principales. Por un lado la incorporación de INR como regularizador explícito en esquemas de reconstrucción basados en ADMM [7], lo que permitiría desacoplar el entrenamiento de la red del operador de adquisición. Por otro la optimización de la codificación de coordenadas de entrada, sustituyendo la inicialización aleatoria de características de Fourier por distribuciones espectrales más representativas de las imágenes a reconstruir.

5. Conclusiones

Este trabajo presenta y evalúa un método no supervisado de reconstrucción de MRI basado en INRs usando tanto datos sintéticos como reales. Se ha observado que integrar la INR como regularizador explícito aporta mejoras sustanciales en PSNR frente a SENSE y posiciona la propuesta híbrida de manera competitiva frente a los métodos supervisados. En conjunto, los resultados respaldan la viabilidad de las INRs como alternativa no supervisada para la reconstrucción MRI y su valor como regularizador en esquemas clásicos.

Agradecimientos

Beca Leonardo de Investigación Científica y Creación Cultural 2025 de la Fundación BBVA. La Fundación BBVA no se responsabiliza de las opiniones, comentarios y contenidos incluidos en el proyecto y/o los resulta-

dos derivados del mismo, los cuales son total y absoluta responsabilidad de sus autores. Proyectos PID2021-129022OA-I00, PID2024-162095OB-I00 financiados por MICIU/AEI/10.13039/501100011033/FEDER, UE.

Referencias

- [1] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger, “SENSE: Sensitivity encoding for fast MRI,” *Magnetic Resonance in Medicine*, vol. 42, pp. 952–962, 1999.
- [2] V. Antun, F. Renna, C. Poon, B. Adcock, and A. C Hansen, “On instabilities of deep learning in image reconstruction and the potential costs of AI,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117(48), p. 30088–30095, Dec 2020.
- [3] L. Shen, J. Pauly, and L. Xing, “NeRP: Implicit neural representation learning with prior embedding for sparsely sampled image reconstruction,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, pp. 770–782, Jan 2024.
- [4] V. Sitzmann, J. N. P. Martel, A. W. Bergman, D. B. Lindell, and G. Wetzstein, “Implicit neural representations with periodic activation functions,” *Proceedings of the 34th International Conference on Neural Information Processing Systems*, vol. 626, pp. 7462–7473, Dec 2020.
- [5] R. Feng, Q. Wu, J. Feng, H. She, C. Liu, Y. Zhang, and H. Wei, “IMJENSE: Scan-specific implicit representation for joint coil sensitivity and image estimation in parallel MRI,” *IEEE Transactions on medical imaging*, vol. 43(4), pp. 1539–1553, 2024.
- [6] F. Knoll, J. Zbontar, A. Sriram, M. J. Muckley, M. Bruno, A. Defazio, and et al., “fastMRI: A publicly available raw k-space and dicom dataset of knee images for accelerated MR image reconstruction using machine learning,” *Radiology: Artificial Intelligence*, vol. 2, no. 1, 2020.
- [7] M. Najaf and G. Ongie, “Accelerated optimization of implicit neural representations for CT reconstruction,” *arXiv*, 2025. <https://arxiv.org/abs/2504.13390>.