

Mejoramiento de la resolución espacial en imágenes dMRI mediante la aplicación de arquitecturas de aprendizaje profundo

E. A. Salas Medina *
J. M. Cuellar Borrero *

*Pontificia Universidad Javeriana Cali, Colombia**
edixonsalasmedina@javerianacali.edu.co
juanma1909@javerianacali.edu.co

Resumen

Las imágenes médicas, como la resonancia magnética por difusión (dMRI), son vitales para el diagnóstico médico. Sin embargo, estas imágenes tienen limitaciones en su resolución espacial y relación señal-ruido debido a las restricciones de tiempo de uso de los equipos de escaneo utilizados. Para abordar este problema, se propone utilizar el Aprendizaje Profundo (Deep Learning) para aplicar la Super Resolución a las imágenes dMRI. Se utilizan tres tipos de redes neuronales: Super-Resolution Convolutional Neural Network (SRCNN), Enhanced Deep Super-Resolution Network (EDSR) y Very Deep Super Resolution Network (VDSR), ampliamente utilizadas en sistemas de visión por computadora. Estas redes neuronales son entrenadas para predecir imágenes de alta resolución a partir de imágenes de baja resolución. El objetivo final es obtener imágenes médicas más nítidas y detalladas, lo que puede mejorar significativamente el diagnóstico clínico y la toma de decisiones médicas. En particular, el uso de la inteligencia artificial en este contexto muestra un gran potencial que se ve reflejado en los resultados para las 3 redes propuestas y destacándose la SRCNN con mejores métricas y tiempos de entrenamiento más cortos.

1. Introducción

En las últimas décadas, los avances tecnológicos han mejorado significativamente la obtención de diagnósticos médicos, gracias al aumento del poder computacional y la capacidad de almacenar grandes volúmenes de datos. Estos avances han impulsado el uso de técnicas como el deep learning para realizar predicciones y clasificaciones de datos, impactando positivamente en el campo de las imágenes médicas.

Las imágenes médicas, como la resonancia magnética, la tomografía computarizada, los ultrasonidos y los rayos X, han revolucionado la medicina al permitir estudiar el cuerpo humano sin procedimientos invasivos. En particular, la imagen de resonancia magnética por difusión (dMRI) es destacada por su excelente contraste en materia blanca y su utilidad en la caracterización de tumores, isquemia cerebral y trastornos neurológicos[1].

Sin embargo, la dMRI presenta limitaciones en cuanto a la resolución espacial y la relación señal-ruido (SNR) debido a las restricciones de hardware en los escáneres actuales, lo que afecta la calidad de los diagnósticos clínicos[2]. Por lo tanto, es necesario mejorar la resolución espacial de estas imágenes para obtener resultados más precisos.

Este proyecto propone una solución basada en la aplicación de métodos de deep learning para mejorar la resolución espacial de las imágenes dMRI y así lograr un mejor diagnóstico clínico y una identificación temprana de diversas afecciones médicas.

2. Fundamentación teórica

- Super Resolución:** Esta se define como el proceso mediante el cual se obtiene una imagen de alta resolución I_{HR} partiendo de una imagen de baja resolución I_{LR} , dicho proceso genera una imagen estimada I_E la cual se aproxima a la imagen de alta resolución I_{HR} . En este caso g sería la función de super resolución y θ serían los parametros recibidos por dicha función, como podría ser el factor de escalamiento[3].

$$g(I_{LR}, \theta) = I_E \approx I_{HR}$$

- DTI:** Es una de las modalidades de dMRI, que existen para modelar la difusión (movimiento browniano) de las moléculas de agua en los tejidos del cuerpo humano, mediante el uso de tensores, los cuales se describen matemáticamente como una matriz simétrica 3×3 positiva definida, la cual esta asociada a un voxel (pixel volumétrico) de la imagen. Esta matriz es capaz de describir la difusión anisotrópica ajustandolo a un elipsoide, donde el eje mayor de difusión es el eje mayor del elipsoide[4].

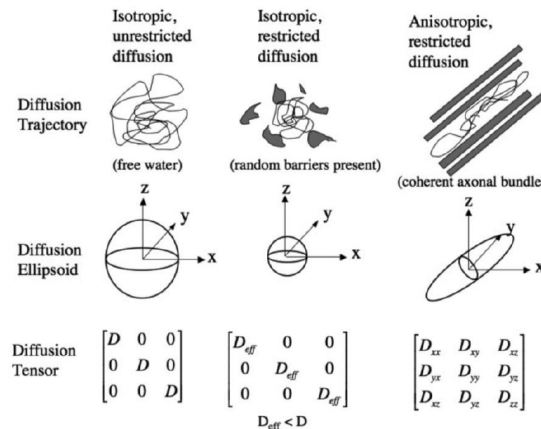


Figura 1: Representación de los tensores de difusión. *Imagen tomada de:* [4]

- Anisotropía fraccional:** Es una medida que permite calcular el grado de direccionalidad intravoxel, utilizando la información que se encuentra en los tensores de difusión[4], dados los 3 ejes principales del tensor $\lambda_1, \lambda_2, \lambda_3$ se puede calcular la anisotropía fraccional (FA) por la formula:

$$FA = \frac{\sqrt{(\lambda_1 - \lambda_2)^2 + (\lambda_2 - \lambda_3)^2 + (\lambda_3 - \lambda_1)^2}}{\sqrt{2}\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}$$

3. Resultados

Para resolver el problema, se utilizaron tres técnicas de deep learning: SRCNN, VDSR y EDSR, todas ellas enfocadas en la tarea de super resolución para aumentar la resolución espacial de imágenes[5, 6]. Se seleccionó un dataset de 858 imágenes, cada una con tamaño 256x256 pixeles en escala de grises. Además, se definió inicialmente un valor de batch size de 32 para aprovechar la GPU que ofrece el servicio de Google Colab, y con cada una de las redes se probó con diferentes números de épocas hasta hallar el valor óptimo que no presentará problemas de overfitting o underfitting. Adicionalmente, se observó que las imágenes contenían mucho ruido fuera del área del cerebro, lo que podría afectar negativamente los resultados de las métricas. a raíz de esto, se decidió crear y aplicar una máscara de recorte para eliminar la información que se encuentra fuera del área de interés, como se puede ver en la figura 2. Los cuadros 1 y 2 muestran las métricas obtenidas sin la utilización del recorte y habiendo aplicado el recorte respectivamente. cabe resaltar que le cuadro 2 tiene las métricas tensoriales FNORM y MSE FA pues fue una vez aplicada la máscara que se evualuaron las imágenes a nivel tensorial.

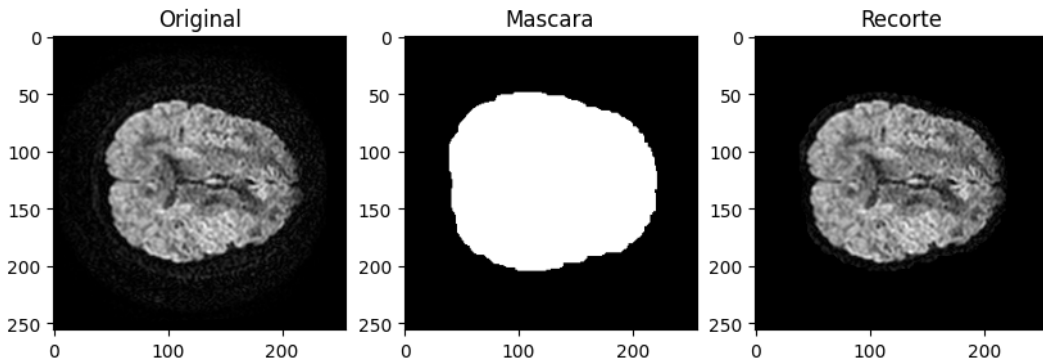


Figura 2: Creación de la máscara

Métrica	SRCNN	VDSR	EDSR
SSIM	$9,63 \times 10^{-1}$	$9,65 \times 10^{-1}$	$7,99 \times 10^{-1}$
MSE	$4,09 \times 10^{-7}$	$3,90 \times 10^{-7}$	$2,35 \times 10^{-1}$
RMSE	$6,23 \times 10^{-4}$	$6,10 \times 10^{-4}$	$4,65 \times 10^{-1}$
PSNR	$1,124 \times 10^2$	$1,13 \times 10^2$	$5,51 \times 10$

Cuadro 1: Comparación de métricas de imágenes para los modelos SRCNN, VDSR y EDSR sin aplicar la mascara de recorte.

3.1. Comparación de los Modelos

Se compararon los resultados de las 3 técnicas utilizando diversas métricas en sus mejores valores de época de entrenamiento: 100 para SRCNN, 500 para VDSR y 2000

Métrica	SRCNN	VDSR	EDSR
SSIM	$9,83 \times 10^{-1}$	$9,85 \times 10^{-1}$	$9,27 \times 10^{-1}$
MSE	$2,05 \times 10^{-7}$	$1,73 \times 10^{-7}$	$1,79 \times 10^{-1}$
RMSE	$4,38 \times 10^{-4}$	$4,05 \times 10^{-4}$	$4,23 \times 10^{-1}$
PSNR	$1,154 \times 10^2$	$1,16 \times 10^2$	$0,562 \times 10^2$
FNORM	$1,16 \times 10^{-4}$	$1,10 \times 10^{-4}$	$7,90 \times 10^{-2}$
MSE FA	$1,34 \times 10^{-2}$	$1,28 \times 10^{-2}$	$1,91 \times 10^{-1}$

Cuadro 2: Comparación de métricas de imágenes para los modelos SRCNN, VDSR y EDSR usando la mascara de recorte

para EDSR. Tanto SRCNN como VDSR mostraron excelentes resultados en términos de calidad de imagen. SRCNN obtuvo un coeficiente SSIM de 0.983 y un MSE de $2,05 \times 10^{-7}$. VDSR logró un coeficiente SSIM ligeramente superior de 0.985 y un MSE de $1,73 \times 10^{-7}$ lo que se traduce en imágenes de alta calidad y con gran similitud en estructura y contenido a las reales. Ambos modelos produjeron imágenes con detalles bien reproducidos, como se puede ver en la figura 3. Además, VDSR tuvo una ligera ventaja numérica con un menor MSE como se puede ver en el cuadro 2. Por otro lado, el modelo EDSR no alcanzó la misma calidad que las otras 2 técnicas. Aunque su coeficiente SSIM fue alto ($9,27 \times 10^{-1}$), su MSE fue significativamente más alto ($1,79 \times 10^{-1}$), lo que indica pérdida de información. En términos de métricas tensoriales, SRCNN y VDSR tuvieron resultados similares como se puede ver en la figura 4, con valores de FNORM de $1,16 \times 10^{-4}$ y $1,10 \times 10^{-4}$, respectivamente.

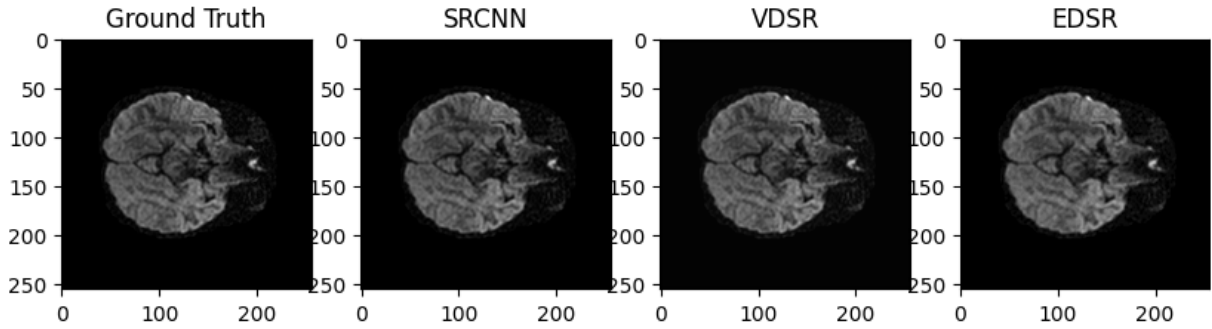


Figura 3: Comparación de resultados de todos los modelos

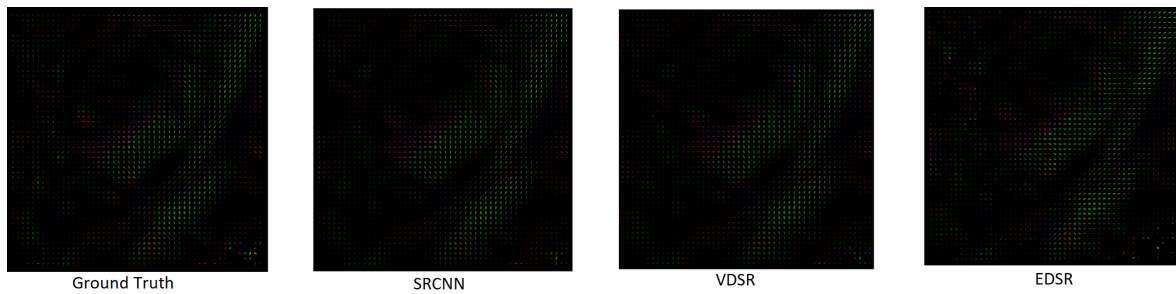


Figura 4: Comparación de resultados mapas de anisotropía fraccional de todos los modelos

SRCNN	VDSR	EDSR
72,804	2316,302	1007,573

Cuadro 3: Comparación de tiempos en segundos de entrenamiento de los modelos SRCNN, VDSR y EDSR.

4. Conclusión

Este proyecto logró abordar un problema altamente estudiado, como lo es la super resolución de imágenes médicas, aplicando modelos que, si bien son muy utilizados en el campo de super resolución, no necesariamente lo son en el ámbito de las imágenes médicas, especialmente con imágenes dMRI, logrando así estudiar su desempeño.

En primer lugar, se evaluaron y compararon tres modelos bastante utilizados en este campo: SRCNN, VDSR y EDSR. A raíz de los resultados obtenidos, se puede concluir que tanto SRCNN como VDSR demostraron ser modelos altamente efectivos en la mejora de la resolución de las imágenes dMRI, obteniendo resultados muy similares tanto a nivel de imagen como a nivel tensorial, además de esto, se logró mostrar empíricamente el hecho de que el utilizar una máscara en las imágenes para aislar la zona de interés, en este caso el cráneo, mejora notablemente las métricas de desempeño de los modelos evaluados, esto gracias a que ya no se toma en cuenta gran parte del ruido generado.

De igual manera, vale la pena resaltar el hecho de la incapacidad humana de evaluar imágenes con gran detalle, pues si bien a simple vista 2 imágenes pueden parecerse, es cuando se va analizar a un nivel mayor de detalle que salen a la superficie los errores de reproducción.

Este proyecto puede tomarse como una base para realizar un estudio más exhaustivo acerca de los diferentes modelos de *Deep Learning* para Super resolución ya existentes, tomando en cuenta una mayor variedad de modelos como los presentados por Bashir en [3], donde se pueden ver modelos de *post-upsampling*, *progressive-upsampling*, entre otros. De igual manera se sugiere usar técnicas de entrenamiento para la EDSR, entrenarla con un conjunto de datos más grande o explorar otras arquitecturas de aprendizaje profundo para mejorar aún más la resolución espacial en imágenes dMRI, así como un posible proceso de optimización de hiperparámetros de las redes mediante el uso de la API de Keras llamada *KerasTuner*, la cual permite definir un modelo de tal forma que sus hiperparámetros sean un espacio de búsqueda, para que a la hora de entrenar se realice una exploración de dicho espacio en diferentes *trials*.

Referencias

- [1] Q. Tian, B. Bilgic, Q. Fan, C. Liao, C. Ngamsombat, Y. Hu, T. Witzel, K. Setsompop, J. R. Polimeni, and S. Y. Huang, “DeepDTI: High-fidelity six-direction diffusion tensor imaging using deep learning,” *NeuroImage*, vol. 219, no. January, p. 117017, 2020.
- [2] H. D. Vargas, “Probabilistic modeling of tensorial data for enhancing spatial resolution in magnetic resonance imaging,” 2018.

- [3] S. M. A. Bashir, Y. Wang, M. Khan, and Y. Niu, “A comprehensive review of deep learningbased single image super-resolution,” *PeerJ Computer Science*, vol. 7, pp. 1–56, 2021.
- [4] A. Mar, E. A. Tutor, R. Estepa, and A. Co, “Protocolo para analizar diffusion tensor imaging (dti) del cerebro humano,” 2016.
- [5] C. Dong, C. C. Loy, K. He, and X. Tang, “Image Super-Resolution Using Deep Convolutional Networks,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295–307, 2016.
- [6] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, “Enhanced Deep Residual Networks for Single Image Super-Resolution,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, vol. 2017-July, pp. 1132–1140, 2017.