

# Red Neuronal Cristalina: aplicación al procesamiento de neuroimágenes ponderadas por difusión

Guido O. Pascariello y Patricio A. Donnelly-Kehoe

CIFASIS, Rosario, Santa Fe, Argentina  
guido.pascariello@gmail.com,  
<http://cifasis-conicet.gov.ar/>

**Resumen** Las imágenes por difusión son un tipo de imágenes por resonancia magnética especialmente sensibles a la dirección de difusión de las moléculas de agua. Utilizándolas es posible determinar el movimiento de dichas moléculas en base a modelos biofísicos. Esta es una técnica muy utilizada para estudiar la microestructura del tejido nervioso y la conectividad cerebral.

En este artículo presentamos un nuevo modelo de difusión, inspirado en las redes neuronales artificiales, que propone una arquitectura de red cristalina. Se presenta el modelo junto a una propuesta de método para ajustarlo. Para evaluar el desempeño se desarrollaron pruebas sobre fantasmas virtuales que imitan posibles disposiciones de fibras nerviosas.

El modelo propuesto convergió en el ajuste sobre los fantasmas y mostró potencial para resolver disposiciones que otros modelos estándares de difusión no pueden, como cruces y curvas en las fibras.

**Keywords:** Imágenes por difusión, Redes neuronales artificiales, Arquitectura de Red Cristalina, Conectividad cerebral

## 1. Introducción

Las imágenes por difusión o *Diffusion Weighted Images* (DWI) son un tipo de imágenes por resonancia magnética (MRI) especialmente sensibles a la dirección de difusión de las moléculas de agua. El resonador magnético adquiere estas imágenes utilizando una secuencia que incluye un par de pulsos de sincronización y resincronización en un plano específico, sensibilizando así la dirección perpendicular a dicho plano. La señal recuperada, dependerá de la dirección en que difundan las moléculas de agua; en los vóxeles en los que el agua se mueve paralela al plano seleccionado, ésta no es afectada por los pulsos, mientras que si el agua se mueve perpendicular a dicho plano la señal es atenuada. Por lo tanto, los vóxeles con difusión en direcciones diferentes a las mencionadas devolverán una señal proporcional a la componente perpendicular al plano [1, 2].

El agua contenida en el tejido experimenta un movimiento Browniano, pero el comportamiento neto puede ser isotrópico o anisotrópico de acuerdo a las restricciones que le imponga la microestructura del tejido [2]. Específicamente

en el cerebro, desde el punto de vista microestructural, se pueden identificar tres tejidos: materia gris, materia blanca y líquido cefalorraquídeo. La materia blanca esta compuesta por los axones neuronales, estructuras tubulares estrechas que restringen el movimiento de las moléculas de agua en la dirección axial de dicho tubo. Utilizando DWI, adquiriendo múltiples imágenes sensibilizadas con diferentes direcciones, es posible obtener una descripción espacial del movimiento de las moléculas de agua en este tejido.

Los parámetros principales de los que depende la secuencia DWI son, el valor  $b$  (o beta) que describe cuán sensible al movimiento del agua es la adquisición y las direcciones de sensibilización. Estos parámetros se utilizan junto con los datos de imágenes en el posprocesamiento [1, 2]. El posprocesamiento estándar que se le realiza a las DWI consiste en ajustar en cada vóxel un modelo físico de difusión del agua, a partir de los datos [2]. Actualmente existen varios modelos, pero el Modelo de Tensor de Difusión (DTI) fue el primero en desarrollarse y es el más simple, por lo que es muy utilizado en la práctica médica [3]. La principal desventaja del modelo DTI es que solo puede representar una población de fibras colineales, por lo que en las zonas donde hay haces que se cruzan el modelo falla. Para paliar esta limitación se han desarrollado otros modelos más complejos como QBall (QB) o *Constrained Spheric Deconvolutions* (cSD) que sí resuelven cruces entre haces [2].

El uso de la DWI es la base para la construcción de tractografías. Este proceso reconstruye la trayectoria que los haces nerviosos recorren dentro del cerebro entre un punto y otro [4]. A partir de la reconstrucción es posible cuantificar la conectividad estructural, es decir qué tan conectados están los distintos núcleos funcionales (materia gris) entre sí [5] [6] [7] y estudiar cómo esta conectividad, al igual que la microestructura, se ve afectada en diferentes patologías [8, 9, 10].

La reconstrucción de tractografías se apoya sobre los modelos de difusión, heredando sus ventajas y limitaciones. Por ejemplo una reconstrucción tractográfica sobre un modelo cSD es capaz de reconstruir cruces entre fibras, mientras que sobre DTI no puede. Recientemente Maier-Hein y sus colaboradores compararon diferentes métodos tractográficos utilizando distintos modelos de difusión y mostraron que mientras el modelo DTI genera una gran cantidad de falsos negativos debido a su incapacidad de representar cruces, los modelos más complejos como cSC generan muchos falsos positivos. Además en dicho trabajo Maier-Hein propone una serie de disposiciones de las fibras que no son cruces, pero que los modelos disponibles al día de hoy confunden [11].

Con el objetivo de superar las limitaciones prácticas mencionadas, aquí se propone explorar un abordaje desde la perspectiva informática. Las redes neuronales artificiales (RNA) son un potente modelo de clasificación que se compone de múltiples unidades de procesamiento conectadas entre sí. Si bien su principal uso es como clasificador, también pueden usarse en regresión [12]. En el último tiempo con la popularización de las redes neuronales profundas, las RNA también se han comenzado a usar como modelos generativos [13, 14]. Lo que proponemos es remplazar los clásicos modelos de difusión por modelos de neuronas artificiales interconectados en una red de arquitectura cristalina, e intentar

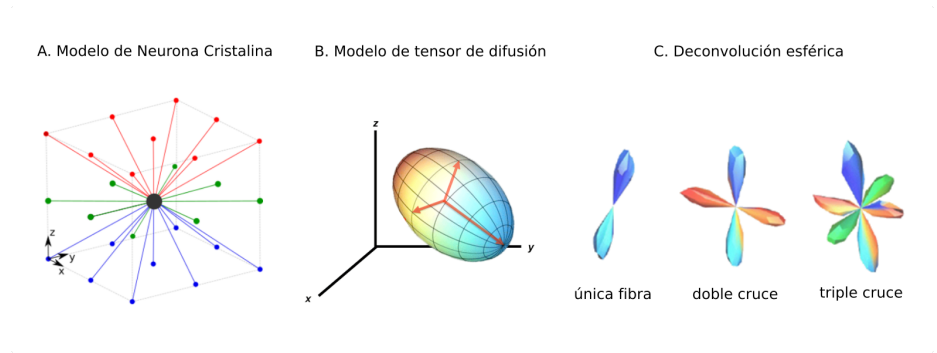
capturar la difusión entre vóxel en los pesos de las neuronas, las aristas del cristal.

En este trabajo, presentamos la Arquitectura de Red Neuronal Cristalina, la extensión del método de Descenso por el Gradiente para realizar el ajuste del modelo, y una serie de experimentos preliminares realizados sobre fantasmas virtuales (simulaciones) de fibras nerviosas, dispuestas en complejidad creciente.

## 2. Arquitectura de Red Neuronal Cristalina

La arquitectura tradicional de redes neuronales, requiere que los datos se presenten como un vector unidimensional de features, por lo que para procesar imagen es necesario “descomponer” la estructura de la misma y por consiguiente se pierde la relación entre píxeles vecinos [12]. La arquitectura de redes convolucionales salvó esta limitación, permitiendo procesar los datos en su espacio nativo y considerar la relación entre elementos vecinos. A pesar de que estas redes convolucionales, es su versión 3D, permitirían analizar los volúmenes de DWI en su espacio original, fueron diseñadas como extractores de descriptores abstractos que terminan alimentando una arquitectura de perceptrón multicapa [13, 14]. Dada la premisa de modelar la difusión de las moléculas de agua en el espacio, enfocamos el análisis específicamente en la relación entre vóxeles (versión tridimensional de un píxel), y proponemos una nueva arquitectura de red, que llamamos cristalina, dada la semejanza de sus conexiones con la estructura de un cristal, por ejemplo la del Cloruro de Sodio.

Este tipo de arquitectura permite procesar paquetes de datos que representan una muestra tridimensional discreta (vóxeles), y obtener un nuevo paquete de datos en el mismo espacio dimensional. El cristal restringe el flujo de información entre los vecinos inmediatos (conectividad rala). Cada neurona en la red recibe información ponderada de sus vecinos inmediatos, y envía información a dichos



**Figura 1.** Representaciones tridimensionales de los modelos de difusión, en un vóxel. A) Neurona artificial conectada con sus veintiséis vecinos inmediatos, B) Elipsoide que representa al modelo DTI y C) Tres glifos que representan al modelo de cSD en tres casos: fibra única, cruce doble y triple perpendiculares.

vecinos con la misma ponderación, ya que el peso de la conexión que vincula dos vóxeles vecinos es compartido (conectividad compartida). Estas características no sólo consideran la relación entre vecinos, sino que la definen como primordial. La Figura 1 compara la representación tridimensional (glifo) de una única neurona en una grilla tridimensional, con la de los otros dos modelos de difusión antes mencionados.

Nuestro objetivo es ajustar las conexiones en la red cristalina, de manera tal que modelen la difusión del agua entre los vóxeles de una DWI. Para lograrlo proponemos disponer en cada vóxel de la imagen 3D una neurona artificial, interconectarlas con sus vecinos inmediatos (aquellos vóxeles que comparten al menos un vértice con él) y ajustar los pesos de las conexiones para que la red sea capaz de predecir el nivel de gris de cada vóxel, combinando linealmente la información de su vecindad. Además para considerar en el ajuste a la serie de direcciones de difusión usadas en la adquisición de la DWI, la combinación lineal será modulada por la proyección de éstas en la base cristalina, como se explica a continuación.

## 2.1. Base Cristalina

Se define como base cristalina a la matriz  $\mathbf{B}$ , cuyas columnas son los versores que especifican las direcciones espaciales en las que se disponen las conexiones de la red cristalina. Para el caso de tres dimensiones, la vecindad de un vóxel se compone de 26 vecinos, dispuestos en 13 direcciones espaciales, uno a cada lado del vóxel central:

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & 1 & 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & 0 & 1 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} \end{pmatrix} \quad (1)$$

Entonces un vector tridimensional  $\mathbf{d}$  (que define una de las direcciones de difusión en la DWI) puede proyectarse a un vector  $\mathbf{D}$  expresado en la base cristalina  $\mathbf{B}$ , de la siguiente manera:

$$\mathbf{D} = \mathbf{d}^T \mathbf{B} \quad (2)$$

## 2.2. Descenso por el Gradiente

Dado un conjunto de datos de entrada  $(I, \mathbf{D})$ , donde  $I$  es una imagen de la serie DWI y  $\mathbf{D}$  es la dirección de difusión en la que fue ponderada, proyectada en la base cristalina  $\mathbf{B}$  (ver ecuación 2). La intensidad predicha en un vóxel ( $I_i^*$ ) se calcula entonces como la combinación lineal de la intensidad de sus vecinos inmediatos en las distintas direcciones ( $j$ ) ponderada por el peso de las conexiones en el cristal ( $W_i$ ) y por la componente correspondiente de la proyección de la dirección de difusión proyectada ( $D_j$ ).

$$I_i^* = f((I_i, \mathbf{D}), W_i) = \sum_{j=1}^{13} (I_i^{(-j)} W_i^{(-j)} + I_i^{(+j)} W_i^{(+j)}) D_j \quad (3)$$

Donde  $W_i$  representa al subconjunto de los pesos que conectan el nodo  $i$  con su vecindad inmediata,  $W_i^{(+j)}$  es la conexión que coincide con la dirección del vector expresado como la columna  $j$  de  $\mathbf{B}$  e  $I_i^{(+j)}$  es la intensidad el vóxel vecino en esa dirección, mientras que el superíndice  $(-j)$  hace referencia al sentido contrario. En la ecuación 3 la función de activación de la neurona es lineal y no está acotada.

El modelo de red neuronal cristalina puede ajustarse utilizando el algoritmo de descenso por el gradiente clásico. Definimos el error cuadrático de predicción para el nodo  $i$  como:

$$E_i = \frac{1}{2} e_i^2 = \frac{1}{2} (I_i - I_i^*)^2 = \frac{1}{2} (I_i - f((I_i, \mathbf{D}), W_i))^2 \quad (4)$$

Y la derivada de dicho error respecto de los pesos de las conexiones de la red:

$$\frac{\partial E_i}{\partial W_i} = (I_i - f((I_i, \mathbf{D}), W_i)) \frac{-\partial f((I_i, \mathbf{D}), W_i)}{\partial W_i} \quad (5)$$

$$\frac{\partial E_i}{\partial W_i} = -e_i \left[ \sum_{j=1}^{13} (I_i^{(-j)} + I_i^{(+j)}) D_j \right] \quad (6)$$

Como los pesos de la red ( $W$ ), se comparten entre los nodos que conectan, la ecuación de actualización debe contemplar el efecto de ambos nodos. Entonces, la actualización para la arista que conecta los nodos  $i$  y  $k$ , vecinos inmediatos en la dirección  $j$  es:

$$\Delta W_i^{(+j)} + \Delta W_k^{(-j)} = -\eta \left( \frac{\partial E_i}{\partial W_i} + \frac{\partial E_k}{\partial W_k} \right) \quad (7)$$

$$\Delta W_i^{(+j)} + \Delta W_k^{(-j)} = \eta \left[ e_i (I_i^{(-j)} + I_k) + e_k (I_i + I_k^{(+j)}) \right] D_j \quad (8)$$

Donde  $\eta$  es un hiperparámetro que regula la velocidad del descenso por el gradiente.

### 2.3. Regularización y Penalidades

Se utilizó una regularización  $l_2$ , también conocida como “*weight decay*”, siendo ésta la solución típica usada en *Machine Learning* para mejorar la convergencia del descenso por el gradiente. A la ecuación 4 se le adiciona el término:

$$E_{l_2} = \gamma \left[ \sum_i \sum_j (W_i^{(-j)})^2 + (W_i^{(+j)})^2 \right] \quad (9)$$

Donde  $\gamma$  es el hiperparámetro de regularización. Este nuevo termino se refleja en la ecuación 8, como:

$$\Delta W_{i,l2}^{(+j)} = -\eta \frac{\partial E_{l2}}{\partial W_i^{(+j)}} = -2\eta\gamma W_i^{(+j)} \quad (10)$$

La regularización propuesta tiende a emparejar los valores de  $W$  uniformemente, por lo que no se generan direcciones preferenciales (las direcciones en las que el agua difunde). Para estimular el crecimiento preferencial de las conexiones alineadas con la a difusión y a la vez la reducción de las que no, se conceptualizó un mecanismo de Penalización Anisotrópica (PA).

Dado un conjunto de datos de entrada  $(I^a, \mathbf{d}^a)$ , se busca un par  $(I^b, \mathbf{d}^b)$  tal que  $\mathbf{d}^a \cdot \mathbf{d}^b \approx 0$ , entonces la penalización anisotrópica se computa para cada vóxel:

$$PA_i = \log\left(\frac{I_i^a + 1}{I_i^b + 1}\right) \quad (11)$$

Al igual que la regularización, la PA se incluye en la ecuación 8 de actualización de los pesos adicionando un nuevo término:

$$\Delta W_{i,PA}^{(+j)} = -\frac{1}{2}\eta\nu(PA_i + PA_k)D_j \quad (12)$$

Donde  $\nu$  es el factor de penalización. De esta manera si el valor medio de  $PA$  entre los dos nodos unidos por una conexión es negativo, el peso de la misma crece, mientras que si es positivo ésta decrece. La idea se fundamenta en que si el valor de un vóxel cualquiera  $I_i^a$  de una imagen ponderada en un dirección  $\mathbf{d}^a$  es menor al valor del mismo vóxel  $I_i^b$  en una imagen ponderada en una dirección  $\mathbf{d}^b$  perpendicular a la primera, entonces este vóxel refleja difusión en la primera dirección  $\mathbf{d}^a$ . En el caso contrario la difusión se refleja mayormente en  $\mathbf{d}^b$ . Finalmente la actualización de  $W$ , queda determinada por la combinación de las ecuaciones 8, 10 y 12:

$$W_i^{(+j)} = W_i^{(+j)} + \Delta W_i^{(+j)} + \Delta W_k^{(-j)} + \Delta W_{i,PA}^{(+j)} + \Delta W_{i,l2}^{(+j)} \quad (13)$$

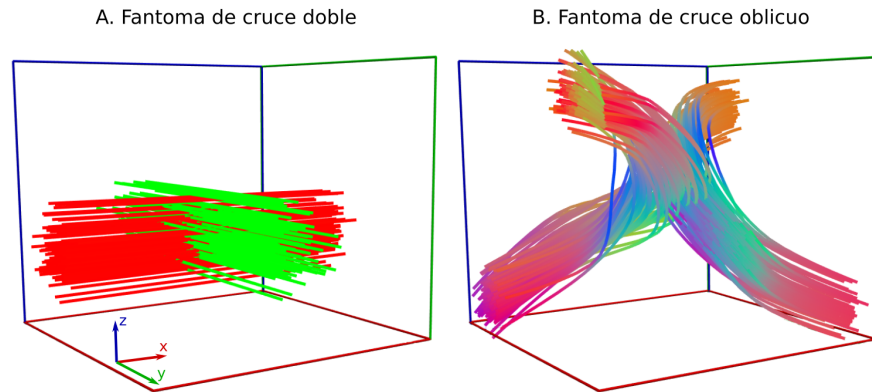
### 3. Materiales y Métodos

#### 3.1. Fantomas digitales

Para probar la factibilidad y robustez de la arquitectura propuesta se desarrollaron fantomas digitales, cuya complejidad aumenta progresivamente. Buscando imitar las características microestructurales de la materia blanca, se simuló DWI a partir de un modelo de fibras en el espacio (ver Figura 2) que representan posibles disposiciones de las fibras nerviosas. Para ello se utilizó un complemento de MITK Difusion llamado Fiberfox [15]. Esta herramienta permite controlar los parámetros de la adquisición de las DWI: cantidad de direcciones, ruido, artefactos de movimiento, resolución de las imágenes y los tiempos que caracterizan a la secuencia de adquisición.

Para testear el modelo se implementaron los siguientes fantomas de complejidad creciente:

1. Fibra única en dirección  $\boldsymbol{x}$
2. Cruce perpendicular de dos fibras en las direcciones  $\boldsymbol{x}$  y  $\boldsymbol{y}$
3. Cruce de dos fibras curvadas con trayectorias diferentes



**Figura 2.** Disposición de las fibras en los fantasmas diseñados para probar el modelo propuesto. A) En rojo la fibra en dirección  $\boldsymbol{x}$  del primer fantoma, en combinación con la fibra en dirección  $\boldsymbol{y}$  en verde (segundo fantoma). B) Las dos fibras curvadas que se cruzan, del tercer fantoma.

Se simularon volúmenes de 111111 vóxeles, isométricos de  $2\text{mm}^3$ , ponderados en 64 direcciones de difusión, no coplanares, con  $b = 1500\text{ s/mm}^2$ ,  $TR = 900\text{ ms}$  y  $TE = 100\text{ ms}$ . Por simplicidad a la simulación no se le agregó ningún tipo de ruido.

### 3.2. Ajuste del modelo

Para favorecer el ajuste del modelo de red cristalina, la intensidad de las DWI se escaló en el rango  $[0 : 1]$ . El ajuste se realizó mediante el clásico algoritmo de descenso por el gradiente (ver sección 2.2). Los tres hiperparámetros del modelo se optimizaron empíricamente y finalmente se establecieron en  $\eta = 0,01$ ,  $\gamma = 0,9$  y  $\nu = 2$ . Para realizar el ajuste se utilizaron los datos de las 64 direcciones de difusión, en grupos de 10 volúmenes por época, seleccionados al azar, e iterando durante 10000 épocas.

La convergencia se supervisó mediante la curva del error cuadrático medio sobre todos los vóxeles. Para controlar el funcionamiento de la regularización se midió la suma cuadrática de todos los pesos de la red ( $W$ ). Y para supervisar el desempeño de la PA se calculó la suma cuadrática de los pesos en las distintas direcciones de la base cristalina ( $W_j$ ). El desempeño final del modelo propuesto se comparó con el de dos modelos usados ampliamente: DTI y cSD.

Tanto el modelo, como los *scripts* de ajuste y las métricas de desempeño se implementaron en Python, utilizando módulos básicos de operaciones matriciales y manejo de imágenes.

## 4. Resultados

En la Figura 3 puede apreciarse la convergencia adecuada de la de red cristalina, para cada uno de los fantasmas (izquierda), y también el efecto de la regularización  $l_2$  en la magnitud de los pesos (medio). A la derecha de la Figura 3 se muestra la suma diferencial de los pesos de las conexiones en cada una de las direcciones de la base cristalina. Puede verse claramente que aquellas conexiones que coinciden con las direcciones principales de fibras en el fantoma adquieren mayor magnitud, mientras que las demás se anulan.

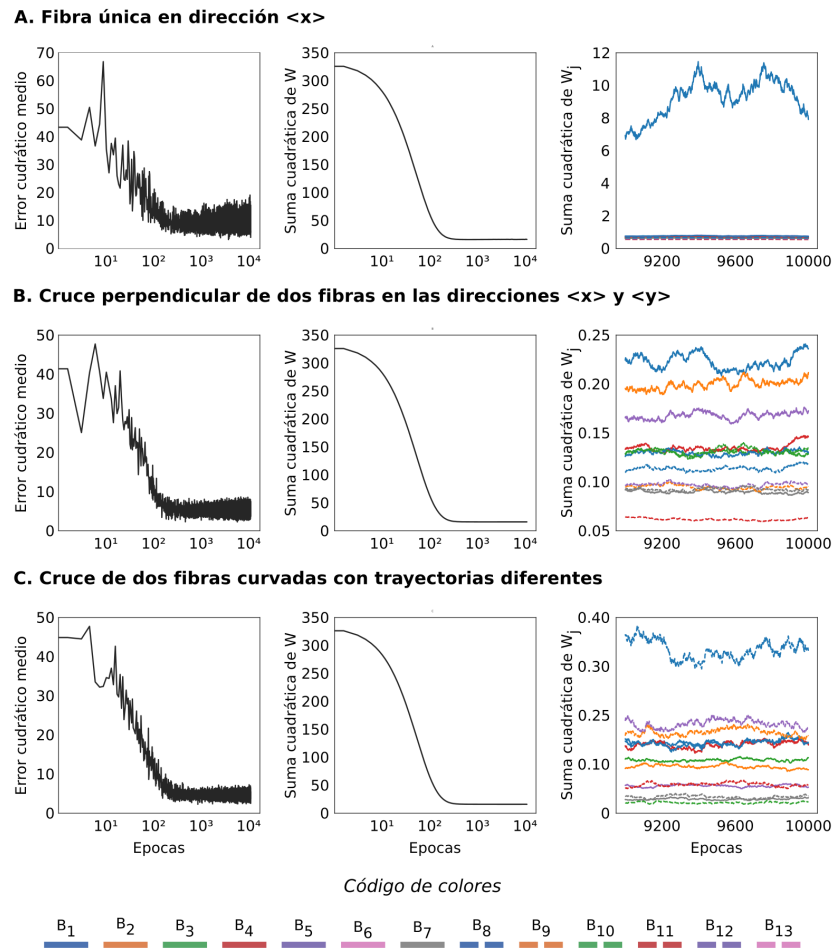
La Figura 4 compara la distribución del error cuadrático medio en cada vóxel (residuos), para reconstruir la DWI original a partir del conjunto de 64 direcciones de difusión. Debido a la diferencia de escala en los residuos de los tres métodos, por razones de visualización, estos fueron escalados entre  $[0,1]$ , para poder comparar cualitativamente su distribución en el espacio de cada fantoma. Es evidente que los tres modelos experimentan dificultades similares para reconstruir la información original. Mientras que el modelo DTI se ve seriamente perjudicado en los cruces de fibras (ver valor apuntado en la Figura 4.B.1), el modelo cSD y la red cristalina pueden resolverlos de mejor manera, logrando errores similares.

Por último en las Figuras 5 y 6 se puede apreciar las reconstrucciones tridimensionales de los distintos modelos. En la Figuras 5 se ve una rebanada (*slice*) del volumen de cada fantoma, al igual que en la figura anteriores. Para los tres fantasmas se aprecia que la red cristalina es capaz de recuperar la fibras, mientras que los pesos de las neuronas ubicadas en vóxeles isotópicos (fuera de la fibra) son prácticamente nulos.

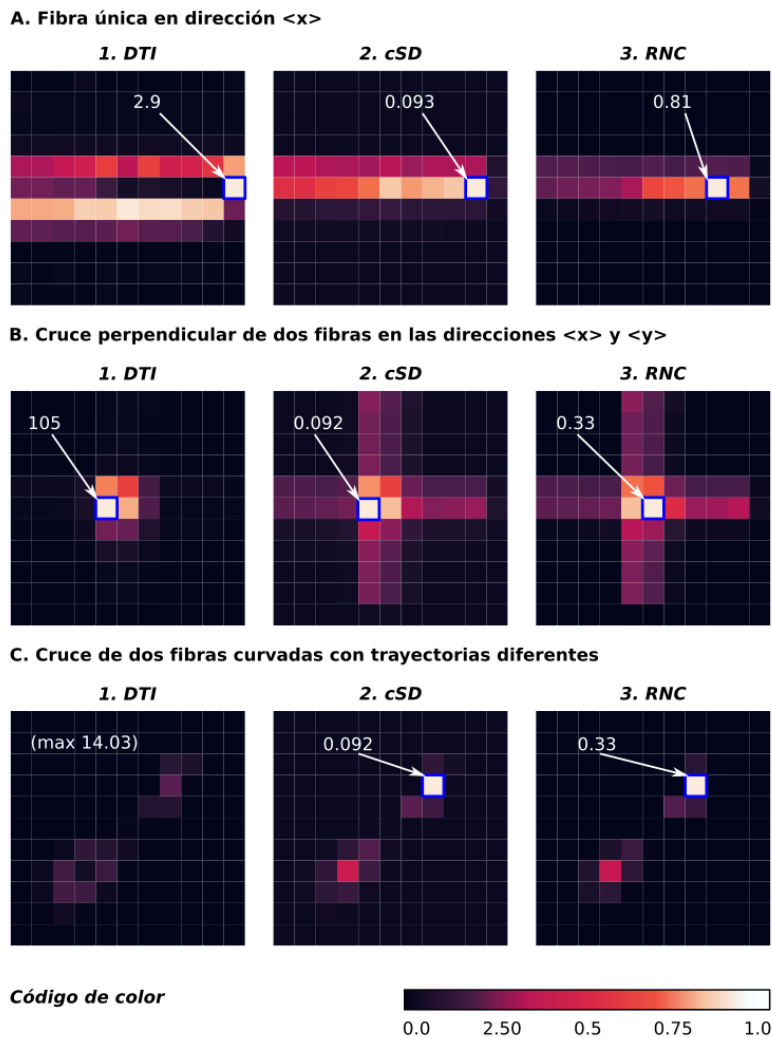
En la Figura 6 se muestra una perspectiva completa del espacio, donde puede apreciarse todos los vóxeles que componen el volumen, de cada fantoma. Puede verse (Figura 6.C) como la red cristalina recupera la estructura curvada de la fibras involucradas en ese fantoma. Es claro que la red cristalina capta la anisotropía de fibras. Para el fantoma de fibra única, el mas simple, la red incluso fue capaz de recuperar la dirección preponderante, dado que las conexiones en dicha dirección son mayores que en las demás (Figura 6.A). Sin embargo puede verse en las imágenes aumentadas (fila inferior de la Figura 6) que aquellas conexiones que no coinciden con las direcciones principales de las fibras, lejos de ser nulas, crecen, conectando vóxeles vecinos y formando una entramado tridimensional. Esto es muy claro para el fantoma de cruce perpendicular (Figura 6.B), donde se ve buena conexión en las dirección de las fibras, pero también en otras direcciones coplanares. Algo destacable es que además de ser muy sensibles a la anisotropía, estos entramados formados por la red, tiene bordes muy bien definidos. Dado que los pesos en las neuronas no tienen porque ser simétricos, a diferencia de los modelos DTI o cSD, las neuronas orientan sus conexiones no nulas hacia los vóxeles vecinos que también forman parte de la fibra, pero no hacia aquellos



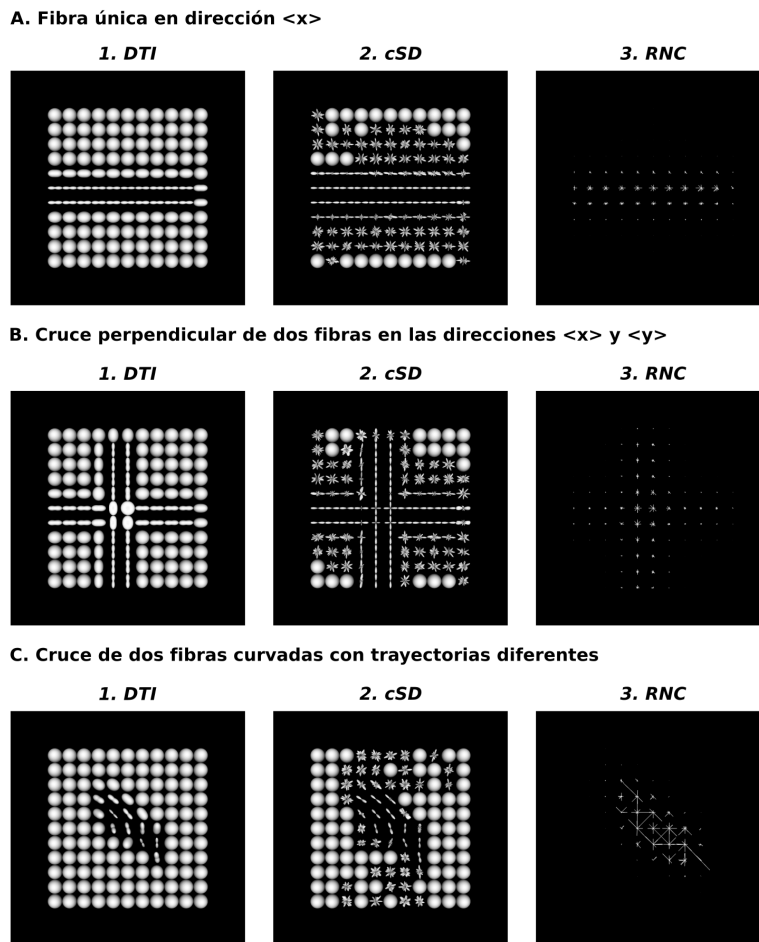
vóxeles anisótropicos en su vecindad. Esto último puede observarse en la tercera imagen aumentada del renglón inferior de la Figura 6.



**Figura 3.** Convergencia del modelo para cada uno de los fantasmas: A) Fibra única en dirección  $x$ , B) Cruce perpendicular de dos fibras en las direcciones  $x$  y  $y$  y C) Cruce de dos fibras curvadas con trayectorias diferentes. A la izquierda puede apreciarse el descenso en el error cuadrático medio sobre todos los vóxeles a lo largo de las épocas, al medio se muestra la reducción en la suma cuadrática media de la matriz  $W$ , a causa de la regularización  $l_2$ , y a la izquierda se puede ver la suma cuadrática de los pesos en las distintas direcciones de la base cristalina  $B$ . Los colores y tipos de línea codifican las diferentes direcciones de la base cristalina (ver sección 2.1).



**Figura 4.** Distribución del error cuadrático medio en cada vóxel. Para conseguir una buena visualización, el error fue escalado entre  $[0,1]$ , de forma independiente para cada ejemplo. Cada fila muestra una única rebanada (*slice*  $z = 5$ ) de un volumen. A) Fibra única en dirección  $\mathbf{x}$ , B) Cruce perpendicular de dos fibras en las direcciones  $\mathbf{x}$  y  $\mathbf{y}$  y C) Cruce de dos fibras curvadas con trayectorias diferentes. De izquierda a derecha se muestran los modelos DTI, cSD y la red neuronal cristalina (RNC). Los vóxeles marcados en azul son los de máximo error, y el valor apuntado es el error cuadrático medio sin escalar (Para el panel C1, el vóxel de máximo error queda fuera del *slice*  $z = 5$ , por lo que solo se apunta el valor sin escalar).



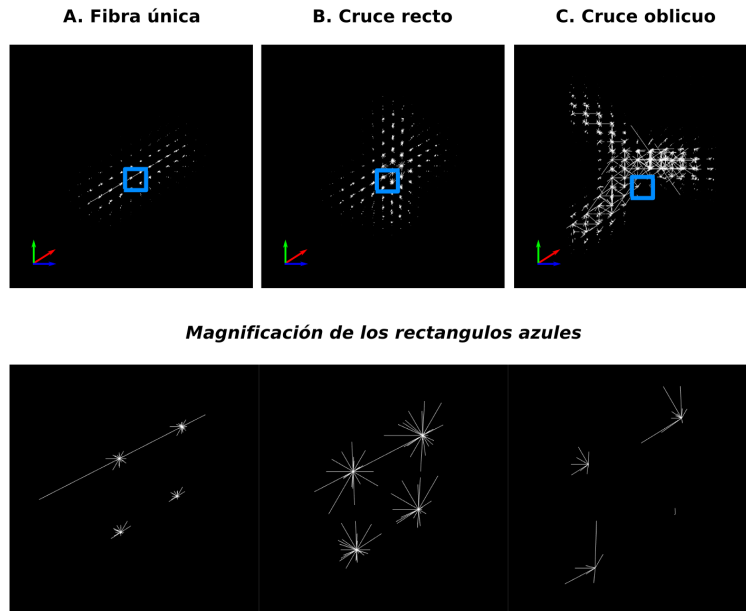
**Figura 5.** Representación tridimensional de los modelos de difusión (ver Figura 1). Cada panel muestra en una única rebanada (*slice*  $z = 5$ ) de un volumen: A) Fibra única en dirección  $\mathbf{x}$ , B) Cruce perpendicular de dos fibras en las direcciones  $\mathbf{x}$  y  $\mathbf{y}$  y C) Cruce de dos fibras curvadas con trayectorias diferentes. En la primer columna (A1, B1 y C1) puede verse el modelo DTI, en la segunda (A2, B2 y C2) el modelo cSD, mientras que en la tercera (A3, B3 y C3) la red neuronal cristalina (RNC).

## 5. Discusión

Hasta aquí se ha propuesto un nuevo modelo de difusión inspirado en las redes neuronales artificiales, que extiende la arquitectura de las mismas propo-

niendo un modelo de red cristalina 3D, generalizable a N dimensiones. Este nuevo modelo se ha probado sobre fantasmas digitales de complejidad creciente, donde mostró ser capaz de recuperar la disposición espacial de las fibras subyacentes.

En general, el modelo converge a soluciones concretas y en cierto punto correctas. Sin embargo no fue capaz de generar diferencias claras entre las conexiones paralelas a las fibras (direcciones predominantes) y las que no. Por ejemplo una solución óptima hubiese sido que las conexiones paralelas a las fibras tuviesen al menos un orden de magnitud de diferencia con respecto al resto. Esta limitación pudo verse más claramente en el caso de los cruces de fibras (Figura 6.B y Figura 6.C), donde la red cristalina no capturó las direcciones predominantes, sino que ajustó sus conexiones en la grilla casi uniformemente. Desde el punto de vista de la reconstrucción de conexiones estructurales (tractografía), esto representa una limitación bastante grande, ya que si todas las conexiones de un nodo (vóxel) son iguales, entonces se podría interpretar a la difusión de esa zona como isotrópica, cuando en realidad no lo es.



**Figura 6.** Reconstrucción tridimensional de la red cristalina completa, en perspectiva. A) fibra única en dirección  $x$ , B) cruce perpendicular de dos fibras en las direcciones  $x$  y  $y$  y C) cruce de dos fibras curvadas con trayectorias diferentes. La fila inferior muestra una vista aumentada de la región delimitada por los rectángulos azules.

El análisis de los resultados, junto al de la información contenida en la DWI, nos permitió concluir que el método de ajuste propuesto es incapaz de distinguir fibras entre sí, ya que relaciona la intensidad entre los distintos vóxeles de un mismo volumen de la serie, por lo que favorece el crecimiento de conexiones perpendiculares a las direcciones predominantes, generando así un entramado, o estructura reticular, que conecta todos los vóxeles por los cuales transcurren las fibras subyacentes. Esta deficiencia se puede resolver repensando el método de ajuste del modelo. Una posible alternativa sería utilizar la información de pares de volúmenes de la serie, ponderados en direcciones diferentes, predecir la intensidad de la segunda en base a la primera, y finalmente aumentar o disminuir  $W_j$  en la dirección de la segunda, en función del signo del error. Esta estrategia tiene el potencial de sensibilizar el método de ajuste para captar la diferencia en los niveles de gris en las imágenes, debidos al cambio en la dirección de ponderación de la DWI.

En cuanto a la regularización  $l_2$ , ésta tiende a preferir una convergencia uniforme de  $W$ , lo cual imposibilitó que se generen direcciones preferenciales. Por su parte la PA demostró ser muy eficaz sensibilizando los nodos que pertenecen a la fibra y atenuando aquellos que no, pero inadecuada para estimular el crecimiento de  $W$  en direcciones preferenciales. Una posibilidad para lograr esto sería utilizar una regularización elástica ( $l_1 + l_2$ ) [16]. Esta regularización aseguraría la convergencia, mientras que permitiría que se generen direcciones preferenciales.

Un punto destacable es que el modelo mostró capacidad para recuperar no solo cruces, sino también curvas (ver Figura 6.C)). La red cristalina no está restringida a modelar la difusión como un fenómeno simétrico, como los otros dos con los que la comparamos en este artículo. Por lo tanto, la red puede captar difusión en dos direcciones diferentes, sin que esto implique un cruce de fibras. Esta es una ventaja con potencial para resolver algunas de las limitaciones planteadas por Maier-Hein [11]. Específicamente, desde el punto de vista de su aplicación a las neurociencias, el modelo de redes presentado tienen un enorme potencial, ya que es el primer *framework* computacional que podría combinar datos de conectividad funcional y estructural en múltiples niveles de organización, desde una perspectiva regional a una global.

## 6. Conclusión

En el presente trabajo se ha desarrollado una arquitectura nueva de redes neuronales que propone patrones de conexión entre las unidades semejantes a un cristal. Esta arquitectura se diseñó con el fin de procesar datos tridimensionales, particularmente la difusión del agua en el cerebro a partir de DWI. También se ha construido un conjunto de fantasmas digitales para poder validar el funcionamiento de los modelos de difusión en un “ambiente” sencillo y controlado.

El análisis e interpretación de los resultados expuestos, permitió entender en profundidad el comportamiento de la nueva arquitectura de redes tridimensional propuesta y plantear mejoras para potenciar sus capacidades en el análisis de datos de DWI.

## Referencias

- [1] A. Hoy, A; Alexander, “Diffusion mri,” in *Brain Mapping: an encyclopedic reference* (A. Toga, ed.), pp. 47–52, London: Academic Press, 2015.
- [2] M. Descoteaux, “High angular resolution diffusion imaging (hardi),” *Wiley Encyclopedia of Electrical and Electronics Engineering*, pp. 1–25, 1999.
- [3] C. Lenglet, “Diffusion tensor imaging,” in *Brain Mapping: an encyclopedic reference* (A. Toga, ed.), pp. 245–251, London: Academic Press, 2015.
- [4] S. Tournier, J-D; Mori, “Fiber tracking with dwi,” in *Brain Mapping: an encyclopedic reference* (A. Toga, ed.), pp. 265–269, London: Academic Press, 2015.
- [5] O. Sporns, G. Tononi, and R. Kötter, “The human connectome: a structural description of the human brain,” *PLoS computational biology*, vol. 1, no. 4, p. e42, 2005.
- [6] P. Hagmann, M. Kurant, X. Gigandet, P. Thiran, V. J. Wedeen, R. Meuli, and J.-P. Thiran, “Mapping human whole-brain structural networks with diffusion mri,” *PloS one*, vol. 2, no. 7, p. e597, 2007.
- [7] O. Sporns, “The human connectome: origins and challenges,” *Neuroimage*, vol. 80, pp. 53–61, 2013.
- [8] I. Ellison-Wright and E. Bullmore, “Meta-analysis of diffusion tensor imaging studies in schizophrenia,” *Schizophrenia research*, vol. 108, no. 1-3, pp. 3–10, 2009.
- [9] G. Nortje, D. J. Stein, J. Radua, D. Mataix-Cols, and N. Horn, “Systematic review and voxel-based meta-analysis of diffusion tensor imaging studies in bipolar disorder,” *Journal of affective disorders*, vol. 150, no. 2, pp. 192–200, 2013.
- [10] F. Kruggel, F. Masaki, A. Solodkin, A. D. N. Initiative, *et al.*, “Analysis of longitudinal diffusion-weighted images in healthy and pathological aging: An admi study,” *Journal of neuroscience methods*, vol. 278, pp. 101–115, 2017.
- [11] K. Maier-Hein, P. Neher, J.-C. Houde, M.-A. Cote, E. Garyfallidis, J. Zhong, M. Chamberland, F.-C. Yeh, Y. C. Lin, Q. Ji, *et al.*, “Tractography-based connectomes are dominated by false-positive connections,” *bioRxiv*, p. 084137, 2016.
- [12] I. A. Basheer and M. Hajmeer, “Artificial neural networks: fundamentals, computing, design, and application,” *Journal of microbiological methods*, vol. 43, no. 1, pp. 3–31, 2000.
- [13] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural networks*, vol. 61, pp. 85–117, 2015.
- [14] I. Arel, D. C. Rose, and T. P. Karnowski, “Research frontier: deep machine learning—a new frontier in artificial intelligence research,” *IEEE Computational Intelligence Magazine*, vol. 5, no. 4, pp. 13–18, 2010.
- [15] P. F. Neher, B. Stieltjes, F. B. Laun, H.-P. Meinzer, and K. H. Fritzsche, “Fiberfox: A novel tool to generate software phantoms of complex fiber geometries,” in *Proceedings of the International Society for Magnetic Resonance in Medicine*, vol. 21, p. 3162, 2013.
- [16] H. Zou and T. Hastie, “Regularization and variable selection via the elastic net,” *Journal of the royal statistical society: series B (statistical methodology)*, vol. 67, no. 2, pp. 301–320, 2005.