



MAIA: SISTEMA FEDERADO DE ENTRENAMIENTO DE MODELOS DE IA PARA ONCOLOGÍA RADIOTERÁPICA

Martínez-Larraz Solís A.¹, Martí Asenjo J.²

Aprendizaje federado

El aprendizaje federado es una forma de entrenamiento de modelos de Inteligencia Artificial compartidos que permite entrenar sin centralizar los datos. En lugar de enviar todos los datos a un servidor central, el aprendizaje federado permite el entrenamiento en dispositivos locales, compartiendo únicamente los parámetros que definen matemáticamente a cada uno de los modelos entrenados. En lugar de consolidar todos los datos en un solo lugar, el aprendizaje federado permite que los modelos se entrenen directamente en los dispositivos locales donde residen los datos. Este enfoque permite el **entrenamiento seguro de modelos compartidos** de Inteligencia artificial, aunque estén basados en datos sensibles y protegidos por leyes de privacidad (LOPD, RGPD), como es el caso de la imagen médica y otras fuentes de información empleadas en Oncología Radioterápica.

maIA

maIA está pensado para su uso en el entorno de la Oncología Radioterápica (OR). Su objetivo es propiciar el avance científico y tecnológico de la especialidad, facilitando el uso de herramientas de inteligencia artificial (IA) al usuario principiante para el entrenamiento de modelos, tanto locales como compartidos con otras instalaciones, garantizando siempre el cumplimiento de las leyes de privacidad.

Generación de Datasets

maIA permite la creación de datasets al trabajar directamente sobre los archivos exportados de los diferentes sistemas de planificación en formato DICOM: imágenes, estructuras, planes y dosis. maIA permite acceder de forma rápida y ágil a grandes volúmenes de datos, del orden de miles de pacientes, para crear datasets de una o más estructuras. La creación de datasets está enfocada a la segmentación, pero también a otras tareas como los modelos predictivos basados en características radiómicas o dosiómicas. Además, permite soslayar problemas habituales como la diversidad en la nomenclatura de las estructuras, tanto dentro de la propia instalación como entre distintas instalaciones.

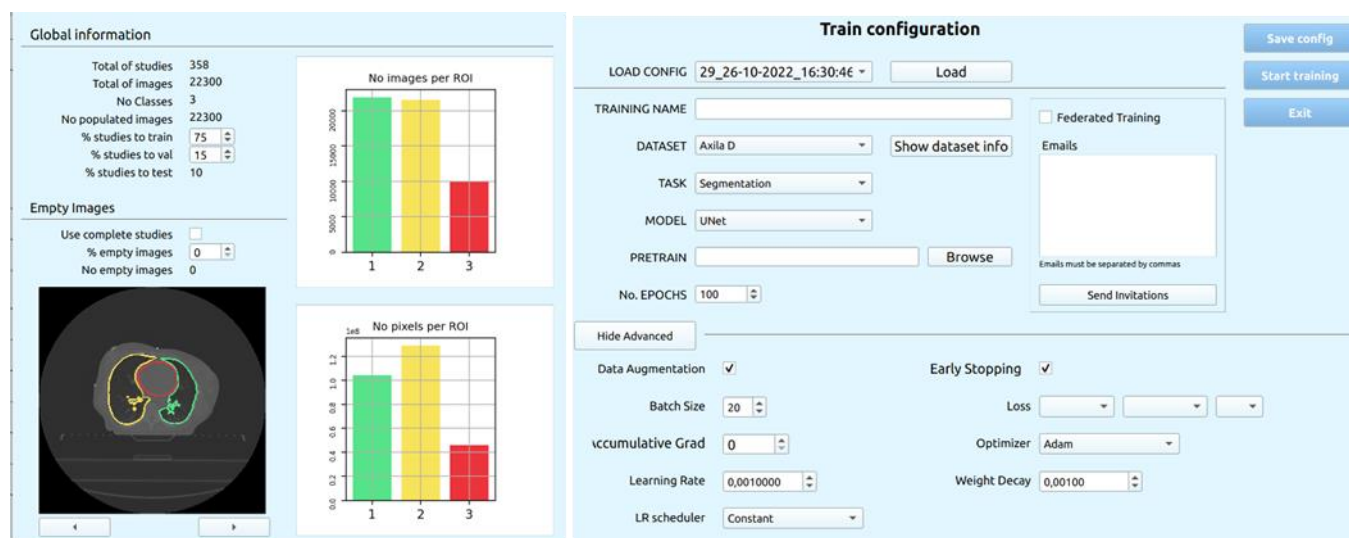


Image 1. a) Creación de un dataset de 3 órganos para segmentación de tórax. b) Panel de configuración del entrenamiento.

Modelos de entrenamiento

El sistema incorpora distintos modelos de entrenamiento enfocados a las distintas tareas disponibles, incluyendo las redes neuronales convolucionales más habituales en imagen médica como la U-Net y sus distintas variantes, incluidos modelos 3D como la UNet3D, así como otras arquitecturas más sofisticados como los llamados *transformers*, o modelos clásicos de Machine Learning para los modelos predictivos como *Random Forest*, SVM o KNN. Todos estos modelos cuentan con una configuración por defecto, aunque está permitida la edición de parámetros, como las funciones de pérdida, *schedulers*, tamaño del *batch*, número de *epochs*...

Métricas y visualización

maIA incorpora una interfaz gráfica para evaluar la evolución del entrenamiento en términos de distintas métricas ya implementadas, como el coeficiente DSC y la distancia Hausdorff, o índices como la sensibilidad, especificidad etc. Permite además la evaluación simultánea de los dataset de entrenamiento y validación, pero también del conjunto de test una vez finalizado el entrenamiento.

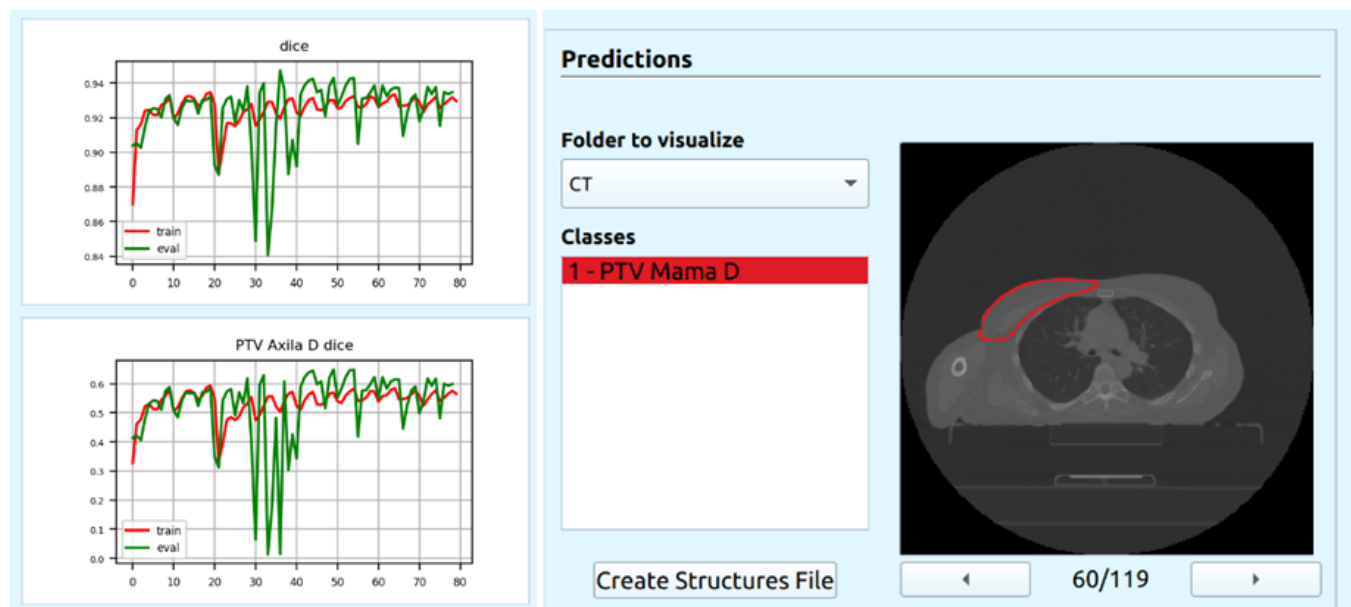


Image 2. a) Estadísticas del parámetro DICE en un entrenamiento. b) Cuadro de inferencia y generación del archivo de estructuras.

Futuro

La Oncología radioterápica es una especialidad íntimamente conectada a la tecnología y los datos. Los datos generados en el flujo de trabajo habitual de los Servicios de ORT, combinados con la Inteligencia Artificial, tienen un gran potencial científico y tecnológico para el futuro. El entrenamiento de modelos compartidos entre instituciones aumenta de forma muy significativa el volumen de datos disponible y disminuye posibles sesgos, resultando en modelos más robustos y precisos. La herramienta maIA está enfocada a facilitar esta tarea, manteniendo la privacidad de los datos compartidos y permitiendo un acceso simple e intuitivo para usuarios sin conocimientos de programación o experiencia en modelos de IA.