

TESIS DE MAESTRÍA EN INGENIERÍA

TOMOGRAFÍA DE PROYECCIÓN ÓPTICA ACELERADA MEDIANTE APRENDIZAJE PROFUNDO

Marcos A. Obando
Maestrando

Dr. Germán Mato
Director

Dr. Teresa Correia
Colaboradora

Miembros del Jurado

Dr. Ariel H. Curiale (Instituto Balseiro)
Dr. Damián G. Hernández (Instituto Balseiro)
Dr. Flavio D. Colavecchia (INTECNUS)

26 de Diciembre de 2022

Departamento de Física Médica – Centro Atómico Bariloche

Instituto Balseiro
Universidad Nacional de Cuyo
Comisión Nacional de Energía Atómica
Argentina

(Biblioteca Leo Falicov CAB-IB)

A quienes tengan problemas

Índice de símbolos

Símbolo	Significado
ANN	<i>Artificial neural network</i> (Red neuronal artificial)
CNN	<i>Convolutional neural network</i> (Red neuronal convolucional)
CT	<i>Computed tomography</i> (Tomografía Computada)
FBP	<i>Filtered backprojection</i> (Retroproyección filtrada)
LIP	<i>Linear inverse problem</i> (Problema lineal inverso)
MRI	<i>Magnetic resonance imaging</i> (Imagen por resonancia magnética)
MoDL	<i>Model-based Deep Learning</i> (Aprendizaje profundo basado en modelo)
OPT	<i>Optical projection tomography</i> (Tomografía de proyección óptica)
PSNR	<i>Peak signal-to-noise ratio</i> (Relación señal a ruido pico)
RNN	<i>Recurrent neural network</i> (Red neuronal recurrente)
SSIM	<i>Structural similarity index metric</i> (Medida del índice de similaridad estructural)
STORM	<i>Smoothness regularization on manifolds</i> (Regularización de suavidad en variedades diferenciables)
TwIST	<i>Two-step Iterative Shrinking/Thresholding</i>
ToMoDL	<i>Tomographic Model-Based Deep Learning</i> (Aprendizaje profundo basado en modelo tomográfico)

Índice de contenidos

Índice de símbolos	v
Índice de contenidos	vii
Índice de figuras	ix
Índice de tablas	xiii
Resumen	xv
Abstract	xvii
1. Tomografía de proyección óptica	5
1.1. Principios ópticos de funcionamiento	5
1.2. Reconstrucción mediante retroproyección filtrada	7
1.3. Artefactos de reconstrucción	9
1.3.1. Artefactos experimentales	10
1.3.2. Artefactos por muestreo	10
2. Teoría de reconstrucción de imágenes	13
2.1. Problema lineal inverso	13
2.2. Aprendizaje profundo	15
2.2.1. Redes convolucionales	15
2.2.2. Redes recurrentes	17
2.3. LIP con red desenrollada: <i>Model-based Deep Learning</i> (MoDL)	18
3. ToMoDL: Propuesta e implementación	23
3.1. Elementos de la propuesta	24
3.2. Datos de entrenamiento	26
3.3. Detalles de implementación	27
4. Experimentos y resultados	31
4.1. Materiales y métodos	31

4.1.1. Variantes utilizadas	31
4.1.2. Hiperparámetros	32
4.1.3. Utilización de los datos de entrenamiento	33
4.2. Algoritmo de alineamiento	34
4.3. Comparación entre métodos	35
4.3.1. Factor de aceleración	36
4.3.2. Impacto de las características del conjunto de datos	37
5. Discusión y conclusiones	43
Bibliografía	45
Agradecimientos	51

Índice de figuras

1.1. Tomografía de proyección óptica. a) El detector colecta la luz resultante de atravesar de un espécimen translúcido. Dicha muestra se halla contenida en un cilindro rotatorio, el cual permite adquirir distintas proyecciones angulares, parametrizadas por γ . b) Fuentes de luz blanca y monocromática pueden utilizarse para extraer información anatómica y/o funcional a partir de la atenuación del tejido y/o emisión de un trazador genético, respectivamente. c) El plano focal esta dispuesto a medio camino entre el eje de rotación y el frente de la muestra. La apertura numérica (NA), relacionada con el ángulo δ , es lo suficientemente pequeña como para que la profundidad de campo comprenda al menos la mitad de la muestra, asegurando su enfoque en el detector.	6
1.2. Transformada de Radon. Para un dado corte transversal al eje de rotación, la proyección $p(I, \gamma)$ es análoga a un corte radial en la transformada de Fourier bidimensional de la imagen. En el espacio de Fourier, las líneas proyectivas adquiridas son un muestreo radial de la transformada de la imagen $\mathcal{F}[f(x, y)](u, v)$, lo cual nos proporciona un método para reconstruirla mediante el teorema de corte.	8
1.3. Artefactos por muestreo. a) muestra como la interpolación entre coordenadas polares a cartesianas en el espacio de Fourier implica un muestreo pobre de altas frecuencias, b) manifestándose en una adquisición afectada por artefactos al aumentar el espaciamiento angular. El mismo puede verse como un factor de aceleración en la adquisición del organismo bajo estudio.	12

2.1. Modelos de visión. a) La respuesta neuronal selectiva del gato ante variaciones de posición y orientación inspiraron a Hubel y Wiesel b) el modelado de células simples y complejas que responden ante regiones locales y globales del estímulo. c) y d) muestran el modelo de red neuronal convolucional (CNN) propuesto por LeCun, donde el mapeo es aprendido mediante el algoritmo de <i>backpropagation</i> para resolver el problema de clasificación de símbolos decimales escritos a mano.	16
2.2. Reconstrucción OPT mediante U-Net. Dos caminos, contractivo y expansivo, son enlazados en las diferentes etapas de muestreo mediante <i>skip connections</i> . Utilizando pares de imágenes OPT reconstruidas mediante FBP, la arquitectura U-Net permite remover artefactos en imágenes obtenidas de un muestreo pobre del corte. Adaptación de [1] y [2].	18
2.3. Recursión. La idea de recurrencia, a) ya capturada por Giotto en su <i>Tríptico Stefaneschi</i> , ha sido explotada en el campo de ANN con la aparición del algoritmo de <i>backpropagation</i> para entrenar redes recurrentes b) . Su aplicación puede verse de manera más sencilla al tratar con su forma desenrollada c)	19
2.4. Esquema de red desenrollada. La imagen obtenida al aplicar el operador lineal hermítico atraviesa un esquema de reconstrucción alternante que resuelve el problema de <i>denoising</i> al aplicar \mathcal{D}_θ y forzar la consistencia de los datos (DC) resolviendo (2.16b).	20
3.1. Bloque residual. El ruido de la imagen de entrada x es estimado a través de una serie de redes convolucionales activadas mediante una función ReLU. Para estimar la imagen sin ruido, una conexión de salto es sumada a la salida de \mathcal{N}_θ	24
3.2. Procesamiento de imágenes OPT a) distintas proyecciones angulares son obtenidas con un muestreo angular equiespaciado. Dichas proyecciones son el resultado de concatenar 4 secciones, como muestra b) de la especie <i>Danio Rerio</i> . c) La red ToMoDL aprende a reconstruir y remover artefactos de forma supervisada, a partir de pares de imágenes retroproyectadas a partir de subconjuntos de muestras (mostradas como puntos) como entrada y reconstrucciones con todas las proyecciones disponibles.	28

3.3. Etapas de reconstrucción iterativa mediante ToMoDL. La imagen x_0 que ingresa a ToMoDL atraviesa las etapas alternantes que resuelven el problema lineal (DC) y el <i>denoising</i> no lineal (\mathcal{D}_θ) con el correspondiente ruido estimado (\mathcal{N}_θ). En la fila inferior se muestra una ampliación de la reconstrucción, mostrando como el ruido estimado evoluciona a una mayor estructura en cada iteración de la red para reconstruir la imagen de la cabeza de un pez cebra.	29
4.1. Exploración de hiperparámetros. Mediante una búsqueda en iteraciones, número de capas y las funciones de pérdida propuestas, a partir de la PSNR de las reconstrucciones del conjunto de evaluación resolvemos por utilizar una arquitectura ToMoDL entrenada de entrada a salida para minimizar la función <i>mean square error</i> con $K = 8$ y $L = 8$	33
4.2. Alineamiento mediante maximización de varianza de la intensidad. a) muestra el corrimiento relativo entre el eje de rotación y el detector, lo cual genera una traslación en el espacio de sinogramas, corregido en b) mediante una transformación de traslación. Para hallarla, buscamos de forma iterativa en una grilla de traslaciones, reduciendo el paso en píxeles hasta hallar la transformación que maximice la varianza de la intensidad de la reconstrucción c). Así, el efecto de <i>doble pared</i> en los bordes de la estructura, observado en d), puede ser corregido. . . .	35
4.3. Desempeño de los métodos de reconstrucción a diferentes factores de aceleración. Observamos como, de manera consistente, la reconstrucción mediante ToMoDL supera los 30 dB para todos los factores de aceleración, con una medida del índice de similaridad estructural mejor al resto de los métodos a partir de un factor de aceleración $\times 10$.	36
4.4. Reconstrucción mediante los distintos métodos. Tomando reconstrucciones con un factor de aceleración $\times 10$, se muestran ejemplos en dos secciones del animal bajo estudio. Observamos que mientras para la cabeza, ToMoDL es el único método que efectivamente remueve artefactos propios del submuestreo de proyecciones al reconstruir.	37
4.5. Reconstrucción por sección. Para cada método propuesto, observamos la distribución de las métricas de reconstrucción PSNR y SSIM para un factor de aceleración $\times 22$, en función de la sección. La cola inferior, la cual presenta gran parte de los cortes sin tejido, es aquel que condiciona el funcionamiento de los métodos. Vemos que ToMoDL presenta un desempeño consistente en el resto de las secciones, incluso superior al estado del arte en reconstrucción OPT.	38

4.6. Desvío estándar de intensidad por reconstrucción <i>ground truth</i>.	
a) Para el volumen B (3 días post fertilización) se muestra el desvío de intensidad σ_I^i de la reconstrucción utilizada como <i>ground truth</i> para cada sección del animal. b) La diferencia entre cortes sucesivos $\Delta\sigma_I^i$ presenta una dispersión mayor c) en las secciones superior e inferior de las colas, observada a través de $P(\Delta\sigma_I^i)$. Estas inhomogeneidades se reflejan en variaciones en el contraste en el conjunto de datos, presentando ciertas dificultades para ser capturadas por los enfoques de aprendizaje supervisado.	39
4.7. Proyecciones y reconstrucción por sección para volumen B.	
a)b) Las reconstrucciones utilizadas muestran una alta variabilidad de contraste entre cortes adyacentes en la sección del cuerpo del espécimen B. c) Aún con esto, la variabilidad de intensidad en función de la PSNR muestra que dicha hipótesis no explica casos como la cola superior, donde la dispersión de $\Delta\sigma_I^i$ no implica una merma en el desempeño de la red.	40
4.8. Reconstrucción por día post fertilización. Para un factor de aceleración $\times 22$, las métricas de cada método de reconstrucción en función del día post fertilización del animal bajo estudio muestran la capacidad de generalización de las reconstrucciones con parámetros entrenables.	40
4.9. Proyecciones por día post fertilización. Observamos una fuerte transición en el desarrollo embrionario de la cabeza entre el espécimen adquirido en el primer y tercer día post fertilización, que explica el pobre desempeño al evaluar la generalización del enfoque supervisado no basado en el modelo de Radon. Por simplicidad, para el volumen 1 día post fertilización consideramos que la única sección disponible corresponde a la categoría 'cabeza'.	41
4.10. Reconstrucción por espécimen. Para un factor de aceleración $\times 22$, mostramos las métricas de cada método de reconstrucción en función del espécimen bajo estudio. Observamos que los volúmenes B y E, que contienen las cuatro secciones del animal, introducen la mayor varianza en las métricas de reconstrucción, seguido del espécimen D, del cual contamos con diferentes días post fertilización.	42

Índice de tablas

3.1. Volúmenes utilizados para la evaluación de métodos supervisados.	27
4.1. Métodos de reconstrucción utilizados.	32

Resumen

Recientemente, el campo de los problemas lineales inversos se ha beneficiado enormemente de técnicas basadas en aprendizaje profundo en la tarea esencial de recuperar una señal a partir de un bajo número de muestras. En el campo de la reconstrucción de imágenes, los recientes abordajes involucran la extensión de la resolución de problemas inversos mediante redes neuronales profundas como funciones promotoras a dominios raros. Por su ventaja computacional, la evaluación de estos entornos, conocidos como redes profundas desenrolladas, usualmente se halla limitada a la reconstrucción de imágenes por resonancia magnética (MRI). En este trabajo, proponemos ToMoDL (*Tomographic Model-based Deep Learning*), una arquitectura desenrollada para tratar el problema de acelerar la adquisición de tomografías de proyección óptica (OPT), una técnica mesoscópica para obtener imágenes de muestras biológicas translúcidas.

Utilizando doce volúmenes de proyecciones angulares de peces cebra (*Danio Rerio*) conteniendo diferentes secciones y días post fertilización de 5 especímenes, nuestro enfoque aborda un conjunto de datos con alta variabilidad en términos de intensidad y estructura, así como un problema usualmente pasado por alto en la aplicación de redes basadas en modelo: tomografía computada [3]. Dado que en su mayoría, los dispositivos OPT son diseñados a medida del laboratorio que hace su uso, las proyecciones crudas han sido tratadas cuidadosamente para evitar artefactos de reconstrucción producidos por desalineación del sistema óptico. Por otro lado, la integración del operador de Radon en un esquema de diferenciación automática ha sido problemática hasta la aparición de recientes avances en soluciones compatibles con librerías de aprendizaje profundo.

Mediante la minimización del error cuadrático medio y la medida del índice de similaridad estructural en un mapeo desenrollado con un número finito de iteraciones, la validación cruzada de ToMoDL muestra resultados de alta calidad incluso con un 5 % de las proyecciones adquiridas. Los mismos usualmente superan el desempeño de reconstrucción que el estado del arte, conformado por métodos de *compressed sensing*, resuelto mediante *Two-Step Iterative Shrinkage/Thresholding* (TwIST) y arquitecturas supervisadas de tipo U-Net. Por último, desarrollamos un análisis extensivo de ciertos inconvenientes en la reconstrucción anatómica de la estructura del pez cebra, para los cuales las soluciones analíticas presentan la mejor alternativa.

Palabras clave: TOMOGRAFÍA DE PROYECCIÓN ÓPTICA, ACELERACIÓN DE ADQUISICIÓN, APRENDIZAJE PROFUNDO EN PROBLEMAS INVERSOS

Abstract

Linear inverse problems have greatly benefited from deep learning techniques in the paramount goal of recovering a signal from a small number of measurements. In the field of image reconstruction, recent approaches involve the augmentation of traditional inverse problem solvers with neural networks as sparsifying functions. Given its computational tractability, gold standards for validating these frameworks, often known as deep unrolled architectures, are usually limited to magnetic resonance imaging (MRI) reconstructions. We propose ToMoDL (*Tomographic Model-based Deep Learning*), a deep unrolled architecture for tackling the problem of accelerating the acquisition of optical projection tomography (OPT), a mesoscopic technique for imaging biological translucent samples.

Using twelve volumes of zebrafish (*Danio Rerio*) angular projections from four longitudinal sections and different days post-fertilisation, our approach deals with an extremely variable dataset in terms of intensity and structure, as well as an often overlooked problem in model-based deep learning: tomography reconstruction. Since many, if not most, OPT devices are custom-built, raw projections have been carefully curated to avoid reconstruction artifacts due to misalignment. On the other hand, integrating Radon operator blocks into an automatic differentiation scheme has been thorny until recent advances in PyTorch-compatible solutions arose.

By minimizing the mean square error and structural similarity index metric in a fixed-iteration unrolled mapping, our cross-validated results show a reasonably high quality reconstruction with even a 5% of the acquired projections, achieving a considerably better performance than compressed sensing methods such as Two-Step Iterative Shrinkage/Thresholding (TwIST) and U-Net architectures. We also analyze problematic reconstruction issues regarding the anatomical structure of zebrafish, where analytical solutions may be chosen instead.

Keywords: OPTICAL PROJECTION TOMOGRAPHY, SCAN ACCELERATION, DEEP LEARNING FOR INVERSE PROBLEMS

Motivación

El término tomografía refiere a la obtención de cortes transversales de un objeto a partir de proyecciones recolectadas al iluminar el mismo desde diferentes direcciones [4]. Si bien su formulación matemática fue provista por Radon [5] a principios del siglo XX, su implementación asistida por computadora ha revolucionado campos de lo más variados [6] [7]. Tomografía computada (CT) y la tomografía por emisión de positrones (PET) son las técnicas más conocidas que utilizan este formalismo para la reconstrucción de cortes de un objeto bajo estudio.

En particular para las ciencias biológicas, analizar la organización del tejido en tres dimensiones resulta de vital importancia para entender el desarrollo embrionario y sus distintas interacciones genómicas [8]. La tomografía de proyección optica (OPT), proporciona una modalidad no invasiva para adquirir imágenes tridimensionales de modelos animales de enfermedad [9]. Sin embargo, la adquisición *in vivo* esta comprometida por su naturaleza fototóxica como por las dosis de anestesia necesarias, por lo que reducir el número de proyecciones necesarias para reconstruir imágenes de alta calidad resulta crucial para las acelerar las técnicas de OPT. La solución analítica del problema tomográfico lineal, conocido como retroproyección filtrada (FBP), requiere de cientos de adquisiciones angulares para cumplir con la cota de muestreo de Nyquist-Shannon [4]. Una reconstrucción analítica con un número de muestras por debajo de la mencionada cota resulta en una imagen ruidosa con una variedad de artefactos que destruyen la información relevante de la muestra adquirida.

Dicha cota presenta una condición suficiente para la reconstrucción ideal de la imagen tomográfica. Trabajar por debajo de la misma constituye la idea fundamental de la teoría de *compressed sensing* [10], cuyo marco propone explotar la condición rala de una señal en un dominio dado para reconstruirla de manera eficiente. Esta estrategia ha sido implementada con éxito para reconstruir de manera iterativa imágenes OPT de peces cebra (*Danio Rerio*) con un bajo número de proyecciones [9]. Bajo esta teoría, definir la transformación adecuada que promueva dicha cualidad rala no resulta trivial, así como regularizar el problema iterativo de manera general sin recurrir a una solución cuyas reconstrucciones pueden tomar del orden de decenas de minutos para un volumen completo.

Una opción diferente se presenta al contar con pares de imágenes tomográficas

con y sin artefactos. Resolver el problema de *denoising* ha sido una de las principales aplicaciones del aprendizaje profundo, donde su formulación supervisada consiste en encontrar el mapeo óptimo entre un conjunto de datos de entrada y salida. Recientemente, la remoción de artefactos en imágenes OPT ha sido encarada entrenando una red U-Net con pares de imágenes con y sin artefactos de submuestreo, logrando interesantes resultados. La necesidad de una elevada cantidad de ejemplos de entrenamiento y la baja interpretabilidad de la operatoria bajo este enfoque resultan desafíos en pie.

La reconstrucción tomográfica puede enmarcarse dentro de la teoría de problemas lineales inversos (LIP), dado que las proyecciones adquiridas se encuentran ligadas con los cortes mediante un modelo lineal que representa la física del sistema. Si bien el tratamiento de LIP utilizando diversos tipos de aprendizaje profundo ha demostrado un enorme potencial teórico en menos de una década [11], su evaluación se ha limitado a modalidades de imágenes cuyos operadores de reconstrucción resultan computacionalmente más sencillos de tratar, ya sea por su amplio espectro de aplicabilidad (resonancia magnética y transformada de Fourier) [11] o su tratabilidad (deconvolución y kernel) [12].

Aún frente a este paisaje de propuestas, encarar el problema de reconstruir imágenes OPT bajo alguna de ellas no resulta trivial por diversos motivos. Tratar con operadores de Radon resulta computacionalmente costoso y su integrabilidad con bibliotecas de aprendizaje automático resulta laboriosa y acotada. Por otra parte, la prevalencia de dispositivos de tomografía de proyección óptica es baja y adaptada por el grupo de investigación que haga uso del mismo, por lo que la variabilidad de los datos adquiridos, ya de por sí escasos, es alta. Sin embargo, el bajo costo de implementación [13] y alta resolución en el rango mesoscópico (cm-mm) convierten a los dispositivos OPT en el método ubicuo para el estudio estructural y funcional del desarrollo de tejidos, por lo que abordar este problema desde un punto de vista computacional presenta un gran interés para estos sistemas.

Bajo estas motivantes premisas, el camino transitado en este trabajo de investigación comprendió:

- Analizar los inconvenientes de la reconstrucción analítica de imágenes de tomografía de proyección óptica, prestando cuidado a diversos efectos indeseados propios del sistema de adquisición *custom*.
- Desarrollar una comprensión del estado del arte en LIP, en particular la transición desde enfoques iterativos tradicionales hacia la incorporación de teoría de aprendizaje en su abordaje.
- Entender las restricciones computacionales del problema de reconstrucción tomográfica, haciendo foco en su integración con bibliotecas de aprendizaje au-

tomático implementadas en GPU.

- Proponer, diseñar y validar una arquitectura de aprendizaje profundo basada en el modelo lineal subyacente al problema de reconstrucción de tomografías de proyección óptica.

Como resultado, introducimos ToMoDL (*Tomographic Model-based Deep Learning*), un entorno de reconstrucción basada en la resolución iterativa de un problema lineal inverso cuya regularización viene dada por una red neuronal convolucional de tipo residual (RCNN), un enfoque inspirado en el marco teórico basado en modelo (MoDL) formulado por Aggarwal et al. [14]. Por primera vez a nuestro entender [3] [15] [16], este problema lineal inverso es tratado mediante esta estrategia, obteniendo reconstrucciones de alta calidad a partir de un bajo número de proyecciones, con resultados comparables e incluso mayor a los métodos utilizados actualmente.

Capítulo 1

Tomografía de proyección óptica

La tomografía de proyección óptica (OPT) proporciona una modalidad no invasiva de imagen tridimensional, obtenida mediante la adquisición de múltiples proyecciones angulares de muestras translúcidas y la posterior reconstrucción de una dada característica de la misma, tales como atenuación (imagen por transmisión) o emisión (mediante fluorescencia). En esta propuesta, análoga a la convencional tomografía computada (CT), la utilización de ciertos solventes orgánicos (como la mezcla de benzoato de bencilo y alcohol de bencilo) son utilizados para ‘transparentar’ ópticamente el tejido de un espécimen y muestrearlo angularmente.

La profundidad de penetración de esta técnica está ligada al concepto de camino libre medio (MFP) del fotón $1/\mu_t$ en el tejido, donde μ_t es el coeficiente de transporte. OPT presenta una resolución mesoscópica de $1 - 10 \mu\text{m}$, principalmente limitada por la penetración de los químicos mencionados y como los mismos reducen el *scattering* de fotones, el principal factor de difusión de la dirección de propagación.

Dentro de las modificaciones al sistema OPT inicialmente propuesto por Sharpe et al. [8] podemos destacar el uso de microscopía de plano de iluminación selectivo (SPIM) [17] como método complementario al mismo, presentado por Bassi et al. [18]. Aquí, el contraste que produce el sistema de SPIM para adquirir un mapeo por fluorescencia puede ser explotado para obtener una reconstrucción tridimensional de la anatomía del espécimen en estudio.

1.1. Principios ópticos de funcionamiento

La adquisición de una proyección consiste en capturar la luz emitida por una fuente LED que atraviesa la muestra, completamente iluminada, siendo esta recogida por un arreglo de detectores. Esta muestra se dispone en un cilindro, lo que permite muestrear distintas vistas en función del ángulo de rotación del mismo, como muestra la Fig. 1.1. Para cada corte a lo largo del eje axial de rotación, el conjunto de líneas proyectadas

en cada ángulo se conoce como sinograma. A partir de los mismos, distintos algoritmos de reconstrucción permiten recuperar un tomograma, el cual informa las características anatómicas (transmisión) y/o funcionales (fluorescencia) del espécimen en el corte correspondiente.

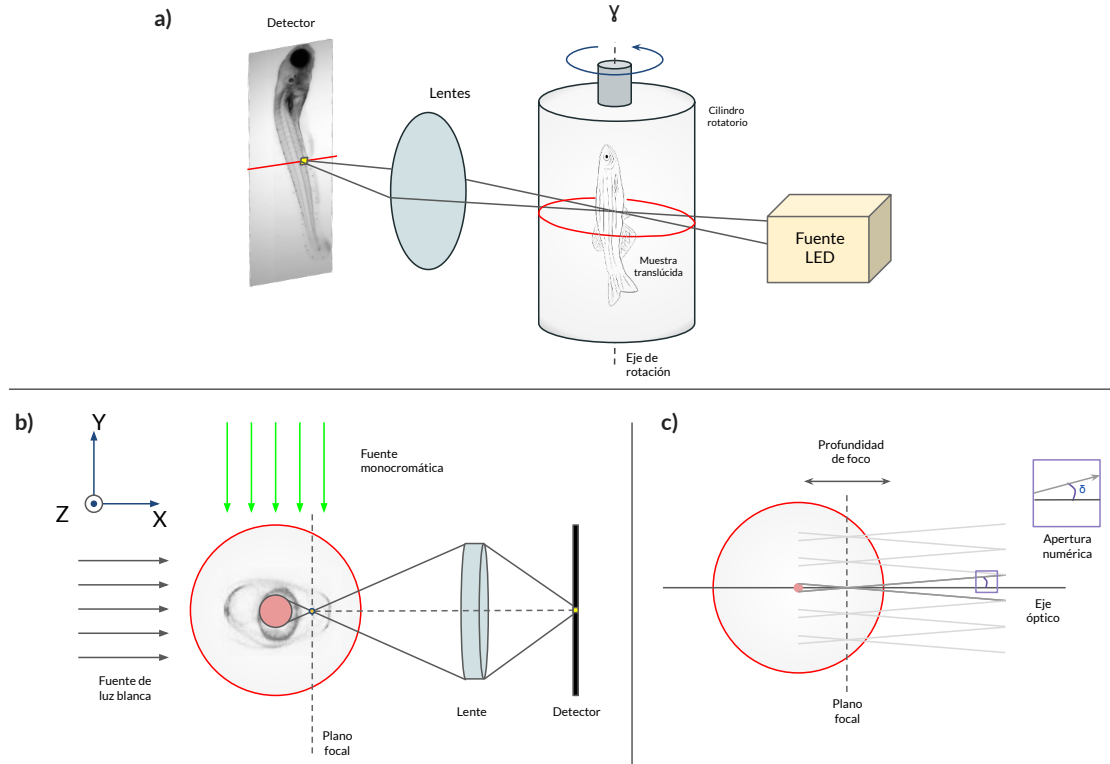


Figura 1.1: **Tomografía de proyección óptica.** **a)** El detector colecta la luz resultante de atravesar de un espécimen translúcido. Dicha muestra se halla contenida en un cilindro rotatorio, el cual permite adquirir distintas proyecciones angulares, parametrizadas por γ . **b)** Fuentes de luz blanca y monocromática pueden utilizarse para extraer información anatómica y/o funcional a partir de la atenuación del tejido y/o emisión de un trazador genético, respectivamente. **c)** El plano focal está dispuesto a medio camino entre el eje de rotación y el frente de la muestra. La apertura numérica (NA), relacionada con el ángulo δ , es lo suficientemente pequeña como para que la profundidad de campo comprenda al menos la mitad de la muestra, asegurando su enfoque en el detector.

Las imágenes OPT son obtenidas mediante un sistema conjugado de objetivo y detector con una fuente no coherente, el cual captura rayos de luz de ancho finito. La resolución mínima Δx y la profundidad de campo, rango de movimiento aceptable del detector o muestra a lo largo del eje óptico, están dadas por:

$$\Delta x = \frac{\lambda}{2\text{NA}} \quad (1.1)$$

$$\text{DOF} = \frac{2\lambda}{\text{NA}^2} \quad (1.2)$$

donde λ es la longitud de onda de la luz incidente y NA es la apertura numérica. Este último caracteriza el rango de ángulos para los cuales el sistema de detección acepta luz incidente.

Como vemos en la Fig. 1.1c, comúnmente el objetivo es colocado de manera que el plano focal se halle a un cuarto de profundidad dentro de la muestra, asegurando que la mitad frontal de la misma se halle en foco [19].

1.2. Reconstrucción mediante retroproyección filtrada

Para recuperar el volumen tridimensional del objeto bajo estudio, una serie de proyecciones ortográficas son utilizadas para reconstruir cortes transversales de la muestra. Generalmente este problema es resuelto de manera analítica mediante el uso del algoritmo de retroproyección filtrada. Describimos los principios matemáticos de este algoritmo analítico, buscando entenderlo posteriormente como parte del conjunto de problemas lineales inversos, desarrollados en el Cap. 2.

Idealmente, en OPT el eje de rotación intersecta el eje óptico del plano de detectores y se halla en paralelo al mismo. Para un dado ángulo de rotación γ , la proyección $p(I, \gamma)$ obtenida es una colección de perfiles de intensidad, con información acerca de la atenuación total $f(x, y)$ de la luz integrada a lo largo de líneas I que atraviesan el objeto. Matemáticamente, esta relación puede describirse mediante una transformada de Radon \mathcal{R} y su producto

$$p(I, \gamma) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \sin \gamma + y \cos \gamma - I) dx dy \quad (1.3)$$

$$(1.4)$$

donde $\delta(x \sin \gamma + y \cos \gamma - I)$ genera un conjunto de líneas paralelas y normales al plano de detección. El conjunto de líneas proyectivas adquiridas en diferentes ángulos γ se conoce como *sinograma*.

¿En qué nos basamos para recuperar información del corte a partir de sus proyecciones? Dado el objeto de interés $f(x, y)$, el teorema de corte de Fourier, ilustrado en la Fig. 1.2 demuestra la equivalencia entre:

- Proyectarla sobre una línea unidimensional y tomar la transformada de Fourier de dicha proyección.
- Realizar una transformada de Fourier para posteriormente tomar un corte a través de su origen, el cual es paralelo a la línea de proyección.

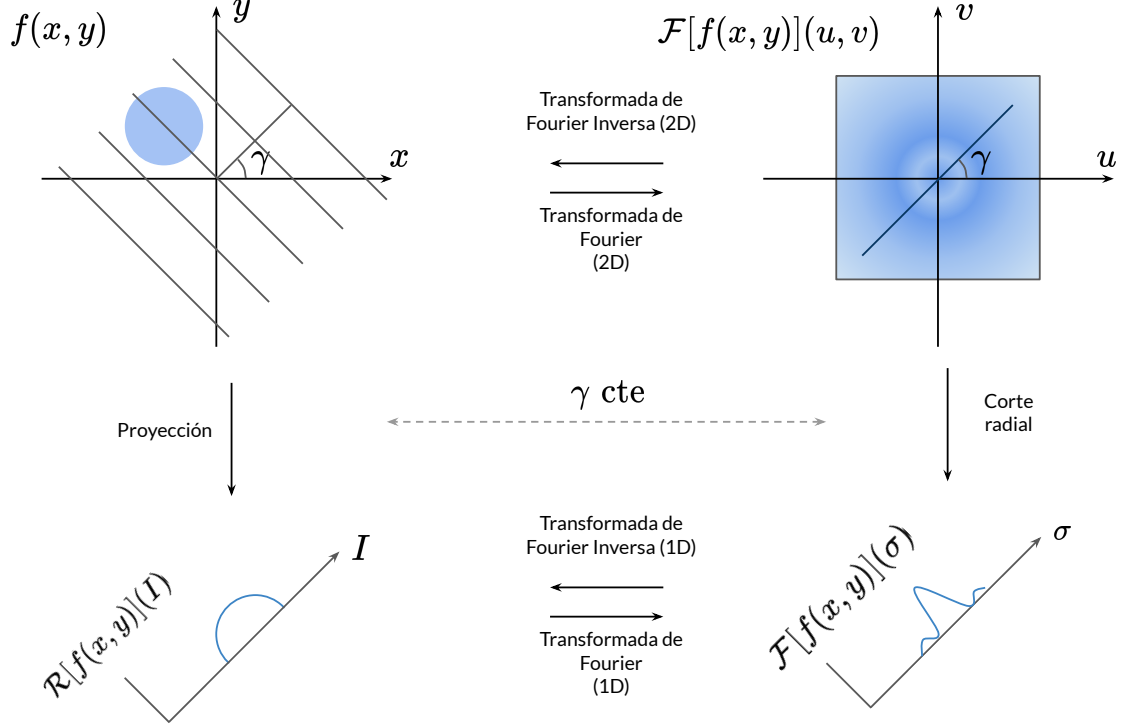


Figura 1.2: **Transformada de Radon.** Para un dado corte transversal al eje de rotación, la proyección $p(I, \gamma)$ es análoga a un corte radial en la transformada de Fourier bidimensional de la imagen. En el espacio de Fourier, las líneas proyectivas adquiridas son un muestreo radial de la transformada de la imagen $\mathcal{F}[f(x, y)](u, v)$, lo cual nos proporciona un método para reconstruirla mediante el teorema de corte.

Recuperar la señal representa un alto costo computacional, dado el muestreo de cortes radiales en el espacio de Fourier, los cuales requieren de ciertos pasos de interpolación intermedios [4]. En este contexto, la reconstrucción por *retroproyección filtrada* (FBP) es un algoritmo rápido y preciso para resolver analíticamente el problema.

El procedimiento consta de dos etapas. En primer lugar, un filtrado que pesa cada proyección en el espacio de Fourier, evitando un sobremuestreo de bajas frecuencias, obteniendo $F(u, v)$

$$F(u, v) = \text{FT}\{P(I, \gamma_n)\} \cdot H(w) \quad (1.5)$$

donde $H(w)$ es el filtro rampa en el espacio de frecuencias u, v , con $u = w \cos \gamma$ y $v = w \sin \gamma$. Finalmente, la antitransformada de $F(u, v)$, $f_{\text{filt}}(x, y)$, es retroproyectada a lo largo de todos los ángulos de la muestra recuperando la imagen $f_{\text{FBP}}(x, y)$:

$$f_{FBP}(x, y) = \int_0^{2\pi} f_{filt}(x \cos \gamma + y \sin \gamma, \gamma) d\gamma \quad (1.6)$$

En el caso discreto, un conjunto de N proyecciones son adquiridas por un detector de Q elementos, definiendo un paso angular $\Delta\gamma_n = \frac{2\pi}{N}$ y una resolución de detección $\Delta x = l/Q$, donde l es el tamaño total del arreglo de detectores. De forma discreta, la transformada de Radon puede expresarse de forma algebraica con una matriz de transformación lineal \mathbf{R} tal que

$$\mathbf{p} = \mathbf{R}\mathbf{f} \quad (1.7)$$

donde $\mathbf{p} \in \mathbb{R}^{N \times Q}$ es el sinograma digitalizado y $\mathbf{f} \in \mathbb{R}^{\sqrt{2}Q \times \sqrt{2}Q}$ es la imagen a recuperar.

Para reconstruir el volumen, resulta útil aplicar una permutación adecuada a la secuencia de proyecciones angulares resultantes de (1.3), permitiendo trabajar en el espacio de sinogramas apilados a lo largo de los cortes definidos por el eje de rotación. Dicha permutación implica operar sobre el volumen de proyecciones, agrupando cada fila horizontal de las proyecciones angulares como parte de un mismo corte transversal, parametrizado por z .

Retomando (1.5) y (1.6), notemos que tanto el enmascarado en el espacio de Fourier de la proyección y su antitransformada son ambas operaciones lineales que resuelven el problema inverso. Dada $\mathbf{R} \in \mathbb{R}^{2Q^2 \times NQ}$, podemos definir la solución \mathbf{f} de (1.7) mediante la pseudoinversa de Moore-Penrose como

$$(\mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \mathbf{p} = \mathbf{f} \quad (1.8)$$

Esta propuesta de pseudo-inversa resulta, por sus autovalores divergentes, no estar acotada. En el capítulo 2 trataremos distintos abordajes a la inversión de problemas lineales con un mal condicionamiento.

1.3. Artefactos de reconstrucción

La técnica de OPT presenta distintos tipos de artefactos que modifican la información de los cortes y degradan su calidad. Por su origen, podemos clasificar los mismos en artefactos experimentales y de muestreo. Mientras que los últimos aparecen a partir de una reconstrucción con un número bajo de muestras, los artefactos experimentales surgen a partir de errores sistemáticos en la adquisición.

1.3.1. Artefactos experimentales

El dispositivo de adquisición OPT, usualmente construido de manera personalizada en el laboratorio por su bajo costo, presenta una variedad de inconvenientes que afectan a la calidad de la imagen reconstruida, que pueden ser corregidos en una etapa de procesamiento previa a la reconstrucción con el método de elección. Para un sinograma cuyo detector desalineado respecto al eje de rotación, presentamos la estrategia utilizada para mitigar su efecto en la imagen reconstruida, causante del artefacto conocido como 'doble pared'.

Si la posición del eje de rotación no ha sido correctamente calibrada de acuerdo a los lineamientos propuestos en la Sec. 1.1, la misma puede ser estimada durante la reconstrucción registrando los sinogramas mediante una transformación de traslación \mathbf{T} a lo largo del eje de los detectores en el sinograma. En el marco de un problema de optimización, la transformación óptima \mathbf{T}_{opt} maximiza una cuantificación de la estructura de la imagen reconstruida, siendo la varianza la elección más sencilla.

$$\mathbf{T}_{opt} = \arg \min_{\mathbf{T}} \text{Var}[\text{FBP}(T(p(x, y)))] \quad (1.9)$$

Esta transformada es aplicada posteriormente al volumen de proyecciones, permitiendo corregir el *blurring* sin necesidad de marcadores fiduciales en la muestra [20]. En la Sec. 4.2 mostramos los resultados de aplicar este algoritmo de corrección de manera iterativa.

Otros artefactos tales como ensombrecimientos causados por iluminación irregular o ruido correlacionado dentro del detector por píxeles 'muertos' o 'calientes' son tratados con mayor detalle por Walls et al. [21].

1.3.2. Artefactos por muestreo

Bajo la teoría analítica de Radon, el teorema de Nyquist-Shannon nos permite afirmar que obtendremos una única reconstrucción si el objeto en cuestión es muestreado con una frecuencia mayor al doble de la frecuencia más alta de detalle en el mismo. En el problema de proyección paralela que encaramos en OPT, para un muestreo de S píxeles por línea proyectada, se requieren N proyecciones para alcanzar la cota propuesta. Si D es el diámetro del objeto a escanear y Δx es la diferencia entre dos puntos de la reconstrucción, luego el número de píxeles en cada línea proyectiva esta dado por:

$$S = \frac{D}{\Delta x} \quad (1.10)$$

Para un escaneo alrededor de 360° ¹, cada punto es nuevamente escaneado luego

¹En el caso de que el objeto en estudio presente homogeneidad y simetría, escanear sobre 180° es suficiente

de un camino equivalente a πD , esto para cada punto situado en la superficie del objeto cilíndrico, la mayor frecuencia espacial. En este caso, el número de proyecciones necesario es:

$$N = \frac{\pi D}{\Delta y} \quad (1.11)$$

donde Δy es la resolución del detector en el espacio de proyecciones. Según el teorema de Nyquist-Shannon, requerimos que el muestreo $\Delta y \geq 2\Delta x$, obteniendo entonces que el número de proyecciones debe ser al menos:

$$N \geq \frac{\pi}{2} S \quad (1.12)$$

Un cálculo aproximado para el sistema OPT abordado en este trabajo muestra que, para el diámetro de un embrión de pez cebra $D = 20\text{mm}$ con un detector de espaciamiento $\Delta x = 1,3 \mu\text{m}$, el número de muestras necesario para muestrear por encima de la cota de Nyquist-Shannon resulta en alrededor de 24000 proyecciones angulares, un valor generalmente impracticable en OPT. La regla de oro en CT indica una reconstrucción razonable con la igualdad entre el número de proyecciones N y el número de detectores Q .

Disminuir el número de proyecciones adquiridas implica una aceleración inversamente proporcional de la adquisición de un volumen completo, el motivo central de este trabajo. Dados los conjuntos de sinogramas, definimos el *factor de aceleración de una reconstrucción* como el cociente entre el número de proyecciones total adquiridas y el número utilizado para la reconstrucción submuestreada. Este factor cuantifica la cantidad de veces que reducimos la información angularmente de manera equiespaciada (un factor de $\times 10$ implica que utilizaremos 1 de cada 10 proyecciones consecutivas adquiridas). En adelante, consideramos este muestreo dado que es el modo ubicuo de implementación en el *hardware* del dispositivo de adquisición.

Para el caso particular de la transformada de Radon, el factor de aceleración de una reconstrucción implica un submuestreo en el espacio de Fourier (Fig. 1.3a) que se manifiesta en un *aliasing* en la imagen reconstruida analíticamente (Fig. 1.3b).

En el siguiente capítulo presentaremos diversas alternativas al planteo del problema de reconstrucción, apuntando a una aceleración de la adquisición que permita disminuir tanto la fototoxicidad del método como los tiempos de anestesia que implica el muestreo de especímenes *in vivo*.

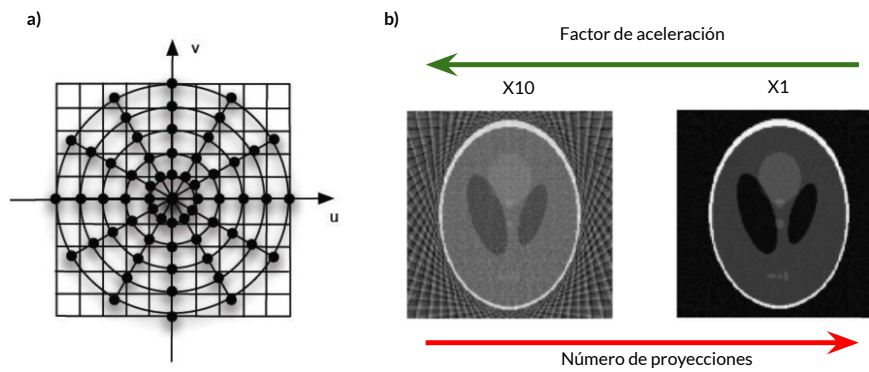


Figura 1.3: **Artefactos por muestreo.** **a)** muestra como la interpolación entre coordenadas polares a cartesianas en el espacio de Fourier implica un muestreo pobre de altas frecuencias, **b)** manifestándose en una adquisición afectada por artefactos al aumentar el espaciamiento angular. El mismo puede verse como un factor de aceleración en la adquisición del organismo bajo estudio.

Capítulo 2

Teoría de reconstrucción de imágenes

2.1. Problema lineal inverso

De manera informal, un problema inverso consiste en utilizar observaciones para calcular el valor de ciertos parámetros, usualmente buscando la mejor solución sugerida por los datos [22]. Resolverlo implica la búsqueda de los parámetros del modelo $\mathcal{A} : \mathbb{R}^M \rightarrow \mathbb{R}^N$ que mapea entre los parámetros a estimar $\mathbf{x} \in \mathbb{R}^M$ y las observaciones $\mathbf{y} \in \mathbb{R}^N$ como:

$$\mathbf{y} = \mathcal{A}(\mathbf{x}) \quad (2.1)$$

Dada la pertinencia a la aplicación, en adelante trabajaremos con variables reales aunque el enfoque puede extenderse a espacios vectoriales complejos.

En el caso de un problema lineal, el mismo puede escribirse en notación matricial como

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad (2.2)$$

En el caso de la transformada de Radon, $\mathbf{A} = \mathbf{S}\mathbf{R}$, donde \mathbf{S} representa el patrón de muestreo. Recuperar \mathbf{x} puede ser un problema mal planteado en el sentido de Hadamard [23], especialmente para \mathbf{A} rectangular. Podemos plantear su reconstrucción \mathbf{x}_{rec} de manera general dentro de un esquema de optimización regularizado

$$\mathbf{x}_{rec} = \arg \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \mathcal{R}(\mathbf{x}) \quad (2.3)$$

donde $\mathcal{R}(\mathbf{x}) : \mathbb{R}^M \rightarrow \mathbb{R}$ puede pensarse como una penalidad a la complejidad de \mathbf{x} y λ es el parámetro de regularización. Entre los regularizadores más utilizados en LIP se encuentra la variación total (TV) [24] y **wavelets** [25].

Bajo este marco, la teoría de *compressed sensing* nos permite afirmar que, si \mathbf{x} tiene una representación rara bajo una transformación dada, podemos reconstruir \mathbf{x} a partir de \mathbf{y} con un número de muestras dramáticamente menor a M [10] utilizando un procedimiento no lineal. Presentaremos una aplicación de este marco al problema de reconstruir una imagen OPT, refiriendo al trabajo de Trémouhéac [26] para un desarrollo teórico preciso del mismo.

Resolución de *compressed sensing* mediante TwIST

Añadamos ruido aditivo η al problema lineal inverso propuesto en la Ec. (2.4):

$$\mathbf{x} = \mathbf{A}\mathbf{x} + \eta \quad (2.4)$$

En adelante trataremos con imágenes bidimensionales, aunque utilizaremos el formalismo propuesto anteriormente para describir el mapeo directo sobre señales unidimensionales. Para no recargar la notación, utilizaremos \mathbf{x} indistintamente para ambas situaciones, asumiendo su representación en cada caso.

Introduzcamos el operador de variación total (TV) para una señal bidimensional \mathbf{x} .

$$\Phi_{\text{TV}}(\mathbf{x}) = \sum_i \sqrt{(D_i^H)^2 + (D_i^V)^2} \quad (2.5)$$

donde D_i^H y D_i^V representan la diferencia finita de primer orden para el píxel i -ésimo de \mathbf{x} en sentido horizontal y vertical, respectivamente. Este operador permite generar una representación rara en el dominio de los gradientes, con la propiedad de preservar bordes de la señal a recuperar. Introduciendo este regularizador en el problema de optimización,

$$\mathbf{x}_{\text{rec}} = \arg \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2^2 + \lambda \Phi_{\text{TV}}(\mathbf{x}) \quad (2.6)$$

La solución de este tipo de problemas de optimización, convexos con regularización no cuadrática, puede encontrarse con algoritmos iterativos de *shrinkage/thresholding* (IST). Para problemas fuertemente mal condicionados, su versión de dos pasos es denominada TwIST (*Two-Step Iterative Shrinkage/Thresholding*) [27]. Esta solución está dada por el esquema:

$$\mathbf{x}_1 = \Gamma_{\lambda}(\mathbf{x}_0) \quad (2.7)$$

$$\mathbf{x}_{k+1} = (1 - \alpha)\mathbf{x}_{k-1} + (\alpha - \beta)\mathbf{x}_k + \beta\Gamma_{\lambda}(\mathbf{x}_k) \quad (2.8)$$

donde α y β son constantes que dependen de los autovalores de $\mathbf{A}^T\mathbf{A}$ y $\Gamma_{\lambda}(\mathbf{x}_k)$ esta

dado por

$$\Gamma_\lambda(\mathbf{x}_k) = \Phi_\lambda(\mathbf{x} + \mathbf{A}^T(\mathbf{Ax} - \mathbf{y})) \quad (2.9)$$

donde Φ_λ es el método de *denoising* de variación total (TV), resuelto iterativamente en [24]. Nuevamente, referimos al trabajo doctoral de Tremoulheac [26] para una introducción a la teoría de operadores proximales y funciones promotoras a dominios ralos tales como Φ_λ y su aplicación a problemas de optimización.

2.2. Aprendizaje profundo

El aprendizaje profundo permite obtener representaciones en múltiples niveles de abstracción mediante la concatenación de múltiples capas de procesamiento [28]. Si bien podemos remontar el deseo quimérico de 'concebir máquinas con la posibilidad de aprender de la experiencia' a trabajos no publicados de Alan Turing a mediados del siglo XX, no fue sino hasta entrada la pasada década que, a hombros de gigantes, el concepto de 'deep learning' se tornó inexorablemente poderoso y accesible para el desarrollo científico y tecnológico. Podemos distinguir ciertos hitos que han concebido la realidad que atraviesa este campo:

- El extenso desarrollo en teoría de redes neuronales, tanto desde un enfoque neurocomputacional como su abstracción a modelos computacionales apuntadas a la resolución de tareas complejas con capacidad de generalización.
- La proposición de un algoritmo que explote la cualidad recursiva de retropropagar errores para el aprendizaje de representaciones [29].
- La aceleración en el entrenamiento y despliegue a partir del uso de *hardware* de procesamiento paralelo y entornos de desarrollo de alto nivel.

Una red neuronal artificial (ANN), el elemento fundacional de este paradigma, es el resultado del aprendizaje de un mapeo óptimo entre un conjunto de elementos de entrada y salida, con capacidad de generalización ante nuevos estímulos. Dentro de este extenso panorama, describimos dos tipos de arquitecturas que resultan ubícuos en nuestra labor: redes convolucionales y redes recurrentes.

2.2.1. Redes convolucionales

En 1959, Hubel y Wiesel echaron luz al estudio de neuronas que responden a pequeñas regiones del campo visual del gato, proponiendo un primer modelo en cascada orientado a explicar su representación de patrones en la corteza estriada (V1). Basado

en el trabajo de Denker [30], el problema de reconocimiento de dígitos manuscritos fue abordado por LeCun (Fig. 2.1) utilizando una ANN mediante *backpropagation*, una implementación eficiente para propagar errores a través de redes multicapa [31].

El concepto subyacente en ambas propuestas, la aplicación en cascada de la operación de convolución permite extraer representaciones de alta dimensionalidad de los datos de entrada para distintas tareas. Un concepto importante en estas arquitecturas es el de *campo receptivo* de una neurona, definido como la región del espacio a la cual la presencia de una entrada afecta a la respuesta de la primera en cuestión. Al aumentar la profundidad de las redes, dicha región aumenta en tamaño y su respuesta en complejidad.

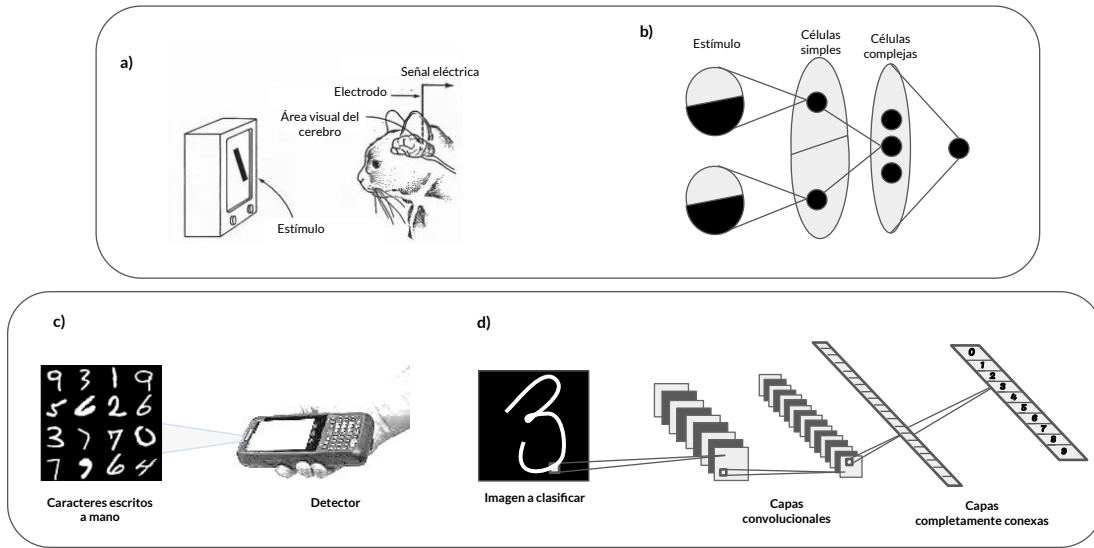


Figura 2.1: **Modelos de visión.** a) La respuesta neuronal selectiva del gato ante variaciones de posición y orientación inspiraron a Hubel y Wiesel b) el modelado de células simples y complejas que responden ante regiones locales y globales del estímulo. c) y d) muestran el modelo de red neuronal convolucional (CNN) propuesto por LeCun, donde el mapeo es aprendido mediante el algoritmo de *backpropagation* para resolver el problema de clasificación de símbolos decimales escritos a mano.

En el caso de una imagen \mathbf{x} , la operación de convolución consiste en la aplicación lineal de un *kernel* \mathbf{h} sobre regiones de cada píxel, pudiendo escribirse como:

$$\mathbf{y}[i, j] = (\mathbf{x} * \mathbf{h})[i, j] = \sum_{m, n} x[i - m, j - n] h[m, n] \quad (2.10)$$

Para una ANN arbitraria, consideraremos una función de pérdida $\mathcal{L}(\theta)$ que cuantifica la idoneidad del conjunto de parámetros θ al resolver una tarea. En el caso de una red neuronal convolucional, estos parámetros componen los *kernels* de la arquitectura mostrada en la Fig. 2.1.d, permitiendo en este caso puntual extraer las características relevantes de la imagen de entrada en capas dispuestas en cascada. En el campo

de reconstrucción de imágenes, donde el objetivo es recuperar una imagen restaurada de una entrada corrupta, dicha función de pérdida cuantifica la similaridad entre el resultado de dicha al atravesar la red y un conjunto de imágenes *ground truth*.

Reconstrucción mediante CNN: U-Net

Diseñada para abordar el problema de segmentación de imágenes, la arquitectura U-Net [2] provee un entorno versátil para la remoción de artefactos. En su descripción original, U-Net se compone de dos redes que, de manera secuencial, contraen y expanden las dimensiones espaciales de la red.

Cada bloque de la red contractiva puede describirse mediante la aplicación de dos convoluciones (*kernels* 3×3), cada una con activación ReLU y un submuestreo mediante una operación *max pool*¹ 2×2 con *stride* de 2, acabando por doblar el espacio de canales. Cada paso en la red expansiva consiste en sobremuestrear la información espacial del mapa de características, luego reduciendo en un factor 2 el número de canales y concatenando con su correspondiente par recortado en la red contractiva, para finalizar aplicando dos convoluciones 3×3 con activación ReLU. Por último, se aplica una convolución 1×1 para obtener un mapa de clases de la imagen con dimensiones espaciales idénticas a la entrada.

Davis et al. [1] proponen su utilización para la tarea de remoción de artefactos en la reconstrucción de imágenes OPT. En un esquema de entrenamiento supervisado, la red convolucional es optimizada enseñando imágenes reconstruidas mediante FBP con un número pobre de muestras (entrada) y su par completamente muestreado (salida) (Fig. 2.2).

2.2.2. Redes recurrentes

En el ámbito de las ciencias de la computación, el concepto de recurrencia emerge de la idea de definir una función por inducción (Fig. 2.3a) [32]. Utilizando esta idea, las redes neuronales recurrentes (RNN) permiten procesar una secuencia de entrada un elemento a la vez, computando un vector de estados que contiene la información acerca de la historia de los elementos pasados de la secuencia de manera implícita (Fig. 2.3b) [28]. De forma manifiesta, la RNN puede expandirse en lo que se conoce como su representación desenrollada (Fig. 2.3c).

De manera general, podemos describir una red neuronal recurrente mediante un mapeo f_{θ}^k tal que

$$\mathbf{x}_{k+1} = f_{\theta}^k(\mathbf{x}_{(k)}; \mathbf{x}_{(0)}) \quad (2.11)$$

¹*max pooling* refiere a la operación de tomar el máximo de una región del campo receptivo

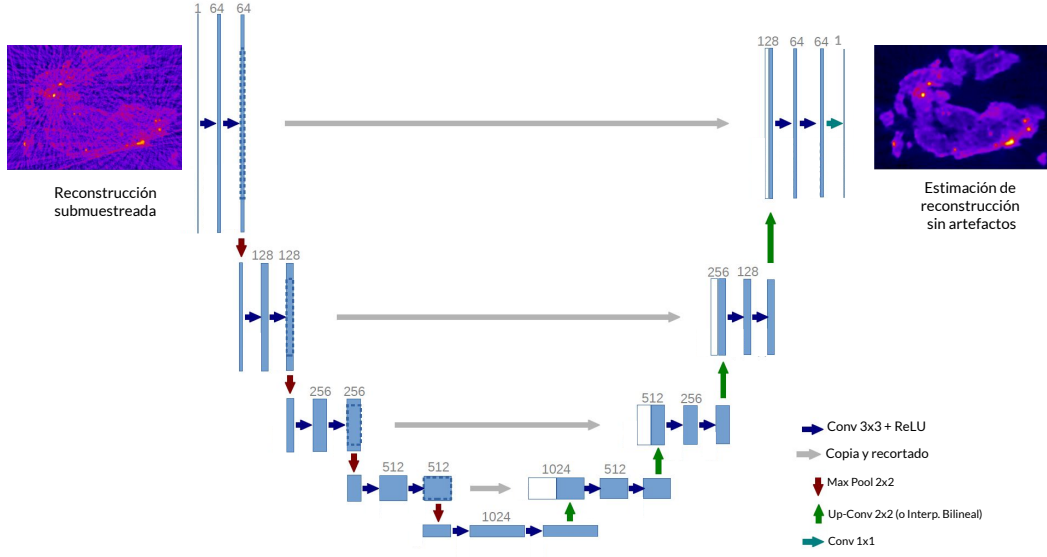


Figura 2.2: **Reconstrucción OPT mediante U-Net.** Dos caminos, contractivo y expansivo, son enlazados en las diferentes etapas de muestreo mediante *skip connections*. Utilizando pares de imágenes OPT reconstruidas mediante FBP, la arquitectura U-Net permite remover artefactos en imágenes obtenidas de un muestreo pobre del corte. Adaptación de [1] y [2].

donde k indica el número de iteraciones y θ los parámetros del modelo. El caso en que los parámetros son compartidos² a través de las iteraciones, tal que $f_{\theta}^k = f_{\theta}$, ha demostrado un desempeño competitivo en diversas tareas [33][34]. En el caso de existir, \mathbf{x}_K para $K \rightarrow \infty$ representa un punto fijo del operador $f_{\theta}(\cdot; \mathbf{x}_0)$ [11].

2.3. LIP con red desenrollada: *Model-based Deep Learning* (MoDL)

A lo largo del presente capítulo introducimos los ingredientes necesarios para construir un método que, bajo la cualidad de aproximador universal de una ANN [35], provea una solución dentro del marco teórico de un problema lineal inverso. Presentamos una arquitectura *model-based deep learning* (MoDL), la propuesta de Aggarwal et al. [14] para lograr este objetivo.

Denotando con \mathcal{N}_{θ} al estimador de ruido que depende de los parámetros θ , podemos escribir la versión sin ruido de x , $\mathcal{D}_{\theta}(x)$ como:

$$\mathcal{D}_{\theta}(\mathbf{x}) = \mathbf{x} - \mathcal{N}_{\theta}(\mathbf{x}) \quad (2.12)$$

²Otra acepción para esta técnica es la de pesos atados o *weight tying*.

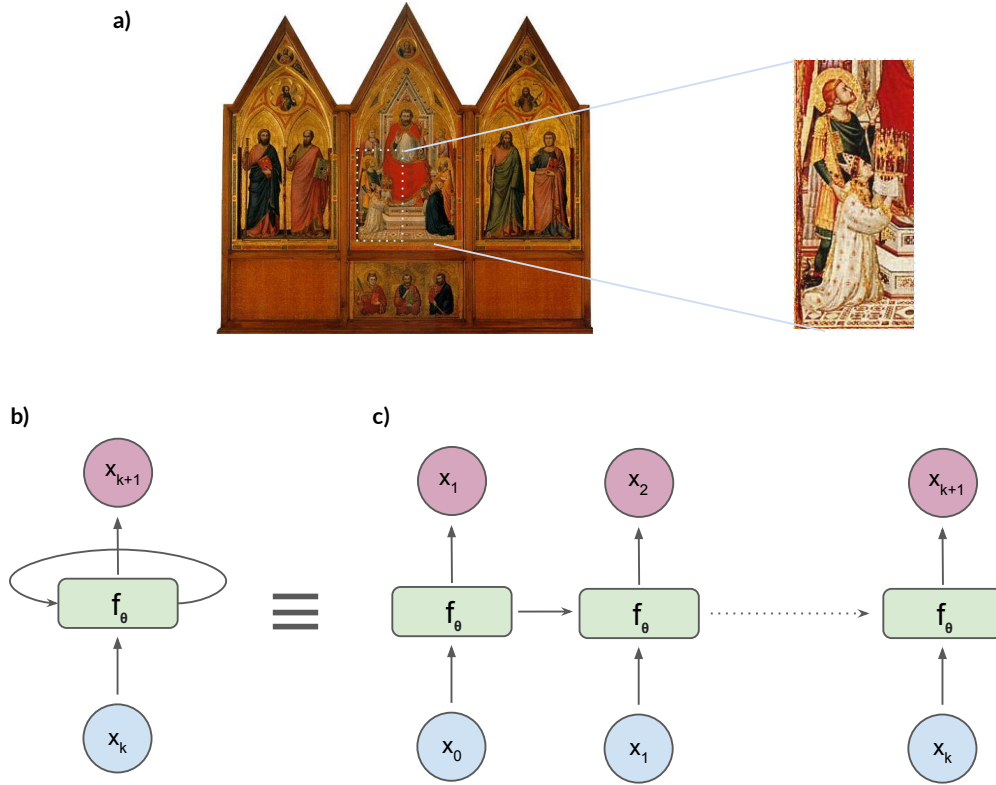


Figura 2.3: **Recursión.** La idea de recurrencia, **a)** ya capturada por Giotto en su *Tríptico Stefaneschi*, ha sido explotada en el campo de ANN con la aparición del algoritmo de *backpropagation* para entrenar redes recurrentes **b)**. Su aplicación puede verse de manera más sencilla al tratar con su forma desenrollada **c)**.

Con \mathcal{N}_θ como función regularizadora en (2.3), podemos obtener una reconstrucción \mathbf{x}_{rec} como:

$$\mathbf{x}_{rec} = \arg \min_{\mathbf{x}} \underbrace{\|\mathbf{Ax} - \mathbf{y}\|_2^2}_{\text{Consistencia de datos}} + \lambda \underbrace{\|\mathcal{N}_\theta(\mathbf{x})\|^2}_{\text{Regularización}} \quad (2.13)$$

Dado el mapeo que resuelve este problema, desarrollemos de manera aproximada alrededor de \mathbf{x}_k la serie de Taylor de primer orden para $\mathcal{D}_\theta(\mathbf{x}_k + \Delta\mathbf{x})$:

$$\mathcal{D}_\theta(\mathbf{x}_k + \Delta\mathbf{x}) \approx \underbrace{\mathcal{D}_\theta(\mathbf{x}_k)}_{\mathbf{z}_k} + \mathbf{J}_k^T \nabla \mathbf{x} \quad (2.14)$$

donde \mathbf{J}_k^T es la matriz jacobiana. Fijando $\mathbf{x} = \mathbf{x}_k + \Delta\mathbf{x}$, el termino de penalidad puede ser aproximado como

$$\|\mathbf{x} - \mathcal{D}_\theta(\mathbf{x}_k + \Delta\mathbf{x})\|^2 \approx \|\mathbf{x} - \mathcal{D}_\theta(\mathbf{x})\|^2 + \|\mathbf{J}_k \Delta\mathbf{x}\|^2 \quad (2.15)$$

donde el segundo término tiende a cero para pequeñas perturbaciones $\Delta\mathbf{x}$. Con esta aproximación en un entorno de \mathbf{x}_n , la solución de (2.13) puede ser obtenida mediante

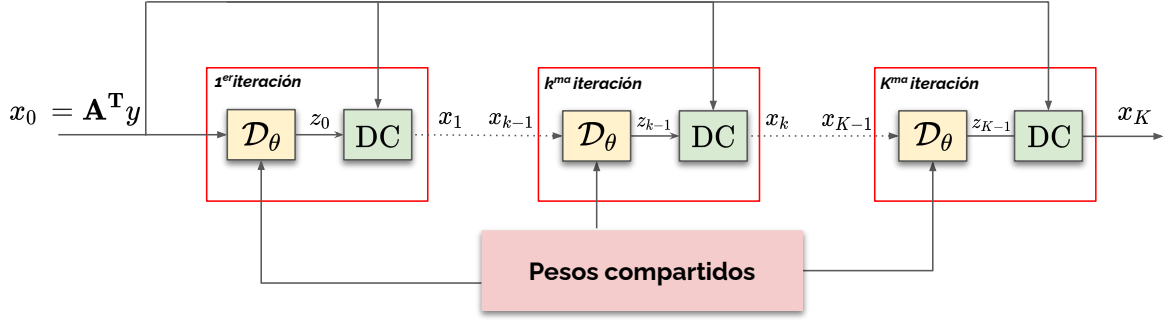


Figura 2.4: **Esquema de red desenrollada.** La imagen obtenida al aplicar el operador lineal hermítico atraviesa un esquema de reconstrucción alternante que resuelve el problema de *denoising* al aplicar \mathcal{D}_θ y forzar la consistencia de los datos (DC) resolviendo (2.16b).

un algoritmo de gradiente proximal [11]:

$$\mathbf{z}_k = \mathcal{D}_\theta(\mathbf{x}_k) \quad (2.16a)$$

$$\mathbf{x}_{k+1} = (\mathbf{A}^H \mathbf{A} + \lambda \mathbf{I})^{-1} (\mathbf{A}^H \mathbf{y} + \lambda \mathbf{z}_k) \quad (2.16b)$$

Recordando la Fig. 2.3, podemos pensar el mapeo iterativo como una red desenrollada, donde se resuelven los problemas lineal y no lineal de manera alternante como muestra la Fig. 2.4. El entrenamiento de esta red consiste en optimizar tanto los pesos de la CNN, \mathcal{D}_θ , como el peso de la regularización λ respecto de una función de costo \mathcal{L} , que cuantifica la calidad de la reconstrucción. Notemos que, tal como presentamos en (2.11), los pesos de \mathcal{D}_θ son compartidos para todas las iteraciones.

El entrenamiento de punta a punta, más conocido como *end-to-end*, permite capturar la estadística de los artefactos introducidos por un submuestreo de \mathbf{y} . Fijando el número de iteraciones K , para una red con pesos compartidos podemos propagar el gradiente como

$$\nabla_\theta \mathcal{L} = \sum_{k=0}^{K-1} \mathbf{J}_\theta^T(\mathbf{z}_k) \nabla_{\mathbf{z}_k} \mathcal{L} \quad (2.17)$$

donde $\mathbf{J}_\theta(\mathbf{z}_k)$ representa la matriz jacobiana con elementos $[\mathbf{J}_\theta(\mathbf{z}_k)]_{i,j} = \partial z_i / \partial w_j$.

Bajo este esquema, resulta crucial remarcar que únicamente los pesos de la CNN son entrenables y que la operación de inversión 2.16b es lineal. Esto implica que la retropropagación de errores a través del bloque de consistencia de datos (DC) es transparente al mismo. Notemos entonces que

$$\nabla_{\mathbf{z}_k} \mathcal{L} = \mathbf{J}_{\mathbf{z}_{k-1}}^T(\mathbf{x}_k) \nabla_{\mathbf{x}_k} \mathcal{L} \quad (2.18)$$

donde $\mathbf{J}_{\mathbf{z}_{k-1}}(\mathbf{x}_k)$ representa la matriz jacobiana con elementos $[\mathbf{J}_{\mathbf{z}_{k-1}}(\mathbf{x}_k)]_{i,j} = \partial \mathbf{x}_i / \partial \mathbf{z}_j$.

Si observamos (2.16b), podemos ver que $x_{k+1} = \mathcal{Q}(\mathbf{A}^H \mathbf{y} + \lambda \mathbf{z}_k)$, donde $\mathcal{Q} = (\mathbf{A}^H \mathbf{A} + \lambda \mathbf{I})^{-1}$. Dada la simetría de \mathcal{Q} , luego

$$\nabla_{\mathbf{x}_k} \mathcal{L} = \lambda \mathcal{Q} \nabla_{\mathbf{z}_k} \mathcal{L} \quad (2.19)$$

En el siguiente capítulo discutiremos nuestra propuesta basada en este enfoque, sus detalles de implementación y detalles asociados al abordaje de la reconstrucción de imágenes OPT.

Capítulo 3

ToMoDL: Propuesta e implementación

Por sus condiciones, la reconstrucción de imágenes OPT esta comprometida por numerosas fuentes de artefactos. En este trabajo de investigación, abordamos con especial énfasis el problema de recuperar un corte tomográfico \mathbf{x} a partir de pocas proyecciones ortográficas. Como notamos en el Cap. 3, el mal condicionamiento del operador de retroproyección filtrada presenta soluciones afectadas por artefactos en caso de un muestreo pobre del objeto bajo estudio. Por otra parte, su tratamiento bajo un esquema iterativo implica una regularización razonable que condicionen las posibles soluciones. Escoger este recurso adecuadamente no es sencillo, resultando en soluciones típicamente engorrosas con limitadas prestaciones a la hora de preservar contraste y textura en la reconstrucción. Más aún, dichas soluciones iterativas resultan en un alto costo computacional por su implementación secuencial en CPU.

Acercándonos a los avances de la pasada década, el campo del aprendizaje automático ha evolucionado vertiginosamente, en particular para el campo de la visión por computador. En la tarea que nos compete, la arquitectura U-Net ha sido reformulada como un mapeo entrenable entre imágenes con y sin artefactos producidos por *aliasing*, utilizando la retroproyección filtrada como método efectivo de reconstrucción de imágenes OPT. Por otra parte, una nueva familia de mapeos entrenables basados en un dado modelo plantean una revisión a la aplicación de redes convolucionales para resolver problemas lineales inversos. Es en este contexto en el que presentamos ToMoDL, una arquitectura *model-based* que, bajo un mapeo iterativo, permite encontrar la función de regularización óptima para la remoción de artefactos producidos por una baja cantidad de proyecciones.

3.1. Elementos de la propuesta

A la vista de los elementos introducidos en la Sec. 2.3, podemos categorizar ToMoDL como una arquitectura desenrollada [11], cuya idea es fijar un número de iteraciones K que corresponda al punto fijo del mapeo general (2.11):

$$\mathbf{x}_K = \underbrace{(\mathbf{A}^T \mathbf{A} + \lambda \mathbf{I})^{-1}}_{\mathcal{Q}} (\mathbf{A}^T \mathbf{y} + \lambda \mathcal{D}_\theta(x_{k-1})) \quad (3.1)$$

donde $k = 0, \dots, K$. \mathbf{A} en nuestro caso es la matriz de Radon \mathbf{R} definida en (1.7), \mathbf{A}^T la matriz real transpuesta \mathbf{R}^T que implementa la *retroproyección*, $\mathbf{x}_0 = \mathbf{A}^T \mathbf{y}$ y $\mathcal{D}_\theta : \mathbb{R}^{\sqrt{2}Q \times \sqrt{2}Q} \rightarrow \mathbb{R}^{\sqrt{2}Q \times \sqrt{2}Q}$ una red convolucional arbitraria encargada de remover artefactos de la imagen $\mathbf{x} \in \mathbb{R}^{\sqrt{2}Q \times \sqrt{2}Q}$ recuperada a partir del sinograma $\mathbf{y} \in \mathbb{R}^{Q \times N}$.

Denoising

La regularización propuesta en la Sec. 2.3 busca obtener un estimador de ruido $\mathcal{N}_\theta(x)$ de la imagen de entrada. En el campo de las redes convolucionales, esta tarea ha encontrado una solución ubícua en un enfoque residual. En el esquema sencillo de la ResNet (*residual network*) propuesta (Fig. 3.1), utilizamos una concatenación de bloques convolucionales con redes de 64 filtros con una función de activación lineal rectificadora (ReLU), basados en la propuesta de Zhang et al. [36]. El paso no lineal descrito en (2.16a) es resuelto al remover el ruido estimado de la imagen de entrada mediante una conexión de salto¹, obteniendo $\mathcal{D}_\theta(x) = x - \mathcal{N}_\theta(x)$.

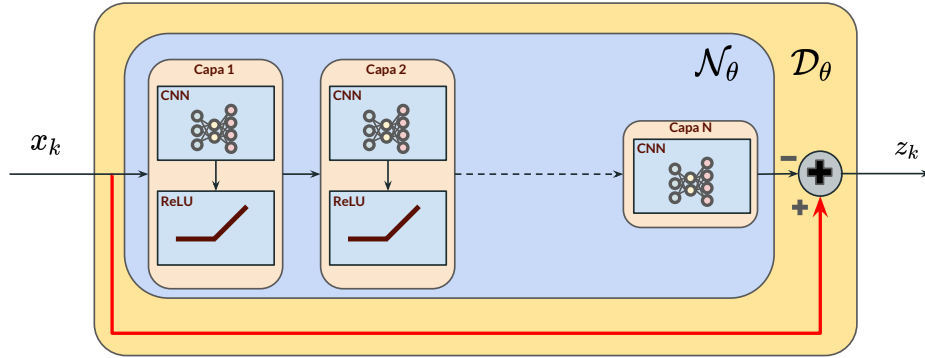


Figura 3.1: **Bloque residual.** El ruido de la imagen de entrada x es estimado a través de una serie de redes convolucionales activadas mediante una función ReLU. Para estimar la imagen sin ruido, una conexión de salto es sumada a la salida de \mathcal{N}_θ .

En contraste con Aggarwal et al. [14], optamos por no utilizar la operación de *batch normalization* en las capas convolucionales. La misma tiene por objetivo realizar una estimación de los primeros momentos de un *batch*, número de muestras a propagar

¹Usualmente conocida como *skip connection*

a través de la red, para reducir el sesgo por covariancia de los datos. Sin embargo, observamos que al utilizar un tamaño pequeño de muestras por *batch*, la operación encuentra complicaciones en aprender estos momentos, refiriendo a [37] para un estudio más detallado de alternativas de normalización en estos casos.

Consistencia de datos

La aplicación del operador \mathcal{Q} se vuelve rápidamente un problema analíticamente intratable, dada por la dimensionalidad de sus operadores (para matrices de orden n , la inversión de matrices densas tiene una complejidad $\mathcal{O}(n^3)$). Este obstáculo fue abordado mediante la resolución acelerada de la transformada de Radon en GPU utilizando *TorchRadon*[38] ha permitido la aplicación numérica de técnicas basadas en el algoritmo de gradientes conjugados.

La librería *TorchRadon*, una implementación en CUDA que opera directamente con tensores PyTorch, resulta crucial para el entrenamiento de una red basada en modelo al problema de OPT. Esto puede verse inmediatamente en (2.19), donde cada retropropagación de errores involucra la aplicación de los operadores de proyección y retroproyección \mathbf{A} y \mathbf{A}^T .

Provista una eficaz manera de operar entre sinogramas e imágenes, resulta imperioso resolver (3.1) sin realizar la computación analítica de \mathcal{Q} , un operador que, aún siendo disperso, impide asimismo descomposiciones del tipo Cholesky por sus dimensiones. Las condiciones de simetría y su definición positiva permiten uso del algoritmo de gradientes conjugados[39], definido en el Alg. 1.

Algoritmo 1 Algoritmo de gradientes conjugados (CG)

Input \mathbf{y}

Initialize $\mathbf{x}_0 = 0$, $\mathbf{r}_0 = \mathbf{y}$, $\mathbf{p}_0 = \mathbf{y}$, $k = 0$, $\beta_0 = 0$, $\alpha_0 = 0$

while $k < K$ or $\mathbf{r}^T \mathbf{r} < 10^{-5}$ **do**

$\alpha_k \leftarrow \mathbf{r}_k^T \mathbf{r}_k / \mathbf{p}_k^T \mathbf{A} \mathbf{p}_k$

$\mathbf{p}_{k+1} \leftarrow \mathbf{x}_k + \alpha_k \mathbf{p}_k$

$\mathbf{r}_{k+1} \leftarrow \mathbf{r}_k - \alpha_k \mathbf{A} \mathbf{r}_k$

$\beta_k \leftarrow \mathbf{r}_{k+1}^T \mathbf{r}_{k+1} / \mathbf{r}_k^T \mathbf{r}_k$

$\mathbf{p}_{k+1} \leftarrow \mathbf{r}_{k+1} + \beta_k \mathbf{p}_k$

$k \leftarrow k + 1$

end while

Output $\hat{\mathbf{x}} = \mathbf{x}_k$

En nuestra implementación, la condición de parada resulta de una norma residual $\mathbf{r}^T \mathbf{r} < 10^{-5}$ o alcanzar un número de iteraciones, fijado en 10. Por las dimensiones del sistema, optamos por no utilizar un preconditionamiento del algoritmo.

Función de pérdida

En este esquema, el entrenamiento de la red se basa en minimizar una dada función de costo $\mathcal{L}(\mathbf{x}_K, \mathbf{t})$ respecto a los parámetros θ del regularizador y el término de peso λ , sumando sobre los pares de imágenes de salida de la red desenrollada \mathbf{x}_K^i y una imagen objetivo \mathbf{t}^i , con $i = 1, \dots, P$ donde P es el número de ejemplos presentados a la red. Para este trabajo proponemos dos opciones de interés en el campo de la reconstrucción como función de pérdida: el error cuadrático medio (MSE),

$$\mathcal{L}_{\text{MSE}}(\mathbf{x}_K, \mathbf{t}) = \frac{1}{2PQ^2} \sum_{i=0}^P \|\mathbf{t}^i - \mathbf{x}_K^i\|^2 \quad (3.2)$$

y la medida del índice de similitud estructural (SSIM), descrita en términos de su función inversa $1 - \text{SSIM}$

$$\mathcal{L}_{\text{SSIM}}(\mathbf{x}_K, \mathbf{t}) = \frac{1}{P} \sum_{i=0}^P 1 - \text{SSIM}(\mathbf{t}^i, \mathbf{x}_K^i) \quad (3.3)$$

donde SSIM se define como el producto de

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \underbrace{\frac{2\mu_{\mathbf{x}}\mu_{\mathbf{y}} + c_1}{\mu_{\mathbf{x}}^2 + \mu_{\mathbf{y}}^2 + c_1}}_{l(\mathbf{x}, \mathbf{y})} \cdot \underbrace{\frac{2\sigma_{\mathbf{x}}\sigma_{\mathbf{y}} + c_2}{\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{y}}^2 + c_2}}_{c(\mathbf{x}, \mathbf{y})} \cdot \underbrace{\frac{2\sigma_{\mathbf{x}\mathbf{y}} + c_3}{\sigma_{\mathbf{x}}\sigma_{\mathbf{y}} + c_3}}_{s(\mathbf{x}, \mathbf{y})} \quad (3.4)$$

donde $\mu_{\mathbf{x}}$, $\mu_{\mathbf{y}}$ son los valores medios de \mathbf{x} y \mathbf{y} , $\sigma_{\mathbf{x}}$, $\sigma_{\mathbf{y}}$ son las desviaciones estándar de \mathbf{x} y \mathbf{y} y las constantes c_1 y c_2 funcionan como estabilizadores relacionados al rango dinámico para el denominador propenso a tender a cero.

Mientras que SSIM suele ser una medida de interés para observar la reconstrucción de las muestras, el error cuadrático medio resulta relativo al rango de intensidades a muestrear. Es por eso que en el campo del procesamiento de imágenes, es común utilizar la relación señal a ruido pico (PSNR) como medida relativa al MSE entre la imagen reconstruida y la objetivo:

$$\text{PSNR} = 10 \log \left(\frac{\text{MAX}_I}{\text{MSE}} \right) \quad (3.5)$$

donde MAX_I es el máximo valor de intensidad que puede tomar un píxel, considerando $I > 0$.

3.2. Datos de entrenamiento

Para el entrenamiento del mapeo recurrente, utilizamos 12 conjuntos de datos de peces cebra (*Danio Rerio*) que representan cuatro secciones, denominadas como *cabeza*, *cuerpo*, *cola inferior* y *cola superior* (ver Fig. 3.3a) y corresponden a 5 especímenes

diferentes, descritos en la Tab. 3.1. Cada uno de estos volúmenes está formado por $N = 720$ proyecciones adquiridas con un sensor de $Q = 800$, obteniendo $Z = 880$ cortes normales al eje de rotación y fueron adquiridos entre 1-5 días post-fertilización (dpf) con una lente magnificadora $\times 10$.

Volumen #	Espécimen	Sección	Días Post Fertilización
1	A	Cabeza	5
2	B	Cola Inferior	3
3	C	Cabeza	5
4	D	Cabeza	3
5	B	Cabeza	3
6	E	Cabeza	5
7	E	Cola Superior	5
8	B	Cuerpo	3
9	D	Cabeza	1
10	B	Cola Superior	3
11	E	Cuerpo	5
12	E	Cola Inferior	5

Tabla 3.1: Volúmenes utilizados para la evaluación de métodos supervisados.

Preprocesamiento

Para todos los volúmenes a utilizar para el entrenamiento, la desalineación en el eje de rotación es corregida mediante el método de maximización de la variancia en la reconstrucción (ver Sec. 1.3.1). Para conservar el tamaño de los elementos detectados, se asignan los valores del borde más cercano a aquellos píxeles indeterminados luego de la registración que corrige el sinograma.

Al operar mediante la transformada de Radon, las retroproyecciones submuestreadas y completamente muestreadas (Fig. 3.3) son normalizadas en el rango de $[0, 1]$ para atravesar la red. Para reducir el costo computacional de los experimentos, el tamaño del sensor es submuestreado mediante una interpolación bilineal a un valor de $Q_{sub} = 141$, dando como resultado una imagen al reconstruir de 100×100 píxeles.

3.3. Detalles de implementación

En la etapa de entrenamiento supervisado, la entrada y la salida se corresponden a la retroproyección sin filtrar submuestreada y su contraparte filtrada y reconstruida a partir de todas las muestras disponibles, respectivamente. En la práctica, los mismos fueron reconstruidos y salvados en memoria previamente para reducir el tiempo de cómputo.

Para un dado corte, el modelo computa una reconstrucción completa a partir de una retroproyección sin filtrar $\mathbf{R}^T \mathbf{y}$, obtenida bajo los pasos mostrados en la Fig. 3.2. La función de costo entre la reconstrucción \mathbf{x}_K y el objetivo \mathbf{t}_i , una reconstrucción FBP con todas las muestras disponibles, permite actualizar los pesos de \mathcal{D}_θ mediante el algoritmo de *backpropagation* en el esquema ToMoDL propuesto. Los parámetros θ y el peso de regularización λ son entrenados utilizando un esquema de optimización *Adam* [40] con tasa de aprendizaje *lr* de 1×10^{-4} . Si bien en [14] se propone un 'pre-calentamiento'² de los pesos con una red de una única iteración, no encontramos que dicha técnica mejore los resultados en nuestro contexto.

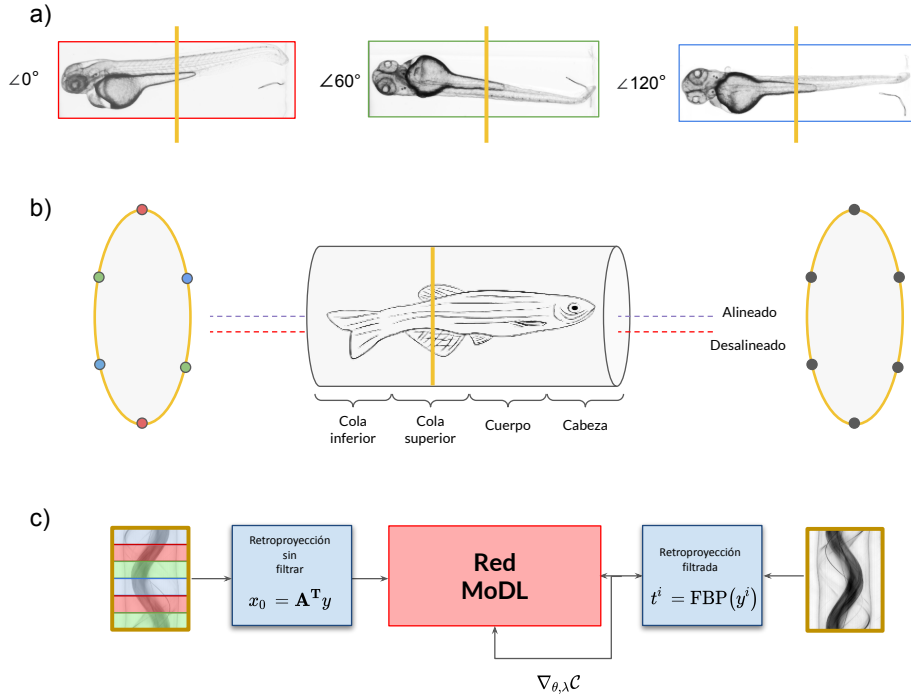


Figura 3.2: **Procesamiento de imágenes OPT** a) distintas proyecciones angulares son obtenidas con un muestreo angular equiespaciado. Dichas proyecciones son el resultado de concatenar 4 secciones, como muestra b) de la especie *Danio Rerio*. c) La red ToMoDL aprende a reconstruir y remover artefactos de forma supervisada, a partir de pares de imágenes retroproyectadas a partir de subconjuntos de muestras (mostradas como puntos) como entrada y reconstrucciones con todas las proyecciones disponibles.

El algoritmo 2 resume los pasos propuestos para el entrenamiento a partir de la obtención de las muestras para un número de épocas E fijado en 40. En cada época, todos los ejemplos de entrenamiento son extraídos de manera aleatoria y presentados a la red en subconjuntos denominados *batches*, cuyo error es utilizado para actualizar los pesos en base al optimizador utilizado. El problema de explosión de gradientes, frecuente en el entrenamiento de redes recurrentes [41], fue paliado mediante el uso de un *gradient*

²Mayormente conocido como *warm start*

clipping, normalizando a la unidad la norma de los gradientes al retropropagar errores.

Algoritmo 2 Algoritmo de entrenamiento para ToMoDL

Input y
Initialize $lr = 0,0001$, $\mathbf{z}_0 = 0$, $k = 0$, $e = 0$, $\theta_0 \sim \mathcal{N}(0, 1/N_{l-1})$
while $e < E$ **do**
 $\mathbf{x}_0 \leftarrow \mathbf{R}^T \mathbf{y}$
 $\mathbf{t} \leftarrow \text{FBP}(\mathbf{y})$
 while $k < K$ **do**
 $\mathbf{z}_k \leftarrow \mathcal{D}_{\theta_e}(\mathbf{x}_k) = \mathbf{x}_k - \mathcal{N}_{\theta_e}(\mathbf{x}_k)$
 $\mathbf{x}_{k+1} \leftarrow (\mathbf{R}^T \mathbf{R} + \lambda \mathbf{I})^{-1}(\mathbf{R}^T \mathbf{y} + \lambda \mathbf{z}_k)$
 $k \leftarrow k + 1$
 end while
 $L_e \leftarrow \nabla_{\theta_e} \mathcal{L}(\mathbf{x}_K, \mathbf{t})$
 $\theta_{e+1} \leftarrow \text{Adam}(\theta_e, L_e)$
end while
Output

La Fig. 3.3 muestra el resultado de desenrollar la red ToMoDL en sus componentes lineal y no lineal. La ampliación del recuadro rojo en la fila inferior muestra cómo la dinámica del mapeo converge a la reconstrucción final al cabo de un número bajo de iteraciones. Observamos como la entrada, un corte de cabeza reconstruido mediante una retroproyección sin filtrar submuestreada, sufre una remoción de sus artefactos utilizando (2.16a) y es reconstruida al resolver (2.16b).

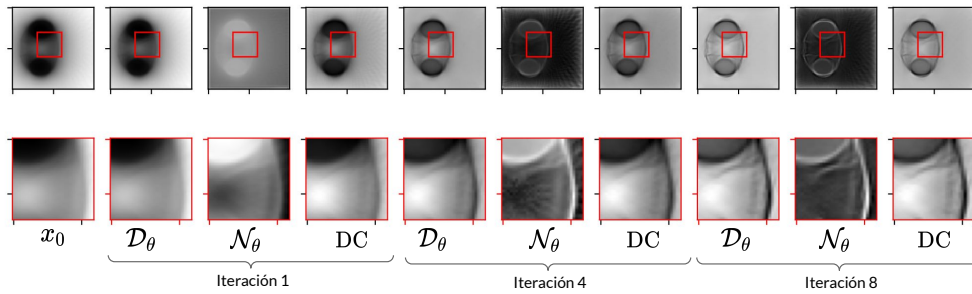


Figura 3.3: Etapas de reconstrucción iterativa mediante ToMoDL. La imagen x_0 que ingresa a ToMoDL atraviesa las etapas alternantes que resuelven el problema lineal (DC) y el *denoising* no lineal (\mathcal{D}_{θ}) con el correspondiente ruido estimado (\mathcal{N}_{θ}). En la fila inferior se muestra una ampliación de la reconstrucción, mostrando como el ruido estimado evoluciona a una mayor estructura en cada iteración de la red para reconstruir la imagen de la cabeza de un pez cebra.

Capítulo 4

Experimentos y resultados

La evaluación de los métodos mencionados y su comparativa con el enfoque ToMoDL presenta resultados interesantes al indagar no solo en su formulación, sino en los datos utilizados para evaluarlos. Una vez formulados los experimentos de búsqueda de hiperparámetros para nuestro método, este capítulo desglosa los resultados del algoritmo de alineamiento para minimizar artefactos del dispositivo experimental. Sobre las variantes propuestas, obtenemos las métricas de reconstrucción PSNR y SSIM en función del factor de aceleración, principal aplicación de las técnicas propuestas. Posteriormente se analizan subconjuntos de estos resultados, marginados por las características de los especímenes muestreados. Por último, investigamos la estadística de los volúmenes disponibles y su correlación con la calidad de reconstrucción correspondiente a cada técnica.

4.1. Materiales y métodos

4.1.1. Variantes utilizadas

Se evaluaron cuatro metodologías para reconstruir imágenes de tomografía de proyección óptica, descritas con mayor detalle en los capítulos 1 y 2, a partir de un submuestreo del conjunto de proyecciones adquirido, cuyas características se ilustran en la Tab. 4.1:

El algoritmo de retroproyección filtrada es aquel que presenta la mayor rapidez en el cómputo de la reconstrucción analítica (para el uso de la librería *TorchRadon*), aunque su desempeño se halla fuertemente limitado por la cota de muestreo de Nyquist-Shannon discutida en la Sec. 1.3.2. Los demás métodos de reconstrucción implementados también utilizan este algoritmo, ya por única vez (U-Net) o de manera iterativa (TwIST, ToMoDL) para resolver un dado problema inverso. Cabe mencionar que existen ciertos algoritmos supervisados que contemplan un aprendizaje del operador directo, como el caso de AUTOMAP [42], aunque su implementación para Radon aún

	Tipo de reconstrucción	Regularización	Parámetros	Tiempo de reconstrucción (por corte 2D)	Parámetros entrenables (Q ~100 píxels)
FBP	Analítica	–	–	~100 ms (CPU) ~1 ms (GPU)	–
TwIST	Iterativa	Variación Total (TV)	$\lambda = 0,01$ [9]	~10 s (CPU) ~30 ms (GPU)	–
U-Net	Aprendizaje supervisado	–	–	~10 ms (GPU)	~10 M
ToMoDL	Aprendizaje supervisado basado en modelo	Red Convolutiva	K = 8 L = 8 λ aprendible[14]	~100 ms (GPU)	~200.000

Tabla 4.1: Métodos de reconstrucción utilizados.

permanece como un problema abierto por su complejidad computacional.

La resolución iterativa ha sido propuesto para OPT mediante TwIST bajo una regularización de variación total (TV) por Correia et al. [9]. El problema de mal condicionamiento del operador $\mathbf{R}^T \mathbf{R}$, señalado por los autores del algoritmo [27], ha sido observado en su implementación como la principal desventaja en materia de tiempos de reconstrucción.

El método supervisado U-Net presenta una rápida remoción de artefactos y resulta consistente ante muestras nuevas. El amplio campo receptivo de la red U-Net proporciona una solución al problema de reconstrucción a expensas de un alto número de parámetros libres. Por otra parte, el operador $\mathbf{R}^T \mathbf{R}$, como se refleja posteriormente en la Sec. 4.2, no presenta una invariancia traslacional (traslaciones en el espacio de sinogramas no resultan en reconstrucciones trasladadas). Esto es un desafío para redes convolucionales entrenadas con datos cuya geometría no haya sido corregida en la etapa de entrenamiento.

Por último, ToMoDL presenta una alternativa basada en el modelo geométrico del problema de reconstrucción OPT. Una regularización basada en una red convolutiva residual con un número de parámetros sustancialmente menor a la red U-Net resulta en una alternativa eficiente en términos de memoria, aunque el tiempo de reconstrucción, por su planteo iterativo, es el mayor de los métodos implementados en una tarjeta gráfica (GPU).

4.1.2. Hiperparámetros

Las gráficas en la Fig. 4.1 muestran el efecto de aumentar el número de capas y el número de iteraciones en la red ToMoDL entrenada bajo las dos funciones de pérdida propuestas, SSIM y PSNR, utilizando datos de validación y reconstruyendo datos con un factor de aceleración $\times 10$. Escogimos tanto un número de iteraciones y de capas de la red residual menores a 10 tanto por un impedimento computacional como por la degradación presentada en distintos trabajos para estas técnicas [14] [11].

En la Fig. 4.1a observamos que el efecto de las capas satura para un número igual o mayor a 8 (utilizando 5 iteraciones), mientras que el número de iteraciones presenta una saturación en $K = 8$. En ambos experimentos observamos una desempeño mayor en redes entrenadas utilizando el error cuadrático medio (MSE) como función de pérdida, por lo que en adelante nos referimos a ToMoDL bajo esta metodología de aprendizaje.

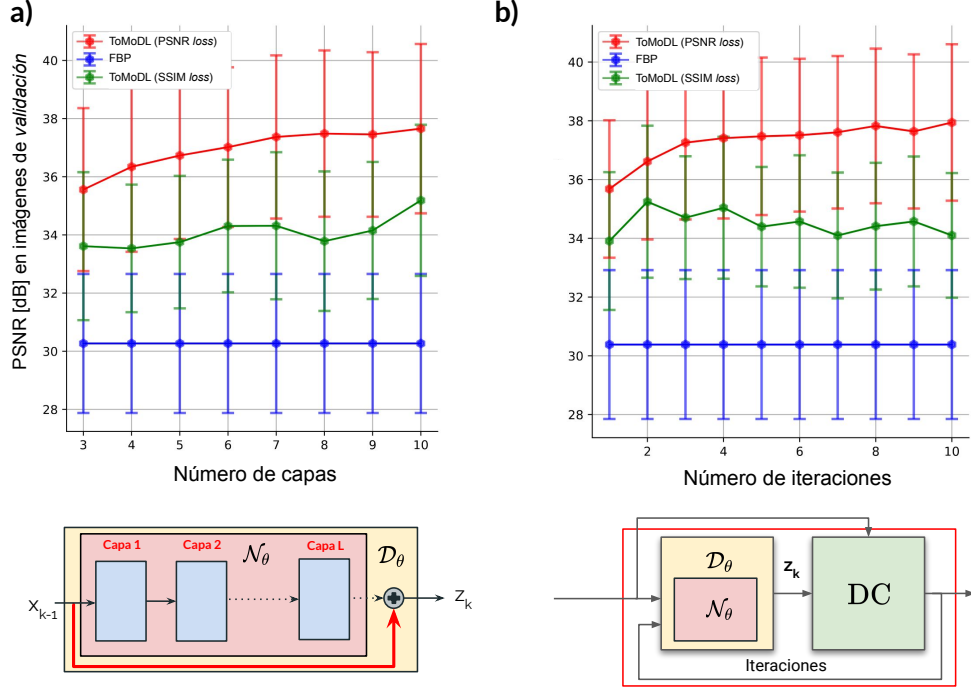


Figura 4.1: **Exploración de hiperparámetros.** Mediante una búsqueda en iteraciones, número de capas y las funciones de pérdida propuestas, a partir de la PSNR de las reconstrucciones del conjunto de evaluación resolvemos por utilizar una arquitectura ToMoDL entrenada de entrada a salida para minimizar la función *mean square error* con $K = 8$ y $L = 8$

4.1.3. Utilización de los datos de entrenamiento

Para estimar la precisión de los modelos supervisados sobre el total de los volúmenes disponibles, utilizamos una validación de K iteraciones (*K-Folding Cross-Validation*), donde los mismos son particionados en $K_{it} = 4$ subconjuntos. En cada iteración, uno de estos subconjuntos es utilizado como fuente de datos de prueba, mientras que el resto es utilizado para entrenar y validar el modelo. Finalmente se computa la media de desempeño de la métrica de evaluación del modelo. La partición los volúmenes de muestra, presentados en la Tab. 3.1, tomó en cada iteración:

- $K_{it} = 1$

- Entrenamiento/Validación: 1, 2, 3, 4, 5, 6, 7, 8, 9
- Evaluación: 10, 11, 12
- $K_{it} = 2$
 - Entrenamiento/Validación: 10, 11, 12, 1, 2, 3, 4, 5, 6,
 - Evaluación: 7, 8, 9
- $K_{it} = 3$
 - Entrenamiento/Validación: 7, 8, 9, 10, 11, 12, 1, 2, 3
 - Evaluación: 4, 5, 6
- $K_{it} = 4$
 - Entrenamiento/Validación: 4, 5, 6, 7, 8, 9, 10, 11, 12
 - Evaluación: 1, 2, 3

Al aplicar esta validación cruzada, todos los volúmenes son utilizados en la etapa de evaluación con el correspondiente modelo entrenado con el conjunto correspondiente. En lo que sigue, haremos una mezcla de estos resultados sin distinguir la partición correspondiente. Adicionalmente, esta técnica nos permite conservar el balance de las clases presentadas en la anterior sección en la etapa de evaluación, conservando las características del conjunto de datos disponibles totales.

4.2. Algoritmo de alineamiento

La corrección de artefactos experimentales de desalineación del eje de rotación respecto al sensor de detección fueron abordados en base al algoritmo descrito en la Sec. 1.3.1. En la gráfica 4.2d observamos la notoria degradación de la imagen para desalineaciones del orden de 10 píxeles, cuantificados en la varianza de la reconstrucción obtenida. La aparición del efecto de 'doble pared' es minimizado mediante una búsqueda iterativa para reducir el tiempo de cómputo, disminuyendo en un orden de magnitud el paso de la transformación y la ventana alrededor del máximo de varianza de reconstrucción. Para conservar el tamaño del sinograma, los píxeles indeterminados luego de la transformación tomaran el valor de intensidad del borde más cercano.

Este algoritmo resulta de interés en métodos de reconstrucción tomográfica donde la transformación geométrica no se aprende y el operador directo $\mathbf{A}^T \mathbf{A}$ no es invariante traslacional.

Dada la geometría del problema, los pares de proyecciones adquiridos a 180° de diferencia son idealmente iguales e invertidos en el caso de que el eje de rotación este

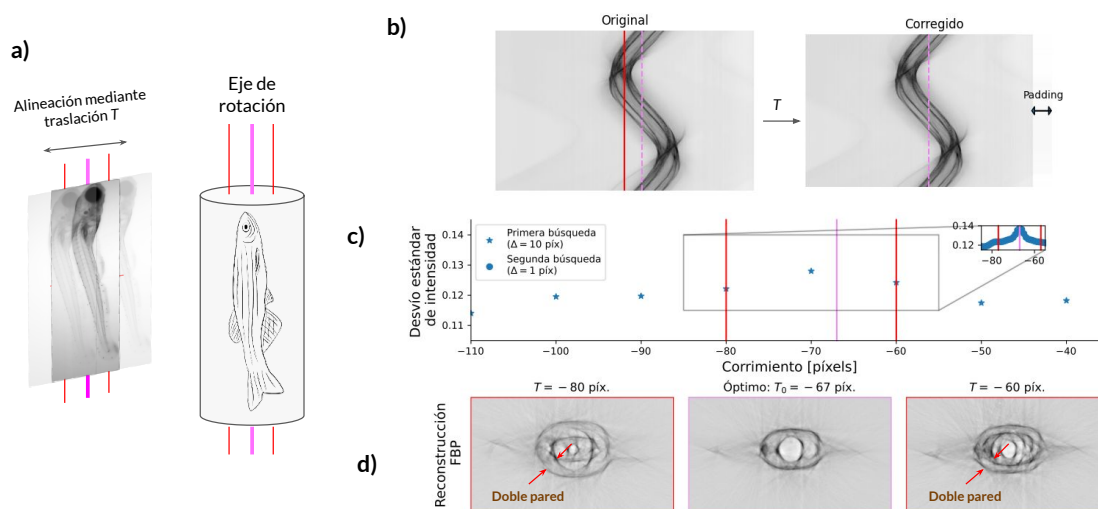


Figura 4.2: **Alineamiento mediante maximización de varianza de la intensidad.** a) muestra el corrimiento relativo entre el eje de rotación y el detector, lo cual genera una traslación en el espacio de sinogramas, corregido en b) mediante una transformación de traslación. Para hallarla, buscamos de forma iterativa en una grilla de traslaciones, reduciendo el paso en píxeles hasta hallar la transformación que maximice la varianza de la intensidad de la reconstrucción c). Así, el efecto de *doble pared* en los bordes de la estructura, observado en d), puede ser corregido.

correctamente emplazado respecto al sensor de detección, encontrando un corrimiento relativo entre ellas en el caso de una desalineación. Una posible alternativa a la corrección sería registrar ambas muestras, buscando la transformación que maximice la similitud entre ambas proyecciones. Este enfoque alternativo resulta problemático tanto por la inhomogeneidad de las proyecciones opuestas y el alto tiempo de cómputo, aunque permite prescindir del método de reconstrucción dado que se trata únicamente de corregir la geometría del problema a resolver.

4.3. Comparación entre métodos

La aceleración del proceso de adquisición de imágenes OPT es el objetivo ulterior de este trabajo. En los capítulos 2 y 3 hemos desarrollado los inconvenientes de reconstruir con pocas muestras bajo el método analítico de retroproyección filtrada y las diversas propuestas para paliar estos efectos. Estudiamos el impacto de aumentar el factor de aceleración y cómo cada método en particular se desempeña al marginalizar sobre condiciones del volumen bajo estudio.

4.3.1. Factor de aceleración

Las gráficas de la Fig. 4.3 muestran en dos paneles el desempeño de los métodos de reconstrucción bajo distintos factores de aceleración a lo largo de todos los datos de evaluación sobre los *folds* disponibles. Para factores de aceleración bajos, observamos un pobre desempeño en términos de PSNR en el uso de métodos iterativos y supervisados, en favor de utilizar FBP para estos casos. Esto implica que los primeros introducen un error de procesamiento que no supera la alta calidad de las imágenes FBP con la mitad o incluso 6 veces menos proyecciones, introduciendo un procesamiento que incluso degrada las mismas.

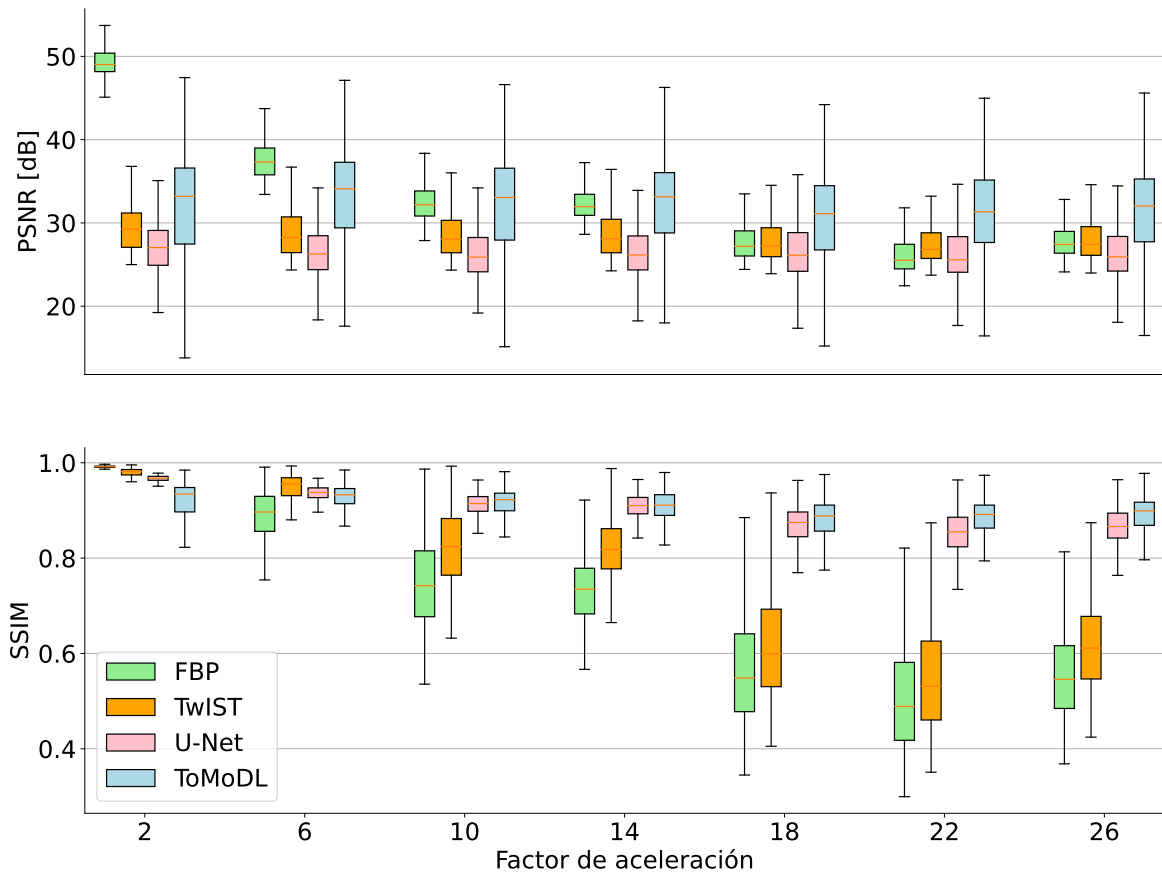


Figura 4.3: **Desempeño de los métodos de reconstrucción a diferentes factores de aceleración.** Observamos como, de manera consistente, la reconstrucción mediante ToMoDL supera los 30 dB para todos los factores de aceleración, con una medida del índice de similitud estructural mejor al resto de los métodos a partir de un factor de aceleración $\times 10$.

La medida del índice de similitud estructural (SSIM) presenta un resultado interesante para los métodos supervisados como U-Net y ToMoDL, donde su desempeño es comparable y mayor en media a 0.9, resultando en una reconstrucción de razonable calidad. En comparación, el desempeño de TwIST resulta comparable al de FBP para factores de aceleración altos. Para el factor de aceleración $\times 10$, en el cual la di-

ferencia de desempeño entre los métodos comienza a hacerse notoria, los paneles de la Fig. 4.4 muestran el resultado de reconstruir cortes de la cabeza y cola superior del animal. Observamos que ToMoDL permite una razonable remoción de artefactos de muestreo respecto a los demás métodos, sin por ello suavizar las regiones con estructura. Debemos remarcar que el método logra estas características utilizando un número de parámetros entrenables dos ordenes de magnitud menor a los utilizados por la red U-Net a costa de incluir implícitamente la resolución del problema lineal inverso.

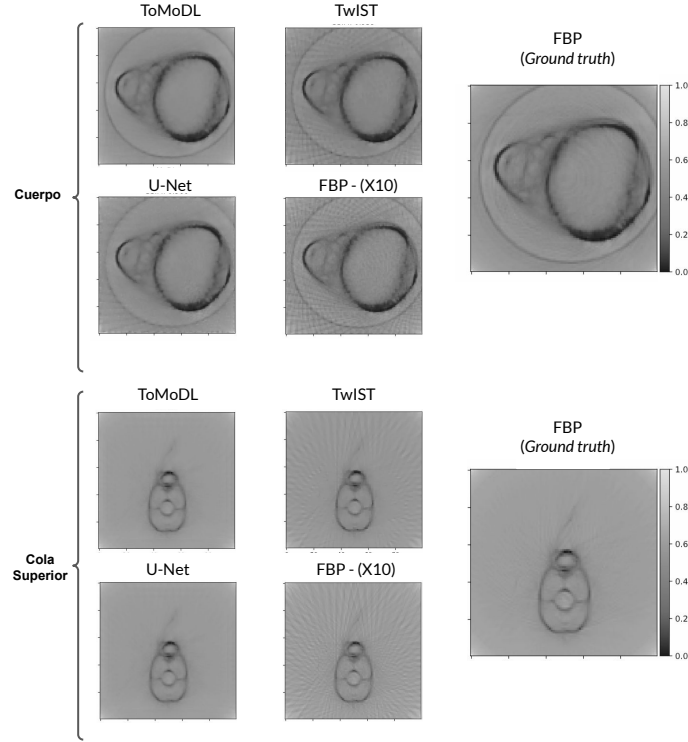


Figura 4.4: **Reconstrucción mediante los distintos métodos.** Tomando reconstrucciones con un factor de aceleración $\times 10$, se muestran ejemplos en dos secciones del animal bajo estudio. Observamos que mientras para la cabeza, ToMoDL es el único método que efectivamente remueve artefactos propios del submuestreo de proyecciones al reconstruir.

Sin embargo, la dispersión en los resultados de reconstrucción de la red ToMoDL plantea ciertas inquietudes respecto a la robustez del método, especialmente al observar la PSNR de reconstrucción. Tomando por caso el factor de aceleración $\times 22$, para el cual la dispersión en todas las métricas es máxima, analizamos como se distribuyen estas medidas en tres características de los datos bajo estudio: sección del animal, días post fertilización y espécimen.

4.3.2. Impacto de las características del conjunto de datos

En su base teórica, la propuesta ToMoDL busca introducir el modelo físico de reconstrucción para mejorar la capacidad de generalización, forzando que la red pro-

puesta funcione como una mapeo a un dominio que promueva una descripción rala de la imagen, en analogía al uso del operador de variación total. Sin embargo, la hipótesis de que la estructura de los objetos utilizados en el entrenamiento condicionen el aprendizaje no está descartada. Para evaluar de manera parcial esta posibilidad, observemos las discrepancias en la distribución de las métricas de reconstrucción al marginar bajo distintas características de los volúmenes disponibles.

Al observar el desempeño de los métodos estudiados en función de la sección de los peces cebrá con un factor de aceleración $\times 22$ (Fig. 4.5), para el cual la degradación de la calidad de reconstrucción analítica es notoria, la dispersión de los mismos para reconstruir la cola inferior del animal es la más evidente. La principal discrepancia cualitativa en estos volúmenes es la alta proporción de cortes que no contienen estructura del animal. En estos casos, dado que las reconstrucciones objetivo provienen del método FBP, el factor de aceleración presenta un impacto menor en las métricas obtenidas mediante el método analítico, mientras que la posibilidad que los métodos supervisados como ToMoDL y U-Net fueren cierta estructura en la reconstrucción es alta.

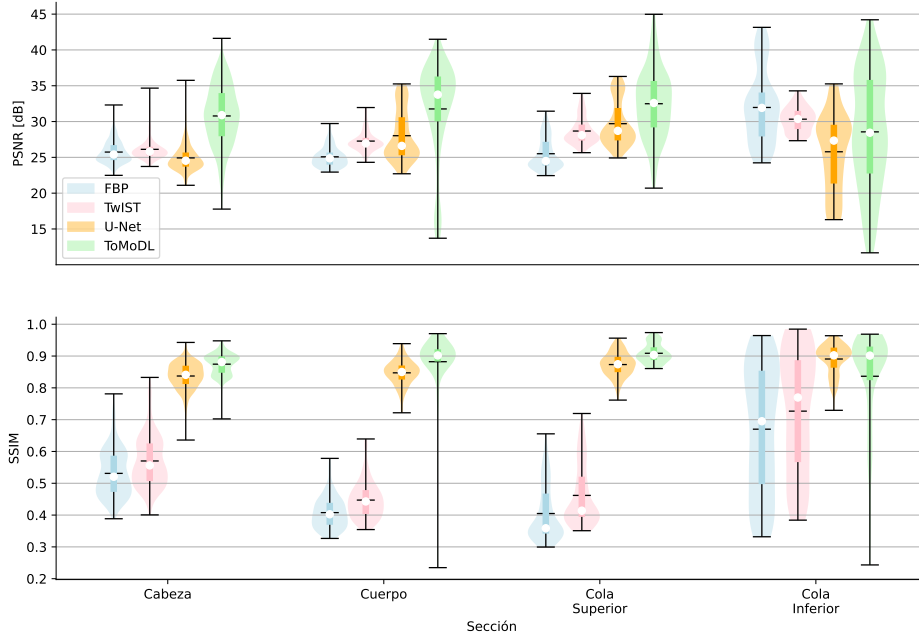


Figura 4.5: **Reconstrucción por sección.** Para cada método propuesto, observamos la distribución de las métricas de reconstrucción PSNR y SSIM para un factor de aceleración $\times 22$, en función de la sección. La cola inferior, la cual presenta gran parte de los cortes sin tejido, es aquel que condiciona el funcionamiento de los métodos. Vemos que ToMoDL presenta un desempeño consistente en el resto de las secciones, incluso superior al estado del arte en reconstrucción OPT.

La gráfica mostrada en la Fig. 4.6a permite analizar la estadística de la reconstrucción de cada corte i a partir del desvío estándar de intensidad σ_I^i , segmentando por sección del animal bajo estudio. Idealmente, dichas variaciones a lo largo del eje

de rotación del dispositivo presentan transiciones suaves, una condición que puede ser estudiada a partir de la variación $\Delta\sigma_I^i$ entre cortes consecutivos mediante la Fig. 4.6b. Sin embargo, distintos artefactos por iluminación inhomogénea, frecuentes en este tipo de dispositivos [19], afectan el contraste entre reconstrucciones, observando que las secciones de la cola inferior y superior son aquellas que muestran una mayor variación entre cortes, como muestra el histograma de la Fig. 4.6c. Artefactos de esta naturaleza han sido abordados recientemente mediante una normalización a partir de una estimación de bajo ruido del perfil de iluminación posterior a la adquisición de estos volúmenes [20].

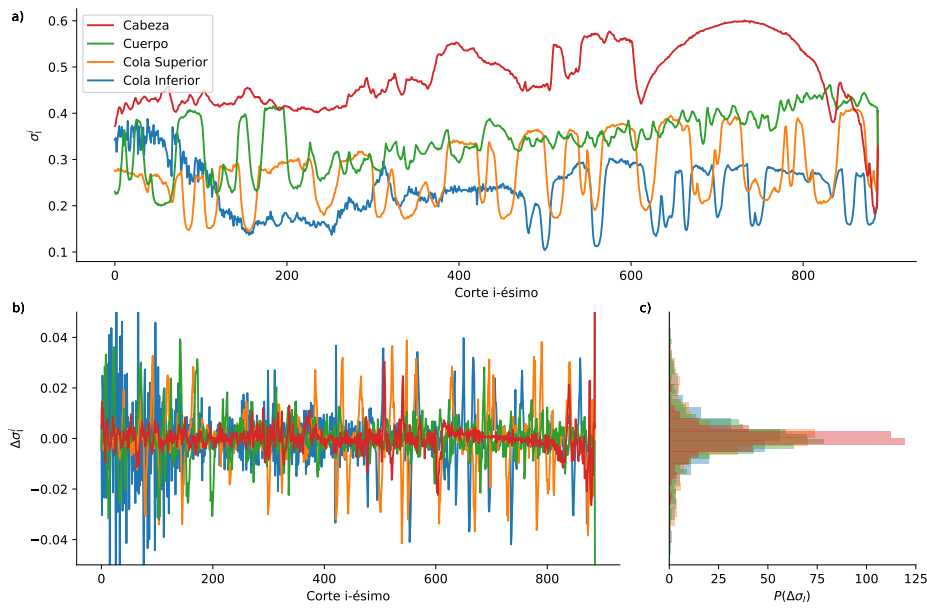


Figura 4.6: **Desvío estándar de intensidad por reconstrucción *ground truth*.** a) Para el volumen B (3 días post fertilización) se muestra el desvío de intensidad σ_I^i de la reconstrucción utilizada como *ground truth* para cada sección del animal. b) La diferencia entre cortes sucesivos $\Delta\sigma_I^i$ presenta una dispersión mayor c) en las secciones superior e inferior de las colas, observada a través de $P(\Delta\sigma_I^i)$. Estas inhomogeneidades se reflejan en variaciones en el contraste en el conjunto de datos, presentando ciertas dificultades para ser capturadas por los enfoques de aprendizaje supervisado.

Al analizar en detalle si estos artefactos de iluminación, ilustrados de manera cualitativa en la Fig. 4.7a y 4.7b, están correlacionados con el desempeño de la red, los resultados no parecen enseñar que esta dispersión entre cortes sea la razón. La Fig. 4.7 discrepa de la hipótesis de que volúmenes con mayor dispersión de $\Delta\sigma_I^i$ entre cortes presenten un menor desempeño. Si bien la

Tomando por caso la dispersión de PSNR para muestras del cuerpo y la cola superior, las primeras presentan una población de reconstrucciones con baja PSNR aún con una baja dispersión de $\Delta\sigma_I^i$. Este argumento propone ciertos interrogantes en la robustez del enfoque ToMoDL ante datos de entrenamiento con similitudes no loca-

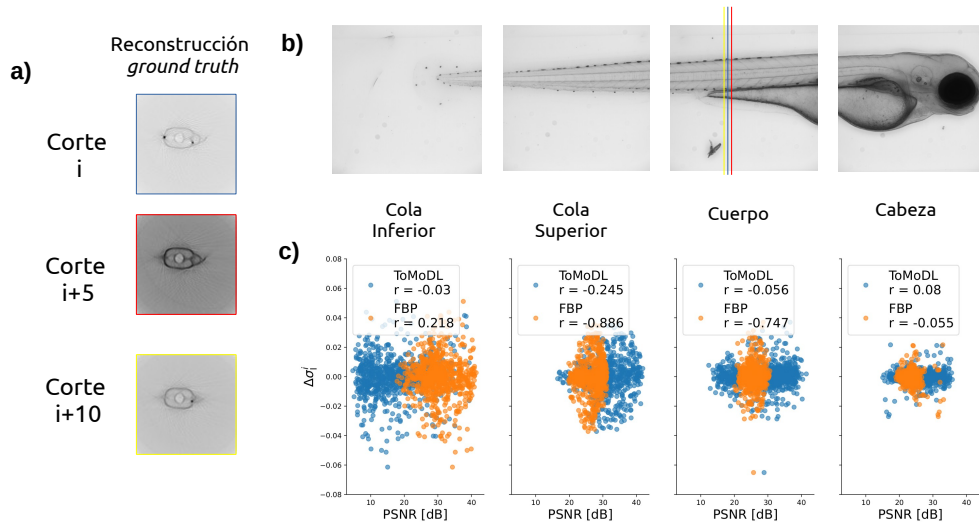


Figura 4.7: **Proyecciones y reconstrucción por sección para volumen B.** a)b) Las reconstrucciones utilizadas muestran una alta variabilidad de contraste entre cortes adyacentes en la sección del cuerpo del espécimen B. c) Aún con esto, la variabilidad de intensidad en función de la PSNR muestra que dicha hipótesis no explica casos como la cola superior, donde la dispersión de $\Delta\sigma_I^i$ no implica una merma en el desempeño de la red.

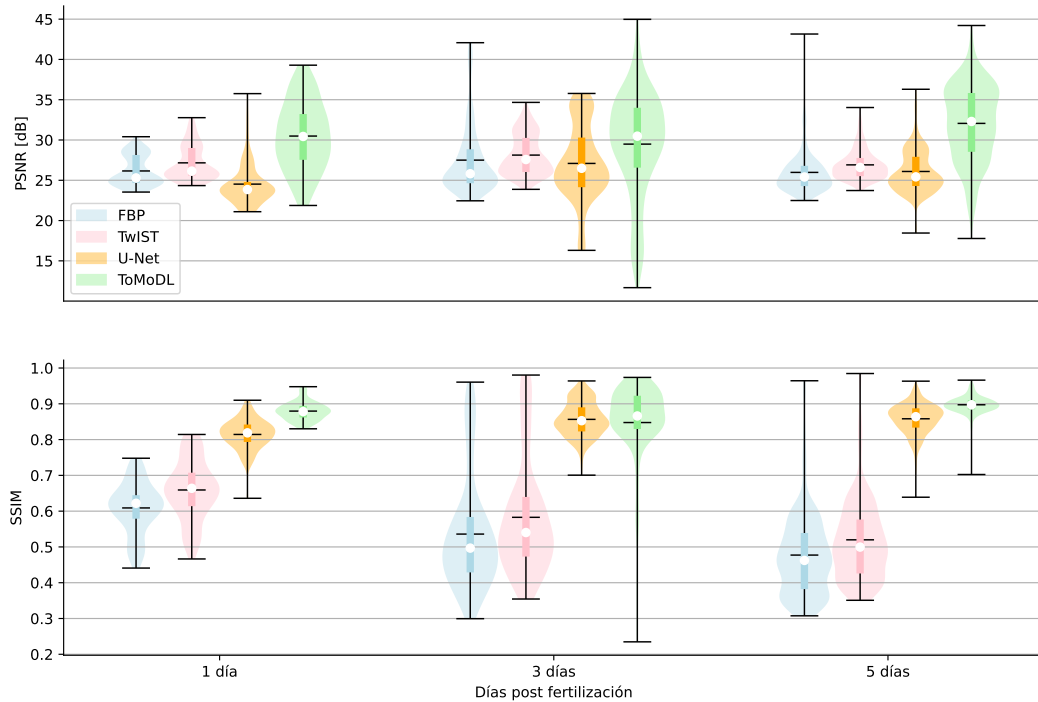


Figura 4.8: **Reconstrucción por día post fertilización.** Para un factor de aceleración $\times 22$, las métricas de cada método de reconstrucción en función del día post fertilización del animal bajo estudio muestran la capacidad de generalización de las reconstrucciones con parámetros entrenables.

les, dependientes de la estructura de cada sección reconstruida. Biswas et al. [43] hacen uso de *priors* para regularizar el entrenamiento de la red desenrollada en base a dentro de una misma variedad (STORM¹) definida por el sujeto bajo estudio.

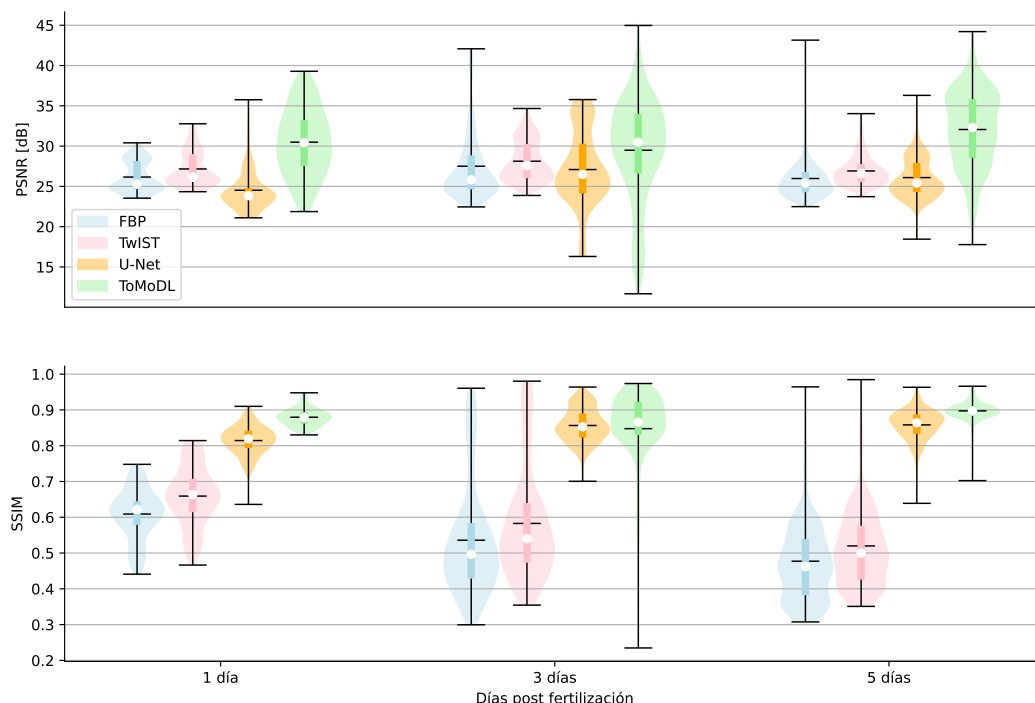


Figura 4.9: **Proyecciones por día post fertilización.** Observamos una fuerte transición en el desarrollo embrionario de la cabeza entre el espécimen adquirido en el primer y tercer día post fertilización, que explica el pobre desempeño al evaluar la generalización del enfoque supervisado no basado en el modelo de Radon. Por simplicidad, para el volumen 1 día post fertilización consideramos que la única sección disponible corresponde a la categoría 'cabeza'.

Por otro lado, observamos que los métodos supervisados presentan una capacidad de generalización razonable ante especímenes en distintos días post fertilización, como se ilustra en la Fig. 4.8. Tomando por caso el único volumen de 1 día post fertilización disponible, cuya anatomía presenta una clara diferencia estructural respecto a los volúmenes, el desempeño de ToMoDL frente a U-Net en SSIM deja entrever que la propuesta de este trabajo podría presentar la remoción de artefactos por submuestreo deseada, descorrelacionada de las estructuras utilizadas en el entrenamiento, como observamos en las proyecciones de la Fig. 4.9.

La dispersión del desempeño de los modelos por espécimen tratado (Fig. 4.10) muestra la capacidad de generalización de los modelos de reconstrucción que contienen todas las secciones del animal bajo estudio. Casos como el volumen D, que contienen diferentes días post fertilización (1 y 3 dpf), muestran que el método ToMoDL, tanto en PSNR como en SSIM, presenta un valor medio superior a los demás métodos de recons-

¹smoothness regularization on manifolds

trucción en el conjunto de imágenes más disímiles. Para el volumen B, cómo mostramos en los paneles de varianza en función del corte (Fig. 4.6), observamos que el algoritmo ToMoDL presenta una mayor sensibilidad ante las variaciones de contraste, reflejada en una amplia dispersión entre los valores máximos y mínimos de reconstrucción de las métricas de reconstrucción.

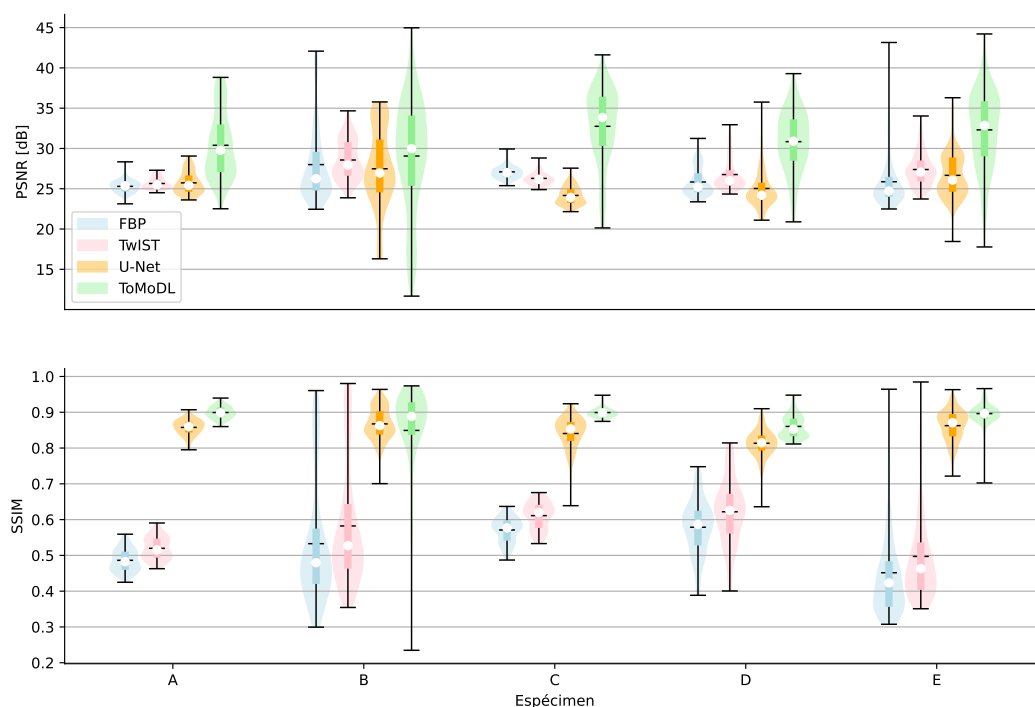


Figura 4.10: **Reconstrucción por espécimen.** Para un factor de aceleración $\times 22$, mostramos las métricas de cada método de reconstrucción en función del espécimen bajo estudio. Observamos que los volúmenes B y E, que contienen las cuatro secciones del animal, introducen la mayor varianza en las métricas de reconstrucción, seguido del espécimen D, del cual contamos con diferentes días post fertilización.

Capítulo 5

Discusión y conclusiones

La propuesta MoDL presenta un poderoso enfoque para la resolución de problemas inversos y su aplicación al campo de reconstrucción de imágenes tomográficas de proyección óptica, ToMoDL, recientemente posibilitada por la integración del operador de Radon a entornos de aprendizaje profundo, ha sido llevada a cabo con razonable éxito. Si bien las ventajas de combinar el modelo subyacente a la medición con una penalización que permita descartar soluciones físicamente implausibles son innegables, su implementación no es una panacea, presentando una heurística que hemos descubierto a base de prueba y error para la aplicación de OPT. Describimos algunos de los interrogantes que surgieron a lo largo de este desarrollo.

La estabilidad y convergencia de redes desenrolladas aplicadas a problemas inversos es un campo de investigación abierto. Enfoques como MoDL, basados en la teoría de operadores proximales, presentan un fenómeno de sobreajuste al aumentar el número de iteraciones de la red [11]. Este problema, abordado recientemente con éxito por Ongie et al. mediante un enfoque correspondiente a un número *infinito* de iteraciones, utiliza teoría de modelos de equilibrio profundo para garantizar la convergencia a un punto fijo [34].

Por otro lado, no existe un gran desarrollo teórico en cuanto a la elección de la arquitectura de *denoising*. Si bien una red multicapa residual resultó la opción más sencilla, hemos probado otras opciones como embeber una arquitectura U-Net como regularización, desistiendo por el alto costo computacional de almacenar gradientes para cada iteración de la dinámica. Sugerimos analizar el uso de bloques de atención espaciales sobre la red convolucional residual [44] [45], un enfoque que presenta ciertas dificultades técnicas en redes recurrentes pero que ha demostrado una mejoría en la capacidad de remoción de ruido para modelos neuronales directos.

La corrección de artefactos en OPT producidos por submuestreo fue abordada para pasos angulares equiespaciados, un procedimiento natural por el funcionamiento de los dispositivos de adquisición. En base a la teoría de variación total, muestrear sinogra-

mas en función de la magnitud de TV [46] se presenta como un potencial siguiente paso. Diversos trabajos en entrenamiento autosupervisado a partir de pocas muestras [47] [48] parecen, a nuestro entender, liderar la pista en explotar la estadística del submuestreo bajo diferentes ángulos iniciales, una metodología que no hemos abordado con detenimiento.

En la misma línea, un naciente campo de investigación referido al análisis de la robustez de enfoques de aprendizaje automático en problemas de reconstrucción plantea los potenciales costos de amplificar pequeñas perturbaciones en los datos adquiridos y la contra intuitiva merma en el desempeño al aumentar las muestras de entrenamiento [49]. Prestar cuidado al uso de estas técnicas y las conclusiones que pueden emerger de reconstrucciones inestables es uno de los principales desafíos en el campo de la reconstrucción mediante aprendizaje automático. Dado el limitado acceso a los datos de dispositivos OPT para contrarrestar estas inestabilidades, diversas propuestas basadas a la agregación de modelos latentes de diferentes fuentes, técnica conocida como aprendizaje federado, se presentan como una potencial herramienta para la colaboración multi-institucional entre usuarios de OPT [50].

Bibliografía

- [1] Davis, S. P., Kumar, S., Alexandrov, Y., Bhargava, A., da Silva Xavier, G., Rutter, G. A., *et al.* Convolutional neural networks for reconstruction of undersampled optical projection tomography data applied to in vivo imaging of zebrafish. *Journal of biophotonics*, **12** (12), e201900128, 2019. [x](#), [17](#), [18](#)
- [2] Ronneberger, O., Fischer, P., Brox, T. U-net: Convolutional networks for biomedical image segmentation. En: International Conference on Medical image computing and computer-assisted intervention, págs. 234–241. Springer, 2015. [x](#), [17](#), [18](#)
- [3] Gilton, D., Ongie, G., Willett, R. Deep equilibrium architectures for inverse problems in imaging. *IEEE Transactions on Computational Imaging*, **7**, 1123–1133, 2021. [xv](#), [3](#)
- [4] Kak, A. C., Slaney, M. Principles of computerized tomographic imaging. SIAM, 2001. [1](#), [8](#)
- [5] Radon, J. On the determination of functions from their integral values along certain manifolds. *IEEE transactions on medical imaging*, **5** (4), 170–176, 1986. [1](#)
- [6] Bhidé, A., Datar, S., Stebbins, K. Case histories of significant medical advances, 2019. [1](#)
- [7] Stewart, R. R. Exploration seismic tomography: Fundamentals. Society of Exploration Geophysicists, 1991. [1](#)
- [8] Sharpe, J., Ahlgren, U., Perry, P., Hill, B., Ross, A., Hecksher-Sørensen, J., *et al.* Optical projection tomography as a tool for 3d microscopy and gene expression studies. *Science*, **296** (5567), 541–545, 2002. [1](#), [5](#)
- [9] Correia, T., Lockwood, N., Kumar, S., Yin, J., Ramel, M.-C., Andrews, N., *et al.* Accelerated optical projection tomography applied to in vivo imaging of zebrafish. *PLOS one*, **10** (8), e0136213, 2015. [1](#), [32](#)

- [10] Donoho, D. L. Compressed sensing. *IEEE Transactions on information theory*, **52** (4), 1289–1306, 2006. [1](#), [14](#)
- [11] Ongie, G., Jalal, A., Metzler, C. A., Baraniuk, R. G., Dimakis, A. G., Willett, R. Deep learning techniques for inverse problems in imaging. *IEEE Journal on Selected Areas in Information Theory*, **1** (1), 39–56, 2020. [2](#), [18](#), [20](#), [24](#), [32](#), [43](#)
- [12] Mur, A. L., Bataille, P., Peyrin, F., Ducros, N. Deep expectation-maximization for image reconstruction from under-sampled poisson data. En: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), págs. 1535–1539. IEEE, 2021. [2](#)
- [13] Watson, T., Andrews, N., Davis, S., Bugeon, L., Dallman, M. D., McGinty, J. Optim: Optical projection tomography integrated microscope using open-source hardware and software. *PLoS One*, **12** (7), e0180309, 2017. [2](#)
- [14] Aggarwal, H. K., Mani, M. P., Jacob, M. Modl: Model-based deep learning architecture for inverse problems. *IEEE transactions on medical imaging*, **38** (2), 394–405, 2018. [3](#), [18](#), [24](#), [28](#), [32](#)
- [15] Huang, Z., Ye, S., McCann, M. T., Ravishankar, S. Model-based reconstruction with learning: from unsupervised to supervised and beyond. *arXiv preprint arXiv:2103.14528*, 2021. [3](#)
- [16] Szczykutowicz, T. P., Toia, G. V., Dhanantwari, A., Nett, B. A review of deep learning ct reconstruction: Concepts, limitations, and promise in clinical practice. *Current Radiology Reports*, págs. 1–15, 2022. [3](#)
- [17] Huiskens, J., Swoger, J., Del Bene, F., Wittbrodt, J., Stelzer, E. H. Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science*, **305** (5686), 1007–1009, 2004. [5](#)
- [18] Bassi, A., Schmid, B., Huiskens, J. Optical tomography complements light sheet microscopy for in toto imaging of zebrafish development. *Development*, **142** (5), 1016–1020, 2015. [5](#)
- [19] Vallejo Ramirez, P. Optical imaging methods for the study of disease models from the nano to the mesoscale. Tesis Doctoral, University of Cambridge, 2021. [7](#), [39](#)
- [20] Vallejo Ramirez, P. P., Zammit, J., Vanderpoorten, O., Riche, F., Blé, F.-X., Zhou, X.-H., *et al.* Optij: Open-source optical projection tomography of large organ samples. *Scientific reports*, **9** (1), 1–9, 2019. [10](#), [39](#)

-
- [21] Walls, J. R., Sled, J. G., Sharpe, J., Henkelman, R. M. Correction of artefacts in optical projection tomography. *Physics in Medicine & Biology*, **50** (19), 4645, 2005. [10](#)
- [22] Tarantola, A. Popper, bayes and the inverse problem. *Nature physics*, **2** (8), 492–494, 2006. [13](#)
- [23] Hadamard, J. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, págs. 49–52, 1902. [13](#)
- [24] Chambolle, A. An algorithm for total variation minimization and applications. *Journal of Mathematical imaging and vision*, **20** (1), 89–97, 2004. [13](#), [15](#)
- [25] Chambolle, A., De Vore, R. A., Lee, N.-Y., Lucier, B. J. Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Transactions on Image Processing*, **7** (3), 319–335, 1998. [13](#)
- [26] Trémouhéac, B. R. Low-rank and sparse reconstruction in dynamic magnetic resonance imaging via proximal splitting methods. Tesis Doctoral, UCL (University College London), 2015. [14](#), [15](#)
- [27] Bioucas-Dias, J. M., Figueiredo, M. A. A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transactions on Image processing*, **16** (12), 2992–3004, 2007. [14](#), [32](#)
- [28] LeCun, Y., Bengio, Y., Hinton, G. Deep learning. *nature*, **521** (7553), 436–444, 2015. [15](#), [17](#)
- [29] Rumelhart, D. E., Hinton, G. E., Williams, R. J. Learning representations by back-propagating errors. *nature*, **323** (6088), 533–536, 1986. [15](#)
- [30] Denker, J., Gardner, W., Graf, H., Henderson, D., Howard, R., Hubbard, W., *et al.* Neural network recognizer for hand-written zip code digits. *Advances in neural information processing systems*, **1**, 1988. [16](#)
- [31] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., *et al.* Backpropagation applied to handwritten zip code recognition. *Neural computation*, **1** (4), 541–551, 1989. [16](#)
- [32] Daylight, E. G. The advent of recursion in programming, 1950s-1960s, 2010. [17](#)
- [33] Bai, S., Kolter, J. Z., Koltun, V. Trellis networks for sequence modeling. *arXiv preprint arXiv:1810.06682*, 2018. [18](#)

- [34] Bai, S., Kolter, J. Z., Koltun, V. Deep equilibrium models. *Advances in Neural Information Processing Systems*, **32**, 2019. [18](#), [43](#)
- [35] Hornik, K., Stinchcombe, M., White, H. Multilayer feedforward networks are universal approximators. *Neural networks*, **2** (5), 359–366, 1989. [18](#)
- [36] Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, **26** (7), 3142–3155, 2017. [24](#)
- [37] Kolarik, M., Burget, R., Riha, K. Comparing normalization methods for limited batch size segmentation neural networks. En: 2020 43rd International Conference on Telecommunications and Signal Processing (TSP), págs. 677–680. IEEE, 2020. [25](#)
- [38] Ronchetti, M. Torchradon: Fast differentiable routines for computed tomography. *arXiv preprint arXiv:2009.14788*, 2020. [25](#)
- [39] Hestenes, M. R., Stiefel, E. Methods of conjugate gradients for solving. *Journal of research of the National Bureau of Standards*, **49** (6), 409, 1952. [25](#)
- [40] Kingma, D. P., Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. [28](#)
- [41] Pascanu, R., Mikolov, T., Bengio, Y. On the difficulty of training recurrent neural networks. En: International conference on machine learning, págs. 1310–1318. PMLR, 2013. [28](#)
- [42] Zhu, B., Liu, J. Z., Cauley, S. F., Rosen, B. R., Rosen, M. S. Image reconstruction by domain-transform manifold learning. *Nature*, **555** (7697), 487–492, 2018. [31](#)
- [43] Biswas, S., Aggarwal, H. K., Jacob, M. Dynamic mri using model-based deep learning and storm priors: Modl-storm. *Magnetic resonance in medicine*, **82** (1), 485–494, 2019. [41](#)
- [44] Zhu, X., Cheng, D., Zhang, Z., Lin, S., Dai, J. An empirical study of spatial attention mechanisms in deep networks. En: Proceedings of the IEEE/CVF international conference on computer vision, págs. 6688–6697. 2019. [43](#)
- [45] Wang, Y., Song, X., Chen, K. Channel and space attention neural network for image denoising. *IEEE Signal Processing Letters*, **28**, 424–428, 2021. [43](#)
- [46] Sanders, T., Dwyer, C. Subsampling and inpainting approaches for electron tomography. *Ultramicroscopy*, **182**, 292–302, 2017. [44](#)

-
- [47] Hendriksen, A. A., Pelt, D. M., Batenburg, K. J. Noise2inverse: Self-supervised deep convolutional denoising for tomography. *IEEE Transactions on Computational Imaging*, **6**, 1320–1335, 2020. [44](#)
 - [48] Hendriksen, A. A., Bühner, M., Leone, L., Merlini, M., Vigano, N., Pelt, D. M., *et al.* Deep denoising for multi-dimensional synchrotron x-ray tomography without high-quality reference data. *Scientific reports*, **11** (1), 1–13, 2021. [44](#)
 - [49] Antun, V., Renna, F., Poon, C., Adcock, B., Hansen, A. C. On instabilities of deep learning in image reconstruction and the potential costs of ai. *Proceedings of the National Academy of Sciences*, **117** (48), 30088–30095, 2020. [44](#)
 - [50] Feng, C.-M., Yan, Y., Wang, S., Xu, Y., Shao, L., Fu, H. Specificity-preserving federated learning for mr image reconstruction. *IEEE Transactions on Medical Imaging*, 2022. [44](#)

Agradecimientos

Encontré dos definiciones de agradecer: una, dada por un beneficio recibido y otra, por corresponder el cuidado y/o atención recibida. Al concluir este trabajo, me gustaría decantar por la última para reconocer a las personas que tomaron partido por mi bienestar y atendieron mi curiosidad por aprender y entender. Haber atravesado este camino, variado en texturas, fue mucho más rico al compartirlo con ustedes.

Yendo a lo concreto, agradezco al Instituto Balseiro por darme todo y más en educación, contención y guiado. Agradezco a Germán por ayudarme a entender desde el día 0, comunicándome ideas con sencillez y templanza y siempre cuidando de avanzar un paso a la vez. A Teresa por hacer que la distancia ni se sienta en nuestra labor, atenta en todo momento a discutir y repensar problemas y soluciones. Ambos me acompañaron en sucesivas iteraciones de mi persona sin condiciones y eso no tiene valor.

La unidad que encontré en mi familia en momentos complicados fue algo que solo puedo agradecer. Gracias a mis viejos por las mil y una, gracias a mis hermanas por apoyarme y escucharme, gracias a mi abuela por ser tremenda crack. A mis amigos por darme su apoyo desde cualquier lugar, sin horarios ni dramas.

No puedo dejar de agradecer a Aguigo por haber crecido juntos de golpe, tomando decisiones de las geniales como de las dolorosas.

Por último, le doy de las gracias completas a Ana por compartirme elegantemente de sus alfajores de maíz violeta y acompañarme en nuevas aventuras. Agradezco compartir la sincronía, tus tarjetas perforadas y tu fantabulosa visión de lo que nos rodea.

