

Seminario PINNs

# A Geometry-Informed Deep Learning Framework for Ultra-Sparse 3D Tomographic Image Reconstruction

Autores:

Liyue Shen, Wei Zhao, Dante Capaldi, John Pauly, and Lei Xing

Sofía Zimmer

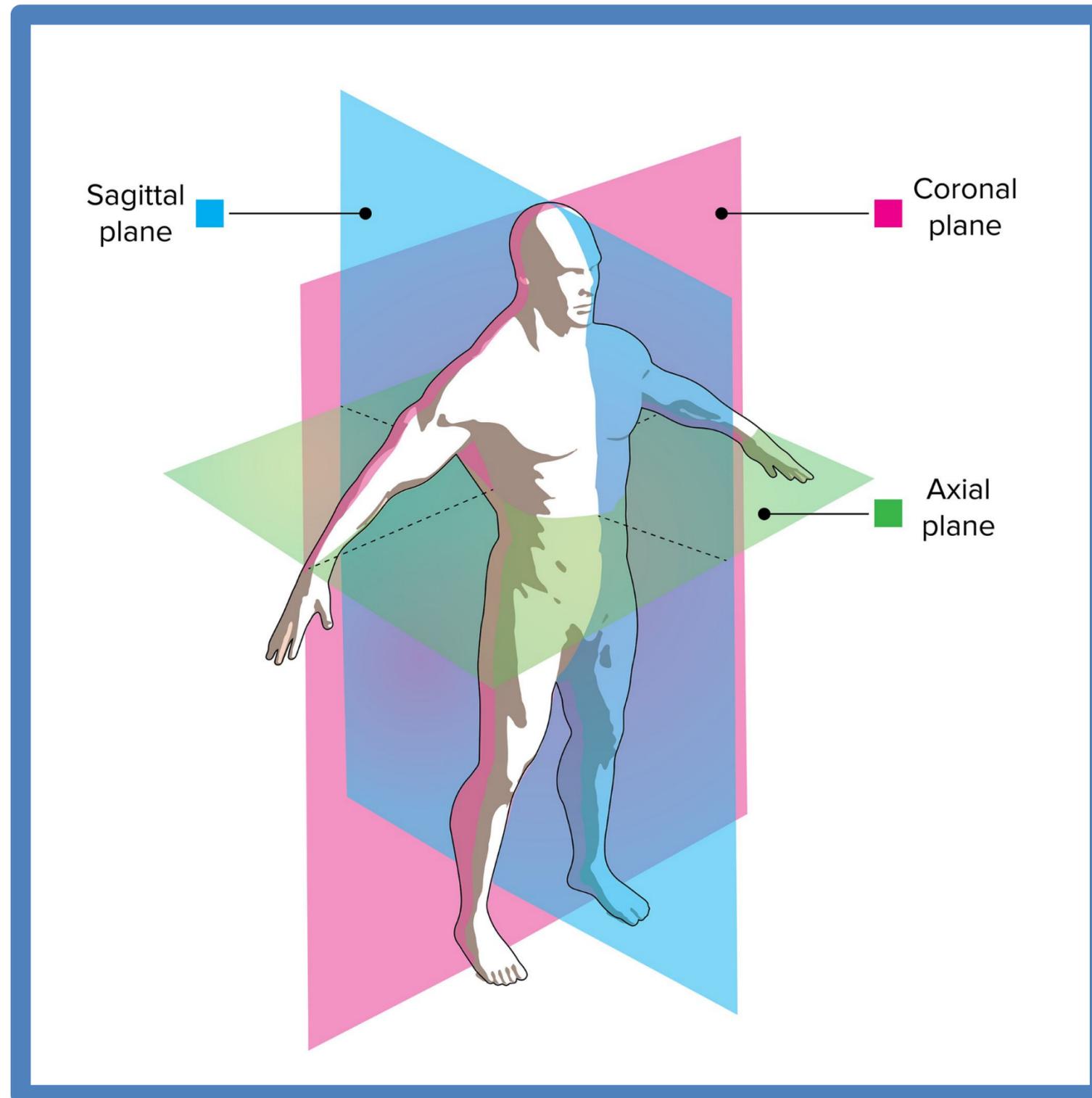
01 de Noviembre 2024

# CONTENIDO

1. Conceptos previos y motivación
2. Arquitectura y modelado
3. Experimentos y resultados

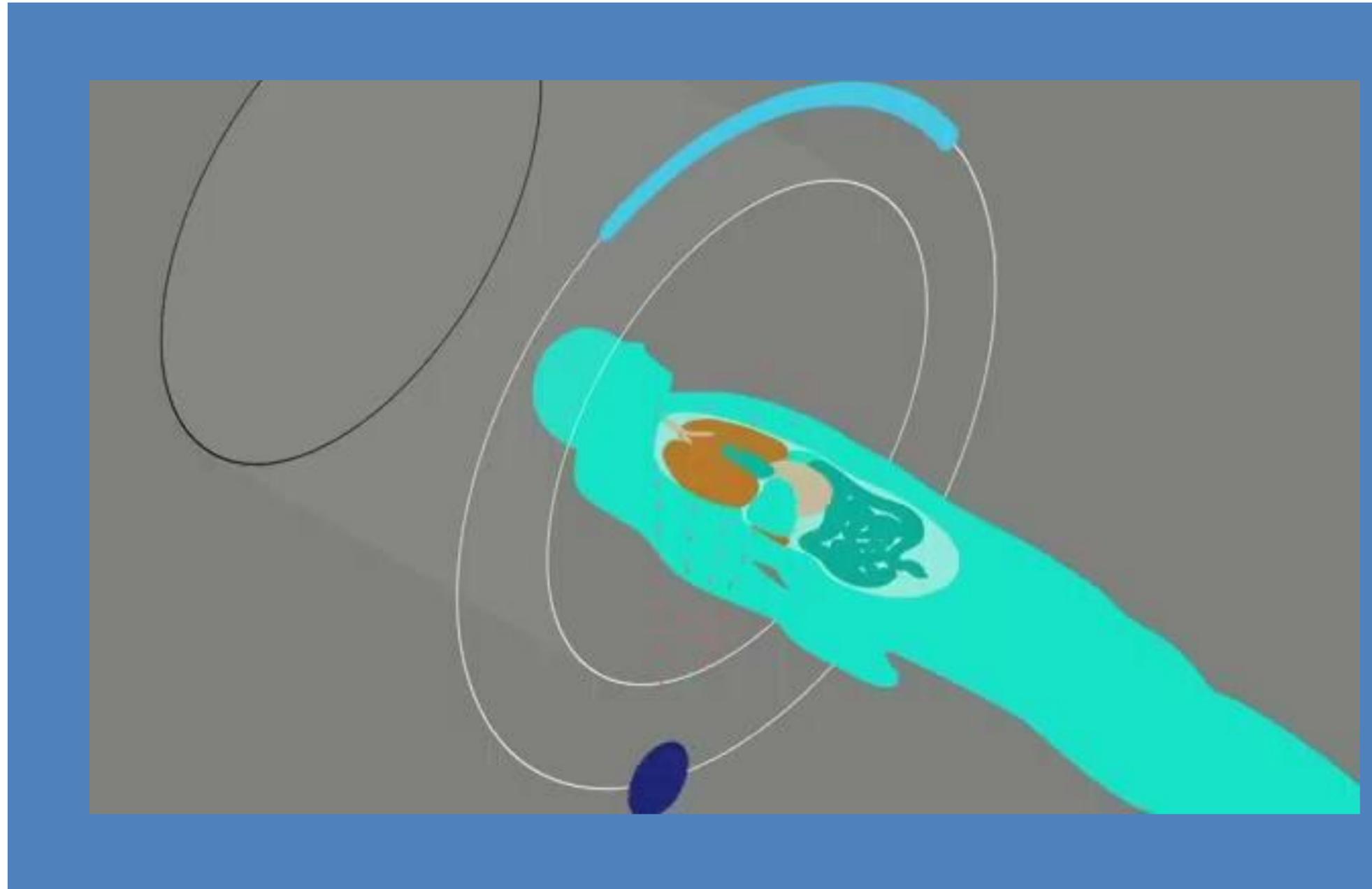
# Conceptos previos y motivación

# Planos de corte anatómicos



# Tomografía

- Cada vez que el tomógrafo da una vuelta completa genera una imagen 2D.
- La imagen resultante es un corte axial del cuerpo, en escala de grises.
- Es posible observar huesos, órganos y otros tejidos.
- Dependiendo de la densidad y la composición, el tono que vamos a observar en la imagen resultante.



Video obtenido de: What is Computed Tomography (CT) and how does it work? Youtube.

# MOTIVACIÓN

Obtener CT en tiempo real para poder utilizar esta práctica en procedimientos clínicos guiados por imágenes, como la radioterapia.

Reducir significativamente la dosis de radiación que reciben los pacientes, sin alterar la calidad de las imágenes.

# Arquitectura y modelado

## Modelado del Problema

Secuencias de proyecciones 2D  $p_1, p_2, \dots, p_n$  donde  $p_i \in \mathbb{R}^{H_{2D} \times W_{2D}}$  y  $1 \leq i \leq n$

Volumen 3D a predecir  $I_{pred} \in \mathbb{R}^{C_{3D} \times H_{3D} \times W_{3D}}$  donde  $I_{truth} \in \mathbb{R}^{C_{3D} \times H_{3D} \times W_{3D}}$  es el ground truth

- 1) 2D-Net  $\phi_1$  con pesos  $W_1$
- 2) Proyección inversa  $\phi_2$  con pesos  $W_2$
- 3) 3D-Net  $\phi_3$  con pesos  $W_3$

$$I_{pred} = \phi_3(\phi_2(\phi_1(p_1, p_2, \dots, p_n; W_1); W_2); W_3)$$

# ARQUITECTURA: Geometry-informed image reconstruction (GIIR)

## Red de generación de proyecciones 2D

Genera nuevas vistas a partir de las imágenes originales.

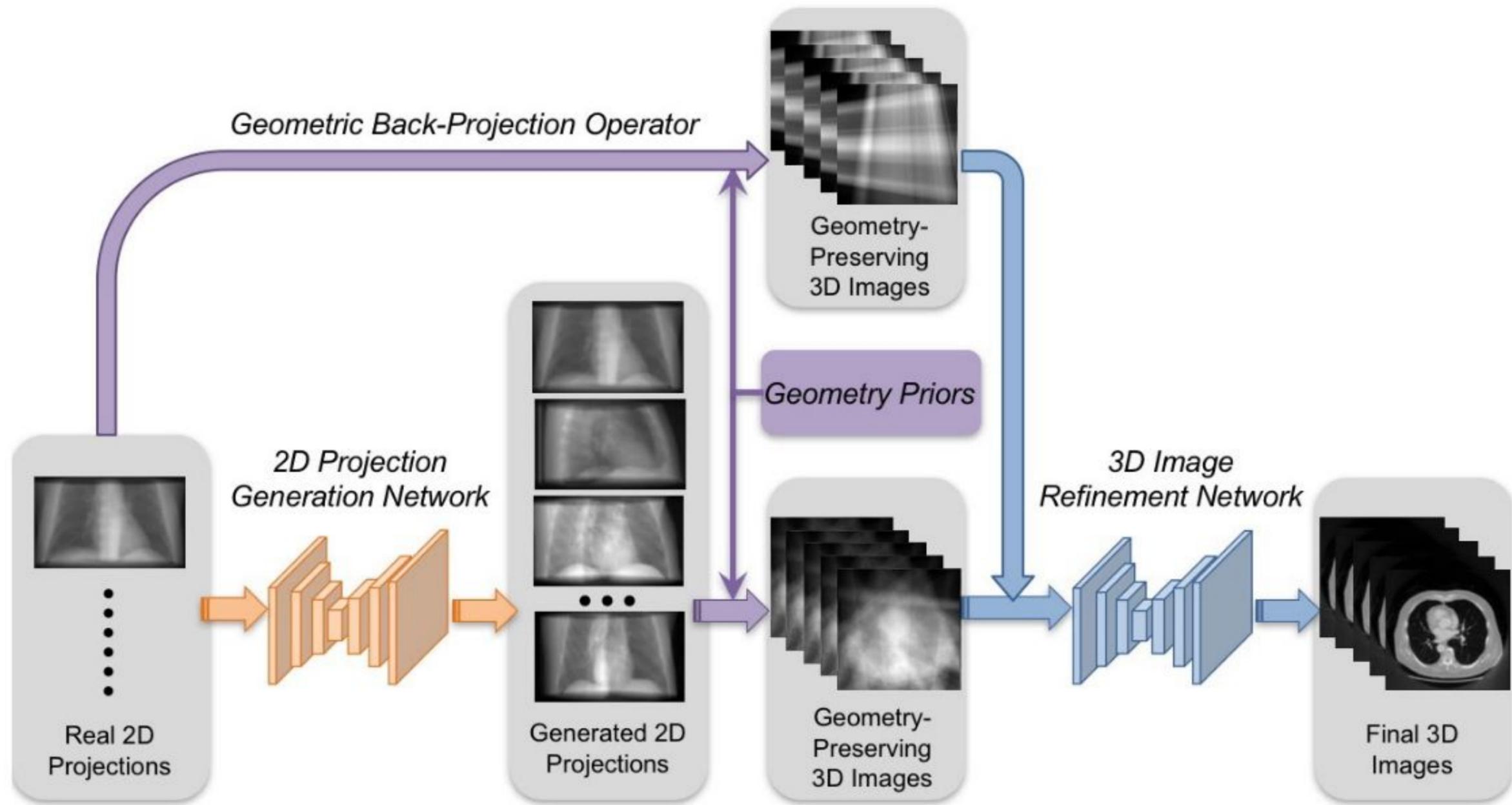
## Operador geométrico de back projection

Transforma las proyecciones 2D en imágenes 3D.

A las imágenes resultantes les llama “*geometric preserving images*” (GPIs).

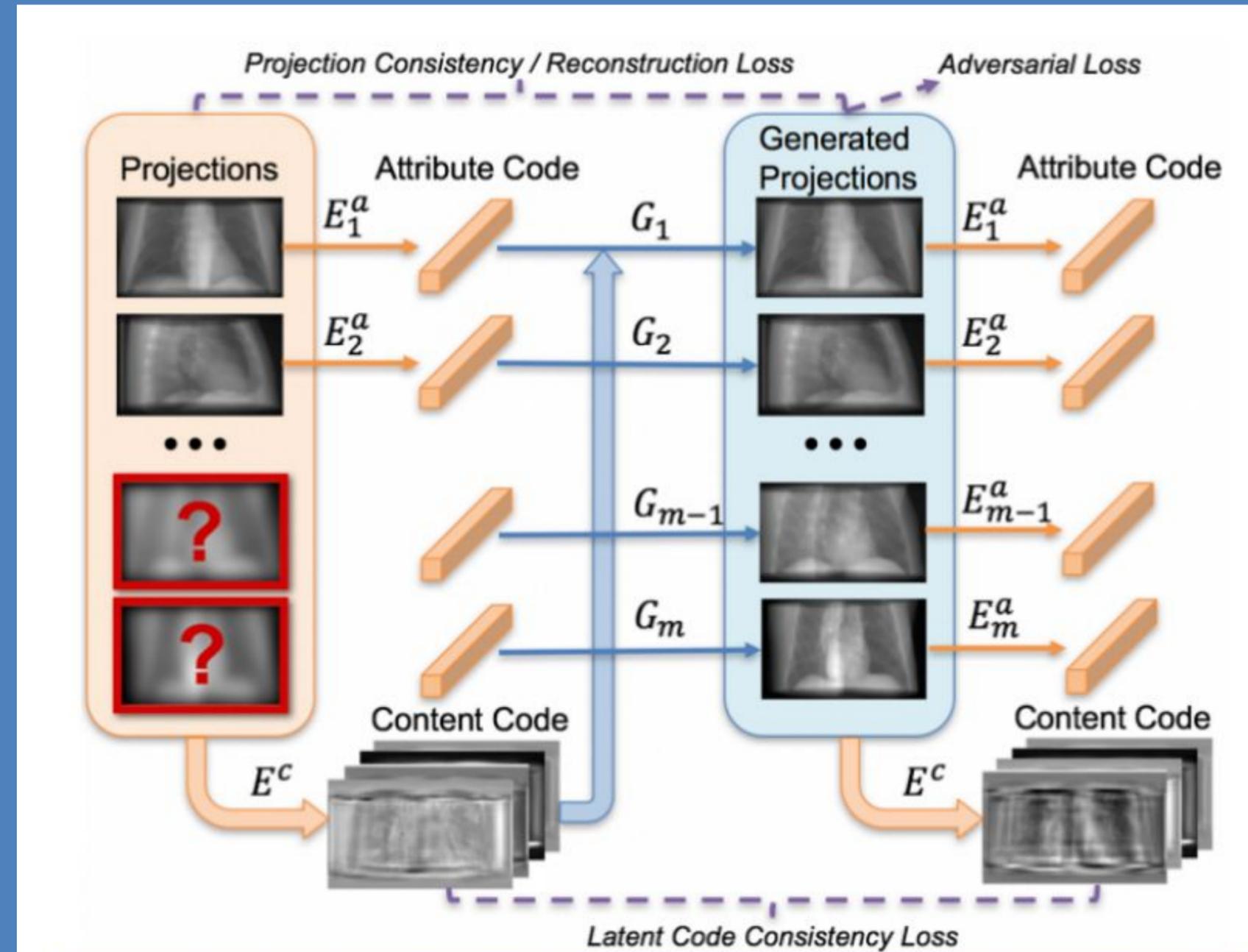
## Red de refinamiento de imágenes 3D

Mejora la reconstrucción de las imágenes 3D de la etapa anterior.



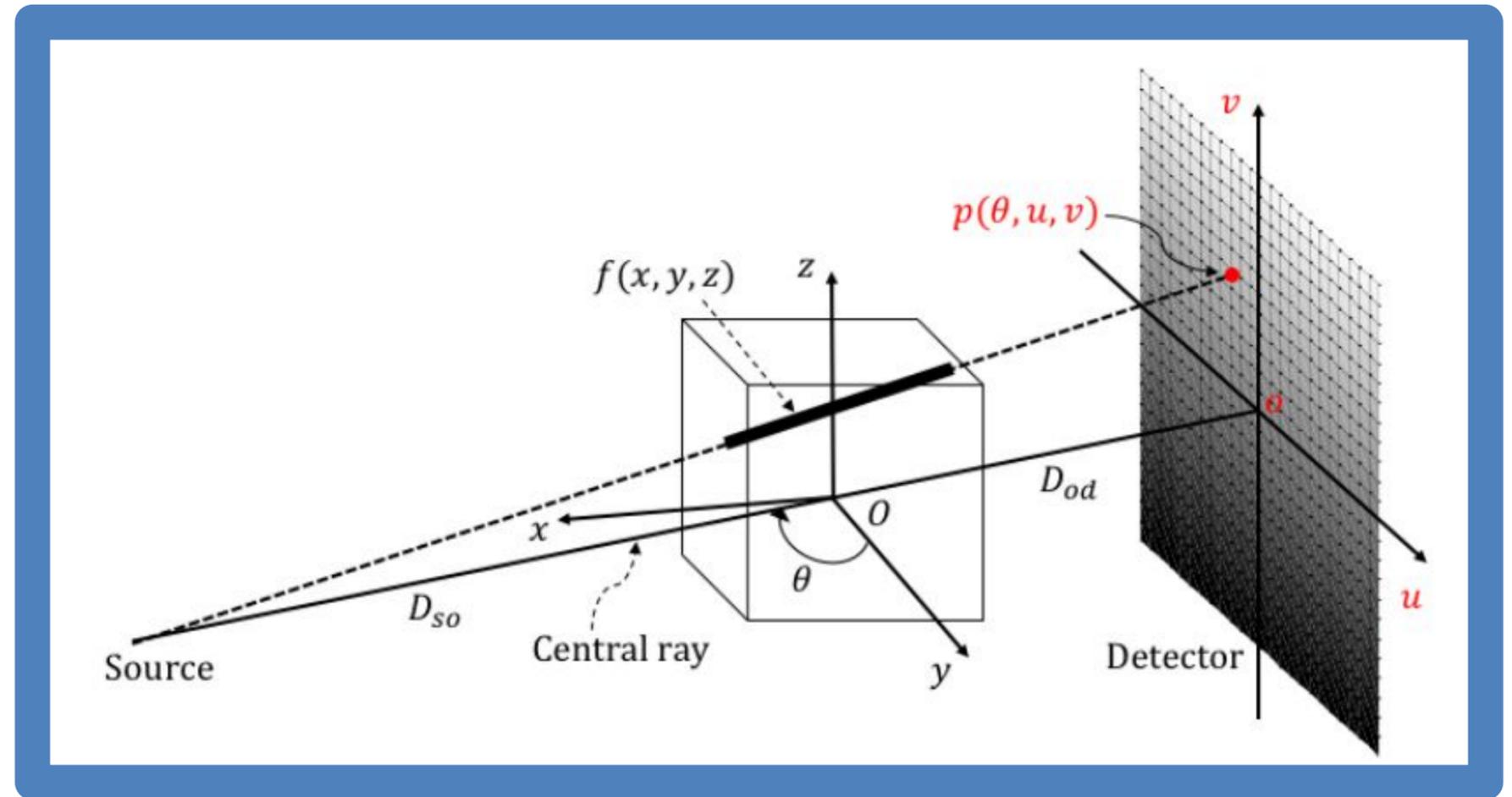
# Red de generación de proyecciones 2D

- Modelo GAN
- Existe cierta información anatómica estructural que le llamaremos “*content code*”.
- Hay otra información que depende del ángulo de visión que se le llama “*attribute code*”.



# Operador geométrico de back projection

- Se agrega la información de todas las líneas de proyección desde diferentes puntos de vista.
- Se trata de una transformación determinística, por lo que no hay parámetros a aprender.

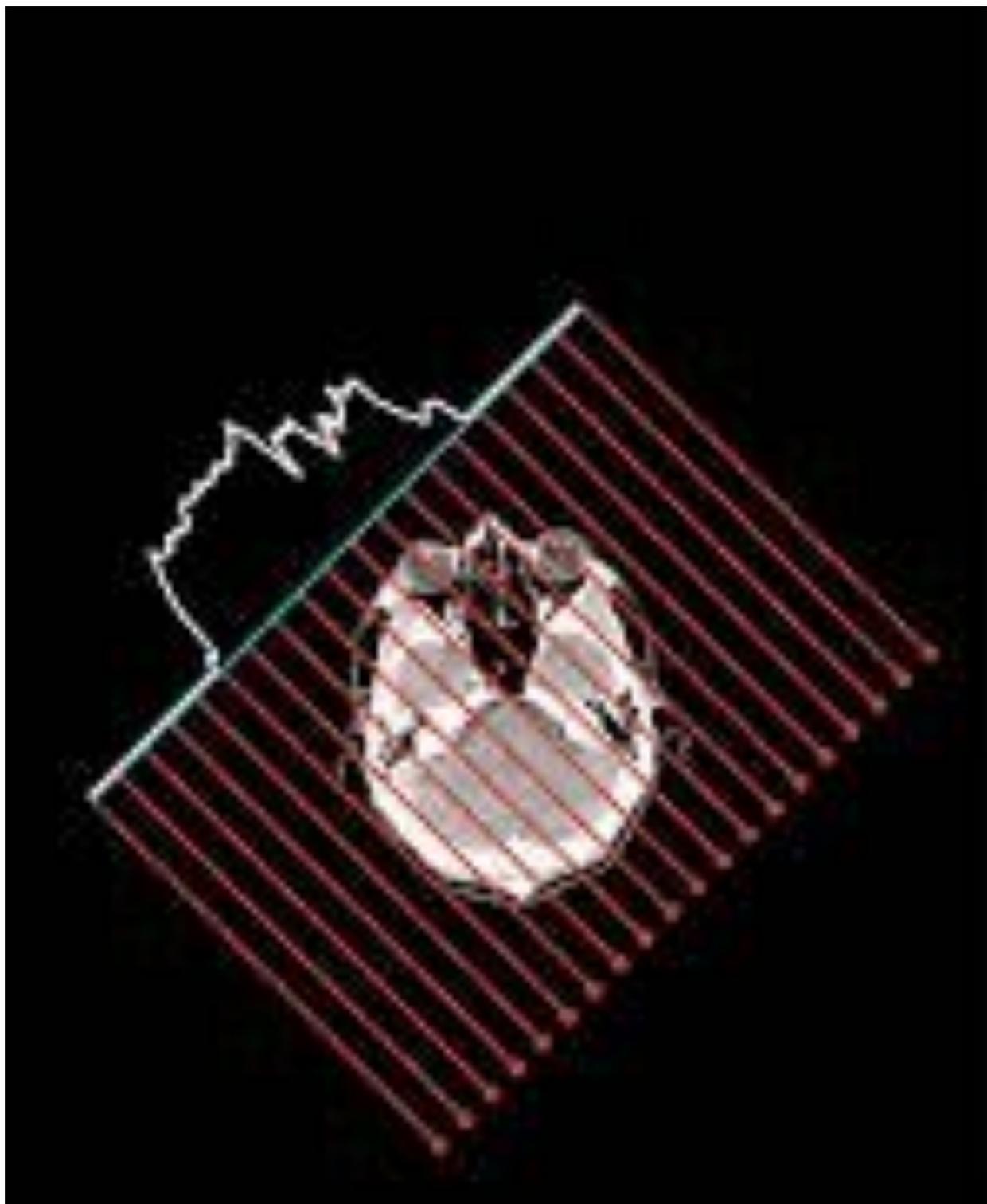


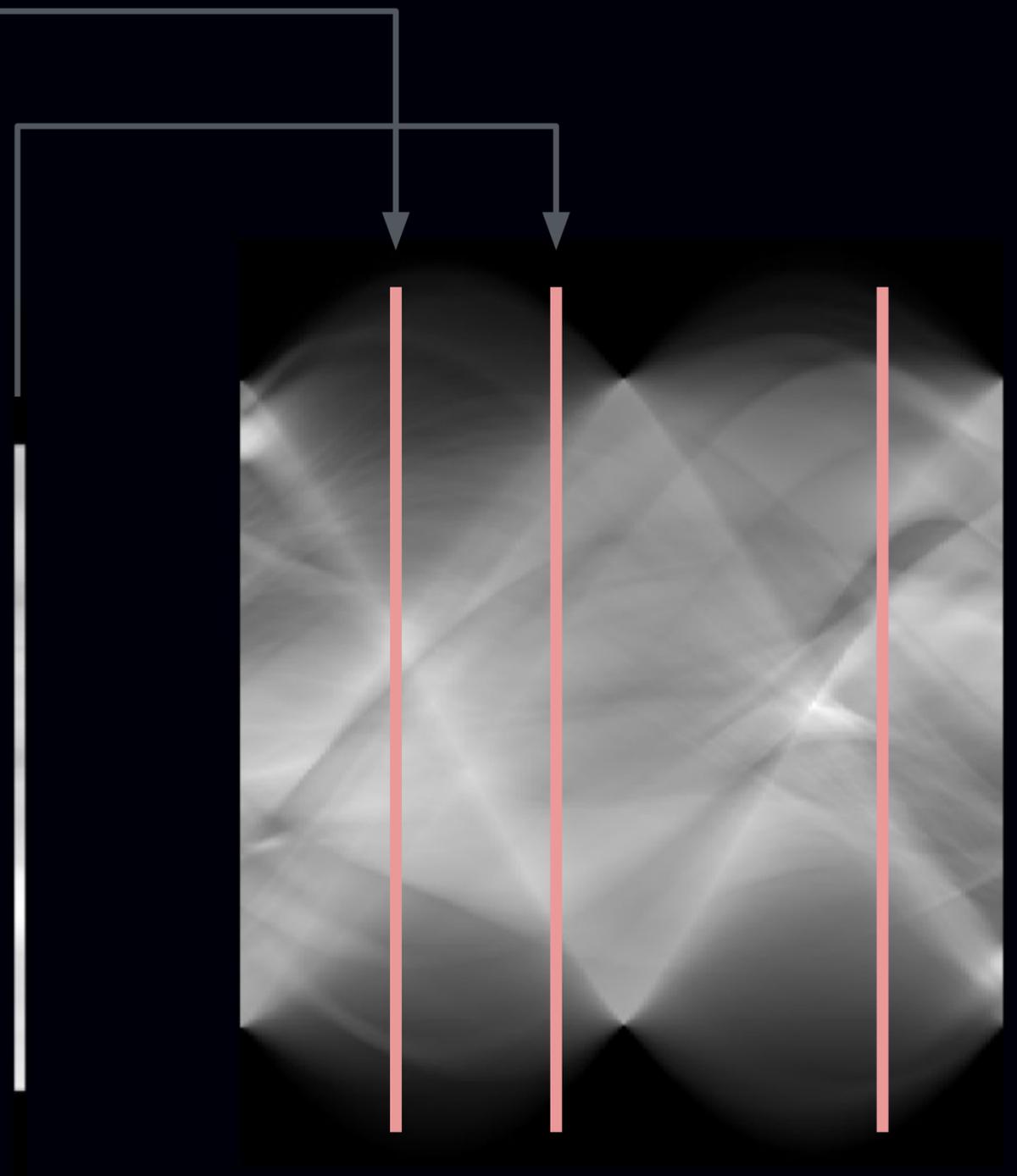
$$I_{GPI}(x, y, z) = \sum_{\theta} p(\theta, u(x, y, \theta), v(x, y, z, \theta))$$

$$u(x, y, \theta) = \frac{x \cos \theta + y \sin \theta}{D_{so} + x \sin \theta - y \cos \theta} D_{sd}$$

$$v(x, y, z, \theta) = \frac{D_{sd}}{D_{so} + x \sin \theta - y \cos \theta} z$$

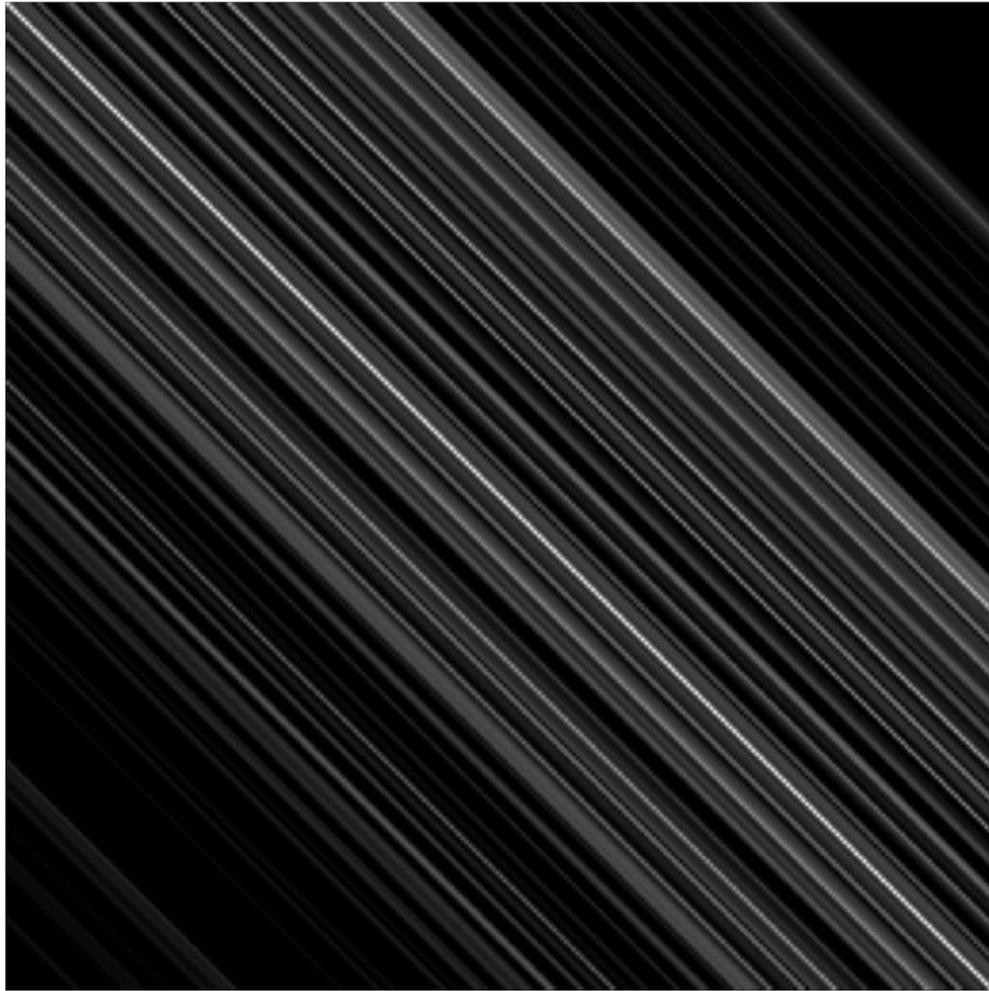
# Sinogram



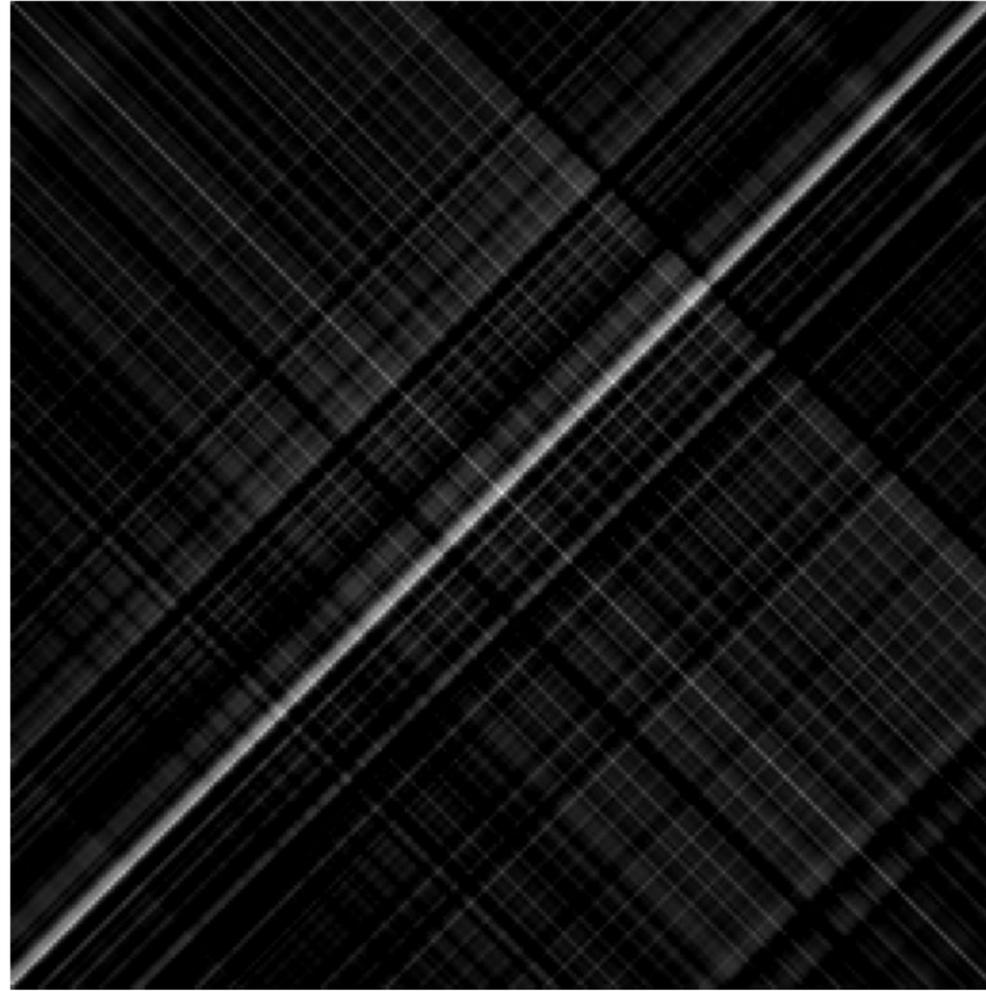


Sinogram

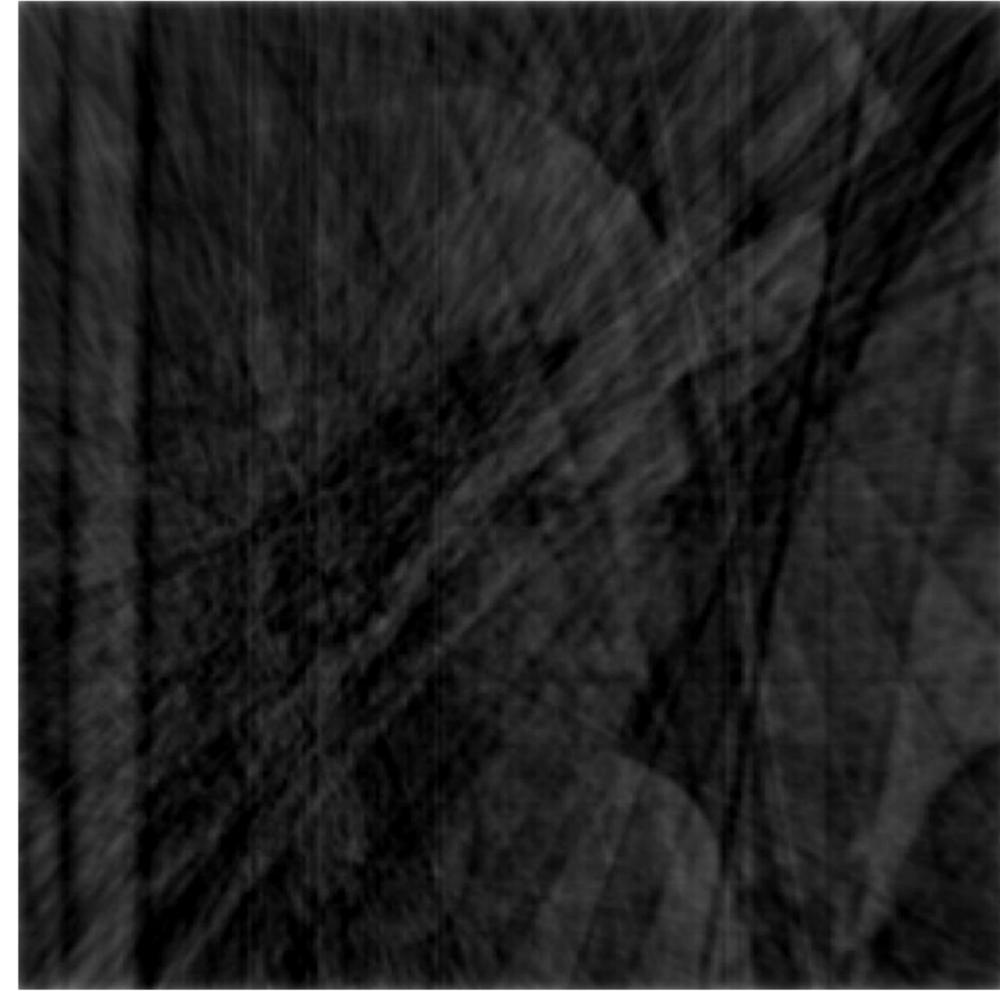
# 3D reconstruction: Weighted Back-Projection



$N = 1$



$N = 2$



$N = 18$

# 3D reconstruction: Weighted Back-Projection



$N = 90$

# U-NET 3D

## Convolución 3D

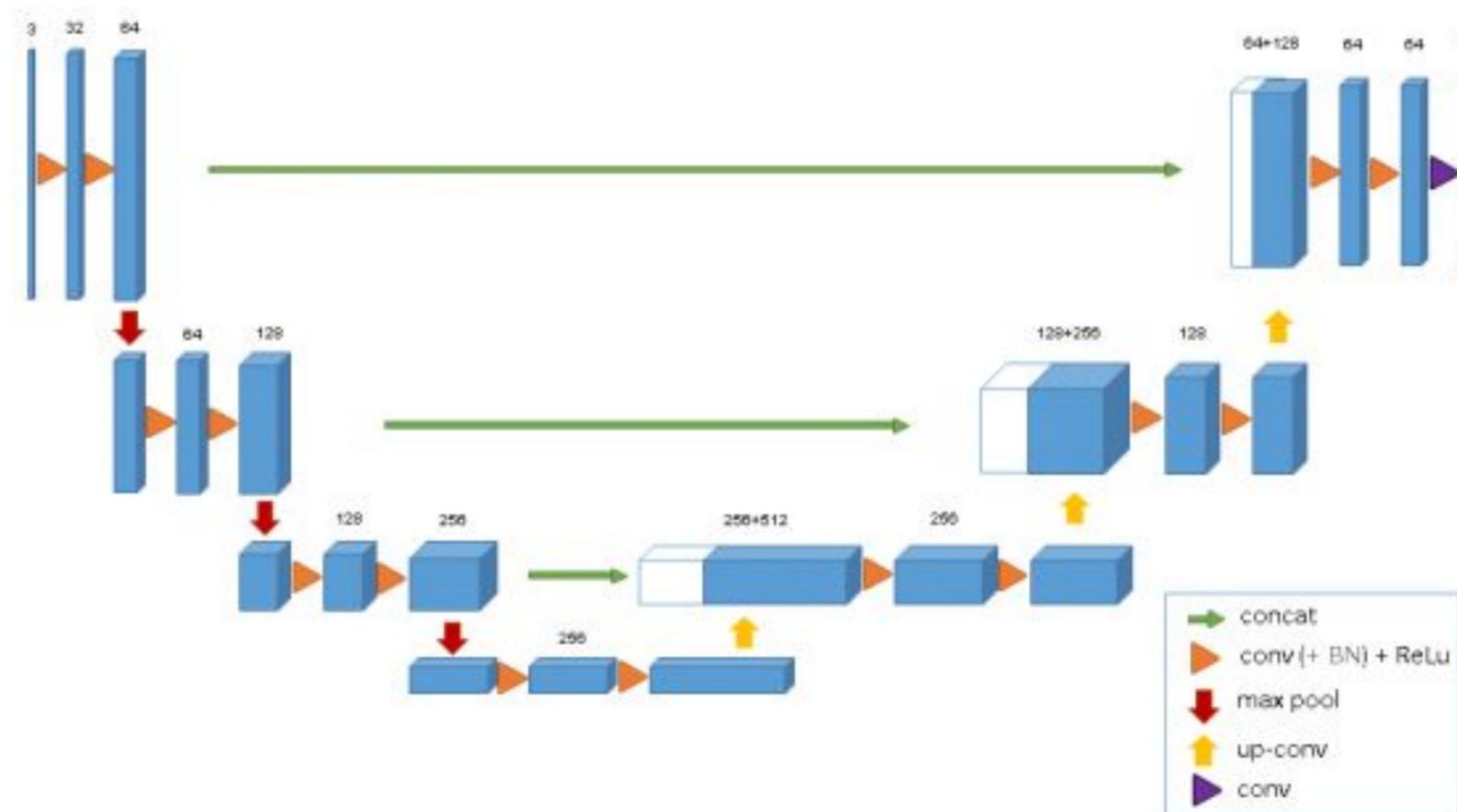
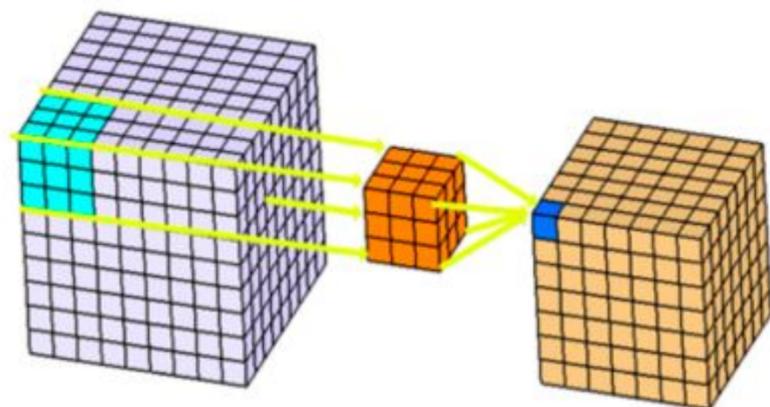


Imagen izquierda obtenida de: 3D Convolutions : Understanding + Use Case. Kaggle.

Imagen derecha obtenida de: 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation.

# Red de refinamiento de imágenes 3D

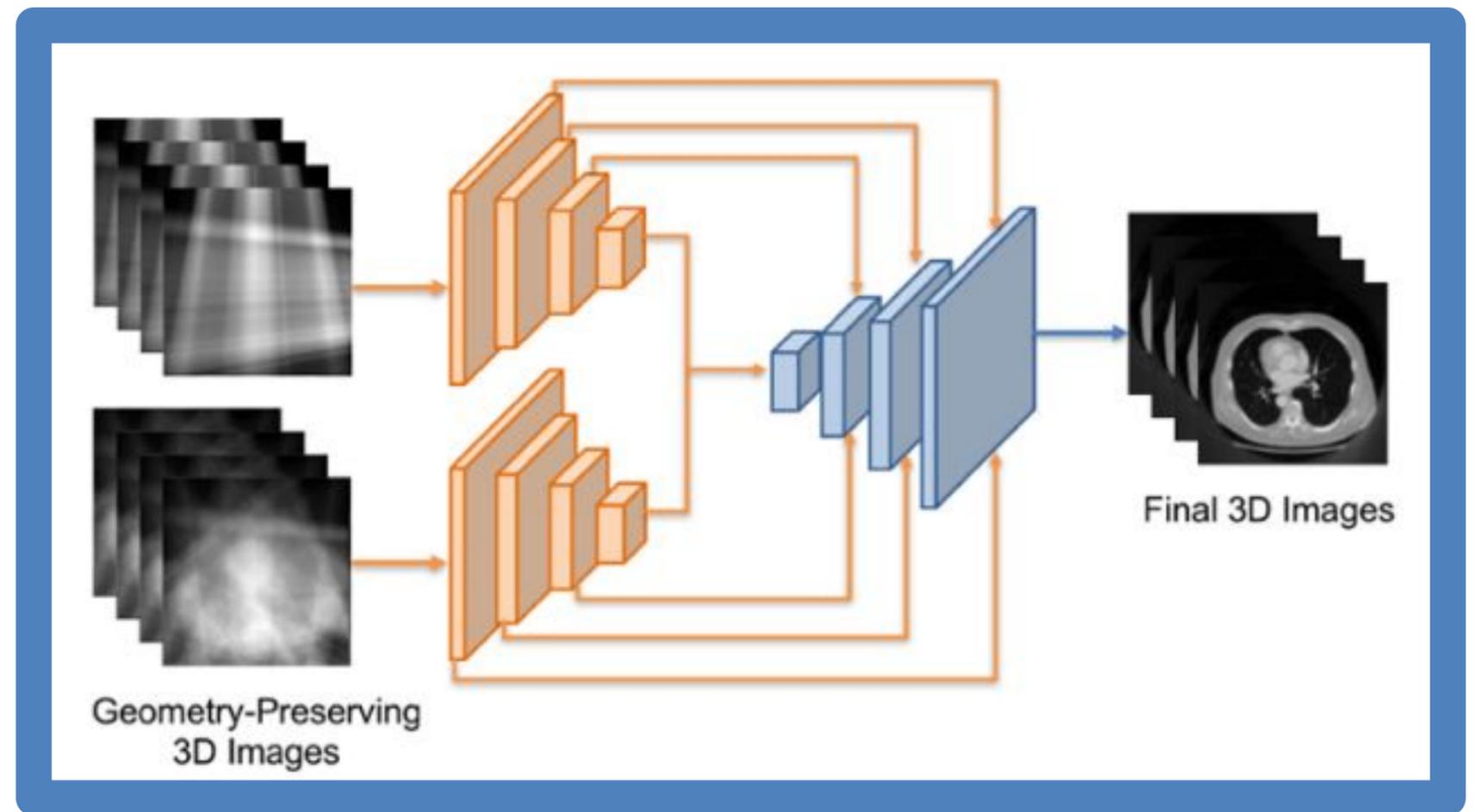
Entradas:

$$I_{GPI}^{src} = \phi_2(p_1, p_2, \dots, p_n; W_2)$$

$$I_{GPI}^{gen} = \phi_2(\phi_1(p_1, p_2, \dots, p_n; W_1); W_2)$$

Queremos:

$$I_{pred} = \phi_3(I_{GPI}^{src}, I_{GPI}^{gen}; W_3)$$



4 bloques de down-sampling / up-sampling:

- 2 convoluciones 3D
- Activación ReLU
- Normalización 3D por grupos
- Max-pooling / Interpolación

Luego:

- Conv 1x1x1
- Activación tangente hiperbólica

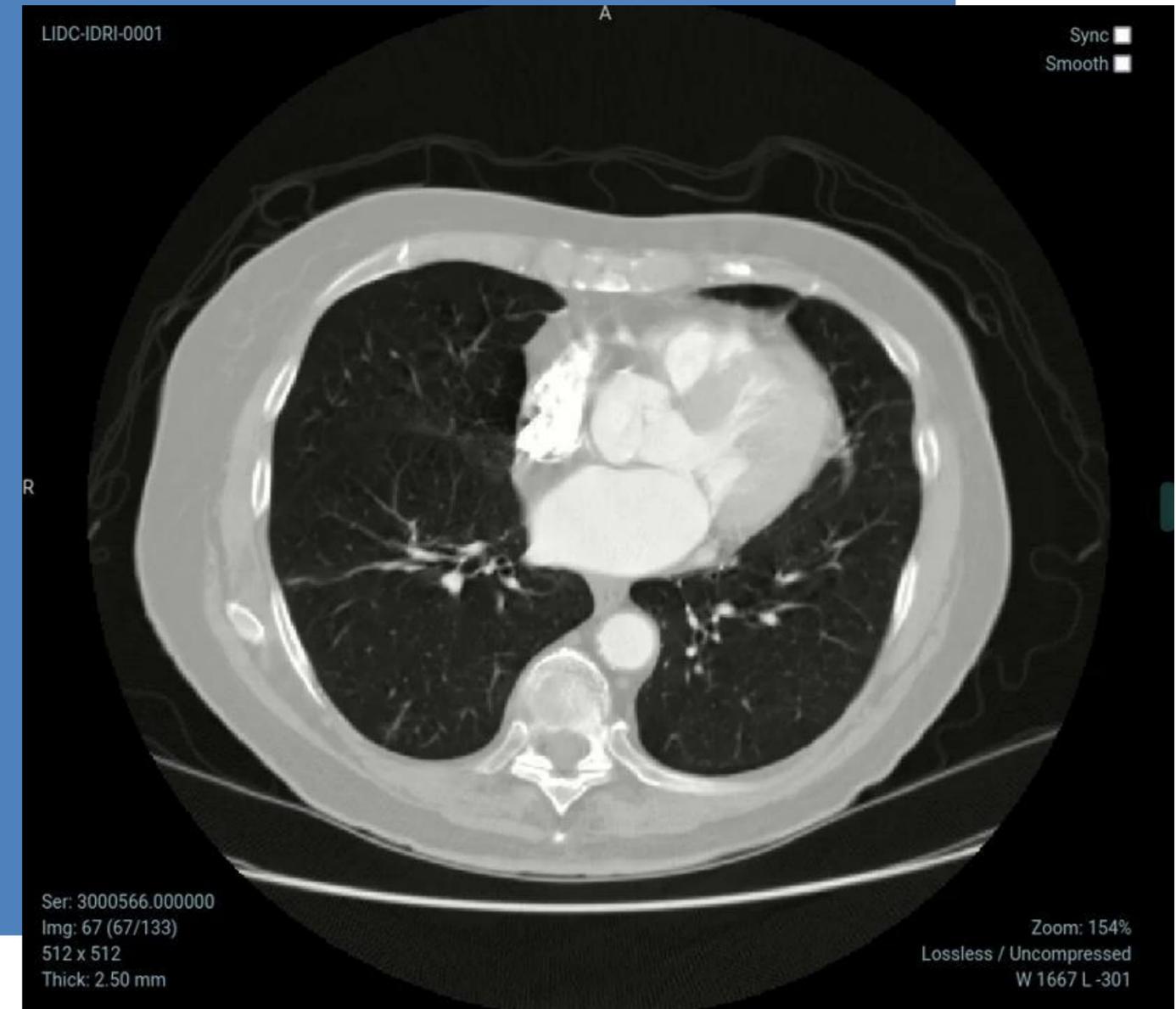
# Experimentos y resultados

# EXPERIMENTOS

Datos: 1018 CT de torax

Se realizan 3 pruebas:

- Red de generación de proyecciones 2D
- GIRR simplificado para reconstrucción 3D
- GIRR completo para reconstrucción 3D a partir de muy pocas imágenes



# Resultados de la red de generación de proyecciones 2D

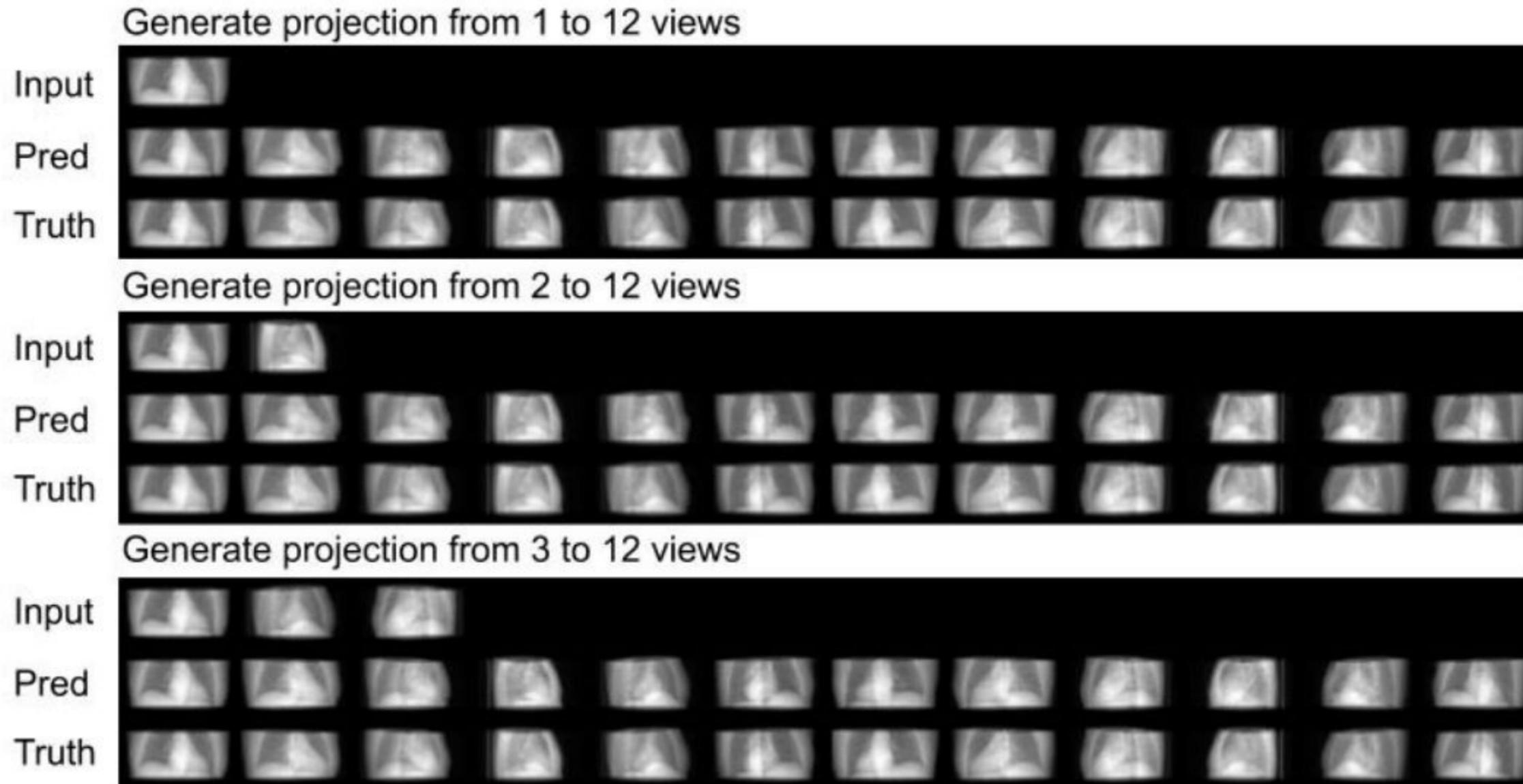


Fig. 6. Results of 2D-Net for generating novel-view projections. Predicted and ground-truth 2D projections (12 angular views evenly distributed over 360 degrees) of a testing sample (with 1-3 input projections respectively). For each group of results, the input projections fed into the network, the predicted novel-view projections output from the network, and the ground-truth of novel-view projections are shown row by row.

## **GIIR simplificado para reconstrucción 3D**

- Se utilizan únicamente el módulo de back projection y el refinamiento 3D.
- Se toman las proyecciones 2D verdaderas y no se generan nuevas vistas.
- Se toman de 10 a 120 proyecciones.

## **GIIR completo para reconstrucción 3D a partir de muy pocas imágenes**

- Entrenamiento del modelo completo.
- Se toman 1, 2 y 3 proyecciones.

# Modelos base

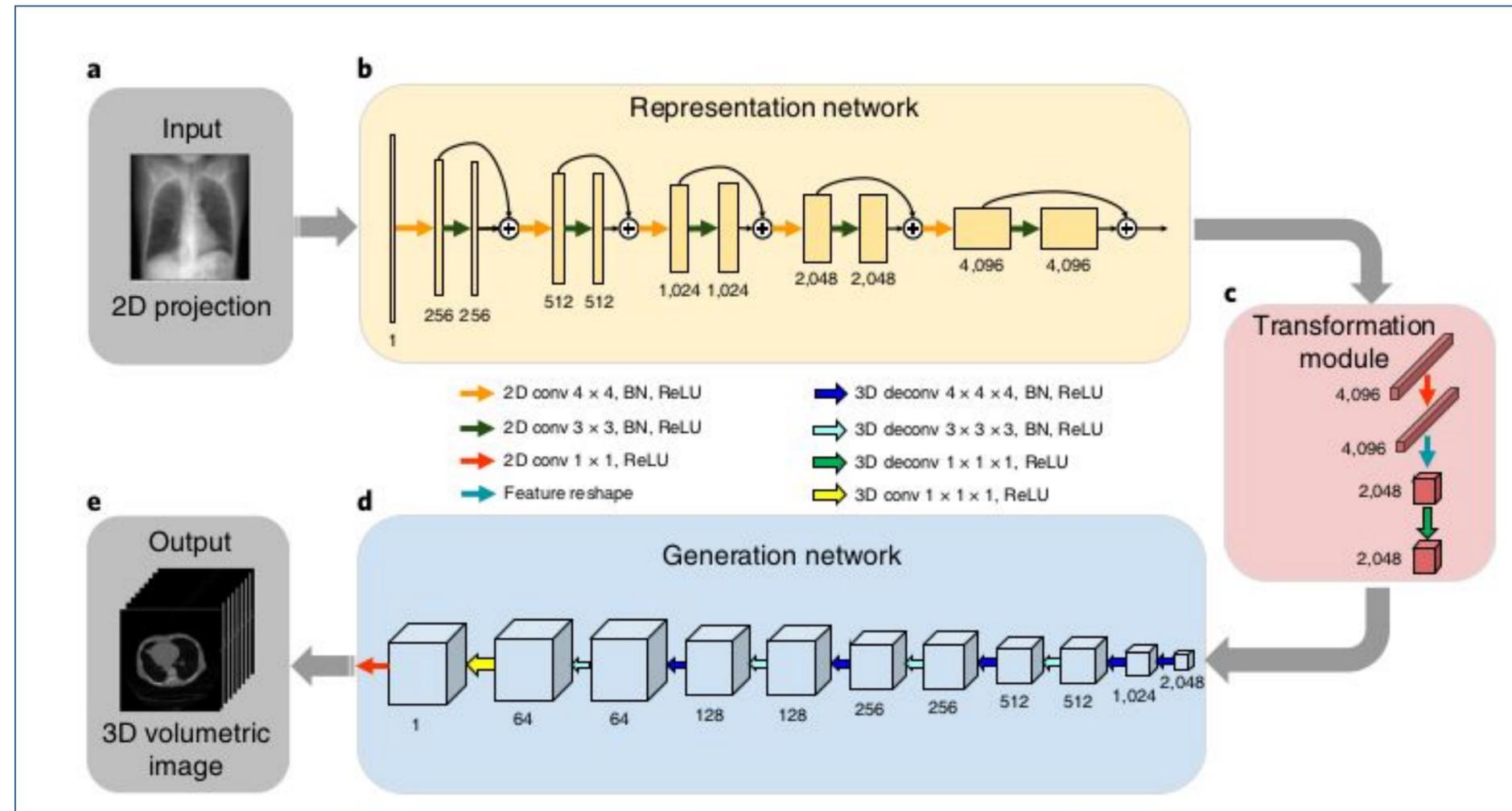


Imagen obtenida de: Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning.

# Comparación entre modelos

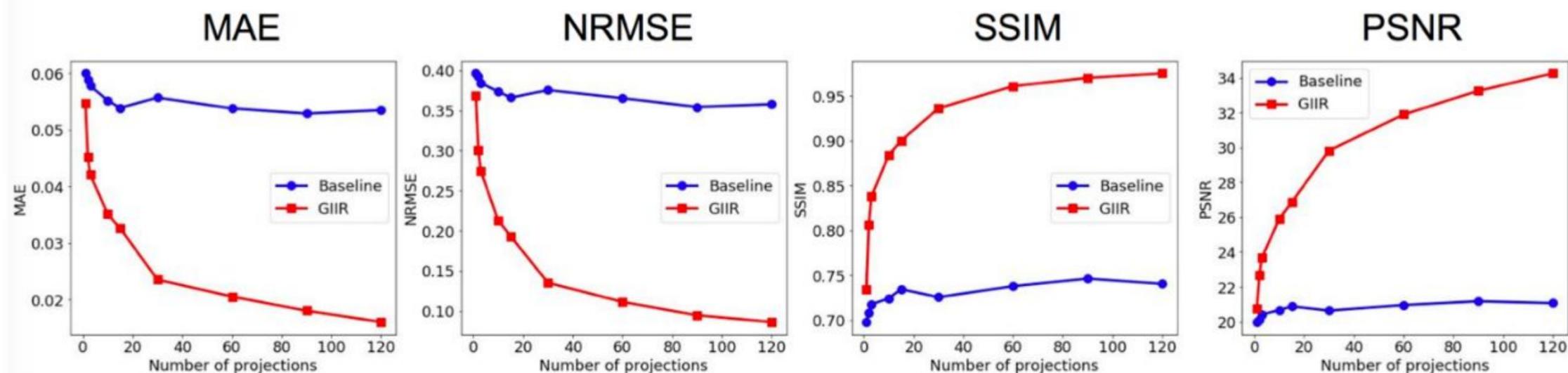


Fig. 7. Evaluation results of “Baseline” model (without inclusion of geometry) and “GIIR” model (with geometry) for different numbers of input projections. The x axis indicates the number of projections given as inputs. The y axis shows the values of evaluation metrics, MAE, NRMSE, SSIM, and PSNR, respectively. For reference, the MAE, NRMSE, SSIM for the fully sampled ground truth images are 0, 0, 1 respectively, while PSNR is not defined.

TABLE I  
RESULTS OF ULTRA-SPARSE-VIEW RECONSTRUCTION

Reconstruction Methods	Single-view			Two-view			Three-view		
	<i>NRMSE</i>	<i>SSIM</i>	<i>PSNR</i>	<i>NRMSE</i>	<i>SSIM</i>	<i>PSNR</i>	<i>NRMSE</i>	<i>SSIM</i>	<i>PSNR</i>
Baseline	0.3961	0.6978	19.9819	0.3927	0.7078	20.1455	0.3841	0.7175	20.4150
X2CT	0.3977	0.7002	20.2308	0.3751	0.7311	20.9120	-	-	-
X2CT(+proj loss)	0.3766	0.7117	20.5705	0.3596	0.7286	20.9924	-	-	-
GIIR	0.3684	0.7341	20.7699	0.3000	0.8067	22.6868	0.2740	0.8378	23.6688

*NRMSE*, normalized root mean squared error; *SSIM*, structural similarity; *PSNR*, peak signal noise ratio.

# Resultados de la red de refinamiento de imágenes 3D

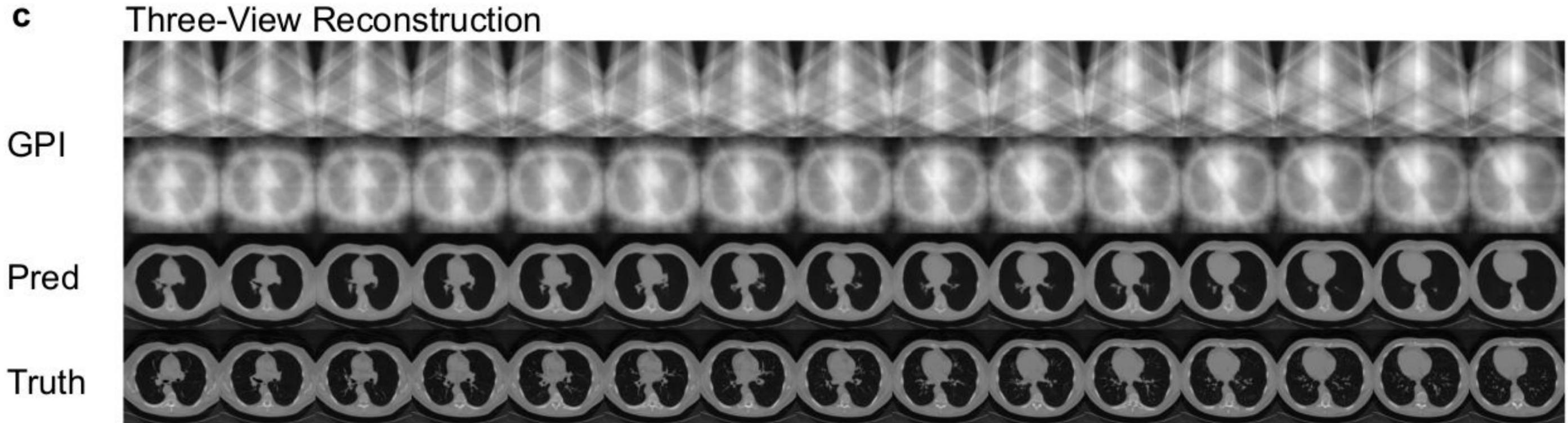


Fig. 8. Results of ultra-sparse reconstruction by GIIIR model. Here, we show the cross-sectional slices of 3D CT images of the same testing sample from (a) single-view reconstruction, (b) two-view reconstruction, (c) three-view reconstruction, respectively. In each row, cross-sectional slices of ground-truth 3D images, predicted final 3D images, and GPIs, are shown when 1-3 projections are inputted to GIIIR model. Two GPIs back-projected from ground-truth projections and generated projections separately serve as inputs to 3D image refinement network.

MUCHAS GRACIAS