

UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingeniería y Sistemas de
Telecomunicación



Contribuciones a la caracterización automática
del tejido cerebral y lesiones asociadas
mediante modelos supervisados en imagen
médica multimodal para diagnóstico asistido

TESIS DOCTORAL

Presentada para optar al título de Doctor por:

Gemma Urbanos García

Máster Universitario en Gestión y Desarrollo de Tecnologías Biomédicas

Madrid, 2025



UNIVERSIDAD POLITÉCNICA DE MADRID
Escuela Técnica Superior de Ingeniería y Sistemas de
Telecomunicación

**Doctorado en Ingeniería de Sistemas y Servicios para la Sociedad
de la Información**

**Contribuciones a la caracterización automática
del tejido cerebral y lesiones asociadas
mediante modelos supervisados en imagen
médica multimodal para diagnóstico asistido**

TESIS DOCTORAL

Presentada para optar al título de Doctor por:

Gemma Urbanos García

Máster Universitario en Gestión y Desarrollo de Tecnologías Biomédicas

Bajo la dirección de:

Dr. Eduardo Juárez Martínez

Dr. Alfonso Lagares Gómez-Abascal

Madrid, 2025

Título: Contribuciones a la caracterización automática del tejido cerebral y lesiones asociadas mediante modelos supervisados en imagen médica multimodal para diagnóstico asistido

Autor: Gemma Urbanos García

Programa de Doctorado: Ingeniería de Sistemas y Servicios para la Sociedad de la Información

Dirección de Tesis:

Dr. Eduardo Juárez Martínez, Catedrático de Universidad, Departamento de Ingeniería Telemática y Electrónica, Universidad Politécnica de Madrid

Dr. Alfonso Lagares Gómez-Abascal, Catedrático de Universidad, Departamento de Cirugía, Facultad de Medicina, Universidad Complutense de Madrid, y Jefe de Servicio de Neurocirugía, Hospital Universitario 12 de Octubre

Revisores Externos:

Tribunal de Tesis:

Fecha de Defensa de Tesis:

Esta tesis ha sido parcialmente financiada por el proyecto NEMESIS-3D-CM (Y2018/BIO-4826), proyecto de la Comunidad de Madrid.

*“La ciencia tiene gran belleza.
Un científico en su laboratorio no es solo un técnico:
es también un niño ante fenómenos naturales
que le impresionan como un cuento de hadas.”*
— Marie Curie

Agradecimientos

En primer lugar, quiero agradecer sinceramente a mis directores, Alfonso Lagares y Eduardo Juárez, por su guía y apoyo a lo largo de este trabajo. Aunque mi recorrido como doctoranda no ha sido el más habitual, vuestra paciencia y confianza han sido determinantes para llegar hasta aquí. Os agradezco también el haberme transmitido vuestra pasión por la investigación, que hoy define en gran parte mi manera de trabajar.

Gracias al personal de la UPM, y en particular a César, por su disponibilidad y su ayuda en todas las gestiones relacionadas con este trabajo.

También quiero agradecer sinceramente a los neurocirujanos del Hospital 12 de Octubre, cuya implicación y visión clínica han aportado un valor esencial a este trabajo y han permitido mantener una orientación aplicada hacia la práctica asistencial.

Extiendo mi agradecimiento a Ana Jiménez y Ángel Alberich por su apoyo y por haberme facilitado los recursos necesarios para realizar este trabajo, así como a todos mis compañeros de Quibim, que, independientemente de su departamento, han contribuido de una u otra forma.

Quiero agradecer especialmente a mis compañeros del departamento de *Research & Frontiers of AI*. Formar parte de un equipo tan preparado y colaborativo ha sido una gran oportunidad de aprendizaje y crecimiento profesional.

Finalmente, más allá de los aspectos técnicos y académicos, este recorrido ha sido también una experiencia personal de transformación y aprendizaje. Echando la vista atrás, da un poco de vértigo pensar en aquella niña que entró por primera vez en el laboratorio del CITSEM y compararla con la persona que soy ahora. Mi evolución, sin duda, se la debo a todas las personas que he ido encontrando por el camino.

Recuerdo con especial cariño aquellos días en que Luisa y yo recorriamos los pasillos del hospital con una maleta y un trípode, buscando carritos de quirófano para llevar a cabo nuestros experimentos. También guardo muy buenos recuerdos con Dani y Raquel, que, además de ofrecerme toda su ayuda como veteranos, se convirtieron en amigos.

De mi etapa en el laboratorio recuerdo una clase de inglés en la que preguntaron: “¿Cuál es tu momento favorito del día?”. Sin pensarlo, respondí: “desayunar con Jaime y Guille”. También quedan para el recuerdo las pocas, pero memorables, mañanas de *gymbros* con Jaime.

Guardo con aprecio las charlas con Manu y los cafés después de comer con Miguel, Gonzalo, Alberto y Pedro. Y, cómo no, las tardes con mi compañero de mesa Guille “CNN”, compartiendo memes, comida preprocesada y alguna que otra clase sobre redes neuronales.

También aprendí mucho y disfruté de las conversaciones con Pallab y Anup, y, sobre todo, de las charlas infinitas con Santi, que cada mañana iba directa a mi mesa con su sonrisa y algún tema nuevo sobre la vida.

Después comenzó mi etapa en Quibim, donde tuve la suerte de coincidir con Carlos a mi llegada.

Poco más tarde, conocí a Ángel y a Carmen, tres compañeros que pronto se convirtieron en buenos amigos. Con vosotros he compartido innumerables risas en la oficina o *summer parties* en casas algo destartaladas pero siempre llenas de buen ambiente. Incluso en los momentos de agobio, siempre encontramos la forma de reírnos de todo.

En este camino no han faltado Josep y Javi, compañeros de máster y amigos con quienes he compartido tantas risas en el *Baby Bull* o donde fuera. Pocos compañeros de viaje tan auténticos como vosotros.

Quiero agradecer a mis amigos los CRPNS por soportar mis charlas sobre mi trabajo, hacer *spam* de mis ponencias y hasta conectaros a verlas. Vuestra amiga “la que cura el cáncer”, por fin se queda sin la excusa de la tesis para no salir.

En todo este tiempo, también ha estado Ester, mi amiga de toda la vida, con quien siempre encuentro paz. Gracias por acompañarme siempre y por escuchar, con una mezcla de asombro y curiosidad, mis intentos de explicar de qué va mi tesis.

Por supuesto, esta tesis se la dedico a mis amigas Ana, Cristina y Lidia. Han sido un apoyo constante, las primeras a las que acudía para compartir cualquier novedad o pedir consejo. Gracias por estar siempre orgullosas de mí; yo también lo estoy de vosotras.

También han sido una parte muy importante de este camino Elo, Jose Manuel, Cristi y la familia Rubio, quienes se han convertido en mi familia. Gracias por vuestro cariño, paciencia y por acompañarme durante todo este proceso.

No puedo dejar de mencionar a mis abuelos, que siempre estuvieron orgullosos de mí y cuya presencia sigo sintiendo en cada paso. Su ejemplo y su cariño incondicional siguen formando parte de lo que soy hoy. Por ellos, y como decía mi abuelo en cada brindis: *por la industria*.

Quiero dar las gracias también a mi hermana María y a Sergio, por su apoyo, su cariño y por estar siempre ahí. Y, por supuesto, a mi sobrino Daniel, nuestro pequeño, que nos llena los días de alegría y nos enseña a ver el mundo de otra manera. Tu tata te quiere mucho.

Nada de esto hubiera sido posible sin mis padres. Gracias por creer siempre en mí, por acompañarme en mis fracasos, en mis victorias, en mis alegrías y en mis tristezas. Por tener siempre un consejo y unos brazos en los que recogerme. Todo lo que soy, os lo debo en gran parte a vosotros.

Y, por último, a ti, mi compañero en todos estos momentos desde hace ya tantos años. Me he desvelado a medianoche y tú te has desvelado conmigo; me he quedado trabajando hasta tarde y siempre has estado a mi lado. Has sido mi apoyo, calma y alegría en este camino, y tu huella también está en estas páginas. Gracias por ayudarme a crecer cada día. Jose Miguel, esta tesis es tan mía como tuya.

Abstract

Clinical practice in neuroscience faces persistent limitations that constrain diagnostic accuracy, reproducibility, and operational efficiency. The visual interpretation of medical images remains an inherently subjective process, influenced by the observer’s experience and interindividual variability, while the use of conventional clinical scales, although useful as reference guidelines, provides only a partial representation of the underlying pathophysiological complexity of neurological diseases. Furthermore, the manual segmentation of anatomical structures or lesions continues to be a time and resource intensive procedure, particularly critical in acute conditions such as aneurysmal subarachnoid hemorrhage (aSAH), where delays in lesion quantification or detection may have significant prognostic consequences.

In this context, the present doctoral thesis comprehensively addresses the development and validation of automated systems for medical image analysis based on *Machine Learning* (ML) techniques, aiming to advance toward a more accurate, reproducible, and efficient diagnosis, to improve the prediction of clinical outcomes, and to support real-time decision-making in neurosurgical settings. Three complementary research axes are proposed: the application of *Hyperspectral Imaging* (HSI) for intraoperative tissue classification in glioblastoma multiforme (GBM); the automatic segmentation of intracranial hemorrhages in computed tomography (CT) scans of patients with aSAH; and the prognostic modeling using radiomic biomarkers derived from both automatic hemorrhage segmentations and the surrounding brain parenchyma.

In the first use case, focused on intraoperative tissue classification using HSI, a prospective cohort of 12 GBM patients undergoing surgery at the Hospital Universitario 12 de Octubre (H12O) has been analyzed. During the interventions, 13 hyperspectral cubes have been acquired using a *snapshot* system capturing 25 bands in the visible and near-infrared range (VNIR, 655–975 nm), enabling simultaneous spatial and spectral information acquisition. Regions of interest have been manually segmented by neurosurgeons into five tissue classes (tumor, healthy tissue, arterial blood, venous blood, and dura mater), from which mean spectral signatures are extracted. Three representative classifiers are implemented: a linear Support Vector Machine (SVM), a *Random Forest* model, and a three-dimensional convolutional neural network (3D-CNN). Two experimental designs are defined to assess both generalization capacity and upper-bound performance: an inpatient scheme with triple stratified cross-validation ($5 \times 5 \times 5$ folds), estimating the maximal achievable performance under controlled conditions, and an interpatient *leave-one-patient-out* nested stratified validation (5×5 folds), evaluating generalization to unseen patients. The inpatient experiments achieve overall accuracies (OACC) above 95 % with errors below 5 %, demonstrating the technical feasibility of HSI-based tissue classification. In the interpatient scenario, performance decreases to approximately 60 % OACC, reflecting biological and technical variability while confirming the potential of this approach in realistic surgical environments. The achieved spectral efficiency (2.4 % OACC/band) exceeds that of prior high-resolution systems, suggesting that more compact and portable configurations could be deployed intraoperatively in real time. The resulting dataset has been integrated into the public SLIMBRAIN repository, an open hyperspectral database focused on neurosurgery, thus promoting multicenter validation and reproducibility.

The second use case addresses the automatic segmentation of intracranial hemorrhages in CT, a key task in the initial management of aSAH and other hemorrhagic lesions, which traditionally requires 15–20 minutes of manual delineation per patient. For model development, a retrospective institutional cohort from H12O including 539 patients diagnosed with aSAH between 2007 and 2024 has been used, of whom 223 have expert manual segmentations of subarachnoid hemorrhage (SAH), intraventricular hemorrhage (IVH), and intraparenchymal hemorrhage (IPH). To increase training diversity and enhance generalization, three public datasets have been integrated: HemSeg-200 (from the RSNA 2019 Brain CT Hemorrhage Challenge, including SAH, IPH, and IVH of various etiologies), BHSD (traumatic brain injuries), and a dataset from Al Hilla Hospital (Iraq) with traumatic hemorrhages, totaling 450 additional CT scans.

Two main experiments have been designed. The **Experiment 1** is trained using only the H12O and HemSeg-200 datasets, representing a controlled scenario with manually segmented aSAH and associated hemorrhages (SAH, IVH, IPH). The **Experiment 2** incorporates BHSD and Al Hilla datasets as additional training sources to increase variability, while maintaining the same test set as in Experiment 1 to enable direct comparison. Models are developed using the **nnU-Net** framework, applying five-fold cross-validation and model ensembling to ensure robustness against partition variability. In the H12O cohort, Experiment 1 achieves a median *Dice score* of 0.75 and an *Average Surface Distance* (ASD) of 1.50 mm, while Experiment 2 improves the *Dice* to 0.81 with an ASD of 1.63 mm. Both models outperform five reference methods from the literature and, in a blinded evaluation by neurosurgeons using the Borda count method, Experiment 2 is ranked first overall. The automatic segmentation reduces the average analysis time from 15 minutes to 23.4 seconds (a 97.4% reduction). Moreover, the automatically derived hemorrhage volume shows prognostic performance equivalent to the manual measurement (AUC = 0.810 vs. 0.812) and superior to the modified Fisher scale (AUC = 0.733), confirming its validity as an objective and reproducible quantitative biomarker.

The third use case focuses on prognostic modeling based on radiomics in a cohort of 444 aSAH patients (403 for model development and 41 for independent validation). Four clinically relevant outcomes are analyzed: mortality, six-month functional outcome according to the *Glasgow Outcome Scale* (GOS), symptomatic vasospasm, and hydrocephalus. Radiomic feature extraction includes first, second, and third-order descriptors. Features are extracted from two complementary regions: the automatically segmented hemorrhage (using both the model from Experiment 2 and an external model from the literature) and the surrounding brain parenchyma, enabling the exploration of whether early diffuse parenchymal injury provides additional prognostic information.

Four modeling strategies are systematically compared: purely clinical models, radiomic models from the hemorrhage, radiomic models from the parenchyma, and combined clinical-radiomic models. Validation is performed using stratified nested cross-validation (5 outer folds \times 3 inner folds) with three random seeds to ensure statistical robustness. Algorithms include *Random Forest*, *ExtraTrees*, *XGBoost*, and *Support Vector Machine*, with hyperparameter optimization, feature selection, outlier removal, and class balancing. In cross-validation, purely clinical and radiomic models achieve comparable performances (AUC = 0.84 and 0.82 for mortality,

respectively), confirming the independent predictive value of radiomics. However, combined models systematically improve the metrics, achieving in the external test set an AUC = 0.89 for mortality and 0.87 for functional outcome, with specificities near 0.98 and sensitivities between 0.75 and 0.88. For vasospasm and hydrocephalus, AUCs of 0.66 and 0.72 are obtained, indicating clinically meaningful discrimination. The interpretability analysis using *SHapley Additive exPlanations* (SHAP) identifies the *World Federation of Neurosurgical Societies* (WFNS) scale at admission and radiomic features from the parenchyma as the most influential variables, reinforcing clinical coherence and suggesting that radiomics capture pathophysiological information not discernible through visual inspection.

Systematic comparisons against established clinical scales robustly confirm the superiority of the proposed models. Specifically, the combined radiomics from parenchyma and clinical data models significantly outperform the *ictWFNS* (AUC 0.50–0.57 in univariable analysis) and the SEBES and SEBES6c (*Subarachnoid Early Brain Edema Score*) scales, whose AUCs range from 0.50–0.57 in univariable settings and 0.59–0.84 when integrated with additional clinical variables such as the modified Fisher scale, age, or WFNS at admission.

As a methodological extension, the “radiomic SEBES” model was developed to objectively reproduce the visual assessment of cerebral edema based on radiomic features. The model closely replicates neurosurgeons’ evaluations (AUC = 0.89), showing that radiomics can capture expert-level perceptual information. However, its prognostic value is lower than that of models optimized for each clinical outcome, indicating that although emulating visual scales enhances interpretability, it does not necessarily improve predictive power.

Although the results confirm the technical feasibility, methodological robustness, and potential clinical value of the developed systems, certain limitations constrain immediate generalization. These include the moderate sample size (particularly in the HSI study), the partially single-center nature of the cohorts, the lack of prospective validation under real clinical conditions, and the sensitivity of image biomarkers to variations in acquisition protocols, resolution, or scanner calibration. Additionally, some outcomes, such as vasospasm, exhibit heterogeneous definitions that introduce additional variability.

Future research directions include multicenter and prospective validation in larger and more heterogeneous cohorts, the application of image or feature harmonization techniques, automated hyperparameter optimization, nonlinear dimensionality reduction, and the incorporation of data drift monitoring mechanisms to ensure model stability after clinical deployment. Likewise, the future integration of these systems into interoperable medical imaging platforms and their evaluation within regulatory frameworks, leading to their certification as medical-grade software, is envisaged.

Overall, this doctoral thesis demonstrates that integrating ML algorithms with hyperspectral and radiomic biomarkers represents a solid and transformative approach to the quantitative analysis of neurological pathology. In the context of clinical neuroscience, these advances lay the groundwork for a new generation of decision-support tools capable of integrating multimodal information and contributing to a more efficient, objective, and evidence-based practice aligned with the principles of translational artificial intelligence in healthcare.

Resumen

La práctica clínica en neurociencia se enfrenta a limitaciones persistentes que condicionan la precisión diagnóstica, la reproducibilidad de las evaluaciones y la eficiencia operativa. La interpretación visual de las imágenes médicas sigue siendo un proceso subjetivo, influido por la experiencia del observador y la variabilidad interindividual, mientras que la aplicación de escalas clínicas convencionales, aunque útiles como guías de referencia, ofrece una representación parcial de la complejidad fisiopatológica de las enfermedades neurológicas. Por otra parte, la segmentación manual de estructuras o lesiones continúa siendo un procedimiento intensivo en tiempo y recursos, especialmente crítico en patologías agudas como la hemorragia subaracnoidea aneurismática (HSAa), donde los retrasos en la cuantificación o detección de la lesión pueden tener consecuencias pronósticas relevantes.

En este contexto, la presente tesis doctoral aborda el desarrollo y validación de sistemas automatizados de análisis de imagen médica basados en técnicas de aprendizaje automático (*Machine Learning*, ML), con el objetivo de avanzar hacia un diagnóstico más preciso, reproducible y eficiente, mejorar la predicción de desenlaces clínicos y facilitar la toma de decisiones en tiempo real en entornos neuroquirúrgicos. Para ello, se plantean tres ejes de trabajo complementarios: la aplicación de imagen hiperespectral (*Hyperspectral Imaging*, HSI) para la clasificación tisular intraoperatoria en glioblastoma multiforme (GBM), la segmentación automática de hemorragias intracraneales en tomografía computarizada (TC) en pacientes con HSAa, y la modelización pronóstica mediante biomarcadores radiómicos derivados de dichas segmentaciones automáticas y del parénquima cerebral adyacente.

En el primer caso de uso, centrado en la clasificación tisular intraoperatoria mediante HSI, se ha empleado una cohorte prospectiva de 12 pacientes con GBM intervenidos en el Hospital Universitario 12 de Octubre (H12O). Durante las cirugías se han adquirido 13 cubos hiperespectrales mediante un sistema *snapshot* que registra 25 bandas en el rango visible e infrarrojo cercano (VNIR, 655–975 nm), permitiendo capturar información espacial y espectral simultáneamente. Las regiones de interés han sido segmentadas manualmente por neurocirujanos en cinco clases tisulares diferenciadas (tumor, tejido sano, sangre arterial, sangre venosa y duramadre), sobre las que se extraen firmas espectrales promedio. A partir de estas firmas se han desarrollado tres clasificadores representativos de paradigmas complementarios: una máquina de vectores soporte (SVM lineal), un modelo basado en *Random Forest* y una red neuronal convolucional tridimensional (3D-CNN). Para evaluar la capacidad de generalización y el límite superior de rendimiento, se han diseñado dos escenarios experimentales: un esquema intrapaciente con validación cruzada estratificada triple ($5 \times 5 \times 5$ folds), que estima el máximo rendimiento alcanzable bajo condiciones controladas, y un esquema interpaciente tipo *leave-one-patient-out* con validación anidada estratificada (5×5 folds), que mide la capacidad de generalizar a nuevos pacientes no vistos. Los resultados intrapaciente muestran precisiones globales (OACC) superiores al 95 % con errores menores del 5 %, evidenciando la viabilidad técnica de la clasificación tisular basada en HSI. En el escenario interpaciente, el rendimiento se reduce a un OACC del 60 %, lo que pone de manifiesto la influencia de la variabilidad biológica y técnica, pero confirma el potencial de la metodología en entornos quirúrgicos reales. La eficiencia espectral alcanzada (2,4 % OACC/banda) supera a la de estudios previos con

cámaras de mayor resolución espectral, lo que abre la posibilidad de implementar sistemas más compactos, rápidos y portátiles para uso intraoperatorio en tiempo real. Los datos derivados de este estudio se han integrado en la base pública SLIMBRAIN, colección hiperespectral abierta centrada en neurocirugía, con el fin de promover la validación multicéntrica y la reproducibilidad futura.

El segundo caso de uso aborda la segmentación automática de hemorragias intracraneales en TC, una tarea esencial en el manejo inicial de pacientes con HSAa y otras lesiones hemorrágicas, que tradicionalmente requiere entre 15 y 20 minutos de delineado manual por paciente. Para el desarrollo de los modelos, se ha empleado una cohorte retrospectiva institucional del H12O compuesta por 539 pacientes diagnosticados de HSAa entre 2007 y 2024, de los cuales 223 disponen de segmentaciones manuales completas llevadas a cabo por neurocirujanos expertos de hemorragia subaracnoidea (HSA), hemorragia intraventricular (HIV) y hemorragia intraparenquimatosa (HIP). Con el propósito de incrementar la diversidad de entrenamiento y mejorar la capacidad de generalización, se han integrado tres bases de datos públicas adicionales: HemSeg-200 (procedente del RSNA 2019 Brain CT Hemorrhage Challenge, que incluye casos de HSA, HIP y HIV de diferentes etiologías), BHSD (centrada en traumatismos craneoencefálicos) y un conjunto del Hospital Al Hilla (Irak) con hemorragias traumáticas, totalizando 450 estudios adicionales que amplían la variabilidad demográfica, técnica y etiológica.

Se han diseñado dos experimentos principales. El **Experimento 1** emplea exclusivamente los datos del H12O y HemSeg-200, representando un escenario controlado con segmentaciones manuales de HSAa y posibles hemorragias derivadas de HSAa (HSA, HIV, HIP). El **Experimento 2** incorpora además las bases BHSD y Al Hilla como estrategia de aumento de datos durante el entrenamiento, manteniendo invariable el conjunto de test del Experimento 1 para permitir una comparación directa y objetiva. Los modelos se desarrollan utilizando el *framework* **nnU-Net**, con validación cruzada de cinco folds y promediado de modelos (*ensembling*) para garantizar robustez frente a la variabilidad de las particiones. En la cohorte H12O, el modelo del Experimento 1 alcanza una mediana de *Dice score* de 0,75 y una distancia media a la superficie (*Average Surface Distance*, ASD) de 1,50 mm, mientras que el modelo del Experimento 2 mejora el *Dice* hasta 0,81 y muestra un ASD de 1,63 mm. Ambos superan las métricas obtenidas por cinco modelos de referencia de la literatura reciente y, en la evaluación ciega por neurocirujanos mediante el método de Borda, el Experimento 2 se ha calificado en primera posición global. La segmentación automática reduce el tiempo medio de análisis de 15 minutos a 23,4 segundos, lo que supone una reducción del 97,4% respecto al proceso manual. Además, el volumen hemorrágico calculado automáticamente muestra capacidad pronóstica equivalente al volumen manual (AUC = 0,810 frente a 0,812) y superior a la escala de Fisher modificada (AUC = 0,733), confirmando su validez como biomarcador cuantitativo objetivo y reproducible.

El tercer caso de uso se centra en la modelización pronóstica basada en radiómica aplicada a una cohorte de 444 pacientes con HSAa, de los cuales 403 se destinan al desarrollo de los modelos y 41 a validación externa independiente. Se analizan cuatro desenlaces clínicamente relevantes: mortalidad, resultado funcional a los seis meses según la *Glasgow Outcome Scale* (GOS), vasoespasmos e hidrocefalia. La extracción de características radiómicas incluye

descriptores de primer, segundo y tercer orden. Las características se extraen desde dos regiones complementarias: la hemorragia segmentada automáticamente (mediante el modelo del Experimento 2 y un modelo externo de la literatura) y el parénquima cerebral adyacente, lo que permite explorar si el daño difuso temprano capturado en el tejido cerebral circundante aporta información pronóstica adicional.

Se comparan sistemáticamente cuatro estrategias de modelización: modelos puramente clínicos, modelos radiómicos derivados de la hemorragia, modelos radiómicos del parénquima y modelos combinados que integran variables clínicas y radiómicas. La validación se lleva a cabo mediante validación cruzada anidada estratificada (5 folds externos \times 3 internos) con tres semillas diferentes para garantizar robustez estadística. Se evalúan algoritmos *Random Forest*, *ExtraTrees*, *XGBoost* y *Support Vector Machine*, aplicando ajuste de hiperparámetros, selección de características, eliminación de *outliers* y balanceo de clases. En el test de la validación cruzada anidada, los modelos clínicos y radiómicos muestran rendimientos comparables (AUC = 0,84 y 0,82 para mortalidad, respectivamente), confirmando el valor predictivo independiente de la radiómica. Sin embargo, los modelos combinados mejoran sistemáticamente las métricas, alcanzando en test independiente AUC = 0,89 para mortalidad y 0,87 para resultado funcional, con especificidades próximas a 0,98 y sensibilidades entre 0,75 y 0,88. Para los desenlaces de vasoespasma e hidrocefalia se obtienen AUC de 0,66 y 0,72 respectivamente, con capacidades discriminativas clínicamente útiles. El análisis de interpretabilidad mediante *SHapley Additive exPlanations* (SHAP) identifica la escala WFNS (*World Federation of Neurosurgical Societies scale*) al ingreso y los descriptores radiómicos del parénquima como las variables más influyentes, lo que refuerza la coherencia clínica y sugiere que la radiómica capta información fisiopatológica no visible en la inspección visual convencional.

Las comparaciones sistemáticas frente a escalas clínicas consolidadas confirman de forma robusta la superioridad de los modelos propuestos. En particular, los modelos combinados de parénquima y variables clínicas superan significativamente a la *ictWFNS* (AUC 0,50–0,57 en análisis univariable) y a las escalas SEBES y SEBES6c (*Subarachnoid Early Brain Edema Score*) cuyos valores oscilan entre 0,50–0,57 en análisis univariable y 0,59–0,84 al integrarse con variables clínicas adicionales como el Fisher modificado, la edad o la WFNS al ingreso.

Como extensión metodológica, se desarrolló el modelo «SEBES radiómico», diseñado para reproducir objetivamente la valoración visual del edema cerebral a partir de características radiómicas. El modelo replica con alta precisión la evaluación de los neurocirujanos (AUC = 0,89), demostrando que la radiómica puede capturar información perceptual del experto. Sin embargo, su valor pronóstico es inferior al de los modelos optimizados para cada desenlace, lo que indica que la emulación de escalas visuales aunque mejore la interpretabilidad, no necesariamente mejora el poder predictivo de los modelos.

Aunque los resultados confirman la viabilidad técnica, la robustez metodológica y el valor clínico potencial de los sistemas desarrollados, se reconocen limitaciones que condicionan su generalización inmediata. Entre ellas destacan el tamaño muestral moderado (especialmente en el estudio HSI), la naturaleza parcialmente unicéntrica de las cohortes, la ausencia de validación prospectiva en condiciones de práctica clínica real y la posible sensibilidad de

los biomarcadores de imagen a las diferencias en protocolos de adquisición, resolución o calibración entre escáneres. Además, algunos desenlaces clínicos, como el vasoespasma, presentan definiciones heterogéneas que introducen variabilidad adicional.

Las líneas futuras de trabajo se orientan hacia la validación multicéntrica y prospectiva de los modelos desarrollados en cohortes amplias y heterogéneas, la incorporación de técnicas de armonización a nivel de imagen o de características, la automatización de la optimización de hiperparámetros, la implementación de estrategias de reducción no lineal de dimensionalidad y la incorporación de mecanismos de monitorización de deriva de datos, que garanticen la estabilidad del rendimiento tras el despliegue clínico. Asimismo, se plantea la futura integración de estos sistemas en plataformas interoperables de imagen médica y su evaluación conforme a marcos regulatorios que conduzcan a su certificación como productos sanitarios.

En conjunto, esta tesis doctoral demuestra que la integración de algoritmos de aprendizaje automático con biomarcadores derivados de imagen hiperespectral y radiómica constituye un enfoque sólido y transformador para el análisis cuantitativo de la patología neurológica. En el contexto de la neurociencia clínica, estos avances sientan las bases para una nueva generación de herramientas de apoyo a la decisión, capaces de integrar información multimodal y contribuir a una práctica más eficiente y científicamente fundamentada, alineada con los principios de la inteligencia artificial traslacional aplicada a la salud.

Tabla de Contenido

Agradecimientos	v
.	vii
Resumen	x
Lista de Figuras	xix
Lista de Tablas	xxv
Abreviaturas y acrónimos	xxx
1 Introducción	1
1.1 Motivación	1
1.2 Objetivos	3
1.3 Estructura del documento	5
2 Estado del arte	7
2.1 Organización y estructura del estado del arte	7
2.2 El reto clínico: diagnóstico y tratamiento en neurociencia	8
2.3 Anatomía del cerebro humano	9
2.4 Imagen médica en trastornos neurológicos: impacto clínico y aplicaciones	10
2.5 Técnicas de imagen y apoyo visual en neuroimagen intraoperatoria y convencional	13
2.5.1 Modalidades convencionales	13
2.5.2 Modalidades intraoperatorias	17
2.6 Biomarcadores de imagen y sus aplicaciones en neurociencia clínica	20
2.6.1 Desarrollo y extracción de biomarcadores de imagen	21
2.6.2 Imagen hiperespectral y firma espectral	24
2.6.3 Radiómica	26
2.7 Inteligencia Artificial en Imagen Médica: Fundamentos, Algoritmos y Aplicaciones	28
2.7.1 Aprendizaje automático (<i>Machine Learning</i>)	30
Regresión logística	30
Máquinas de vectores soporte (<i>Support Vector Machines</i>)	31
Árboles de decisión	31
2.7.2 Aprendizaje profundo (<i>Deep Learning</i>)	34
Estructura y Entrenamiento de Redes Neuronales Artificiales	35
Redes Neuronales Convolucionales	38
Arquitectura U-Net	40
2.7.3 Aplicaciones de la Inteligencia Artificial en neurociencia clínica	42

2.8	Casos de estudio: contextualización y relevancia clínica	43
2.8.1	Caso de estudio: Clasificación automática de tejidos intraoperatorios (NEMESIS-3D-CM)	45
2.8.2	Caso de estudio: Segmentación automática y predicción de pronóstico en pacientes con hemorragia subaracnoidea aneurismática	50
	Escalas clínicas y radiológicas de evaluación inicial	53
3	Materiales y métodos	63
3.1	Diseño del estudio y fuentes de datos	65
3.1.1	Planteamiento metodológico	65
	Integración de adquisición hiperespectral y análisis espectral para la clasificación de tejidos intraoperatorios en glioblastoma	66
	Integración de segmentación automática y radiómica para la caracterización y predicción de desenlaces en Hemorragia Subaracnoidea Aneurismática	67
3.1.2	Bases de datos y criterios de selección de pacientes	69
	Criterios de inclusión y exclusión	70
3.1.3	Procedimientos éticos y anonimización	71
	Procedimientos éticos y de anonimización en glioblastoma (Imagen Hiperespectral)	71
	Procedimientos éticos y de anonimización en Hemorragia Subaracnoidea Aneurismática (TC)	71
3.1.4	Proceso de adquisición de datos	72
	Imagen y datos espectrales en glioblastoma mediante imagen hiperespectral	73
	Sistema de captura de imagen hiperespectral intraoperatorio	73
	Imagen y datos clínicos en Hemorragia Subaracnoidea Aneurismática mediante TC	75
	Protocolos de adquisición de imagen	75
	Fuentes de datos para las variables clínicas	75
3.2	Preparación y anotación de datos	76
3.2.1	Definición de las clases objetivo	76
	Clases tisulares en glioblastoma (Imagen Hiperespectral)	76
	Desenlaces pronósticos en Hemorragia Subaracnoidea Aneurismática (TC)	77
3.2.2	Construcción del <i>ground truth</i>	79
	Etiquetado tisular en Imagen Hiperespectral de glioblastoma	79
	Segmentación manual de hemorragia	81
	Segmentación automática del parénquima cerebral	81
3.2.3	Preprocesado de datos	83
	Preprocesado de imagen	83
	Preprocesado de datos clínicos	86
3.3	Extracción de características	87
3.3.1	Firmas espectrales (HSI)	87
3.3.2	Características radiómicas (TC)	88

	Tipos de características extraídas	88
	Características derivadas mediante filtrado (radiómica de tercer orden).	89
3.4	Modelado y evaluación	90
3.4.1	Clasificación tisular en HSI GBM	90
	Diseño experimental	91
	Estrategias de entrenamiento	93
3.4.2	Segmentación automática de Hemorragia Subaracnoidea Aneurismática	96
	Diseño experimental	98
	Arquitectura y configuración de modelos	99
	Comparativa con modelos externos	100
	Impacto clínico : relación con escala de Fisher modificada	100
3.4.3	Modelos predictivos basados en radiómica	101
	Diseño experimental	102
	Integración de información multimodal	104
	Interpretabilidad	104
	Comparación con escalas clínicas	105
3.4.4	Métricas de evaluación	106
4	Resultados	109
4.1	Caracterización de las bases de datos	109
4.1.1	Clasificación tisular en glioblastoma (Imagen Hiperespectral)	110
4.1.2	Segmentación automática de Hemorragia Subaracnoidea Aneurismática (TC)	112
	Tamaño muestral	112
	Información a nivel de imagen	113
	Información clínica	117
	Variables clínicas y Desenlaces pronósticos	117
4.1.3	Modelos predictivos radiómicos en HSAa (TC)	118
	Información a nivel de imagen	119
	Información clínica	121
4.2	Extracción de características	125
4.2.1	Firmas espectrales en glioblastoma (Imagen Hiperespectral)	126
4.2.2	Características radiómicas en Hemorragia Subaracnoidea Aneurismática (TC)	126
4.3	Modelado y evaluación	131
4.3.1	Clasificación tisular en glioblastoma (Imagen Hiperespectral)	131
	Experimento A	131
	Experimento B	135
4.3.2	Segmentación automática en Hemorragia Subaracnoidea Aneurismática	138
	Resultados cuantitativos	138
	Resultados cualitativos	143
	Eficiencia temporal: segmentación manual frente a automática	147
	Análisis comparativo con escalas clínicas convencionales	147
4.3.3	Modelos predictivos radiómicos	149
	Interpretabilidad de los modelos	155

	Comparación con escalas clínicas	157
5	Discusión	163
5.1	Rendimiento técnico y comparación con la literatura	163
5.1.1	Clasificación tisular de glioblastoma en Imagen Hiperespectral	163
5.1.2	Segmentación automática de Hemorragia Subaracnoidea Aneurismática en TC	168
5.1.3	Modelos predictivos radiómicos en Hemorragia Subaracnoidea Aneurismática	171
5.2	Implicaciones clínicas	176
5.2.1	Clasificación tisular de glioblastoma en Imagen Hiperespectral	177
5.2.2	Segmentación automática de Hemorragia Subaracnoidea Aneurismática en TC	178
5.2.3	Modelos predictivos radiómicos en Hemorragia Subaracnoidea Aneurismática	180
	Interpretabilidad y factores que condicionan el rendimiento de los modelos	180
	Comparación con escalas clínicas tradicionales	181
5.3	Limitaciones	185
5.3.1	Clasificación tisular de glioblastoma en Imagen Hiperespectral	185
5.3.2	Segmentación automática de Hemorragia Subaracnoidea Aneurismática en TC	186
5.3.3	Modelos predictivos radiómicos en Hemorragia Subaracnoidea Aneurismática	187
6	Conclusiones y líneas futuras	189
6.1	Conclusiones	190
6.1.1	OBJ-S1: Clasificación y segmentación automática de lesiones cerebrales	190
	Clasificación tisular intraoperatoria mediante HSI (OBJ-S1.1)	190
	Segmentación automática de hemorragias cerebrales en TC (OBJ-S1.2)	191
6.1.2	OBJ-S2: Integración de biomarcadores avanzados en modelos predictivos	192
	Firmas espectrales en HSI para clasificación tisular (OBJ-S2.1)	192
	Características radiómicas en HSAa para predicción de desenlaces (OBJ-S2.2)	192
6.1.3	OBJ-S3: Validación clínica del sistema integrado	193
	Aplicabilidad clínica de la HSI en cirugía de GBM	193
	Biomarcadores pronósticos en HSAa: volumen hemorrágico	194
	Biomarcadores pronósticos en HSAa: radiómica	194
6.1.4	Marco metodológico y contribución global	195
6.1.5	Contribuciones a la caracterización automática del tejido cerebral y lesiones asociadas mediante modelos supervisados en imagen médica multimodal para diagnóstico asistido	195
6.2	Líneas futuras	196
6.2.1	Imagen hiperespectral aplicada a neurocirugía	196
6.2.2	Segmentación automática de Hemorragia Subaracnoidea Aneurismática en TC	199

6.2.3	Modelos predictivos basados en radiómica para Hemorragia Subaracnoidea Aneurismática en TC	200
6.2.4	Líneas futuras transversales	202
6.3	Impacto	204
6.3.1	Conferencias y congresos internacionales	204
6.3.2	Publicaciones en revistas	205
6.3.3	Otras publicaciones	206
6.3.4	Premios y reconocimientos	206
6.3.5	Trabajos cosupervisados	206
6.3.6	Proyectos de investigación	206
Referencias		207
Anexos		227
A	Configuración técnica y experimental del sistema HSI en glioblastoma	227
A.1	Especificaciones técnicas del sistema HSI intraoperatorio	227
A.2	Interfaz y visualización del sistema de anotación de las imágenes HSI intraoperatorias	228
A.3	Configuración de hiperparámetros de los modelos de clasificación tisular	229
B	Configuración y resultados complementarios de los modelos aplicados a Hemorragia Subaracnoidea Aneurismática en TC	230
B.1	Modelos de segmentación automática de hemorragia	230
	Interfaz y funcionalidades de segmentación en <i>QUIBIM Precision</i> [®]	230
	Hiperparámetros utilizados en el modelo de segmentación	231
B.2	Modelos basados en radiómica	232
	Espacios de búsqueda y configuraciones de hiperparámetros	236
	Comparativa de modelos SEBES/SEBES6c y radiómicos por desenlace	247
C	Métricas de evaluación	250
C.1	Métricas derivadas de la matriz de confusión	250
C.2	Métricas específicas de segmentación	252
C.3	Métricas estadísticas para regresión logística	253
D	Detalles de implementación	255

Lista de Figuras

1.1	Diagrama de bloques de los objetivos generales y específicos de la tesis doctoral	4
2.1	Estructura organizativa del capítulo de estado del arte, mostrando la progresión lógica desde el contexto clínico hasta la integración de inteligencia artificial en imagen médica.	8
2.2	Representación esquemática de la anatomía cerebral: a la izquierda, capas meníngeas, sustancia gris y blanca, así como la organización en lóbulos y principales regiones encefálicas (Imagen adaptada de Cleveland Clinic, 2024; a la derecha, el sistema ventricular y la circulación del LCR en relación con el espacio subaracnoideo y la columna vertebral (Imagen adaptada de Equipo de la Torre, 2021).	11
2.3	Representación comparativa de modalidades convencionales de neuroimagen. Arriba: TC y RM; abajo: RMA y PET/RM. Imagen adaptada de (Aryasinghe et al., 2016; San Diego Brain Injury Foundation, n.d. Waldron et al., 2018). . .	17
2.4	Ejemplos de biomarcadores en neuroimagen. En la fila de arriba se representan los biomarcadores anatómicos y de radiómica extraídos de una región de interés; en la fila de abajo biomarcadores funcionales y de conectividad cerebral. . . .	22
2.5	Diagrama del flujo de desarrollo de un biomarcador de imagen, desde la definición conceptual hasta su validación clínica, codificado por área disciplinar.	24
2.6	Extracción de la firma espectral de cada píxel de un cubo hiperespectral. . .	25
2.7	Clasificación de características radiómicas: morfológicas, de primer orden, de segundo orden y de orden superior.	27
2.8	Visualización del clasificador SVM lineal: hiperplanos, margen maximizado y vectores soporte	32
2.9	Árboles de decisión: comparación entre modelo único y técnicas de ensamblado (RF, ExtraTrees, XGBoost)	33
2.10	Funciones de activación comunes en redes neuronales artificiales	36
2.11	La red neuronal procesa la información a través de capas (entrada, ocultas y salida), calcula la predicción, evalúa el error con una función de pérdida y ajusta los pesos mediante retropropagación y tasa de aprendizaje. Este ciclo corresponde a una época de entrenamiento.	36
2.12	Operación de convolución: un filtro 3×3 se aplica localmente sobre la imagen para generar un mapa de activación.	39

2.13	Ejemplo de <i>max pooling</i> 2×2 con <i>stride</i> 2: en cada región local se conserva el valor máximo, reduciendo así la resolución del mapa de activación.	39
2.14	Esquema del flujo de datos a través de una CNN	40
2.15	Arquitectura U-Net: estructura simétrica con codificador-decodificador y conexiones puente	41
2.16	Ejemplo de resultado del proyecto NEMESIS-3D-CM. La primera imagen muestra una vista RGB del cerebro obtenida durante la cirugía. La segunda corresponde a la HSI adquirida intraoperatoriamente. Finalmente, se presenta el mapa de clasificación generado a partir de un modelo de ML entrenado para distinguir diferentes tipos de tejido cerebral.	49
2.17	Representación esquemática de los principales tipos de HIC según su localización anatómica en TC. A) HSA, B) HIP, C) HIV, D) HED, E) HSD	51
3.1	Estructura metodológica general del estudio, organizada en cuatro bloques principales (3.1–3.4) con tareas específicas para HSI en GBM y TC en HSAa.	64
3.2	Bloque 3.1 — Diseño del estudio y fuentes de datos.	65
3.3	Flujo de trabajo de la metodología: a partir de una TC, se segmentan la parénquima y las lesiones. Se extraen las características radiómicas y se combinan con los datos clínicos para desarrollar modelos de IA (Urbanos et al., 2025).	69
3.4	Interfaz general de la plataforma QUIBIM Precision® utilizada para la gestión de estudios. Se muestra el listado de casos almacenados con identificadores pseudonimizados y la estructura jerárquica de organización de los datos.	73
3.5	Bloque 3.2. Preparación y anotación de datos: generación del <i>ground truth</i> , preprocesado de imagen y variables clínicas.	76
3.6	Ejemplo de imagen real intraoperatoria con identificación numérica de las cinco clases tisulares utilizadas. (1) tejido sano, (2) tejido tumoral, (3) vaso venoso, (4) vaso arterial, (5) duramadre.	78
3.7	Segmentación del tejido cerebral utilizando CTseg en dos pacientes ejemplo. La columna derecha de cada paciente muestra la imagen original con la segmentación del cerebro. En la columna izquierda también se representa la segmentación de la hemorragia (Urbanos et al., 2025).	82
3.8	Bloque 3.3. Extracción de características: obtención de descriptores espectrales (HSI) y radiómicos (TC).	87
3.9	Flujo simplificado de radiómica: entrada, preprocesado, extracción y salida.	89
3.10	Bloque 3.4. Modelado y evaluación: desarrollo, validación e interpretación de modelos para clasificación, segmentación y predicción radiómica.	90
3.11	Esquema del procedimiento de validación cruzada anidada doble utilizado para el entrenamiento de los modelos. Por simplicidad, únicamente se ilustran los bucles externo e interno correspondientes a la combinación $K = 1$; el resto de combinaciones K_n siguen un procedimiento análogo (Urbanos et al., 2021).	92
3.12	Esquema de los dos diseños experimentales. En el Experimento A (clasificación intrapaciente), el modelo se entrena y evalúa dentro de cada paciente, estimando el rendimiento en condiciones óptimas. En el Experimento B (clasificación entre pacientes), se aplica un enfoque <i>leave-one-patient-out</i> para valorar la capacidad de generalización a nuevos individuos (Urbanos et al., 2021).	94

3.13	Esquema de la arquitectura 3D-CNN empleada para la clasificación hiperespectral de tejidos cerebrales.	95
3.14	Flujos de procesamiento y validación cruzada aplicados a la clasificación tisular con imágenes hiperespectrales. (a) División clásica del conjunto de datos en entrenamiento (80 %) y test externo (20 %), seguida de validación cruzada triple (5×5×5) en los clasificadores SVM, RF y CNN. (b) Estrategia <i>leave-one-out</i> , utilizando 12 imágenes para entrenamiento y 1 para test externo, con validación cruzada doble (5×5) en los mismos clasificadores Urbanos et al., 2021.	97
3.15	Esquema del flujo de validación cruzada de 5 <i>folds</i> para el entrenamiento y evaluación del modelo de segmentación automática, mostrando la TC craneal como entrada y la segmentación de hemorragia como salida.	98
3.16	Esquema de validación cruzada anidada	103
4.1	Imagen RGB sintética generada a partir de cubos hiperespectrales y mapas de <i>ground truth</i> con un color por clase para todos los pacientes incluidos.	111
4.2	Ejemplos de segmentaciones manuales en cada uno de los datasets usados en el estudio. En la primera fila, segmentaciones del H12O (en rojo HSA, en verde IVH); en la segunda fila el dataset HemSeg200 (IPH y IVH); en la tercera fila el dataset BHSD (en cyan HED; en rojo HSD; en verde IPH; en amarillo HSA; en azul marino HIV); en la última fila el dataset de Al Hilla (HIV, IPH, HSD, HED y HSA).	114
4.3	Esquema de división de los conjuntos de datos en los distintos experimentos, detallando las particiones aplicadas para las fases de entrenamiento, validación y evaluación del modelo.	115
4.4	Histogramas de las características técnicas y de anotación de la base de datos de imagen del H12O. (a) Distribución del tamaño de píxel (mm) (b) Distribución del grosor de corte (mm) (c) Distribución del número total de cortes por exploración (d) Distribución del número de cortes con segmentaciones manuales	116
4.5	Distribución del número de estudios según el tipo de hemorragia en los tres conjuntos analizados. Se muestran los casos disponibles en el conjunto H12O (A), en el Experimento 1 (B) y en el Experimento 2 (C).	117
4.6	Diagramas de flujo de pacientes según los diferentes desenlaces (resultado funcional, mortalidad, vasoespasmos e hidrocefalia). a) Conjunto de desarrollo del modelo (entrenamiento, validación y test) y b) Conjunto de test independiente	120
4.7	Distribución de los parámetros de adquisición en las cohortes utilizadas para los modelos predictivos de desenlaces clínicos. Histogramas del tamaño de píxel (izquierda) y grosor de corte (derecha) en a) Conjunto de desarrollo del modelo y b) Conjunto de test independiente.	121
4.8	Porcentaje de valores nulos en la base de datos de variables clínicas relevantes del H12O.	122
4.9	Firma media espectral por clase, representada de manera independiente para cada tipo de tejido, incluyendo el intervalo de variabilidad (área sombreada).	127
4.10	Comparación de la firma media espectral para las cinco clases en un mismo gráfico.	128

4.11	Gráfico de Bland–Altman global comparando las mediciones de volumen de hemorragia obtenidas con los algoritmos automáticos nuestro Experimento 2 y con SMCCH.	129
4.12	Error en % para SVM (azul), RF (rojo) y CNN (marrón) tras clasificar cada imagen de paciente en el experimento A.	132
4.13	ACC medio para todas las predicciones de pacientes y todos los tejidos para los algoritmos SVM, RF y CNN obtenidos en el experimento A.	133
4.14	Sens medio para todas las predicciones de pacientes y todos los tejidos para los algoritmos SVM, RF y CNN obtenidos en el experimento A.	134
4.15	Spec medio para todas las predicciones de pacientes y todos los tejidos para los algoritmos SVM, RF y CNN obtenidos en el experimento A.	135
4.16	RGB sintéticos, <i>ground truths</i> y resultados de mapas de clasificación de los pacientes ID33 e ID51 obtenidos en el experimento A tras ser clasificados por RF, SVM y CNN.	136
4.17	ACC medio para todas las predicciones de pacientes y todos los tejidos para los algoritmos SVM, RF y CNN obtenidos en el Experimento B.	137
4.18	Sens medio para todas las predicciones de pacientes y todos los tejidos para los algoritmos SVM, RF y CNN obtenidos en el experimento B.	138
4.19	RGB sintéticos, verdades fundamentales y mapas de clasificación de los pacientes ID33 e ID51 obtenidos en el experimento B tras ser clasificados por RF, SVM y CNN.	139
4.20	Gráfico de Bland–Altman comparando los volúmenes hemorrágicos obtenidos mediante segmentación automática frente a la segmentación manual en el conjunto de test. A la izquierda se muestran los resultados del Experimento 1 y a la derecha los del Experimento 2.	142
4.21	Diagrama de dispersión del Dice Score en función del volumen de la hemorragia en el conjunto de test. A) Conjunto completo (HemSeg+ H12O), B) Conjunto HemSeg, C) Conjunto H12O.	143
4.22	Segmentaciones obtenidas a partir del <i>ground truth</i> , modelos automáticos de la literatura y los modelos desarrollados en este trabajo: (A) casos favorables; (B) casos desfavorables.	144
4.23	Media e IC95 % de las puntuaciones de los neurocirujanos según cinco criterios: cobertura anatómica, sobresegmentación, infrasegmentación, consistencia intercorte y valoración global.	146
4.24	Distribución del resultado funcional en función del volumen de hemorragia. A) Volumen obtenido con segmentación manual, B) con segmentación automática	149
4.25	Comparación del rendimiento entre modelos clínicos y modelos radiómicos derivados de distintas ROIs para los diferentes desenlaces: azul para los modelos clínicos, verde para radiómica del parénquima, naranja para radiómica del sangrado obtenida con SMCCH y marrón para radiómica del sangrado correspondiente al Experimento 2.	151

4.26	Comparación del rendimiento de modelos radiómicos y radiómico-clínicos en distintas ROIs: verde (radiómica parénquima), morado (radiómica+clínica parénquima), naranja (radiómica sangrado SMCCH), rojo (radiómica+clínica sangrado SMCCH), marrón (radiómica sangrado Exp. 2) y cyan (radiómica+clínica sangrado Exp. 2).	152
4.27	Rendimiento de los modelos en el conjunto de test externo basados en clínica + radiómica usando las distintas ROIs. En azul, radiómica extraída del parénquima, en naranja del sangrado con SMCCH y en verde del sangrado con el Experimento 2.	154
4.28 (a)	Valores SHAP en modelo clínico	156
4.28 (b)	Valores SHAP en modelo radiómico	156
4.28 (c)	Valores SHAP en modelo radiómico + clínico	157
4.28	Gráficos de valores SHAP de los modelos entrenados sobre el parénquima: A) clínico, B) radiómico y C) combinado clínica + radiómica.	157
4.29	Visualización de la predicción de SEBES mediante LASSO incremental. Izquierda: curva del rendimiento del modelo en función del número de características radiómicas incluidas. Derecha: gráfico de barras con las métricas de clasificación (Sens, Spec, F1-score y balanced accuracy) diferenciadas para SEBES 0–2 y SEBES 3–4.	160
A.1	Interfaz de la herramienta de etiquetado en HSI, mostrando los tres paneles principales: imagen RGB de referencia (izquierda), máscara de selección por similitud espectral (centro) y acumulación de etiquetas (<i>ground truth</i>) (derecha).	228
B.1	Capturas del visor <i>QUIBIM Precision</i> [®] que muestran ejemplos de segmentaciones en dos pacientes distintos. A la izquierda se representan las regiones correspondientes a HSAa (amarillo) y HIV (azul); a la derecha, la HSAa (amarillo) y HIP (verde).	230
B.2	Rendimiento de los modelos combinando clínica y radiómica a partir de distintas regiones, estratificado por grupos de edad (<70 y ≥70 años).	238
B.3	Curvas de calibración de los modelos en los distintos folds de la validación cruzada por experimento: usando variables clínicas, radiómicas o su combinación. La radiómica fue extraída del parénquima cerebral.	240
B.4	Matrices de confusión normalizadas para los modelos que combinan radiómica y clínica por cada desenlace usando la radiómica extraída de la parénquima, hemorragia SMCHH y hemorragia del Experimento 2.	243
B.5	Gráficas SHAP obtenidas por cada desenlace clínico de los modelos que solo usan variables clínicas, radiómica de la parénquima o combinados (radiómica + clínica).	244
B.6	Frecuencia de selección de variables en los cinco folds de validación cruzada en los modelos de GOS funcional que usan variables clínicas, radiómicas y combinadas.	245

Lista de Tablas

2.1	Comparación entre modalidades de imagen médica y apoyo visual en neurociencia clínica	20
2.3	Resumen de estudios sobre el uso de imagen hiperespectral (HSI) en neurocirugía.	48
2.5	Resumen de las principales escalas clínicas y radiológicas empleadas en la HSAa.	56
2.6	Modelos recientes de segmentación automática de hemorragias intracraneales.	58
2.8	Estudios recientes que emplean radiómica del sangrado cerebral para predecir desenlaces clínicos.	61
3.1	Resumen de desenlaces utilizados en el estudio, con su codificación binaria, fuente de datos y criterios temporales de evaluación.	80
4.1	Conjunto de datos con el número total de píxeles etiquetados y su porcentaje por clase y paciente.	112
4.2	Distribución de casos con segmentaciones manuales por fuente de datos y tipo de hemorragia.	113
4.3	Características principales de los conjuntos de las bases de datos utilizadas para la segmentación del sangrado.	116
4.4	Distribución de las principales variables clínicas asociadas a los pacientes del conjunto H12O (n=223) para el desarrollo del algoritmo de segmentación de hemorragia.	118
4.5	Variables clínicas incluidas en el conjunto final tras aplicar el criterio de exclusión por datos faltantes.	123
4.7	Características clínicas y comparación entre cohortes.	124
4.9	Características radiómicas más correlacionadas con cada desenlace en las diferentes cohortes (coeficiente de correlación de Spearman y p-valor).	130
4.10	Resultados de OACC para los algoritmos SVM, RF y CNN en el Experimento A por paciente. Se incluyen la media y la DE.	133
4.11	Resultados de OACC para los algoritmos SVM, RF y CNN tras clasificar cada imagen de paciente en el Experimento B.	137
4.12	Métricas de rendimiento en el conjunto de test (N=89) utilizando cinco algoritmos de la bibliografía y dos desarrollados en este trabajo, con respecto a la segmentación manual como referencia. Las métricas se presentan como la mediana y su IQR (Q1-Q3). En negrita se subrayan los valores medianos mas altos por métrica respecto a los modelos evaluados.	141

4.13	Matriz de comparación por pares según solapamiento de IC95 % de la mediana. “Equivalente” indica IC solapados (no diferencia concluyente); “Diferente” indica IC no solapados (diferencia relevante).	142
4.14	Ranking de los modelos de segmentación de hemorragia en función de la suma de puntuaciones en los cinco criterios de evaluación (la posición en el ranking se indica entre paréntesis).	146
4.16	Comparación de la eficiencia temporal entre segmentación manual y automática	147
4.17	Resultados de la regresión logística uni- y multivariable para la predicción de resultado funcional favorable (N=223).	148
4.19	Rendimiento predictivo de la ictWFNS y de los modelos radiómicos cerebrales en la cohorte de test (n=41).	158
4.20	Resumen comparativo del rendimiento predictivo (AUC y BalAcc con IC 95 %) de modelos basados en SEBES/SEBES6c y radiómica del parénquima para los cuatro desenlaces evaluados. Los mejores valores para cada desenlace aparecen en negrita.	159
4.21	Top 5 características radiómicas correlacionadas con SEBES y SEBES6c. Nota: FO = first order; GLCM = gray-level co-occurrence matrix; GLRLM = gray-level run length matrix; σ = log-sigma.	160
4.22	Variables radiómicas comunes identificadas en los modelos de edema (SEBES) y los basados en radiómica para predecir desenlaces clínicos.	161
4.23	Rendimiento predictivo (AUC) del modelo de regresión logística entrenado con las características radiómicas seleccionadas para reproducir SEBES (“SEBES radiómico”) al aplicarse a los desenlaces clínicos.	161
5.1	Comparación entre los resultados de OACC de enfoques llevados a cabo en el estado del arte y los propuestos en este estudio.	167
5.2	Rendimiento comparativo de los modelos radiómicos frente a escalas clínicas tradicionales en la predicción de resultados en HSAa	182
A.1	Estructura de archivos y contenido asociado a cada adquisición hiperespectral.	227
A.2	Resumen de especificaciones del sistema HSI intraoperatorio.	228
A.3	Configuración de los hiperparámetros empleados para cada algoritmo supervisado.	229
B.1	Configuración detallada de los parámetros arquitectónicos y de entrenamiento utilizados en los Experimentos 1 y 2 con <i>nnU-Net</i>	231
B.2	Resumen de variables clínicas categóricas y continuas incluidas en el estudio, con sus valores o descripción correspondiente.	232
B.3	Valores fisiológicos de referencia de las variables clínicas incluidas.	233
B.4	Definiciones y frecuencias de las comorbilidades en la base de datos clínica.	233
B.5	Cumplimiento de la guía TRIPOD para la notificación transparente del desarrollo y evaluación de modelos predictivos.	236
B.5	Espacio de búsqueda de hiperparámetros para RF.	236
B.6	Espacio de búsqueda de hiperparámetros para <i>Extra Trees</i>	237
B.6	Espacio de búsqueda de hiperparámetros para <i>XGBoost</i>	237
B.7	Resultados de predicción por grupos clínicos en los distintos modelos y desenlaces.	239
B.8	Modelos seleccionados para Mortalidad (HSAa, Experimento 2).	241

B.9 Modelos seleccionados para GOS funcional (HSAa, Experimento 2).	241
B.10 Modelos seleccionados para Hidrocefalia (HSAa, Experimento 2).	241
B.11 Modelos seleccionados para Vasoespasmo (HSAa, Experimento 2).	241
B.12 Modelos seleccionados para Mortalidad (Parénquima).	241
B.13 Modelos seleccionados para GOS funcional (Parénquima).	242
B.14 Modelos seleccionados para Hidrocefalia (Parénquima).	242
B.15 Modelos seleccionados para Vasoespasmo (Parénquima).	242
B.16 Estadísticos descriptivos de las escalas radiológicas empleadas para calcular el ictWFNS en el conjunto de test ($n = 41$).	246
B.17 Distribución del SEBES en el conjunto de test ($n = 41$).	246
B.18 Métricas comparativas de rendimiento de los modelos predictivos para mortali- dad: aproximaciones basadas en SEBES frente a radiómica.	247
B.19 Métricas comparativas de rendimiento de los modelos predictivos para mal resultado funcional: aproximaciones basadas en SEBES frente a radiómica.	247
B.20 Métricas comparativas de rendimiento de los modelos predictivos para vaso- espasmo: aproximaciones basadas en SEBES frente a radiómica.	248
B.21 Métricas comparativas de rendimiento de los modelos predictivos para hidroce- falia: aproximaciones basadas en SEBES frente a radiómica.	248
B.22 Coeficientes obtenidos con regresión logística LASSO para las principales características radiómicas seleccionadas para predecir SEBES. Se muestran el nombre de la característica y el coeficiente estimado (<i>coef</i>).	249

Abreviaturas y acrónimos

3DCNN Red Neuronal Convolutacional 3D, por sus siglas en inglés *3D Convolutional Neural Network*

5-ALA Ácido 5-aminolevulínico

Acc Exactitud, por sus siglas en inglés *Accuracy*

ADASYN *Adaptive Synthetic Sampling* (acrónimo en inglés)

ASD Distancia media de superficie, por sus siglas en inglés *Average Surface Distance*

AUC Área bajo la curva, por sus siglas en inglés *Area Under the Curve*

BalAcc Exactitud balanceada, por sus siglas en inglés *Balanced Accuracy*

CEIm Comité de Ética de la Investigación con medicamentos

CNNs Redes Neuronales Convolucionales, por sus siglas en inglés *Convolutional Neural Networks*

CTA Angiografía por Tomografía Computarizada, por sus siglas en inglés *Computed Tomography Angiography*

DICOM *Digital Imaging and Communications in Medicine* (acrónimo en inglés)

DL Aprendizaje profundo, por sus siglas en inglés *Deep Learning*

DSA Angiografía digital por sustracción, por sus siglas en inglés *Digital Subtraction Angiography*

DSC Coeficiente de Dice, por sus siglas en inglés *Dice Similarity Coefficient*

ExtraRF *Extremely Randomized Trees* (acrónimo en inglés)

fMRI Resonancia Magnética Funcional, por sus siglas en inglés *functional Magnetic Resonance Imaging*

FN Falsos Negativos, por sus siglas en inglés *False Negatives*

FP Falsos Positivos, por sus siglas en inglés *False Positives*

GCS Glasgow Coma Scale

GBM Glioblastoma

GLCM *Gray-Level Co-occurrence Matrix* (acrónimo en inglés)

GLRLM *Gray-Level Run Length Matrix* (acrónimo en inglés)

GLSZM *Gray-Level Size Zone Matrix* (acrónimo en inglés)

GOS Glasgow Outcome Scale

H12O Hospital Universitario 12 de Octubre

HCE Historia Clínica Electrónica

HD Distancia de Hausdorff, por sus siglas en inglés *Hausdorff Distance*

HED Hemorragia Epidural

HIC Hemorragia Intracraneal

HIP Hemorragia Intraparenquimatosa

HIV Hemorragia Intraventricular

HSA Hemorragia Subaracnoidea

HSAa Hemorragia Subaracnoidea Aneurismática

HSD Hemorragia Subdural

HSI Imagen Hiperespectral, por sus siglas en inglés *Hyperspectral Imaging*

HU Unidades Hounsfield, por sus siglas en inglés *Hounsfield Units*

IA Inteligencia Artificial

ICC Coeficiente de correlación intraclase, por sus siglas en inglés *Intraclass Correlation Coefficient*

ioUS Ecografía intraoperatoria, por sus siglas en inglés *intraoperative Ultrasound*

LCR Líquido cefalorraquídeo

MAD Desviación Mediana Absoluta, por sus siglas en inglés *Median Absolute Deviation*

MICE *Multiple Imputation by Chained Equations* (acrónimo en inglés)

ML Aprendizaje automático, por sus siglas en inglés *Machine Learning*

MRA Angiografía por Resonancia Magnética, por sus siglas en inglés *Magnetic Resonance Angiography*

MSE Error cuadrático medio, por sus siglas en inglés *Mean Squared Error*

NEMESIS clasificación intraoperatoria de tumores cerebrales mediante modelos InnerSivos

NIfTI *Neuroimaging Informatics Technology Initiative* (acrónimo en inglés)

NGTDM *Neighbourhood Gray-Tone Difference Matrix* (acrónimo en inglés)

NIR Infrarrojo cercano, por sus siglas en inglés *Near InfraRed*

nnU-Net *no-new U-Net* (acrónimo en inglés)

OACC Precisión global, por sus siglas en inglés *Overall Accuracy*

PET Tomografía por emisión de positrones, por sus siglas en inglés *Positron Emission Tomography*

ReLU *Rectified Linear Unit* (acrónimo en inglés)

RF Bosques aleatorios, por sus siglas en inglés *Random Forest*

RGPD Reglamento General de Protección de Datos

RM Resonancia Magnética

RM iO Resonancia Magnética Intraoperatoria

RSNA Radiological Society of North America

SAM *Spectral Angle Mapper* (acrónimo en inglés)

SEBES *Score for the Early Brain Edema Evaluation Scale* (acrónimo en inglés)

Sens Sensibilidad

SGD *Stochastic Gradient Descent* (acrónimo en inglés)

SHAP *SHapley Additive exPlanations* (acrónimo en inglés)

SNC Sistema Nervioso Central

Spec Especificidad

SVM Máquinas de Vectores Soporte, por sus siglas en inglés *Support Vector Machines*

Tanh Tangente hiperbólica

TC Tomografía Computarizada

TCi Tomografía Computarizada Intraoperatoria

TCE Traumatismo craneoencefálico

TP Verdaderos Positivos, por sus siglas en inglés *True Positives*

TN Verdaderos Negativos, por sus siglas en inglés *True Negatives*

- TPR** Tasa de Verdaderos Positivos, por sus siglas en inglés *True Positive Rate*
- TRIPOD** *Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis* (acrónimo en inglés)
- UPM** Universidad Politécnica de Madrid
- UV** Ultravioleta
- viT** *Vision Transformer* (acrónimo en inglés)
- VNIR** Rango visible e infrarrojo cercano, por sus siglas en inglés *Visible and Near InfraRed*
- WFNS** *World Federation of Neurosurgical Societies* (acrónimo en inglés)
- XGBoost** *Extreme Gradient Boosting* (acrónimo en inglés)
- MAV** Malformación arteriovenosa
- IQR** Rango intercuartílico, por sus siglas en inglés *Interquartile Range*
- DE** Desviación Estándar
- OR** Razón de probabilidades, por sus siglas en inglés *Odds Ratios*

Capítulo 1

Introducción

1.1 Motivación

En el ámbito de la neurociencia clínica, la caracterización precisa del sistema nervioso central (SNC) es fundamental para comprender mejor su funcionamiento, detectar alteraciones funcionales o estructurales y apoyar la toma de decisiones clínicas. Ante alteraciones del SNC, como tumores, accidentes cerebrovasculares o lesiones traumáticas, una evaluación detallada de la anatomía y la función cerebral resulta crucial no solo para delimitar con precisión la extensión del daño, sino también para guiar intervenciones clínicas más seguras y eficaces, adaptadas a las particularidades de cada paciente.

En la práctica clínica, las herramientas de imagen médica avanzadas, como la tomografía computarizada (TC) y la resonancia magnética (RM), constituyen pilares esenciales para el diagnóstico de lesiones cerebrales. Estas modalidades permiten visualizar las estructuras anatómicas y detectar alteraciones morfológicas sutiles. No obstante, el diagnóstico sigue dependiendo en gran medida de la interpretación subjetiva de los profesionales de la salud, lo que introduce una fuente de variabilidad interobservador como intraobservador. Esta dependencia del criterio humano no solo puede generar sesgos diagnósticos, sino que también implica una elevada carga de trabajo para los clínicos, con posibles repercusiones sobre la precisión diagnóstica, la planificación terapéutica y, en última instancia, sobre la evolución clínica del paciente.

En este contexto, una de las tareas más demandantes es la segmentación de regiones de interés en imagen médica. La segmentación manual, realizada por expertos en neuroimagen, permite delinear con precisión estructuras patológicas como tumores, hematomas o áreas isquémicas. A partir de estas segmentaciones es posible extraer información cuantitativa como el volumen, la forma, la localización espacial de las lesiones, y su relación con estructuras cerebrales funcionalmente relevantes, información útil para el diagnóstico, la planificación quirúrgica y el seguimiento longitudinal del paciente. Esta aproximación permite también monitorizar la progresión de la enfermedad y evaluar la respuesta al tratamiento en estudios de seguimiento.

Sin embargo, la segmentación manual es un proceso intensivo en tiempo y recursos, que exige la intervención directa de clínicos con elevada experiencia. Dado que la mayoría de las

imágenes adquiridas son tridimensionales, con decenas o incluso cientos de cortes, este trabajo puede suponer una carga significativa para los clínicos, siendo un trabajo tedioso que limita su disponibilidad para otras tareas clínicas de igual o mayor relevancia. Además, la variabilidad en la interpretación de los contornos, especialmente en casos donde las fronteras entre tejidos son poco definidas o difusas, puede comprometer la reproducibilidad de los análisis y generar sesgos en los estudios clínicos y en la planificación quirúrgica. Por tanto, existe una necesidad de automatizar este proceso mediante herramientas objetivas y reproducibles que reduzcan la carga de trabajo y aumenten la fiabilidad de los resultados.

En el contexto intraoperatorio, estos desafíos se acentúan. Durante la cirugía cerebral, el neurocirujano debe tomar decisiones rápidas y precisas para extirpar tejido patológico preservando al máximo el tejido sano y funcional. Sin embargo, fenómenos intraoperatorios como el desplazamiento cerebral (*brain shift*), causado por la liberación de presión intracraneal, la pérdida de líquido cefalorraquídeo o el efecto de agentes anestésicos, pueden alterar significativamente la posición de las estructuras cerebrales respecto a las imágenes preoperatorias.

En este escenario, la imagen hiperespectral (HSI, por sus siglas en inglés *Hyperspectral Imaging*) ha emergido como una técnica prometedora. La HSI proporciona información espectral de los tejidos al adquirir imágenes en múltiples bandas del espectro electromagnético, lo que permite identificar distintos tipos de tejidos cerebrales, principalmente diferenciar el tejido sano del tejido tumoral. Esta capacidad resulta especialmente relevante en patologías como el glioblastoma (GBM), uno de los tumores cerebrales más agresivos, donde una resección precisa y lo menos invasiva posible puede tener un impacto significativo en la evolución del paciente. A diferencia de otras modalidades de imagen, la HSI puede integrarse en el campo quirúrgico, ofreciendo información en tiempo real y posibilitando la generación de mapas funcionales que podrían guiar la resección con mayor precisión y, en consecuencia, reducir el riesgo de daño neurológico, un riesgo presente de forma inherente en la neurocirugía convencional .

Para abordar estos desafíos, en los últimos años se ha avanzado significativamente en el desarrollo de modelos de inteligencia artificial (IA) aplicados a la imagen médica. Combinando algoritmos de aprendizaje automático (*Machine Learning*, ML, por sus siglas en inglés) y de aprendizaje profundo (*Deep Learning*, DL, por sus siglas en inglés) con técnicas de procesamiento de imágenes, se han diseñado soluciones capaces de analizar grandes volúmenes de datos clínicos de forma automatizada, precisa y en tiempos reducidos. Estas herramientas han demostrado ser útiles en distintas tareas, entre las que se incluyen:

- Segmentación automática de lesiones y estructuras anatómicas en imágenes médicas, disminuyendo la variabilidad entre observadores y mejorando la objetividad del análisis.
- Reducción de los tiempos de diagnóstico e intervención, aumentando la eficiencia operativa y permitiendo una respuesta más rápida en contextos clínicos relacionados con la neurociencia.
- Predicción del pronóstico clínico a partir de datos multimodales, con el objetivo de anticipar la evolución de la enfermedad y apoyar la toma de decisiones terapéuticas individualizadas.

La incorporación progresiva de estas tecnologías basadas en IA en la práctica clínica tiene el

potencial de transformar de manera sustancial el abordaje diagnóstico y terapéutico de las enfermedades neurológicas. Su aplicación en el entorno hospitalario no solo tiene el potencial de mejorar la precisión diagnóstica y la planificación quirúrgica, sino también contribuir a una medicina más personalizada, eficiente y centrada en el paciente. A medida que estas herramientas se integren en los flujos de trabajo clínicos, podrían facilitar el apoyo de decisiones en tiempo real, reducir la carga cognitiva de los profesionales sanitarios y, potencialmente, mejorar los resultados clínicos y la calidad asistencial en el ámbito de la neurociencia clínica.

En esta tesis doctoral se propone el desarrollo de sistemas automáticos que combinan técnicas de imagen médica avanzada, como la TC y HSI, junto con algoritmos de IA especializados en tareas de clasificación y segmentación. Estos sistemas buscan optimizar la identificación de los tejidos afectados, contribuyendo a la mejora de los tiempos de intervención y a la reducción de errores en entornos clínicos. Además, se evaluará su aplicabilidad en contextos reales, analizando su impacto en la práctica médica y su potencial para apoyar una medicina más personalizada, eficiente y basada en datos.

1.2 Objetivos

El objetivo principal de esta tesis doctoral [OBJ-P] es:

Objetivo principal – [OBJ-P]

Desarrollar y validar un sistema automatizado basado en técnicas avanzadas de procesamiento de imágenes médicas y algoritmos de ML, con el fin de mejorar la precisión diagnóstica, predecir la evolución clínica y apoyar la toma de decisiones en el ámbito de la neurociencia clínica.

Con este fin, el objetivo principal se desglosa en los siguientes objetivos específicos [OBJ-S]:

[OBJ-S1]

Diseñar e implementar modelos de clasificación y segmentación automática para la detección precisa de lesiones cerebrales en imágenes médicas de pacientes en el contexto de la neurociencia clínica, con un rendimiento superior a los métodos manuales y a las técnicas previas descritas en la literatura. Este objetivo se aplica a dos casos de uso:

[OBJ-S1.1]

Desarrollar un sistema de clasificación intraoperatoria en tiempo real basado en HSI para la diferenciación entre tejido sano y patológico en cirugías de GBM.

[OBJ-S1.2]

Implementar modelos de segmentación profunda para la segmentación automática de hemorragias cerebrales en TC, con el objetivo de optimizar el flujo de trabajo clínico mediante la reducción de tiempos de análisis y la obtención de segmentaciones precisas.

[OBJ-S2]

Explorar e integrar biomarcadores derivados de imagen médica avanzada en modelos de apoyo diagnóstico y pronóstico en enfermedades neurológicas. Este objetivo se aplica a dos casos de uso:

[OBJ-S2.1]

Incorporar firmas espectrales extraídas de HSI en los modelos de segmentación para la clasificación tisular intraoperatoria en GBM (OBJ-S1.1).

[OBJ-S2.2]

Integrar características radiómicas obtenidas de segmentaciones en TC (OBJ-S1.2) como entrada a modelos predictivos para estimar la progresión de la enfermedad en pacientes con hemorragia subaracnoidea aneurismática (HSAa).

[OBJ-S3]

Validar el rendimiento y utilidad clínica de los sistemas desarrollados, entendidos como la integración de los modelos de análisis de imagen (OBJ-S1) y los modelos predictivos (OBJ-S2). La validación incluirá comparación con anotaciones expertas, evaluación cuantitativa, cualitativa y análisis de su impacto potencial en el ámbito de la neurociencia clínica.

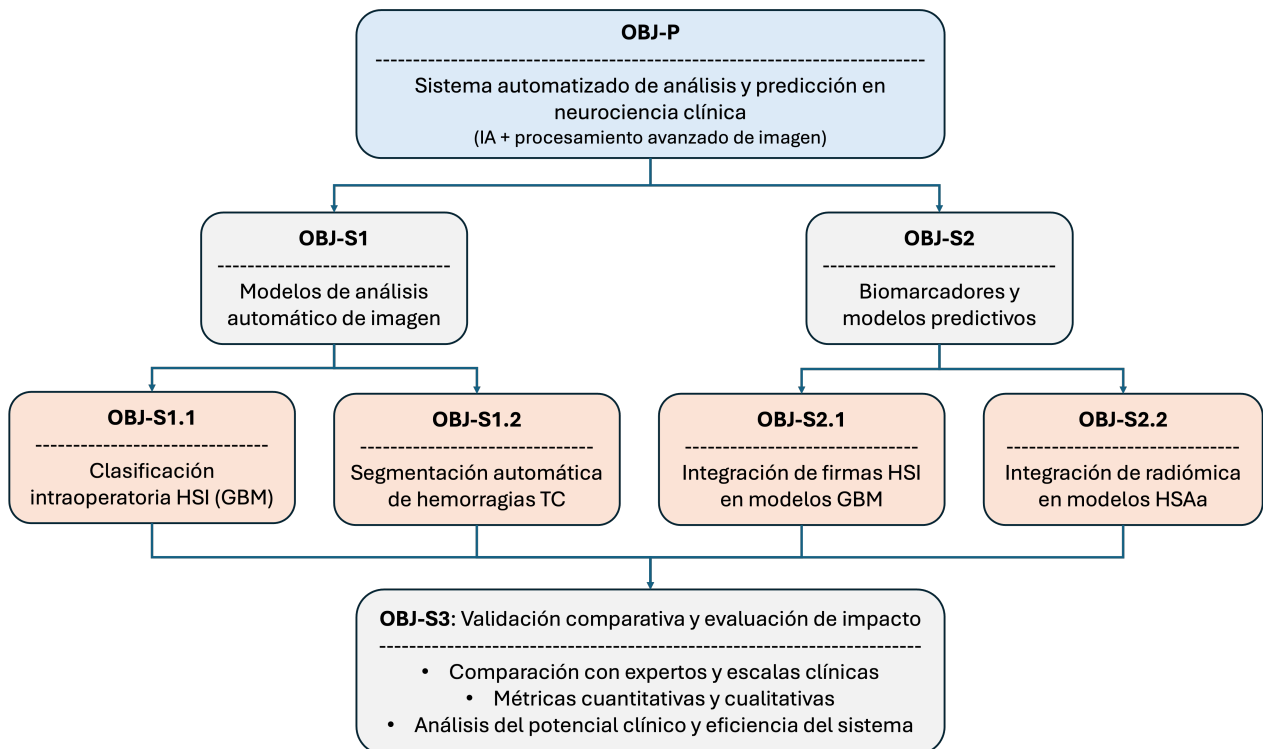


Figura 1.1: Diagrama de bloques de los objetivos generales y específicos de la tesis doctoral

1.3 Estructura del documento

El presente documento se organiza en seis capítulos principales, complementados con anexos, con el propósito de guiar al lector de forma progresiva desde la motivación inicial y el estado del arte, hasta los resultados obtenidos, su discusión crítica y las conclusiones finales.

En el **Capítulo 1 (Introducción)** se exponen la motivación de la investigación, los objetivos generales y específicos de la tesis doctoral, así como una descripción de la organización del manuscrito. Se contextualiza el problema clínico en el ámbito de la neurociencia y se justifica la relevancia de explorar estrategias de caracterización automática del tejido cerebral y de las lesiones mediante técnicas avanzadas de imagen y aprendizaje automático.

El **Capítulo 2 (Estado del arte)** presenta una revisión detallada de la literatura en torno a los aspectos fundamentales de la tesis. En primer lugar, se describen los retos clínicos actuales en el diagnóstico y tratamiento de patologías neurológicas, así como los fundamentos anatómicos necesarios para comprender el problema abordado. A continuación, se revisan las principales modalidades de imagen médica convencionales e intraoperatorias, y se introducen los biomarcadores de imagen y sus aplicaciones en neurociencia clínica. De manera específica, se profundiza en los conceptos de HSI, la radiómica y la IA aplicada a imagen médica, cubriendo tanto el aprendizaje automático clásico como el DL y las arquitecturas basadas en U-Net. Finalmente, se incluyen dos casos de estudio representativos de la literatura: la clasificación automática de tejidos intraoperatorios y la segmentación y predicción de desenlaces en HSAa.

El **Capítulo 3 (Materiales y métodos)** describe la estrategia metodológica empleada. Se detallan las cohortes de estudio utilizadas, los criterios de inclusión y exclusión, y los procedimientos éticos y de anonimización. Se explican las fases de adquisición de datos, tanto hiperespectrales en GBM como clínicos y radiológicos en HSAa, así como la descripción de las bases de datos y sus características principales. Posteriormente, se presentan los procedimientos de preparación y anotación, incluyendo la construcción de los *ground truth* para segmentación y clasificación. El capítulo también aborda las técnicas de extracción de características, diferenciando entre firmas espectrales y radiómicas, así como los modelos de evaluación propuestos en cada ámbito. Se incluyen asimismo las métricas de evaluación empleadas y los detalles de implementación.

El **Capítulo 4 (Resultados)** recoge de manera ordenada los hallazgos derivados de los experimentos realizados. Primero, se caracterizan las bases de datos empleadas, seguidas de la extracción de características específicas. Posteriormente, se presentan los resultados de los experimentos de clasificación tisular en GBM mediante HSI, la segmentación automática de HSAa en TC y los modelos predictivos radiómicos desarrollados. Para cada experimento, se incluyen resultados cuantitativos y cualitativos, comparaciones con escalas clínicas, análisis de eficiencia temporal e interpretabilidad de los modelos.

El **Capítulo 5 (Discusión)** constituye un análisis crítico de los resultados. Se evalúa el rendimiento técnico en comparación con la literatura, se discuten las implicaciones clínicas potenciales de los hallazgos y se reconocen las limitaciones de los experimentos desarrollados. Además, se ofrece una visión transversal sobre la caracterización automática, integrando las tres áreas de estudio.

El **Capítulo 6 (Conclusiones y líneas futuras)** sintetiza los principales logros de la tesis en relación con los objetivos iniciales y expone de forma organizada las direcciones futuras de investigación. Se distinguen líneas específicas para cada ámbito (HSI en neurocirugía, segmentación automática de HSAa en TC y modelos predictivos radiómicos), así como líneas transversales futuras de este trabajo. El capítulo termina con una sección dedicada al impacto académico y científico de la tesis, donde se recogen publicaciones, comunicaciones en congresos, trabajos co-supervisados y contribuciones a proyectos de investigación.

Finalmente, se incluyen los Anexos, que recogen información complementaria de carácter técnico y metodológico. En concreto, el Anexo A describe la configuración técnica y experimental del sistema HSI empleado en glioblastoma; el Anexo B detalla la configuración experimental y los resultados complementarios de los modelos aplicados a la HSAa; el Anexo C presenta las métricas de evaluación empleadas; y el Anexo D resume los aspectos relacionados con la implementación y ejecución de los experimentos.

Capítulo 2

Estado del arte

2.1 Organización y estructura del estado del arte

La revisión del estado del arte presentada en este capítulo sigue una estructura progresiva que refleja el camino lógico desde la comprensión del problema clínico hasta la aplicación de soluciones tecnológicas avanzadas. Esta organización responde directamente a la motivación central de esta tesis doctoral (Sección 1.1): desarrollar sistemas automatizados que mejoren la precisión diagnóstica y apoyen la toma de decisiones en neurociencia clínica mediante la integración de técnicas avanzadas de imagen médica e IA.

Como se muestra en la Figura 2.1, el capítulo se estructura en cuatro bloques fundamentales que guían al lector desde los aspectos más básicos hasta las aplicaciones más avanzadas:

1. **Contexto clínico:** Se inicia con una descripción de los retos actuales en el diagnóstico y tratamiento de patologías neurológicas, así como los fundamentos anatómicos necesarios para comprender el ámbito de aplicación. Esta base contextual establece la necesidad de herramientas objetivas y cuantitativas que superen las limitaciones de la evaluación subjetiva y manual.
2. **Base diagnóstica y de imagen:** Se revisan las principales modalidades de imagen médica convencionales e intraoperatorias disponibles para el diagnóstico en neurociencia clínica. Este bloque conecta las necesidades clínicas identificadas con las herramientas tecnológicas actuales, mostrando tanto sus capacidades como sus limitaciones.
3. **Extracción de información cuantitativa:** Se introduce el concepto de biomarcadores de imagen y se profundiza en dos técnicas fundamentales para esta tesis: la HSI y la radiómica. Estas metodologías permiten extraer información cuantitativa y reproducible que trasciende la evaluación visual tradicional.
4. **Integración de la IA en imagen médica:** Se presenta el marco teórico y práctico de la IA aplicada a imagen médica, incluyendo algoritmos de ML y DL, así como sus aplicaciones específicas en neurociencia clínica. Este bloque muestra cómo la IA puede transformar los biomarcadores de imagen en herramientas de apoyo a la decisión clínica.

Esta estructura progresiva responde a la lógica inherente al problema abordado: partir del contexto clínico y los desafíos reales (Sección 1.1), comprender las herramientas disponibles, explorar formas de extraer información cuantitativa de estas herramientas, y finalmente aplicar técnicas avanzadas de IA para transformar esa información en soluciones prácticas. Esta organización facilita la comprensión de cómo cada componente contribuye al objetivo final de desarrollar sistemas automatizados que mejoren la práctica clínica en neurociencia.

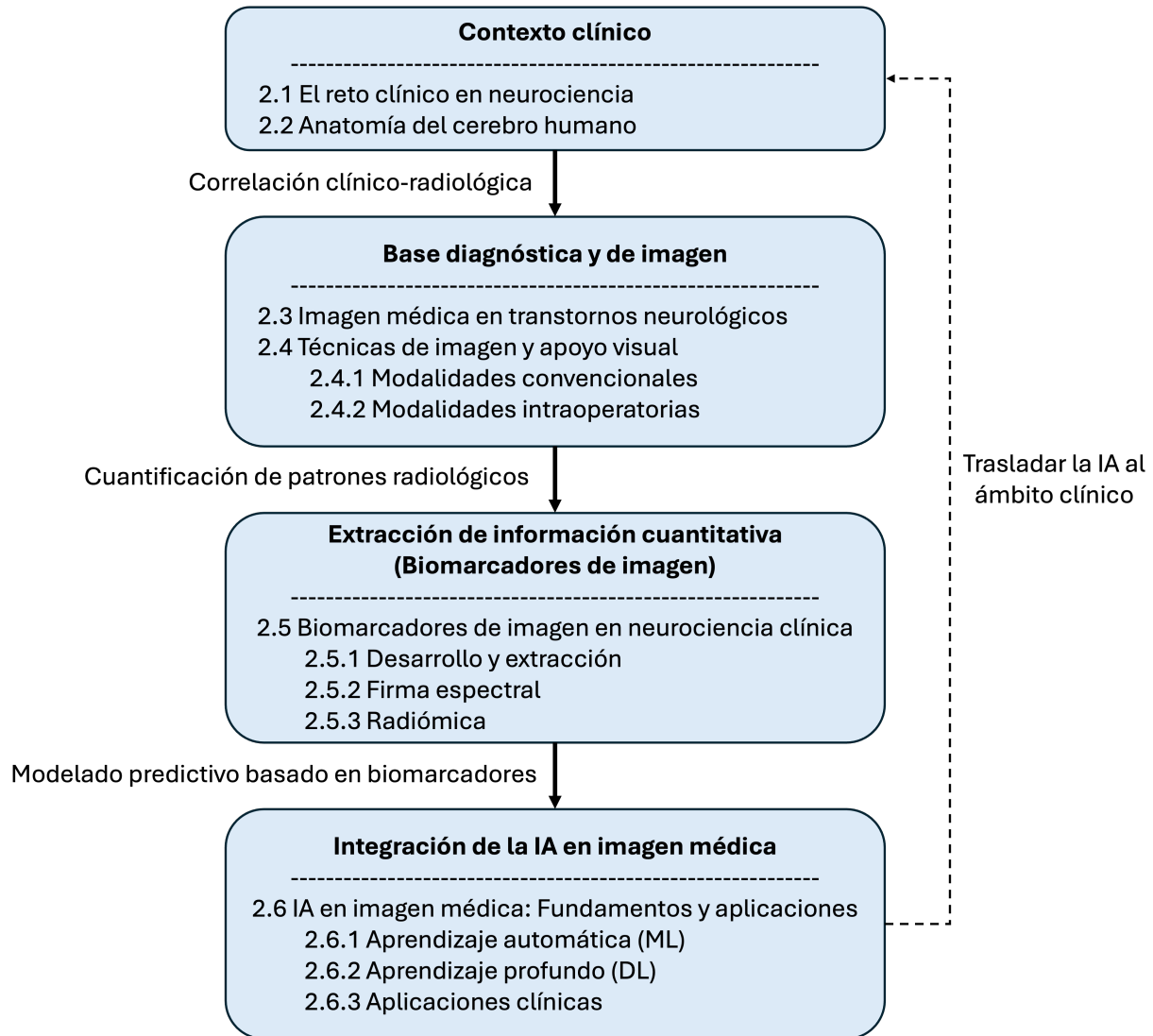


Figura 2.1: Estructura organizativa del capítulo de estado del arte, mostrando la progresión lógica desde el contexto clínico hasta la integración de inteligencia artificial en imagen médica.

2.2 El reto clínico: diagnóstico y tratamiento en neurociencia

El sistema nervioso central (SNC) constituye una de las regiones más complejas del cuerpo humano tanto desde el punto de vista funcional como anatómico. El SNC está constituido

por millones de neuronas interconectadas en redes dinámicas que regulan funciones motoras, sensoriales, cognitivas y emocionales. Esta complejidad se refleja en la práctica clínica, donde los síntomas que presentan los pacientes suelen ser inespecíficos y pueden observarse en distintas condiciones neurológicas. Manifestaciones como cefalea, mareo, alteraciones visuales, trastornos cognitivos o de conducta, pueden darse tanto en cuadros benignos como en patologías potencialmente graves, lo que complica el diagnóstico clínico inicial.

Además, la ausencia de biomarcadores específicos y validados para muchas enfermedades neurológicas, así como la subjetividad de la evaluación de síntomas complejos como el dolor o alteraciones cognitivas, limita la objetividad y reproducibilidad de los diagnósticos. Esto refleja la necesidad de enfoques multidisciplinarios que engloben aspectos neurobiológicos, clínicos y tecnológicos.

Por otro lado, la prevención y la personalización del tratamiento siguen siendo retos importantes. La evolución de las enfermedades neurológicas, los diferentes desenlaces clínicos y la respuesta a tratamientos varían considerablemente entre pacientes debido a factores genéticos, moleculares y anatómicos. Por ello, se requieren estrategias que integren múltiples fuentes de información para una planificación terapéutica personalizada, que maximicen la eficacia clínica y minimicen los efectos adversos.

El diagnóstico depende en gran medida de herramientas que permitan reducir la incertidumbre y proporcionar información objetiva y cuantificable. En este contexto, la imagen médica es una de las técnicas más determinantes, ya que ofrece una representación detallada y precisa del SNC. De esta forma, la imagen constituye una herramienta fundamental tanto para el diagnóstico como para el seguimiento evolutivo de las enfermedades neurológicas.

Además, más allá de su utilidad clínica inmediata, las imágenes contienen patrones complejos y características sutiles que a menudo pasan desapercibidos en la evaluación convencional, pero que pueden revelar información oculta relevante relacionada con el funcionamiento y las alteraciones del SNC. Esta información representa una oportunidad para avanzar en la investigación en neurociencia clínica, tanto en la comprensión de los mecanismos patológicos como identificar nuevos perfiles biológicos.

2.3 Anatomía del cerebro humano

El cerebro humano es la estructura central del SNC, caracterizada por una alta complejidad morfológica y funcional. Se encuentra protegido por el cráneo y recubierto por tres membranas denominadas meninges (duramadre, aracnoides y piamadre), que delimitan compartimentos anatómicos y condicionan la circulación del líquido cefalorraquídeo (LCR) (Brodal, 2010).

Desde una perspectiva macroscópica, el encéfalo se divide en tres grandes regiones: los hemisferios cerebrales, el tronco encefálico y el cerebelo. Los hemisferios cerebrales constituyen la porción más voluminosa y se organizan en lóbulos diferenciados (frontal, parietal, temporal y occipital) que concentran funciones motoras, sensoriales, cognitivas y perceptivas. La superficie cortical, altamente plegada en surcos y circunvoluciones, incrementa la superficie de la corteza, permitiendo una mayor capacidad de procesamiento.

La corteza está formada principalmente por sustancia gris, compuesta por somas neuronales

y células gliales que posibilitan la actividad sináptica y el procesamiento de información. Por debajo se dispone la sustancia blanca cerebral, constituida por haces de axones mielinizados que conforman tractos de proyección, asociación y comisurales, cuya función es integrar y transmitir la información entre la corteza cerebral y las estructuras subcorticales. Ambas conforman en conjunto el parénquima cerebral, es decir, el tejido funcional del encéfalo encargado de la transmisión y procesamiento de la información nerviosa.

El tronco del encéfalo constituye el nexo con la médula espinal y alberga núcleos responsables del control de funciones vitales como la respiración o el ritmo cardíaco, además de la mayoría de los nervios craneales. El cerebelo, ubicado en la fosa posterior, asegura la coordinación motora, la postura y el equilibrio.

El sistema ventricular, formado por los ventrículos laterales, el tercer ventrículo y el cuarto ventrículo, mantiene la circulación del LCR, que fluye hacia el espacio subaracnoideo. Este líquido cumple funciones de amortiguación mecánica, eliminación de desechos metabólicos y homeostasis del microambiente neuronal.

El cerebro recibe sangre a través de dos grandes sistemas: el carotídeo y el vertebrobasilar. Ambos se unen en el polígono de Willis, una red de vasos que permite mantener el flujo sanguíneo incluso si alguna arteria se obstruye. Las bifurcaciones arteriales son zonas delicadas, porque en ellas pueden formarse aneurismas o malformaciones. La sangre venosa se recoge mediante venas superficiales y profundas, que drenan en los senos duros y de ahí pasan a la vena yugular interna.

En conjunto, la organización anatómica del cerebro proporciona la base estructural sobre la que se desarrollan múltiples funciones superiores, pero también determina los patrones de propagación y afectación de diversas patologías.

2.4 Imagen médica en trastornos neurológicos: impacto clínico y aplicaciones

La imagen médica es una de las técnicas más usadas en la práctica neurológica y neuroquirúrgica contemporánea. Más allá de su papel diagnóstico, es una herramienta transversal que acompaña al paciente en distintas fases del proceso asistencial: desde la evaluación inicial hasta la planificación del tratamiento, el seguimiento longitudinal y la investigación traslacional. Su capacidad de aportar información objetiva y reproducible explica su impacto tanto en la toma de decisiones clínicas inmediatas como en el diseño de estrategias terapéuticas personalizadas.

En los escenarios agudos, la neuroimagen se presenta como un recurso crítico. En el ictus, diferenciar con rapidez un infarto isquémico de una hemorragia cerebral determina de manera directa la elección del tratamiento, cuyo éxito depende de que se administre en las primeras horas del cuadro clínico. Un error diagnóstico en este contexto puede conducir a terapias inadecuadas con consecuencias potencialmente graves (Frade et al., 2022; Luff et al., 2024). De forma análoga, en crisis epilépticas de inicio reciente, la detección inmediata de lesiones subyacentes como tumores o malformaciones vasculares permite orientar un manejo precoz y más eficaz (Adamczyk et al., 2021). La combinación de rapidez en el acceso a la imagen y precisión en la interpretación es, en estos casos, determinante para mejorar el pronóstico.

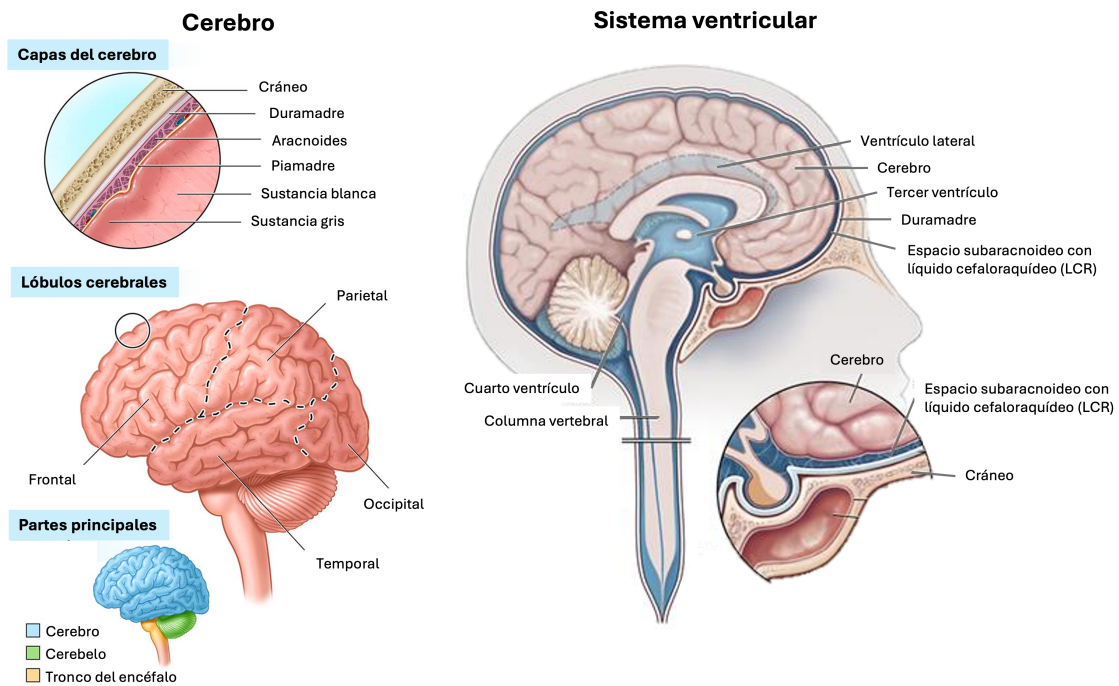


Figura 2.2: Representación esquemática de la anatomía cerebral: a la izquierda, capas meníngeas, sustancia gris y blanca, así como la organización en lóbulos y principales regiones encefálicas (Imagen adaptada de Cleveland Clinic, 2024; a la derecha, el sistema ventricular y la circulación del LCR en relación con el espacio subaracnoideo y la columna vertebral (Imagen adaptada de Equipo de la Torre, 2021).

Fuera del ámbito urgente, la neuroimagen se ha consolidado como un instrumento para la detección temprana de enfermedades neurológicas crónicas. En patologías como la enfermedad de Alzheimer o el Parkinson, la identificación de patrones de atrofia cerebral o de alteraciones metabólicas precede en ocasiones a la aparición de síntomas clínicos, aportando biomarcadores de riesgo con potencial para anticipar el diagnóstico y favorecer intervenciones en fases iniciales (Ali, 2022; Petrella et al., 2003).

Una vez establecido el diagnóstico, la imagen desempeña un papel esencial en la planificación terapéutica. La caracterización precisa de la localización, el volumen y la extensión de una lesión, así como su relación con estructuras funcionales críticas, permite evaluar la viabilidad de la intervención y elegir la estrategia más segura. En tumores cerebrales o aneurismas, esta información condiciona el tipo de abordaje quirúrgico y contribuye a minimizar complicaciones asociadas (Algethami et al., 2021; Papacci et al., 2015). En las lesiones cerebrales traumáticas, la detección precoz de complicaciones como edema o isquemia facilita actuaciones inmediatas que pueden modificar de forma sustancial la evolución clínica (Vande Vyvere et al., 2024).

Los avances tecnológicos han extendido este valor al procedimiento quirúrgico. La imagen intraoperatoria permite actualizar la localización y los márgenes de una lesión en tiempo real, adaptando la resección al contexto anatómico específico de cada paciente. En neurocirugía oncológica, esta posibilidad resulta especialmente relevante: permite maximizar la resección tumoral sin comprometer áreas funcionalmente críticas, mejorando tanto los resultados oncológicos como la preservación de la calidad de vida (Bin-Alamer et al., 2023; Dixon et al., 2022).

Durante el seguimiento, la neuroimagen ofrece un medio objetivo para evaluar la respuesta a los tratamientos y la progresión de la enfermedad. En la esclerosis múltiple, el número y la distribución de nuevas lesiones sirven como indicadores de la eficacia terapéutica (Granziera et al., 2021). En oncología, los criterios de respuesta basados en imagen constituyen un estándar para valorar la progresión tumoral y ajustar la estrategia de tratamiento en función de la evolución (Eisenhauer et al., 2009).

Más allá de la asistencia clínica, la neuroimagen ha sido clave para profundizar en la fisiopatología de las enfermedades neurológicas. Ha permitido estudiar la conectividad cerebral, los mecanismos de plasticidad neural y los procesos de reorganización funcional tras lesiones, aportando bases científicas para nuevas estrategias de rehabilitación (Grefkes y Fink, 2011). El desarrollo de biomarcadores derivados de imagen, además, ha favorecido la estratificación de pacientes en subgrupos con distinto pronóstico y respuesta terapéutica, lo que impulsa enfoques más personalizados (Jack et al., 2010; Weiner et al., 2017). Estas posibilidades se ven reforzadas por la implementación de protocolos estandarizados y técnicas de armonización, que reducen la variabilidad intercentros y mejoran la comparabilidad de estudios multicéntricos (Fortin et al., 2018; Keator et al., 2013).

En conjunto, la imagen médica se ha consolidado como un componente esencial de la medicina neurológica y neuroquirúrgica. Su capacidad para guiar intervenciones inmediatas, anticipar diagnósticos, orientar la planificación terapéutica, monitorizar resultados y aportar biomarcadores de investigación convierte a la neuroimagen en una herramienta insustituible, con impacto directo tanto en la práctica clínica como en el avance científico.

2.5 Técnicas de imagen y apoyo visual en neuroimagen intraoperatoria y convencional

Las técnicas de adquisición de imagen en neurología han evolucionado significativamente, distinguiéndose entre modalidades no intraoperatorias, fundamentales para el diagnóstico y la planificación, y las modalidades intraoperatorias, que han revolucionado la precisión quirúrgica en tiempo real. Este análisis explora ambos enfoques, destacando sus aplicaciones clínicas, ventajas y limitaciones técnicas, así como su integración en la práctica clínica contemporánea. Las tecnologías más recientes como la HSI están mostrando un potencial considerable para mejorar la precisión en la delimitación de tejidos, mientras que técnicas establecidas como la TC continúan ampliando sus capacidades mediante nuevas secuencias y aplicaciones funcionales.

2.5.1 Modalidades convencionales

Las modalidades de imagen no intraoperatorias representan la base del diagnóstico neurológico, ya que ofrecen información clave tanto en la fase previa a la intervención como durante el seguimiento postoperatorio. Estas técnicas han evolucionado considerablemente, alcanzando un elevado grado de sofisticación y estando plenamente integradas en la práctica clínica diaria.

1. **Tomografía Computarizada:** La TC es una herramienta fundamental en la evaluación neurológica de urgencia, especialmente en el contexto de hemorragias intracraneales, traumatismos craneoencefálicos y otras situaciones en las que se requiere una valoración rápida y precisa del encéfalo. Su funcionamiento se basa en la emisión de rayos X desde múltiples ángulos alrededor del cráneo del paciente. Estos haces atraviesan los tejidos y son detectados en función de la atenuación que sufren, lo que permite reconstruir imágenes axiales del cerebro mediante algoritmos computacionales avanzados. Posteriormente, estas imágenes pueden reorganizarse en múltiples planos o incluso en volúmenes tridimensionales, proporcionando una representación detallada tanto del parénquima cerebral como de las estructuras óseas del cráneo.

En la TC, la intensidad de cada vóxel se expresa en unidades Hounsfield (HU), un sistema de medida que cuantifica la atenuación de los rayos X respecto a dos referencias fijas: el agua destilada y el aire, a presión y temperatura estándar. Por definición, el agua tiene un valor de 0 HU y el aire de -1000 HU. Los tejidos y materiales se sitúan en esta escala en función de su densidad y composición: la grasa suele presentar valores entre -100 y -50 HU, el tejido blando como el parénquima cerebral oscila en torno a $30-45$ HU, la sangre puede situarse entre 50 y 70 HU, y el hueso cortical alcanza valores superiores a $+1000$ HU. Esta escala es lineal respecto al coeficiente de atenuación de los rayos X, lo que significa que diferencias iguales en HU corresponden a diferencias proporcionales en la capacidad de los tejidos para atenuar la radiación.

Los parámetros de adquisición desempeñan un papel determinante en la calidad y utilidad de la TC. La resolución espacial depende del tamaño de píxel en el plano axial, definido por el campo de visión (FOV) y la matriz de reconstrucción; en neuroimagen suele oscilar entre $0,3$ y $0,6$ mm. El grosor de corte (*slice thickness*), habitualmente

entre 0,5 y 5 mm, condiciona la capacidad de detectar lesiones pequeñas y la precisión de las reconstrucciones tridimensionales. Asimismo, la separación entre cortes (*spacing between slices*) debe mantenerse constante para evitar distorsiones volumétricas en la reconstrucción. Estos factores, junto con la colimación del haz y la calidad del detector, determinan tanto la resolución espacial como el nivel de ruido de la imagen.

La evolución de la TC hacia sistemas multidetectores ha supuesto una mejora considerable en la calidad de imagen. La adquisición helicoidal rápida y la posibilidad de obtener cortes finos permiten generar reconstrucciones multiplanares de alta resolución que son especialmente útiles en el entorno neuroquirúrgico. Gracias a estas capacidades, la TC no solo permite confirmar o descartar hallazgos agudos, sino que también facilita la planificación quirúrgica preoperatoria, definiendo con mayor precisión el trayecto de acceso y la relación espacial de la lesión con estructuras críticas (Ashraf et al., 2020; Boone, 2006).

Además de su valor anatómico, la TC puede ofrecer información funcional cuando se aplican técnicas específicas como la TC de perfusión, que analiza la distribución del contraste en tiempo real para estimar parámetros hemodinámicos como el flujo y el volumen sanguíneo cerebral (Wintermark, 2005). Incluso en estudios sin contraste, la TC permite realizar valoraciones cuantitativas con relevancia clínica. En el caso de las hemorragias cerebrales agudas, por ejemplo, el volumen total de sangre observado en la imagen inicial se ha identificado como un marcador pronóstico importante. En un estudio clásico, (Lagares et al., 2001) demostraron que el volumen hemorrágico en pacientes con HSAa se relaciona con la evolución clínica del paciente, lo que demuestra el valor predictivo de la información obtenida de esta técnica más allá de la detección de sangrado.

Sin embargo, la TC presenta limitaciones. Su sensibilidad para detectar lesiones isquémicas precoces es menor que la de la RM, y su uso implica exposición a radiación ionizante, factor de riesgo en estudios repetidos o en poblaciones vulnerables. Asimismo, el uso de contraste intravenoso puede suponer un problema de salud en determinados pacientes. Aun así, por su velocidad, disponibilidad y calidad diagnóstica, la TC sigue siendo una técnica indispensable en el manejo inicial del paciente neurológico y una base sólida para la planificación de muchas intervenciones neuroquirúrgicas.

2. **Resonancia Magnética:** La RM se ha consolidado como una de las modalidades de imagen más completas y precisas para el diagnóstico neurológico, proporcionando información anatómica y funcional detallada sin la necesidad de utilizar radiación ionizante (Lerch et al., 2017). Las distintas secuencias que pueden adquirirse (T1, T2, FLAIR, difusión, perfusión) permiten caracterizar con gran precisión diversas patologías cerebrales, desde tumores hasta enfermedades neurodegenerativas o inflamatorias.

Una de sus aplicaciones es la resonancia magnética funcional (fMRI, de sus siglas en inglés *functional Magnetic Resonance Imaging*), que ha permitido un enfoque más detallado de la actividad cerebral, especialmente en la planificación prequirúrgica en pacientes con tumores cerebrales. La fMRI permite localizar áreas elocuentes (como regiones motoras o del lenguaje) del cerebro que deben ser preservadas durante la

cirugía. Diversos estudios, como el de (Castellano et al., 2017) demuestran cómo la incorporación rutinaria de fMRI en la planificación de la cirugías de gliomas puede mejorar significativamente los resultados funcionales postquirúrgicos. Este enfoque también ha mostrado ser eficaz en la planificación de la cirugía de epilepsia, ayudando a la localización precisa de focos epilépticos (Benjamin et al., 2018).

En el ámbito de la neurooncología, las secuencias avanzadas como la espectroscopia por RM han permitido diferenciar entre distintos tipos de tumores cerebrales según su perfil metabólico. Se ha demostrado que la combinación de secuencias avanzadas de RM, incluidas la espectroscopia y la difusión, mejora significativamente la sensibilidad y especificidad en la caracterización de gliomas de alto grado en comparación con las metástasis cerebrales (Arévalo-Sáenz et al., 2022; Ishimaru et al., 2001). Además, las técnicas de RM para evaluar la perfusión y la permeabilidad de la barrera hematoencefálica, como el contraste con gadolinio, han sido cruciales para caracterizar gliomas malignos, ayudando en la evaluación de la agresividad tumoral (Heye et al., 2014).

Además, la RM funcional y estructural combinada con la imagen por tractografía ha permitido estudiar con precisión las conexiones cerebrales, lo cual es relevante en cirugías de tumores en áreas motoras o del lenguaje. Los avances en la tractografía por difusión están permitiendo a los neurocirujanos planificar intervenciones quirúrgicas con mayor precisión, reduciendo el riesgo de daños en tractos nerviosos importantes (Tae et al., 2018).

A pesar de sus múltiples ventajas, la RM presenta diversas limitaciones. Se trata de una técnica más costosa y menos accesible que la TC, con tiempos de adquisición más largos que pueden dificultar su uso en situaciones de urgencia. La calidad de imagen puede verse afectada por el movimiento del paciente, lo que puede ser problemático en niños, pacientes no colaborativos o en estado crítico. Además, su uso está contraindicado o limitado en personas portadoras de ciertos dispositivos metálicos, como marcapasos o clips aneurismáticos no compatibles. El uso de gadolinio como medio de contraste, aunque generalmente seguro, requiere precaución en pacientes con insuficiencia renal, y su depósito en algunos tejidos sigue siendo objeto de investigación. Por último, interpretar adecuadamente las imágenes de RM, especialmente cuando se utilizan secuencias funcionales o metabólicas, requiere una curva de aprendizaje más exigente y experiencia especializada.

3. **Tomografía por Emisión de Positrones:** La tomografía por emisión de positrones (PET, del inglés *Positron Emission Tomography*) es una técnica de imagen funcional que permite visualizar procesos metabólicos y bioquímicos mediante la administración de radiofármacos. A diferencia de técnicas como la TC o la RM, que muestran principalmente la estructura del cerebro, la PET se centra en su actividad, lo que resulta especialmente útil cuando los cambios funcionales aparecen antes que las alteraciones visibles en la anatomía (Fink et al., 2015; Tuan et al., 2025).

En enfermedades neurodegenerativas como el Alzheimer, la PET ha demostrado ser muy útil para detectar zonas de hipometabolismo típicas de esta patología incluso en etapas iniciales, ayudando a llegar a un diagnóstico más temprano y fiable (Chapleau

et al., 2022; Chételat et al., 2020). En el caso de tumores cerebrales, esta técnica permite identificar regiones con mayor actividad metabólica, lo que puede reflejar una mayor agresividad del tumor o cambios asociados a la progresión. (Chapleau et al., 2022; W. Chen, 2007).

Además, el uso de sistemas híbridos como PET/TC o PET/RM permite obtener en una sola exploración tanto imágenes anatómicas como funcionales. La PET/TC es especialmente útil en oncología, mientras que la PET/RM ofrece una mejor visualización del cerebro sin aumentar la exposición a radiación, lo que resulta ventajoso en estudios neurológicos complejos o en población pediátrica (Boss et al., 2010; Garibotto et al., 2013).

A pesar de su alto valor clínico, la PET presenta limitaciones que dificultan su uso en la práctica diaria. Entre ellas destacan su elevado coste, la limitada disponibilidad de equipos, la necesidad de instalaciones especializadas para la producción y manejo de radiofármacos, y la exposición a radiación ionizante. Estas restricciones hacen que su uso quede habitualmente reservado a casos en los que la información funcional aporte un valor diagnóstico diferencial claro.

4. **Angiografía:** La angiografía es una técnica especializada en la visualización del árbol vascular cerebral, que permite evaluar con alta precisión la anatomía y el estado funcional de las arterias y venas intracraneales. A diferencia de otras modalidades como la TC o la RM que se centran en la morfología general del tejido cerebral, la angiografía se utiliza específicamente para caracterizar alteraciones vasculares como estenosis, aneurismas, malformaciones arteriovenosas o trombosis. Esta técnica puede llevarse a cabo mediante distintos enfoques, entre los que destacan la angiografía por TC (CTA, de sus siglas en inglés *Computed Tomography Angiography*), la angiografía por RM (MRA, de sus siglas en inglés *Magnetic Resonance Angiography*), y la angiografía digital por sustracción (DSA, de sus siglas en inglés *Digital Subtraction Angiography*), que emplea rayos X junto con la administración intraarterial de contraste para obtener imágenes dinámicas de alta resolución (Jansen et al., 2017; Lasjaunias et al., 2013).

En el contexto neurovascular, la angiografía desempeña un papel clave tanto en el diagnóstico como en la planificación terapéutica. La CTA por su rapidez y disponibilidad, se ha consolidado como una herramienta esencial en la fase aguda del ictus isquémico, al permitir la detección de oclusiones en grandes vasos y guiar decisiones sobre la administración de tratamiento endovascular (Turc et al., 2023). Por otro lado, la DSA se utiliza cuando se requiere una caracterización detallada de la vasculatura cerebral, y es indispensable en procedimientos de neuroquirúrgicos, como la embolización de aneurismas o malformaciones arteriovenosas (Muir et al., 2017; Wong et al., 2012). Además, la angiografía ha mostrado utilidad en el estudio de hemorragias cerebrales espontáneas, especialmente cuando se sospechan causas secundarias, como fistulas durales o pequeños aneurismas que podrían pasar desapercibidos en otras técnicas (Velthuis et al., 1999).

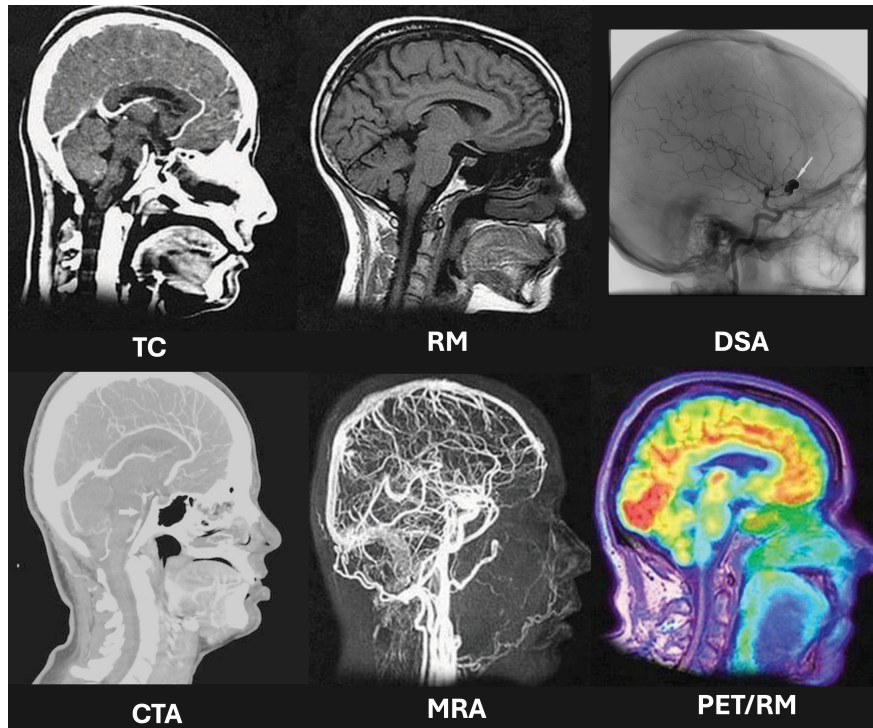


Figura 2.3: Representación comparativa de modalidades convencionales de neuroimagen. Arriba: TC y RM; abajo: RMA y PET/RM. Imagen adaptada de (Aryasinghe et al., 2016; San Diego Brain Injury Foundation, n.d. Waldron et al., 2018).

2.5.2 Modalidades intraoperatorias

Las modalidades intraoperatorias permiten la visualización en tiempo real del cerebro y estructuras relacionadas, proporcionando información crucial durante los procedimientos quirúrgicos. Estas técnicas están diseñadas para asistir a los neurocirujanos en la toma de decisiones en tiempo real, facilitando una resección tumoral más precisa, la delimitación de márgenes patológicos y la identificación de estructuras anatómicas clave que deben preservarse.

1. **Ecografía Intraoperatoria:** La ecografía intraoperatoria (ioUS, de sus siglas en inglés *intraoperative ultrasound*) es una modalidad de imagen que emplea ondas ultrasónicas para proporcionar visualización en tiempo real durante procedimientos quirúrgicos. Aunque presenta una resolución espacial inferior a la de técnicas como la RM o la TC, su portabilidad, bajo coste y la ausencia de radiación ionizante la convierten en una herramienta especialmente valiosa en el entorno quirúrgico, donde la rapidez y la accesibilidad son esenciales (Dixon et al., 2022).

En neurocirugía, la ioUS se ha incorporado de forma generalizada en procedimientos como la resección de tumores cerebrales, la evacuación de hematomas intracraneales y la localización de estructuras neuroanatómicas críticas. Su principal ventaja radica en la capacidad de guiar al cirujano en tiempo real, permitiendo una navegación dinámica y adaptativa durante la intervención. Este aspecto es especialmente relevante frente al fenómeno del desplazamiento cerebral o *brain shift* por la despresurización del líquido

cefalorraquídeo, que puede comprometer la precisión de las imágenes preoperatorias (Shetty et al., 2021).

Diversos estudios han demostrado que la ioUS mejora la identificación de los márgenes tumorales, favoreciendo resecciones más completas y reduciendo el riesgo de dejar tejido tumoral residual (Dixon et al., 2022). Además, el uso de ecografía Doppler en contexto vascular permite monitorizar la hemodinámica cerebral durante intervenciones como el clipaje de aneurismas o la resección de malformaciones arteriovenosas, contribuyendo a minimizar complicaciones intraoperatorias (Ganau et al., 2019).

Entre las aplicaciones avanzadas de la ioUS se encuentra la ecografía intraoperatoria navegada, que combina la imagen ecográfica en tiempo real con sistemas de neuronavegación, lo que permite una localización más precisa y una resección más controlada de gliomas (Villegas-Gomez et al., 2022). También destaca la ecografía con contraste, que mediante el uso de agentes de contraste mejora la visualización de la vascularización tumoral y la perfusión cerebral durante la cirugía, aportando información funcional adicional que facilita la toma de decisiones intraoperatorias (Ganau et al., 2019). Otra técnica relevante es la fusión de imágenes, que permite superponer la ecografía intraoperatoria con estudios preoperatorios de RM, mejorando la orientación espacial del cirujano y compensando el desplazamiento cerebral que ocurre durante el procedimiento (Ganau et al., 2019). Además, la ioUS está adquiriendo un papel creciente en cirugía de columna, donde se ha utilizado con éxito para identificar con precisión los niveles vertebrales para guiar a la cirugía sin exposición a radiación, con estudios recientes que avalan su seguridad y efectividad como alternativa a técnicas basadas en fluoroscopia o TC intraoperatoria (Omar II y Starreveld, 2022).

Aunque la ioUS es una herramienta accesible y versátil que ofrece imágenes en tiempo real durante la cirugía, presenta algunas limitaciones importantes. Su resolución y contraste son menores que otras técnicas de imagen, lo que puede dificultar la detección precisa de márgenes tumorales o lesiones pequeñas. Además, la calidad de las imágenes depende mucho de la experiencia del operador y puede verse afectada por artefactos como aire o sangre. Por último, la ioUS proporciona una visión localizada y en cortes específicos, lo que puede extender la duración de la intervención.

2. **TC y RM intraoperatoria:** La tomografía computarizada intraoperatoria (TCi) y la resonancia magnética intraoperatoria (RM iO) son herramientas que permiten al neurocirujano evaluar en tiempo real hasta dónde se ha extirpado un tumor, detectar restos tumorales y localizar posibles complicaciones antes de terminar la cirugía (Nimsky et al., 2006; Senft et al., 2011). Esto es especialmente útil cuando se trabaja en zonas profundas o cerca de áreas críticas para funciones esenciales del cerebro (Nimsky et al., 2006).

La TCi destaca por su mayor disponibilidad, menor coste y rápida de obtener. Además, gracias a que muchas veces se puede usar con unidades móviles dentro del quirófano, no es necesario mover al paciente durante la intervención. Esto facilita la seguridad y la organización durante la cirugía (Bot et al., 2017). Sin embargo, su capacidad para distinguir tejidos es menor que la de la RM iO.

La RM iO, por su parte, ofrece una imagen más detallada y permite diferenciar mejor el tejido tumoral del sano, algo fundamental en tumores infiltrativos como el GBM. Además, ayuda a guiar la cirugía con mayor precisión. Pero esta técnica requiere equipamiento más complejo y costoso, aumenta el tiempo quirúrgico y necesita que el quirófano esté especialmente adaptado para funcionar con equipos compatibles con campos magnéticos (Senft et al., 2011; Wach et al., 2024).

Ambas modalidades también pueden interferir con la organización habitual del quirófano, afectando la disposición del equipo y la esterilidad, por lo que su introducción requiere una planificación cuidadosa y protocolos específicos. Además, requiere la presencia continua de técnicos de imagen capacitados para manejar el equipo durante la cirugía.

3. **Imagen Hiperespectral:** La HSI es una técnica óptica que combina información espacial y espectral, permitiendo capturar imágenes en múltiples bandas del espectro electromagnético. Cada píxel de una imagen hiperespectral contiene una firma espectral continua que refleja las propiedades ópticas del tejido, lo que permite su caracterización con una precisión superior a la imagen convencional (Lu y Fei, 2014).

En el ámbito de la neurocirugía, la HSI es una herramienta prometedora para la identificación intraoperatoria de márgenes tumorales, especialmente en tumores infiltrativos como los gliomas de alto grado. Esta técnica puede detectar diferencias espectrales sutiles entre tejido tumoral y parénquima cerebral sano, sin necesidad de agentes de contraste exógenos (Akbari et al., 2012; Fabelo et al., 2019). Diversos estudios han demostrado que la HSI permite visualizar zonas de infiltración tumoral no visibles por el ojo humano o en otras modalidades de imagen (Jansen-Winkel et al., 2021), lo que contribuye a una resección más precisa y a una menor tasa de recurrencia. Además, se están investigando técnicas de fusión de imágenes hiperespectrales con RM preoperatoria y otros datos multimodales, con el objetivo de enriquecer la experiencia intraoperatoria mediante herramientas avanzadas como la realidad aumentada y la navegación quirúrgica asistida (Villa et al., 2024).

Sin embargo, la HSI todavía está en una fase temprana de desarrollo clínico, siendo empleada principalmente en entornos de investigación. Entre sus principales desafíos se encuentran el alto coste del equipamiento, la necesidad de procesamiento computacional a tiempo real y la formación especializada del equipo médico. Por otro lado, los avances recientes en algoritmos de análisis de datos, incluyendo técnicas de ML, están facilitando la clasificación automática de tejidos en tiempo casi real (Halicek et al., 2018; Lu y Fei, 2014).

La Tabla 2.1 presenta una comparativa de las distintas modalidades de imagen y apoyo visual empleados en el ámbito de la neurociencia clínica. Se incluyen tanto las tecnologías convencionales como técnicas de imagen intraoperatorias anteriormente descritas. Los criterios de comparación consideran aspectos técnicos (resolución espacial, capacidad de contraste, velocidad de adquisición), así como su aplicabilidad práctica en entornos de urgencia o en cirugía cerebral.

Tabla 2.1: Comparación entre modalidades de imagen médica y apoyo visual en neurociencia clínica

Modalidad	Resolución espacial	Sensibilidad tisular	Tiempo adquisición	Captura en urgencias	Tipo de información	Invasividad
CT	Alta	Baja	Rápido	Sí	Estructural	No
RM	Alta	Muy alta	Lento	No	Estructural y funcional	No
Angiografía	Alta	Alta (vascular)	Variable	Sí	Vascular anatómica	Sí
HSI	Media	Alta (espectral)	Rápido	No	Espectral (óptica)	No
PET	Baja	Metabólica	Lento	No	Molecular / funcional	Sí (requiere trazador)
ioUS	Media	Baja	Rápido	No	Estructural (tiempo real)	No

2.6 Biomarcadores de imagen y sus aplicaciones en neurociencia clínica

En el ámbito de la neurociencia clínica, la mera identificación visual de una alteración en la imagen médica no resulta suficiente. Es necesario llevar a cabo una cuantificación precisa de su extensión, determinar con exactitud su localización y analizar el impacto sobre regiones cerebrales funcionalmente relevantes.

El avance tecnológico en imagen médica digital ha posibilitado la adquisición en neuroimagen de datos anatómicos y funcionales con una gran calidad y resolución. Paralelamente, el desarrollo de métodos computacionales ha facilitado la extracción de información cuantitativa de estas imágenes. En este contexto, los biomarcadores de imagen emergen como características medibles extraídas a partir de imágenes médicas, capaces de reflejar diferentes niveles de información, desde aspectos morfológicos macroscópicos hasta relaciones moleculares implícitas. Estos biomarcadores actúan como indicadores objetivos de procesos biológicos, patológicos o de respuesta terapéutica. Su versatilidad les permite obtenerse de múltiples modalidades de imagen, proporcionando en neurociencia clínica una herramienta no invasiva, cuantitativa y reproducible para el estudio cerebral (O’connor et al., 2017).

Se pueden distinguir, en líneas generales, cuatro grandes categorías de biomarcadores de imagen según la naturaleza de la información que aportan y su aplicación clínica:

1. **Biomarcadores anatómicos:** Comprenden características morfológicas cuantificables de las imágenes cerebrales, tales como tamaño, forma, volumen o integridad de estructuras específicas. Son esenciales para identificar alteraciones neuroanatómicas asociadas a diversas patologías y permiten realizar evaluaciones objetivas y cuantitativas del estado cerebral, así como su seguimiento longitudinal. Por ejemplo, la atrofia del hipocampo o de la corteza cerebral constituye un marcador consolidado en la enfermedad de Alzheimer (Jack et al., 2010). Asimismo, la detección y cuantificación precisa de lesiones estructurales, como tumores o infartos cerebrales mediante TC o RM, es clave

para el diagnóstico y la planificación terapéutica en ictus o gliomas. Técnicas como la morfometría basada en vóxeles permiten analizar con detalle variaciones en volumen cerebral y grosor cortical, con utilidad demostrada en enfermedades como la esclerosis múltiple (Pelletier et al., 2004).

2. **Biomarcadores radiómicos:** Los biomarcadores radiómicos son descriptores cuantitativos extraídos de cualquier modalidad de imagen médica, los cuales reflejan la heterogeneidad espacial, morfológica y textural de los tejidos. La radiómica es una tecnología emergente que se basa en la premisa de que las imágenes médicas contienen información biológica subyacente que no es perceptible a simple vista, pero que puede ser cuantificada a través de algoritmos matemáticos y utilizada para la caracterización no invasiva de enfermedades (Gillies et al., 2016). Este tipo de biomarcadores se encuentra más extensamente explicado en la Sección 2.6.3.
3. **Biomarcadores funcionales:** Evalúan parámetros relacionados con la actividad cerebral y el flujo sanguíneo, ofreciendo información sobre metabolismo neuronal y perfusión. Por ejemplo, el flujo sanguíneo cerebral, medido mediante PET o RM de perfusión, es ampliamente utilizado para diagnosticar ictus y deterioro cognitivo vascular (Jezzard et al., 2018). Los mapas de coeficiente de difusión aparente, derivados de imágenes de difusión por RM, reflejan la integridad tisular y resultan relevantes para el diagnóstico precoz del ictus y el estudio de enfermedades neurodegenerativas (Mori y Zhang, 2006). Asimismo, las firmas espectrales obtenidas mediante HSI proporcionan biomarcadores funcionales no invasivos basados en la respuesta espectral del tejido. Estas últimas se encuentran más detalladas en la Sección 2.6.2.
4. **Biomarcadores de conectividad:** Se centran en la evaluación de las interacciones entre distintas áreas cerebrales, aportando información sobre la conectividad estructural y funcional. Son especialmente útiles en el estudio de trastornos que afectan las redes neuronales, como diversas enfermedades neuropsiquiátricas. Por ejemplo, la resonancia magnética funcional en estado de reposo (resting-state fMRI) permite analizar redes neuronales como la red por defecto, cuya alteración se ha asociado a patologías como esquizofrenia, autismo y depresión mayor (Cattarinussi et al., 2023; Itahashi et al., 2015; Sheline et al., 2009).

En conjunto, estos biomarcadores complementan el diagnóstico convencional al aportar datos cuantitativos que permiten caracterizar con mayor detalle lesiones hemorrágicas, tumorales o procesos degenerativos, facilitando la evaluación objetiva de su evolución y respuesta terapéutica (Smits, 2021). En la Figura 2.4 se observan representaciones ejemplo de estos cuatro biomarcadores que se pueden extraer en neuroimagen.

2.6.1 Desarrollo y extracción de biomarcadores de imagen

El desarrollo de un biomarcador de imagen clínicamente útil requiere un proceso meticuloso que asegure su coherencia conceptual, validez técnica, reproducibilidad y relevancia clínica (Bonmatí et al., 2012). En la Figura 2.5 se resumen las principales etapas de este proceso, que abarcan desde la definición del concepto biológico hasta su validación y posible incorporación en la práctica clínica.

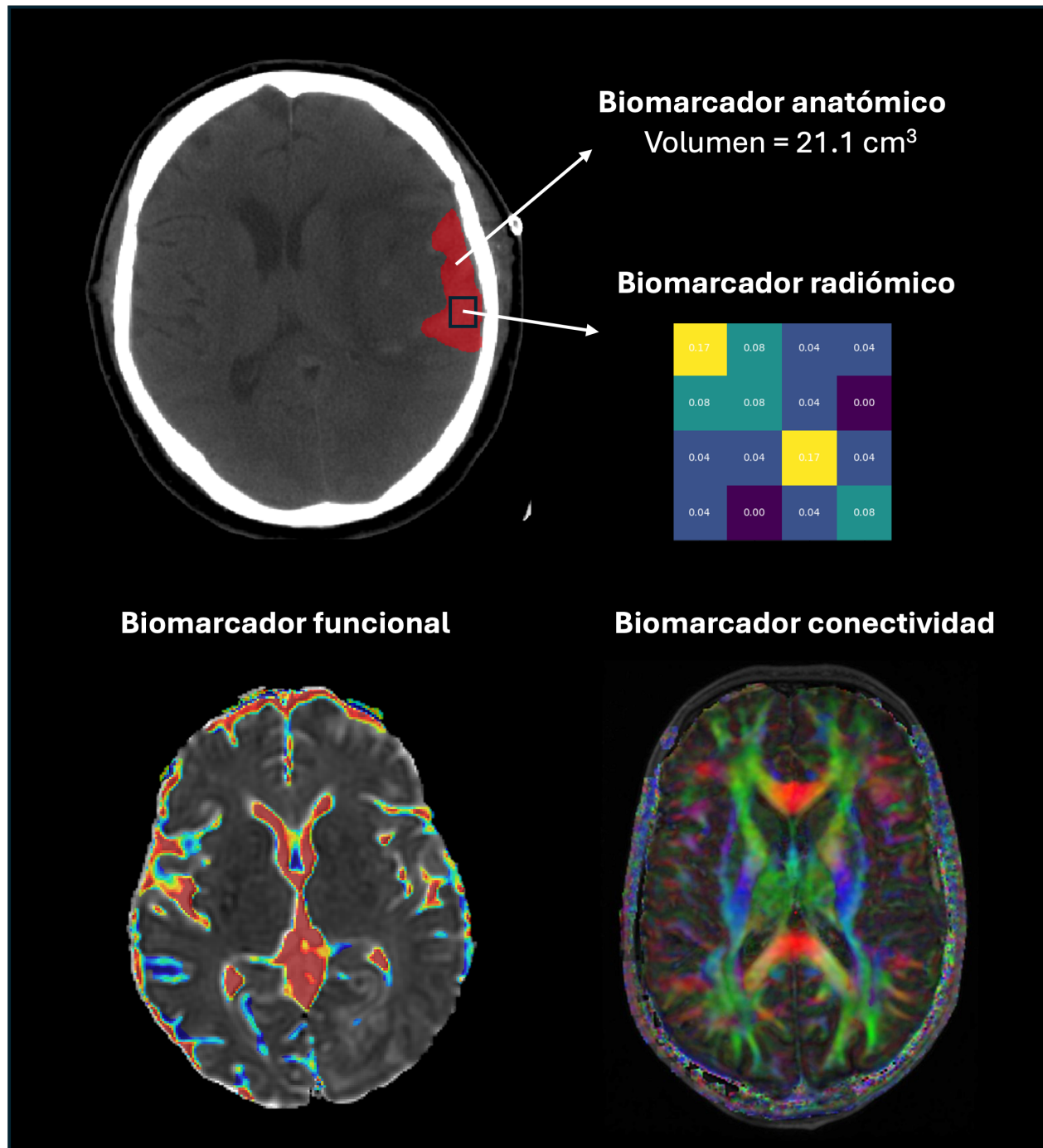


Figura 2.4: Ejemplos de biomarcadores en neuroimagen. En la fila de arriba se representan los biomarcadores anatómicos y de radiómica extraídos de una región de interés; en la fila de abajo biomarcadores funcionales y de conectividad cerebral.

1. Definición del proceso biológico de interés. El primer paso consiste en identificar con claridad el fenómeno fisiopatológico que se pretende cuantificar, el cual debe ser observable mediante técnicas de imagen adecuadas. La prueba de concepto debe demostrar que dicho proceso es viable mediante métodos cuantitativos basados en imágenes (Bonmatí et al., 2012).

2. Asociación clínica. Una vez definido el concepto, es necesario demostrar que el biomarcador se relaciona de manera sólida con variables clínicas o biológicas de interés. Por ejemplo, sería esperable que el volumen de hemorragia medido automáticamente se correlacione con el pronóstico funcional, o que la textura de una lesión tumoral esté relacionada con su grado histológico.

3. Calidad de imagen y preprocesamiento. La adquisición de imágenes con calidad suficiente es un requisito indispensable. Factores como la resolución espacial y temporal, la técnica empleada o los parámetros del equipo afectan directamente a la precisión del biomarcador. Posteriormente, las imágenes suelen someterse a una serie de pasos de preprocesamiento que incluyen corrección de artefactos, normalización de intensidades y reducción de ruido, con el objetivo de mejorar la comparabilidad entre estudios y minimizar la variabilidad técnica (Dietrich et al., 2007).

4. Segmentación de la región de interés. La identificación precisa de las áreas anatómicas o patológicas relevantes es esencial para la extracción focalizada de biomarcadores. Este paso puede realizarse de forma manual, semiautomática o automática, siendo estas últimas cada vez más frecuentes gracias al uso de algoritmos de DL, como redes convolucionales, que ofrecen mejoras sustanciales en precisión y reproducibilidad (Heller et al., 2021; Isensee et al., 2021).

5. Extracción de características. Los biomarcadores de imagen se derivan a partir de distintas medidas cuantitativas, en función del tipo de información que se desea capturar. Entre ellas se incluyen características morfológicas (forma, tamaño), anatómicas (localización, relaciones espaciales), intensidades de señal (valores de píxeles o vóxeles) y métricas funcionales asociadas a procesos fisiológicos, como el flujo sanguíneo, la perfusión o la difusión (Aerts et al., 2014). En modalidades avanzadas, como la espectroscopia o la imagen hiperespectral, es posible extraer además biomarcadores espectrales o metabólicos, que ofrecen información molecular o bioquímica adicional (Calin et al., 2014; Lu y Fei, 2014).

6. Selección de variables y reducción dimensional. Dado que la extracción puede generar cientos de variables, es crucial aplicar técnicas estadísticas y de ML para seleccionar las más relevantes, evitar redundancias y reducir el riesgo de sobreajuste. Esta fase también facilita la interpretabilidad clínica del biomarcador (Parmar et al., 2017; Zwanenburg et al., 2020).

7. Modelización y validación clínica. Los biomarcadores seleccionados se integran en modelos matemáticos, estadísticos o de inteligencia artificial, orientados al diagnóstico, pronóstico o evaluación terapéutica. La validación rigurosa —mediante validación cruzada, cohortes independientes o ensayos clínicos— es esencial para demostrar su utilidad real en entornos clínicos (Collins et al., 2015a; Kumar et al., 2012).

8. Criterios de adopción clínica. Para que un biomarcador sea incorporado en la práctica médica, debe cumplir con estándares exigentes: ser tecnológicamente estable, reproducible,

estandarizable, seguro, coste-efectivo y, sobre todo, correlacionarse de forma sólida con variables clínicas o biológicas relevantes. Asimismo, debe mostrar una sensibilidad y especificidad suficientes para apoyar la toma de decisiones médicas (Group et al., 2001).

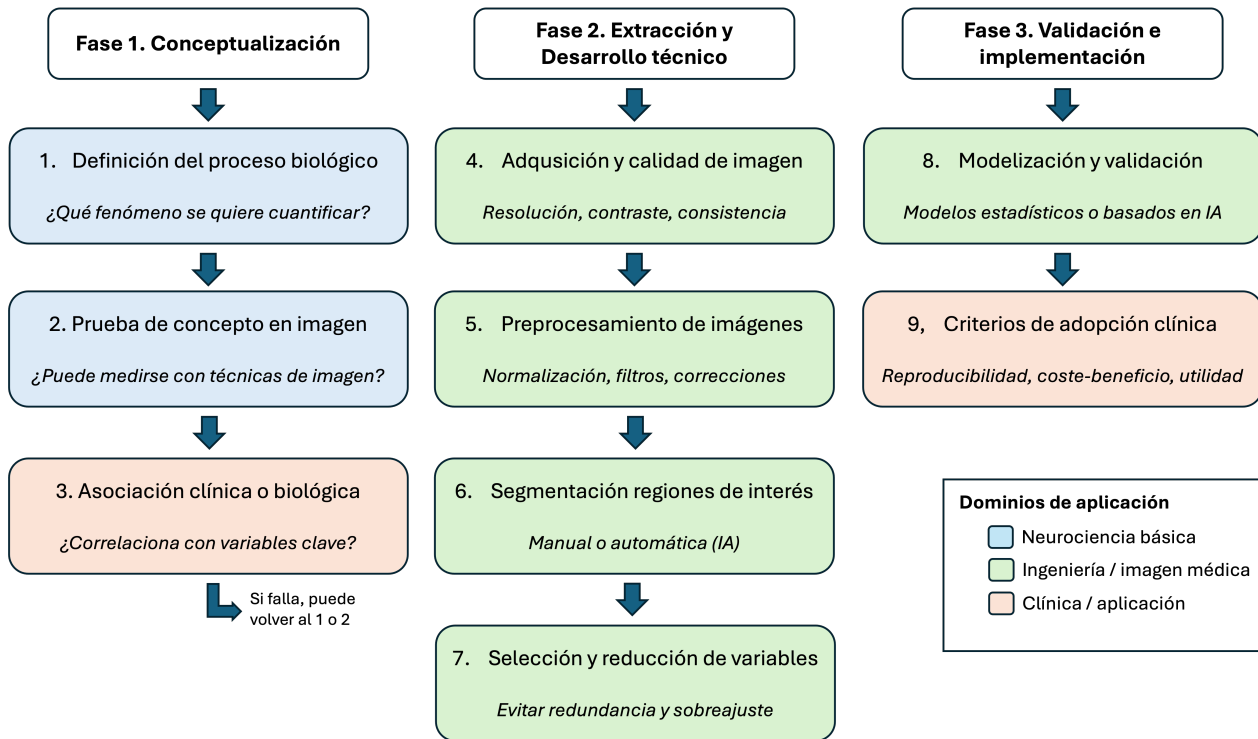


Figura 2.5: Diagrama del flujo de desarrollo de un biomarcador de imagen, desde la definición conceptual hasta su validación clínica, codificado por área disciplinar.

2.6.2 Imagen hiperespectral y firma espectral

La HSI es una técnica óptica emergente que combina simultáneamente información espacial y espectral, ofreciendo una representación tridimensional de la imagen capturada en forma de cubo hiperespectral, en el que dos dimensiones corresponden a las coordenadas espaciales (x , y) y la tercera a las longitudes de onda (λ). A diferencia de las imágenes RGB o en escala de grises, que capturan solo una fracción limitada del espectro electromagnético, la HSI captura cientos de bandas espectrales en un amplio rango de longitudes de onda, que puede abarcar desde el ultravioleta (UV) hasta el infrarrojo cercano (NIR, de sus siglas en inglés *Near InfraRed*), típicamente entre 400 y 1000 nm en entornos biomédicos.

La firma espectral es el vector que representa la intensidad de reflectancia de un tejido a lo largo de las distintas longitudes de onda. Esta curva espectral es única para cada tipo de tejido, lo que permite utilizarla como biomarcador óptico para identificar, diferenciar o clasificar estructuras biológicas.

Como se observa en la Figura 2.6, cada píxel del cubo hiperespectral contiene un espectro continuo de reflectancia (firma espectral), que refleja las propiedades ópticas intrínsecas de los tejidos capturados.

Esta diferenciación se debe a que los distintos tejidos contienen diferentes proporciones de componentes como la hemoglobina, la oxihemoglobina, los citocromos, la melanina o el agua, cada uno con un comportamiento específico frente a la absorción de luz en distintos rangos del espectro. En tumores cerebrales, como los gliomas, se producen alteraciones en la vascularización, metabolismo y densidad celular, que suponen modificaciones medibles de la firma espectral respecto al parénquima normal. (Fabelo, Ortega, Lazcano et al., 2018) demostraron que dichas firmas pueden capturarse de forma fiable durante la cirugía cerebral, y utilizarse como entrada para algoritmos de ML destinados a clasificar los distintos tipos de tejido con alta precisión.

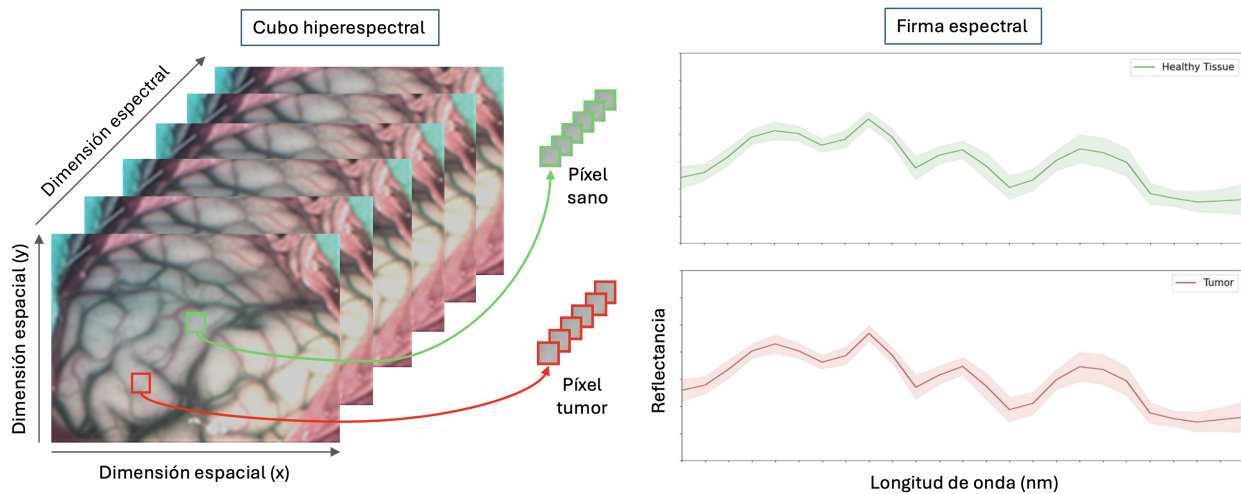


Figura 2.6: Extracción de la firma espectral de cada píxel de un cubo hiperespectral.

Los sistemas de imagen hiperespectral constan típicamente de:

- Una fuente de iluminación uniforme y estable, que proporcione una irradiancia homogénea sobre el campo visual. En entornos quirúrgicos se utilizan frecuentemente fuentes halógenas o LED con espectro amplio.
- Un mecanismo de dispersión espectral, como un prisma, una red de difracción o un interferómetro acoplado a un sistema óptico, que permite descomponer la luz reflejada o transmitida por el tejido en sus diferentes componentes espectrales.
- Un sensor de detección, habitualmente una cámara CCD o CMOS sensible a las longitudes de onda de interés, acoplado a un sistema de adquisición que puede funcionar bajo distintos modos, según cómo se capture la información espectral y espacial. Los principales son:
 - Escaneo espacial (pushbroom y whiskbroom): en el modo pushbroom, se registra una línea espacial completa a través de todas las longitudes de onda, y el desplazamiento del sistema permite construir el cubo hiperespectral. En el modo whiskbroom, menos frecuente en aplicaciones clínicas, se captura un único punto del campo visual a todas las longitudes de onda mediante un barrido óptico puntual, repitiendo el proceso para cada posición.

- Escaneo espectral (wavelength scanning): se adquieren imágenes bidimensionales completas a distintas longitudes de onda de forma secuencial.
- Captura instantánea (snapshot): se obtiene todo el cubo hiperespectral en un solo disparo, gracias a sensores especializados, con la ventaja de alta velocidad pero una resolución espectral generalmente inferior que las técnicas anteriores.

Cada uno de estos modos implica compromisos entre resolución espacial, espectral, velocidad de adquisición y complejidad técnica. En entornos quirúrgicos, donde el tiempo real y la estabilidad son cruciales, los sistemas basados en escaneo espectral rápido o snapshot tienen mayor aplicabilidad clínica (Fabelo et al., 2019).

La obtención de firmas espectrales fiables en entornos clínicos implica superar múltiples desafíos técnicos, ya que la señal hiperespectral puede verse comprometida por diversos factores. Entre ellos se encuentran la variabilidad en la iluminación (incluyendo problemas como la no uniformidad o parpadeos de la fuente lumínica), el ruido electrónico del sensor, los artefactos generados por el movimiento del paciente o del instrumental quirúrgico, y la presencia de fluidos biológicos, como sangre o edema, que alteran la reflectancia del tejido y distorsionan el espectro capturado. Para mitigar estos efectos, el preprocesado de la imagen constituye una etapa fundamental, ya que permite corregir las distorsiones sistemáticas, eliminar el ruido y normalizar la información para que las firmas espectrales extraídas sean consistentes, reproducibles y clínicamente interpretables.

En conjunto, la imagen hiperespectral y el análisis de firmas espectrales pueden usarse para el desarrollo de biomarcadores funcionales cuantitativos, al permitir la caracterización detallada de los tejidos en función de sus propiedades ópticas intrínsecas.

2.6.3 Radiómica

La radiómica es una disciplina centrada en la extracción y análisis cuantitativo de datos de imagen médica, con el objetivo de descubrir patrones que sirvan como biomarcadores no invasivos. Estos biomarcadores radiómicos permiten describir de manera objetiva propiedades fenotípicas del tejido, como su textura, morfología o heterogeneidad, que pueden relacionarse con procesos biológicos subyacentes, y que podrían asociarse con diagnóstico, pronóstico, respuesta terapéutica o características moleculares (Gillies et al., 2016).

A diferencia de la interpretación visual convencional, limitada a descripciones cualitativas, la radiómica se basa en análisis matemático reproducible, que transforma las imágenes en vectores de características estructuradas. Estas variables pueden integrarse posteriormente en modelos predictivos, junto con datos clínicos, genómicos o histopatológicos, dentro de estrategias de medicina de precisión.

Las características radiómicas se agrupan en categorías que describen diferentes aspectos del volumen o región de interés, y cada grupo captura una dimensión distinta de la imagen biomédica:

- **Características de forma:** Describen la forma, el tamaño y la compacidad del objeto segmentado. Incluyen métricas como volumen, superficie, esfericidad, elongación, relaciones superficie/volumen, y bounding box. Son útiles para caracterizar la arquitectura

de lesiones o estructuras anatómicas.

- **Características de primer orden:** Representan la estadística de la distribución de intensidades dentro de la región segmentada, sin tener en cuenta la disposición espacial de los voxeles. Incluyen medidas como la media, desviación estándar (DE), curtosis, asimetría (skewness), entropía, energía, y percentiles. Estas variables reflejan la intensidad global y su dispersión.
- **Características de segundo orden:** Capturan la organización espacial de las intensidades de los voxeles, lo que proporciona información sobre la heterogeneidad interna del tejido. Se derivan de matrices estadísticas como:
 - GLCM (Gray-Level Co-occurrence Matrix): evalúa la frecuencia con la que se repiten pares de intensidades con cierta distancia y orientación.
 - GLRLM (Gray-Level Run Length Matrix): cuantifica corridas consecutivas de voxeles con igual intensidad.
 - GLSZM (Gray-Level Size Zone Matrix): analiza zonas de tamaño homogéneo con igual intensidad.
 - NGTDM (Neighbourhood Gray-Tone Difference Matrix): mide la diferencia entre la intensidad de un voxel y la media de sus vecinos.
- **Características de orden superior:** Se extraen tras aplicar filtros (por ejemplo, wavelet, Laplacian of Gaussian, log-sigma) sobre la imagen original, con el fin de resaltar patrones multiescalares o estructuras específicas. Estas características suelen aumentar la sensibilidad del modelo a variaciones sutiles de textura o bordes.

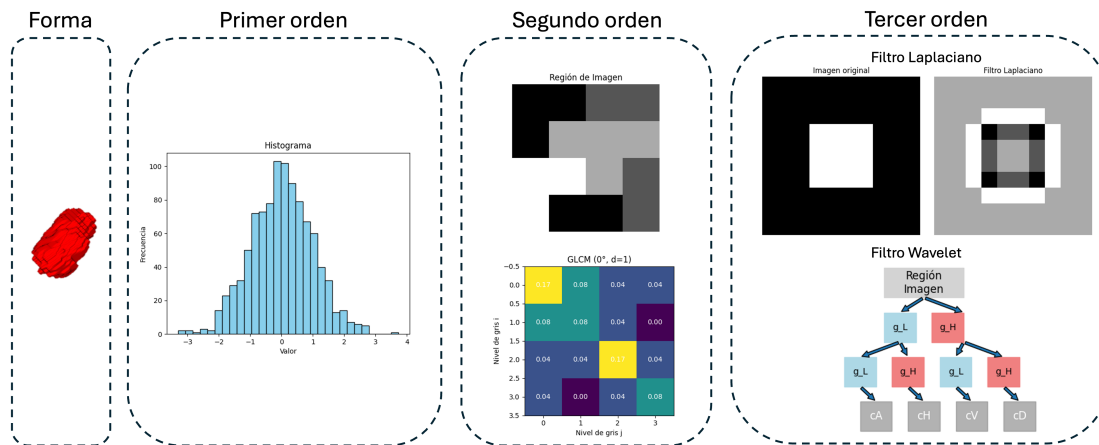


Figura 2.7: Clasificación de características radiómicas: morfológicas, de primer orden, de segundo orden y de orden superior.

Estas categorías, en conjunto, permiten construir representaciones del tejido, creando firmas radiómicas cuantitativas de la región de interés segmentada.

En el ámbito de la neurociencia clínica, los biomarcadores radiómicos han demostrado un amplio potencial. En neurooncología, por ejemplo, se han utilizado para caracterizar

tumores cerebrales como los gliomas mediante imágenes de RM (incluyendo secuencias FLAIR, T1 con contraste, T2 y mapas de difusión), identificando firmas radiómicas asociadas con alteraciones moleculares como la mutación IDH, la metilación del promotor de MGMT o el subtipo transcriptómico (Aerts et al., 2014; Doniselli et al., 2024; Peng et al., 2021). Estos biomarcadores han permitido además la predicción no invasiva de la supervivencia global y la respuesta al tratamiento con precisión comparable a la de modelos basados en histología (Aftab et al., 2022; Lao et al., 2017).

En enfermedades neurodegenerativas, se han identificado firmas radiómicas capaces de discriminar entre sujetos sanos, pacientes con deterioro cognitivo leve y enfermedad de Alzheimer, a partir del análisis textural y volumétrico de regiones como el hipocampo o la corteza entorrinal en imágenes de RM estructural (Feng y Ding, 2020; X. Liu et al., 2018).

La radiómica también ha mostrado aplicaciones en la predicción de la evolución funcional tras ictus, así como en la clasificación automatizada de lesiones cerebrales (Eghbali et al., 2025; H. Wang et al., 2021). En este sentido, los biomarcadores radiómicos actúan como un puente entre la imagen médica y la inteligencia artificial, proporcionando una representación cuantitativa de características radiológicas cerebrales y que puede integrarse con otros dominios de datos clínicos, genómicos o funcionales.

2.7 Inteligencia Artificial en Imagen Médica: Fundamentos, Algoritmos y Aplicaciones

La IA ha dejado de ser un concepto abstracto para convertirse en una herramienta real, con impacto directo en múltiples disciplinas científicas y en la práctica clínica cotidiana. En los últimos años, hemos sido testigos de una evolución sin precedentes en las capacidades de la IA, impulsada por tres factores clave: la disponibilidad creciente de datos, el aumento de la capacidad computacional y el desarrollo de nuevas técnicas de ML. Este avance ha abierto la puerta a una medicina más precisa, personalizada y basada en datos.

La aplicación de la IA en medicina presenta retos singulares. La información médica es compleja, heterogénea y está cargada de implicaciones clínicas. Por eso, la integración de modelos de IA en el entorno sanitario no se limita a una cuestión técnica, sino que exige una comprensión profunda del contexto clínico y un diálogo constante entre ingenieros, científicos y profesionales de la salud. El objetivo de la IA debe aportar valor al clínico, reducir la incertidumbre diagnóstica, mejorar la eficiencia en su trabajo y, en última instancia, beneficiar al paciente.

Dentro de las múltiples aplicaciones clínicas de la IA, el análisis de imágenes médicas destaca por su alto potencial y grado de desarrollo. No solo por el volumen de datos que genera, sino también por la riqueza de información contenida en cada estudio. Las imágenes médicas ofrecen representaciones detalladas anatómicas y funcionales, cuyo análisis puede beneficiarse enormemente de técnicas capaces de detectar patrones, segmentar estructuras o cuantificar alteraciones sutiles.

En el ámbito de la imagen médica se usan distintos tipos de aprendizaje de métodos de IA: supervisado, no supervisado, semi-supervisado y auto-supervisado (Cheplygina et al., 2019;

Litjens et al., 2017; Lundervold y Lundervold, 2019).

- **Aprendizaje supervisado:** El aprendizaje supervisado es, con diferencia, el más utilizado en aplicaciones clínicas. En este enfoque, el modelo se entrena a partir de un conjunto de datos en el que cada entrada está asociada a una etiqueta o salida conocida. El objetivo del algoritmo es aprender la relación entre los datos de entrada y las salidas esperadas, de modo que pueda generalizar esa relación a nuevos casos no vistos. Este tipo de aprendizaje es especialmente útil para tareas como clasificación, segmentación semántica de estructuras anatómicas o lesiones y predicción de desenlaces clínicos. Sin embargo, su principal limitación se basa en la necesidad de grandes cantidades de datos etiquetados, lo cual implica trabajo experto, tiempo y, en ocasiones, ambigüedad diagnóstica (Litjens et al., 2017; Lundervold y Lundervold, 2019).
- **Aprendizaje no supervisado** A diferencia del anterior, el aprendizaje no supervisado se basa en datos sin etiquetas. En este caso, el modelo intenta descubrir por sí mismo estructuras subyacentes, agrupamientos o relaciones dentro del conjunto de datos. Es un enfoque exploratorio, útil cuando no se dispone de una verdad conocida o cuando se desea comprender mejor la organización interna de los datos. Aunque su uso es menos común, se utiliza para agrupamiento de características, reducción de dimensionalidad o detección de datos anómalos (*outliers*). Este tipo de aprendizaje puede servir también como paso previo al aprendizaje supervisado (Bengio et al., 2013).
- **Aprendizaje semi-supervisado** Estrategias intermedias que combinan las ventajas del aprendizaje supervisado y no supervisado. El aprendizaje semi-supervisado utiliza un pequeño conjunto de datos etiquetados junto con una gran cantidad de datos sin etiquetar para mejorar el rendimiento del modelo. Esto resulta especialmente útil en medicina, donde la disponibilidad de etiquetas puede ser limitada (Cheplygina et al., 2019).
- **Aprendizaje auto-supervisado** Este aprendizaje permite a los modelos aprender representaciones útiles sin etiquetas explícitas, formulando tareas auxiliares (como predecir partes de una imagen o el orden de secuencias). Este tipo de preentrenamiento ha demostrado ser muy eficaz como paso previo al entrenamiento supervisado en conjuntos de datos pequeños (Z. Zhou et al., 2019).

Dentro de estos tipos de aprendizaje, dos enfoques han adquirido especial relevancia en el análisis de imagen médica: el ML y el DL.

Ambos enfoques permiten a los sistemas aprender de los datos para abordar tareas complejas en imagen médica, pero difieren en cómo procesan la información. El ML tradicional se basa en una etapa previa de ingeniería de características, en la que expertos definen manualmente las variables más relevantes que luego se utilizan como entrada para los modelos. Este enfoque ofrece interpretabilidad, pero su rendimiento depende en gran medida de la calidad de las características seleccionadas.

En contraste, el DL utiliza redes neuronales profundas que extraen automáticamente representaciones jerárquicas y abstractas directamente a partir de los píxeles de una imagen médica. Esta capacidad elimina la necesidad de definir manualmente las características, pero conlleva

una elevada carga computacional, una alta complejidad de hiperparámetros y una necesidad mucho mayor de datos etiquetados para evitar sobreajuste y garantizar una generalización adecuada.

Este documento aborda los fundamentos y principales algoritmos de ML y DL desde una perspectiva técnica y aplicada. Se busca ofrecer una visión estructurada y crítica sobre su impacto en el análisis de imágenes médicas.

2.7.1 Aprendizaje automático (*Machine Learning*)

El ML agrupa un conjunto de técnicas que permiten construir modelos que capturan patrones estadísticos complejos dentro de los datos de entrenamiento, de forma que puedan generalizar ese conocimiento a datos nuevos. Su eficacia depende, en gran medida, de la calidad de los datos y de la adecuada selección de características relevantes que representen el problema (Bishop y Nasrabadi, 2006).

A continuación, se describen los algoritmos de ML utilizados en esta tesis doctoral. La elección responde a dos criterios principales: su uso frecuente en el ámbito biomédico y la complementariedad de sus enfoques. Así, se incluyen desde modelos sencillos y muy interpretables, como la regresión logística, hasta métodos capaces de capturar relaciones más complejas, como las máquinas de vectores soporte y los árboles de decisión (y sus variantes en ensamblado). De este modo, se ofrece una visión equilibrada entre interpretabilidad, flexibilidad y rendimiento predictivo, aspectos importantes cuando se busca trasladar estas técnicas al contexto clínico.

Regresión logística

La regresión logística es uno de los modelos más sencillos y ampliamente utilizados para problemas de clasificación binaria. A diferencia de la regresión lineal, que estima una variable continua, la regresión logística estima la probabilidad de pertenencia a una clase a partir de una combinación lineal de variables independientes (Hosmer Jr et al., 2013).

Matemáticamente, se basa en la función sigmoide:

$$P(y = 1 | \mathbf{x}) = \frac{1}{1 + e^{-(\boldsymbol{\beta}^\top \mathbf{x})}}, \quad \text{donde } \boldsymbol{\beta}^\top \mathbf{x} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p$$

Esta función transforma cualquier valor real en un rango entre 0 y 1, interpretado como una probabilidad. Si el resultado supera un umbral (habitualmente 0,5), se asigna la clase positiva.

Aunque su capacidad de modelado es limitada en problemas no lineales, su simplicidad e interpretabilidad la hacen muy útil en entornos clínicos, donde se valora la transparencia del modelo. Además, permite analizar la influencia individual de cada variable predictora sobre la probabilidad del evento, lo que facilita su aplicación en estudios biomédicos (Steyerberg y Vergouwe, 2014).

Máquinas de vectores soporte (*Support Vector Machines*)

Las **máquinas de vectores soporte** (*Support Vector Machines*, SVM) son algoritmos supervisados de clasificación ampliamente utilizados en problemas de análisis de datos. Su objetivo fundamental es encontrar un hiperplano óptimo que divida el espacio de características de forma que las clases queden separadas con el mayor margen posible. Este margen se define como la distancia entre el hiperplano y los puntos más cercanos de cada clase, conocidos como *vectores soporte*. Cuanto mayor sea este margen, mejor será la capacidad del modelo para generalizar a nuevos datos no vistos. Esto está representado en la Figura 2.8 (Cortes y Vapnik, 1995).

Formalmente, si se tiene un conjunto de entrenamiento (x_i, y_i) con $x_i \in \mathbb{R}^n$ y $y_i \in \{-1, 1\}$, el objetivo es encontrar un vector \mathbf{w} y un sesgo b tal que el hiperplano definido por la ecuación $\mathbf{w} \cdot \mathbf{x} + b = 0$ separe las clases y maximice el margen, que es igual a $\frac{2}{\|\mathbf{w}\|}$. Esta optimización se formula como un problema cuadrático convexo con restricciones, resoluble mediante técnicas de optimización dual (Smola y Schölkopf, 2004).

En muchos casos reales, los datos no son completamente linealmente separables. Para abordar esto, las SVM introducen términos de holgura que permiten violaciones del margen, penalizando estos errores en la función objetivo. Además, se introduce un parámetro de regularización que controla el equilibrio entre maximizar el margen y minimizar los errores de clasificación (Bishop y Nasrabadi, 2006).

Cuando no es posible separar las clases en el espacio original, las SVM pueden utilizar la llamada técnica del núcleo o (*kernel trick*), que permite proyectar los datos a un espacio de mayor dimensión en el que la separación sea factible. Esta transformación se realiza de forma implícita mediante funciones núcleo, entre las que destacan el kernel lineal, el polinómico y el de base radial, según la naturaleza del problema (Schölkopf y Smola, 2002).

Gracias a esta capacidad de adaptación a relaciones no lineales complejas, las SVM son especialmente potentes en contextos con conjuntos de datos pequeños o de alta dimensionalidad. Su formulación teórica robusta y su capacidad de control sobre la complejidad del modelo permiten una buena generalización, incluso en presencia de ruido (Cyran et al., 2013).

No obstante, las SVM presentan también algunas limitaciones prácticas. Por un lado, su entrenamiento puede volverse computacionalmente costoso en conjuntos de datos grandes, ya que requiere resolver un problema de optimización con tantas restricciones como muestras haya. Por otro lado, la interpretabilidad de los modelos se reduce notablemente cuando se utilizan kernels no lineales, lo que dificulta su aplicación en contextos donde se requiere explicar el razonamiento del modelo.

Árboles de decisión

Los árboles de decisión son modelos jerárquicos que dividen recursivamente el espacio de características en regiones homogéneas en cuanto a la variable objetivo (Kingsford y Salzberg, 2008). En cada nodo interno del árbol, se selecciona una característica y un umbral de corte que permiten separar los datos.

Su estructura es intuitiva, donde las decisiones se representan mediante nodos, y las posibles

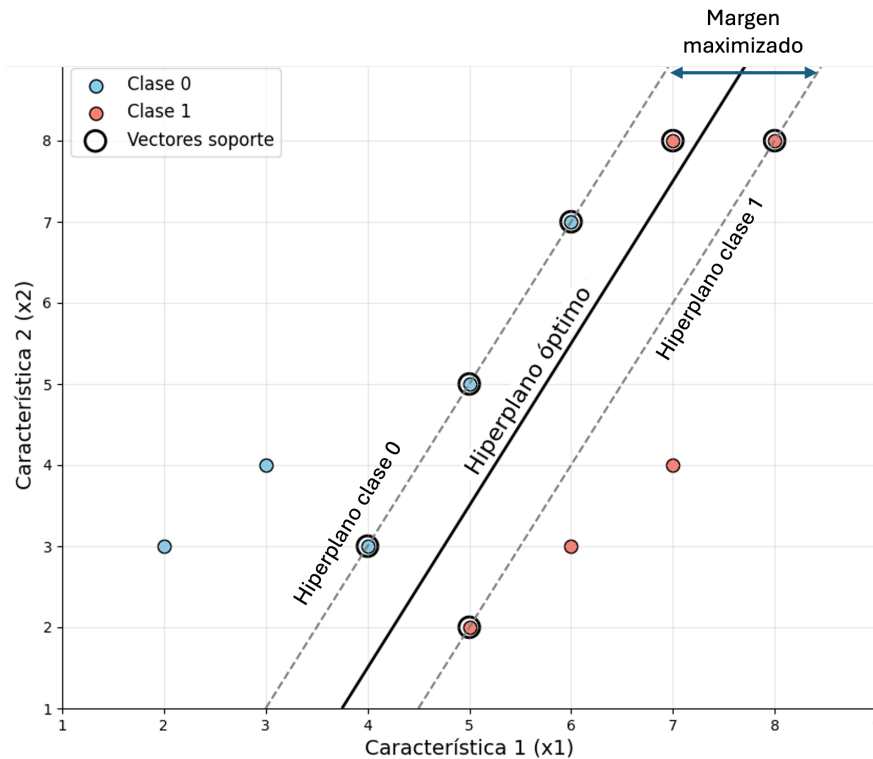


Figura 2.8: Visualización del clasificador SVM lineal: hiperplanos, margen maximizado y vectores soporte

salidas mediante ramas, lo que facilita la interpretación por parte de usuarios no expertos. En el contexto clínico, esto puede ser una gran ventaja, ya que permite visualizar de manera clara cómo se llega a una determinada predicción (Larranaga et al., 2006).

No obstante, los árboles de decisión individuales tienden a sobreajustar los datos de entrenamiento, especialmente si no se aplican técnicas de poda o control de profundidad. Por ello, en muchos casos se recurre a métodos de ensamblado que combinan múltiples árboles para mejorar la estabilidad y el rendimiento predictivo, tal y como se muestra en la Figura 2.9 (Rokach, 2010). Entre los árboles de decisión más utilizados se encuentran Bosques Aleatorios o *Random Forest* (RF), *Extremely Randomized Trees* (ExtraTrees) y *Extreme Gradient Boosting* (XGBoost), cada uno con características particulares:

- **RF** es un algoritmo de ensamblado basado en árboles de decisión. Su funcionamiento se basa en entrenar múltiples árboles sobre subconjuntos aleatorios de los datos y promediar sus predicciones o aplicar una votación mayoritaria (Breiman, 2001).

Además, para aumentar la diversidad entre los árboles, en cada nodo se selecciona un subconjunto aleatorio de características sobre las que se evalúa la mejor división. Esta aleatorización adicional reduce la correlación entre árboles y mejora la generalización del modelo.

RF presenta varias ventajas que lo hacen especialmente adecuado para su aplicación en biomedicina. Es capaz de manejar datos con alta dimensionalidad, lo que resulta útil

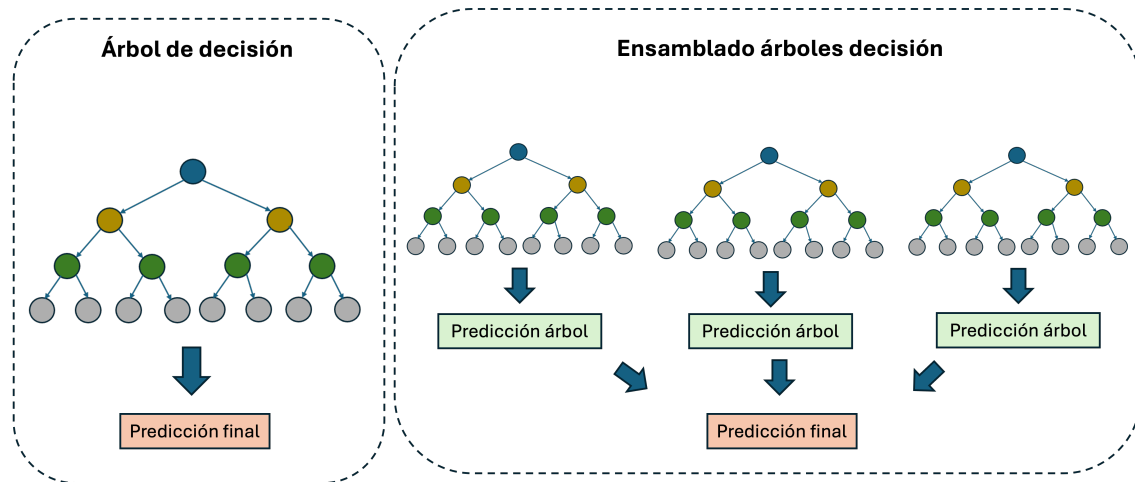


Figura 2.9: Árboles de decisión: comparación entre modelo único y técnicas de ensamblado (RF, ExtraTrees, XGBoost)

en contextos donde se dispone de un gran número de variables. Además, es robusto al sobreajuste y tolera bien tanto el ruido como la presencia de valores perdidos en los datos. Otra característica destacada es su capacidad para estimar la importancia de cada variable en la predicción, lo que facilita la interpretación clínica (Cutler et al., 2007). Gracias a estas propiedades, RF se ha consolidado como uno de los algoritmos más empleados en el desarrollo de biomarcadores de imagen, la selección de variables y la construcción de modelos clínicos predictivos (Bzdok y Meyer-Lindenberg, 2018).

- **ExtraTrees** es una variante del algoritmo RF que introduce aún más aleatoriedad en el proceso de construcción de los árboles. Al igual que RF, se trata de un método de ensamblado basado en múltiples árboles de decisión, pero con dos diferencias clave. En primer lugar, no utiliza un muestreo aleatorio con reemplazo del conjunto de datos, sino que entrena cada árbol utilizando todo el conjunto de entrenamiento o un subconjunto fijo, sin replicación. En segundo lugar, introduce una mayor aleatoriedad en la construcción de los árboles: en cada nodo, en lugar de buscar la mejor división posible entre un subconjunto de características, selecciona aleatoriamente tanto la característica como el umbral de corte, lo que puede mejorar la generalización en ciertos escenarios y reducir el sobreajuste, especialmente en conjuntos de datos ruidosos (Geurts et al., 2006). Aunque esta aleatorización adicional puede hacer que el modelo sea algo menos interpretativo, en la práctica suele ofrecer un rendimiento comparable o incluso superior al de RF, con tiempos de entrenamiento más cortos.
- **XGBoost** es una implementación eficiente y optimizada del algoritmo de *gradient boosting*, una técnica de ensamblado que combina múltiples árboles de decisión secuencialmente. A diferencia de RF, donde los árboles se entrenan en paralelo e independientemente, en XGBoost cada nuevo árbol intenta corregir los errores del anterior, minimizando una función de pérdida mediante gradientes (T. Chen y Guestrin, 2016).

Este enfoque permite construir modelos altamente expresivos y precisos, con un fuerte rendimiento en tareas de clasificación, regresión y predicción clínica. XGBoost incorpora

además mecanismos de regularización (L1 y L2), lo que lo hace más resistente al sobreajuste.

2.7.2 Aprendizaje profundo (*Deep Learning*)

El DL es una rama del ML que se caracteriza por el uso de redes neuronales artificiales con múltiples capas jerárquicas. Estas redes son capaces de aprender representaciones complejas directamente a partir de los datos, sin necesidad de una ingeniería manual previa de características. Esta capacidad las hace especialmente potentes para trabajar con imágenes médicas, donde los patrones relevantes suelen ser sutiles, multidimensionales y difíciles de definir explícitamente (Litjens et al., 2017; X. Liu et al., 2021).

En el núcleo del DL se encuentran las redes neuronales artificiales (*Artificial Neural Networks*, ANN), que constituyen el fundamento matemático y computacional de estos modelos. Las ANN están inspiradas en la estructura y funcionamiento de las neuronas biológicas y se constituyen por un conjunto de unidades computacionales llamadas neuronas artificiales, organizadas en capas y conectadas entre sí mediante pesos sinápticos que modulan la transmisión de información.

La **neurona artificial**, unidad básica de procesamiento de estas redes, puede entenderse como un sistema que recibe múltiples señales de entrada. En este contexto, las entradas corresponden a las variables que describen cada ejemplo: desde valores simples como la edad o una medida clínica, hasta intensidades de píxeles o características radiómicas extraídas de una imagen. Cada entrada tiene una importancia relativa expresada mediante un peso asociado. La neurona realiza una combinación lineal ponderada de todas estas entradas, añade un término independiente o sesgo (*bias*), y finalmente aplica una función no lineal, denominada función de activación, para producir una salida.

En este contexto, para una neurona con entradas $\mathbf{x} = (x_1, x_2, \dots, x_n)$, pesos $\mathbf{w} = (w_1, w_2, \dots, w_n)$ y sesgo b , la salida y se define como:

$$y = f(z)$$

donde

$$z = \sum_{i=1}^n w_i x_i + b$$

representa la combinación lineal y $f(\cdot)$ es la función de activación.

Este proceso, aunque sencillo, es la base para construir sistemas capaces de aproximar funciones complejas y no lineales, ya que la función de activación introduce la no linealidad necesaria para romper la limitación de las combinaciones lineales.

Funciones de activación

La **función de activación** es un componente esencial en las redes neuronales ya que permite a la red aprender y modelar relaciones no lineales entre las variables de entrada y salida.

Algunas funciones de activación clásicas y ampliamente utilizadas se representan en la Figura 2.10 (Lederer, 2021; Szandała, 2020).

Entre las funciones clásicas más empleadas se encuentran:

Las más utilizadas son:

- **Sigmoide**: transforma los valores en un rango entre 0 y 1, útil cuando se quiere interpretar la salida como probabilidad.

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

- **Tangente hiperbólica (tanh)**: similar a la sigmoide, pero sus valores van de -1 a 1, lo que centra la salida en torno a cero.

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

- **Escalón**: devuelve 0 o 1 según el signo de la entrada. Fue la primera función usada en perceptrones, pero hoy apenas es utilizada.

$$f(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

- **ReLU (*Rectified Linear Unit*)**: muy usada en redes profundas por su sencillez y eficiencia. Devuelve 0 para valores negativos y el propio valor cuando es positivo.

$$f(x) = \max(0, x)$$

Estructura y Entrenamiento de Redes Neuronales Artificiales

La estructura básica de una red neuronal artificial está compuesta por un conjunto de capas organizadas de manera secuencial, representado en la Figura 2.11 y detallado a continuación:

- **Capa de entrada**: recibe las variables o características del problema. En el contexto de la imagen médica, por ejemplo, estas características pueden corresponder a los valores de intensidad de píxeles o vóxeles.
- **Capas ocultas**: una o varias capas de neuronas que procesan la información de forma jerárquica y progresiva. En estas capas, cada neurona puede estar conectada a todas las neuronas de la capa anterior (red totalmente conectada) o solo a un subconjunto de ellas, según la arquitectura diseñada.
- **Capa de salida**: proporciona la respuesta final del modelo, que puede ser la probabilidad asignada a una clase en tareas de clasificación o la predicción de una variable continua en problemas de regresión.

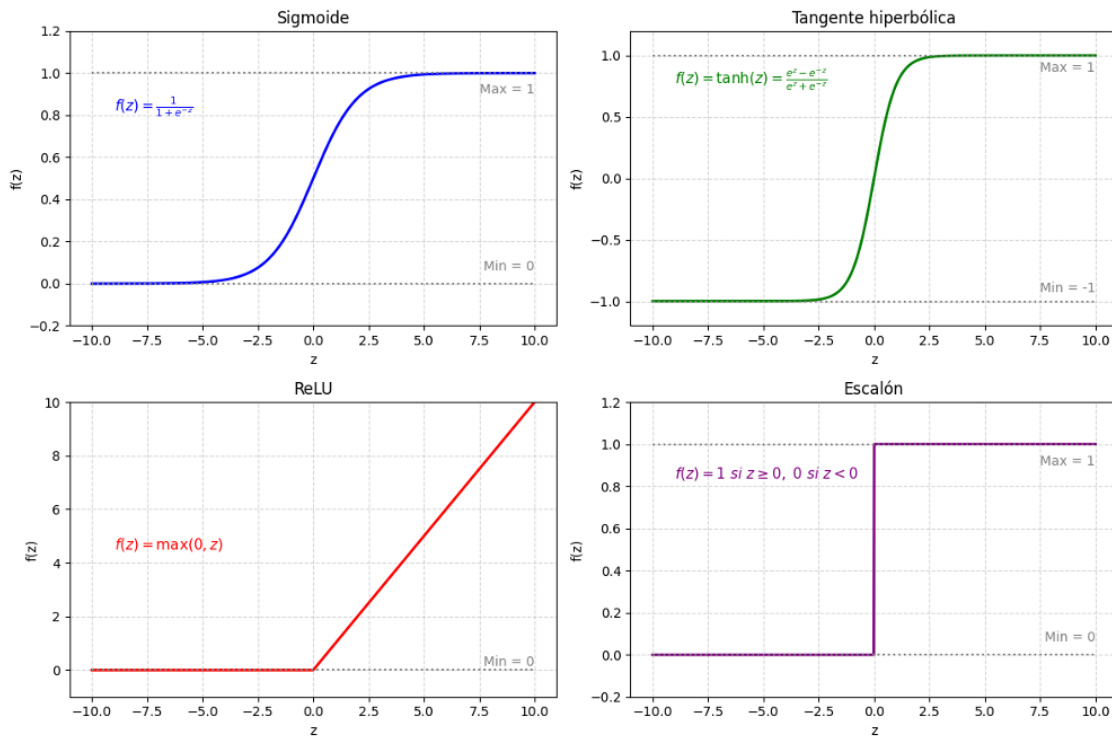


Figura 2.10: Funciones de activación comunes en redes neuronales artificiales

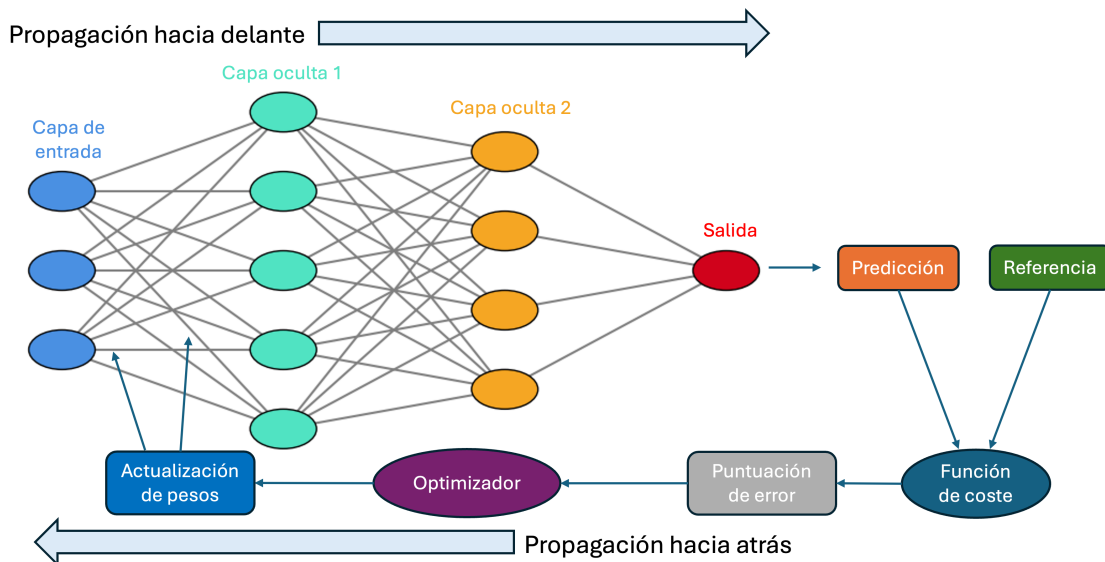


Figura 2.11: La red neuronal procesa la información a través de capas (entrada, ocultas y salida), calcula la predicción, evalúa el error con una función de pérdida y ajusta los pesos mediante retropropagación y tasa de aprendizaje. Este ciclo corresponde a una época de entrenamiento.

Dos hiperparámetros fundamentales que condicionan la capacidad representacional y el poder generalizador del modelo son la **profundidad**, que se refiere al número de capas ocultas, y la **anchura**, que corresponde al número de neuronas en cada capa.

El entrenamiento de una red neuronal se basa en un proceso iterativo que busca optimizar los parámetros internos (pesos y sesgos) para minimizar el error cometido en la predicción. Este proceso consta de tres fases principales: propagación hacia adelante (*forward propagation*), retropropagación del error (*backpropagation*) y actualización de parámetros mediante optimización. A continuación, se describen en detalle:

1. Propagación hacia adelante (Forward propagation):

El vector de entrada \mathbf{x} se propaga a través de la red desde la capa de entrada hasta la capa de salida. En cada capa l , se calcula la activación $\mathbf{a}^{(l)}$ como:

$$\mathbf{a}^{(l)} = f(\mathbf{W}^{(l)}\mathbf{a}^{(l-1)} + \mathbf{b}^{(l)})$$

donde $\mathbf{a}^{(l-1)}$ es el vector de activaciones de la capa anterior (o la entrada \mathbf{x} para $l = 1$), $\mathbf{W}^{(l)}$ la matriz de pesos, $\mathbf{b}^{(l)}$ el vector de sesgos, y $f()$ es la función de activación. La salida final de la red, $\hat{y} = \mathbf{a}^{(L)}$, representa la predicción generada por la red neuronal.

2. Cálculo de la función de pérdida:

La diferencia entre la predicción \hat{y} y la etiqueta real y se cuantifica mediante una función de pérdida $\mathcal{L}(\hat{y}, y)$, que puede ser:

- *Error cuadrático medio (MSE, de sus siglas en inglés Mean Squared Error)*: para tareas de regresión, definida como

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2$$

- *Entropía cruzada (cross-entropy)*: para clasificación, que mide la distancia entre distribuciones de probabilidad reales y predichas.

3. Retropropagación del error (Backpropagation):

Se calculan los gradientes de la función de pérdida con respecto a cada parámetro $\mathbf{W}^{(l)}$ y $\mathbf{b}^{(l)}$ utilizando la regla de la cadena del cálculo diferencial:

$$\frac{\partial \mathcal{L}}{\partial \mathbf{W}^{(l)}}, \quad \frac{\partial \mathcal{L}}{\partial \mathbf{b}^{(l)}}$$

Estos gradientes indican cómo modificar los pesos y sesgos para reducir el error, propagando el gradiente desde la salida hacia las capas iniciales de forma eficiente.

4. Actualización de parámetros (optimización):

Los parámetros se ajustan mediante un algoritmo de optimización, típicamente basado en descenso del gradiente, según la regla:

$$\theta_{t+1} = \theta_t - \eta \frac{\partial \mathcal{L}}{\partial \theta_t}$$

donde θ representa un peso o sesgo, η la tasa de aprendizaje y t la iteración actual. Existen variantes adaptativas como Adam, RMSprop o Adagrad que mejoran la convergencia y estabilidad.

Este ciclo iterativo (propagación hacia adelante, cálculo de pérdida, retropropagación y actualización) se repite durante numerosas épocas, procesando lotes de datos (*batches*) para optimizar la función de pérdida y ajustar el modelo. El objetivo es que la red aprenda representaciones útiles para generalizar correctamente a nuevos datos.

Es importante considerar que la profundidad y anchura de la red, así como las funciones de activación y la elección del algoritmo de optimización, influyen directamente en la capacidad del modelo para capturar patrones complejos y evitar problemas como el sobreajuste, el desvanecimiento o explosión de gradientes, los cuales se abordan mediante técnicas de regularización y arquitecturas especializadas (X. Liu et al., 2021; Salahuddin et al., 2022).

Redes Neuronales Convolucionales

Las redes neuronales convolucionales (CNNs, por sus siglas en inglés *Convolutional Neural Networks*) constituyen un tipo de arquitecturas dentro del DL, diseñadas principalmente para el procesamiento y análisis de datos con estructura espacial, como las imágenes (LeCun et al., 2002). A diferencia de las redes neuronales artificiales tradicionales, en las que cada neurona está completamente conectada con todas las neuronas de la capa anterior y posterior, las CNNs aprovechan la estructura espacial de los datos visuales aplicando filtros convolucionales y técnicas de submuestreo que permiten extraer patrones jerárquicos de forma eficiente, permitiendo la extracción jerárquica de características espaciales con alta eficiencia computacional (Rawat y Wang, 2017).

El principio fundamental que sustenta las CNNs es la convolución, una operación matemática que permite aplicar un filtro o *kernel* sobre una imagen para extraer patrones (bordes, texturas, formas, etc.) conservando la correlación espacial entre los píxeles. Estos filtros se aprenden automáticamente durante el proceso de entrenamiento. En las primeras capas, los filtros tienden a detectar estructuras simples (por ejemplo, bordes horizontales o verticales), mientras que en capas más profundas, las representaciones se vuelven progresivamente más abstractas y semánticamente complejas (formas anatómicas, lesiones, órganos) (Goodfellow et al., 2016). El funcionamiento de esta operación se representa en la Figura 2.12, donde se muestra cómo un kernel de 3×3 se desliza sobre una imagen de entrada para generar un mapa de activación. Cada valor resultante refleja la intensidad de la presencia del patrón representado por el filtro en esa región.

Una red convolucional típica está formada por una secuencia de capas con distintas funciones:

- **Capas convolucionales:** aplican múltiples kernels de tamaño reducido (por ejemplo, 3×3 o 5×5) que se deslizan sobre la imagen para generar mapas de activación. Un mapa de activación es la nueva imagen que se obtiene tras aplicar un filtro a la entrada. En él, cada valor indica en qué medida aparece el patrón que busca ese filtro (como un borde o una textura) en cada zona.
- **Funciones de activación no lineales:** tras la convolución, se aplica una función de

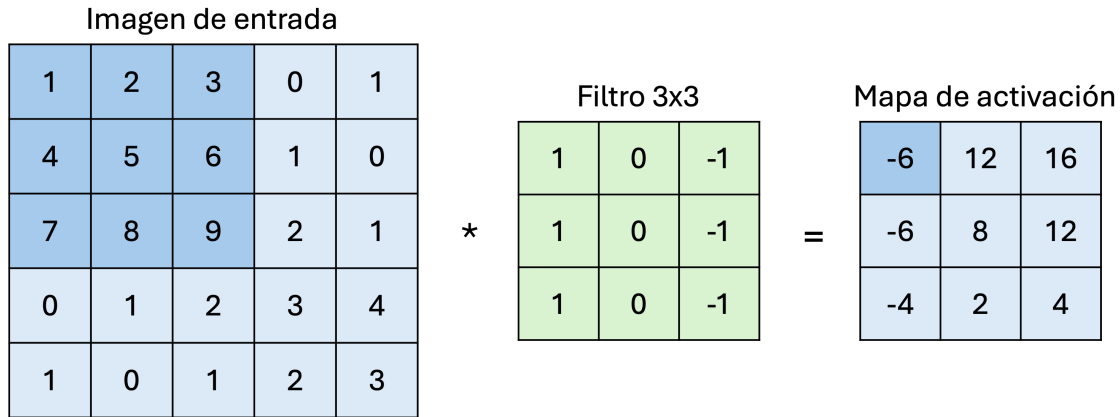


Figura 2.12: Operación de convolución: un filtro 3×3 se aplica localmente sobre la imagen para generar un mapa de activación.

activación para introducir capacidad de modelado no lineal. Esta operación ayuda a que la red pueda aprender funciones más complejas que la mera combinación lineal de los datos.

- **Capas de pooling (submuestreo):** reducen la resolución espacial de los mapas de activación conservando la información más relevante. El tipo más común es el *max pooling*, que selecciona el valor máximo dentro de una ventana local. Al conservar solo la información más significativa en cada región, se consigue una representación más robusta a pequeñas variaciones espaciales y se disminuye el coste computacional. El desplazamiento de esta ventana a lo largo del mapa de activación está determinado por el *stride*, un parámetro que indica cuántos píxeles se avanza en cada paso del barrido. La Figura 2.13 representa el funcionamiento del max pooling.

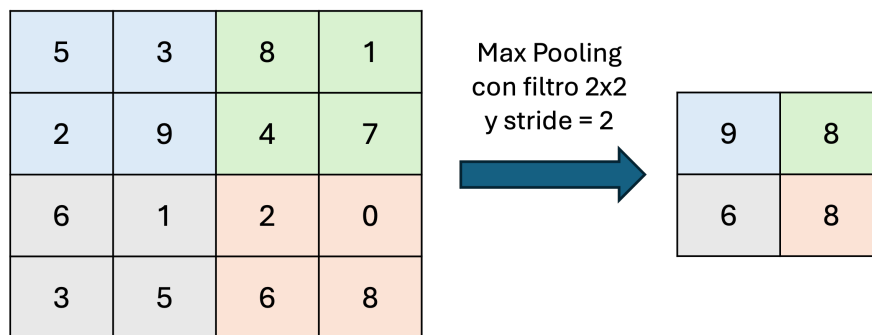


Figura 2.13: Ejemplo de *max pooling* 2×2 con *stride* 2: en cada región local se conserva el valor máximo, reduciendo así la resolución del mapa de activación.

- **Capas completamente conectadas (*fully connected layers*):** aunque se utilizan en la etapa final de muchas CNNs clásicas para realizar la clasificación o regresión, en los últimos años han sido progresivamente reemplazadas o minimizadas en favor de arquitecturas más eficientes, especialmente en tareas de segmentación o detección donde se requiere preservar la resolución espacial (Long et al., 2015).

Una de las propiedades clave de las CNNs es su capacidad para compartir pesos, lo que reduce drásticamente el número de parámetros del modelo en comparación con las redes densas. Este aspecto no solo disminuye el riesgo de sobreajuste, sino que permite escalar la arquitectura a imágenes de gran tamaño.

La Figura 2.14 muestra una representación esquemática del flujo típico de una CNN, desde la imagen de entrada hasta la salida del modelo. En ella se observa cómo las capas convolucionales, de activación, submuestreo y totalmente conectadas se integran de manera secuencial para construir una representación jerárquica de los datos.

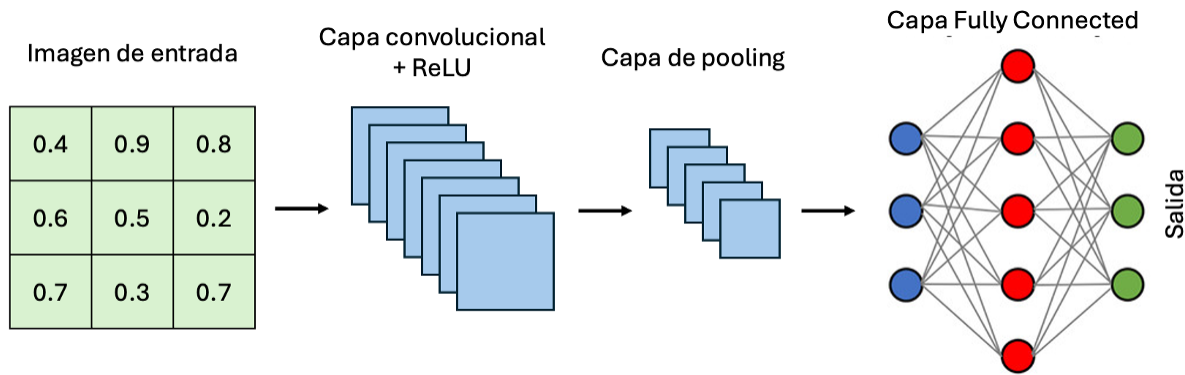


Figura 2.14: Esquema del flujo de datos a través de una CNN

Esquema general de una CNN: la imagen de entrada pasa por capas convolucionales (con filtros 3×3), funciones de activación ReLU, capas de pooling, y finalmente capas densas para producir la salida de clasificación.

A lo largo de los últimos años, se han desarrollado múltiples variantes y extensiones de las CNNs, adaptadas a tareas específicas. Por ejemplo, las redes totalmente convolucionales (*Fully Convolutional Networks*) permiten realizar segmentaciones pixel a pixel sin capas densas finales (Long et al., 2015), mientras que arquitecturas como U-Net han sido diseñadas específicamente para segmentación médica (Ronneberger et al., 2015).

Arquitectura U-Net

La arquitectura U-Net es un estándar en segmentación médica y de las contribuciones más influyentes al campo de la segmentación. Fue introducida por (Ronneberger et al., 2015), originalmente diseñada para la segmentación de imágenes de microscopía biomédica. Desde su publicación, U-Net se ha convertido en un referente dentro del DL aplicado al ámbito médico, dada su capacidad para segmentar estructuras con alta precisión incluso en escenarios con datos escasos y variabilidad morfológica significativa.

La arquitectura U-Net (Figura 2.15) recibe su nombre de su estructura en forma de “U”, que consta de dos caminos principales: un camino de contracción (*encoder*) y un camino de expansión (*decoder*), unidos mediante conexiones puente (*skip connections*). Esta forma simétrica permite combinar eficientemente la información semántica global con los detalles espaciales locales, algo esencial en tareas de segmentación donde se requiere tanto comprensión contextual como precisión en los contornos (Siddique et al., 2020).

1. **Camino de contracción (*encoder*):** El encoder sigue una estructura típica de red convolucional, con bloques que combinan capas convolucionales, seguidas de funciones de activación no lineales y operaciones de submuestreo (*max pooling*). Este camino actúa como extractor de características jerárquicas, reduciendo progresivamente la resolución espacial mientras se preserva y amplifica la información de alto nivel.
2. **Camino de expansión (*decoder*):** El decoder revierte el proceso del encoder mediante capas de *upsampling* (normalmente convoluciones transpuestas o interpolaciones seguidas de convoluciones) que aumentan la resolución espacial de las representaciones intermedias. Cada bloque del decoder recibe, además de la información de la capa anterior y la salida correspondiente del encoder a través de una conexión puente.
3. **Conexiones puente (*skip connections*):** Una de las innovaciones clave de U-Net es el uso sistemático de estas conexiones entre capas simétricas del encoder y decoder. Gracias a ellas, la red puede acceder simultáneamente a información de contexto (del *encoder* profundo) y a detalles locales de alta resolución (del *encoder* superficial), lo que mejora la localización precisa de los objetos segmentados. Para ello se realiza una fusión (normalmente mediante concatenación de canales) entre los bloques del *encoder* y *decoder* al mismo nivel, permitiendo que la red recupere detalles espaciales finos que podrían haberse perdido durante el proceso de submuestreo, y por lo tanto mejorando la precisión de los bordes segmentados.
4. **Salida final:** La última capa de U-Net es una convolución 1×1 que proyecta los canales intermedios en un mapa de probabilidad por clase para cada píxel (en 2D) o vóxel (en 3D). Este mapa puede posteriormente umbralizarse o refinarse para obtener una máscara binaria o multicategoría.

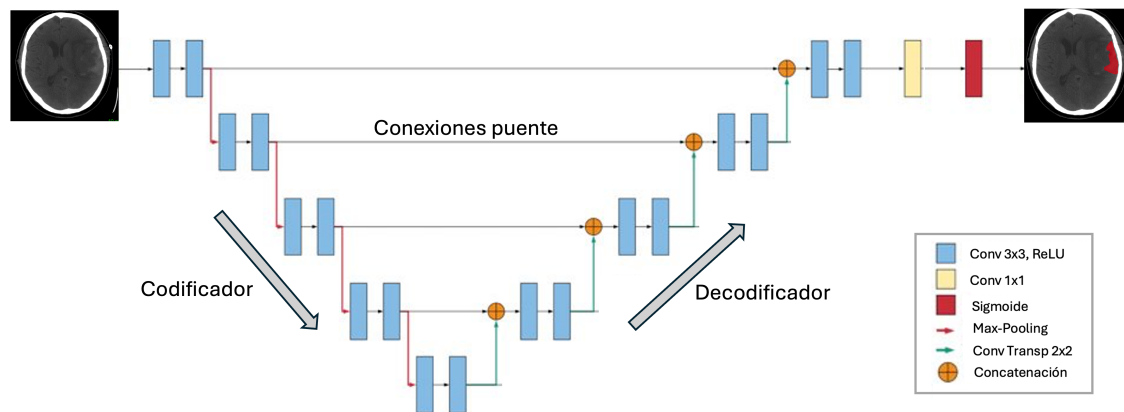


Figura 2.15: Arquitectura U-Net: estructura simétrica con codificador-decodificador y conexiones puente

Desde su propuesta original, U-Net no solo se ha consolidado como un modelo clave en segmentación médica, sino que ha dado lugar a numerosas variantes que buscan adaptarse mejor a distintos contextos clínicos. Por ejemplo, la *Residual U-Net* integra bloques residuales inspirados en *ResNet*, lo que permite entrenar redes más profundas sin que se degrade el rendimiento. La *Attention U-Net*, por su parte, incorpora mecanismos de atención que ayudan

al modelo a centrarse en las regiones más relevantes de la imagen, algo especialmente útil cuando las estructuras de interés son pequeñas o están parcialmente ocultas. Otras evoluciones, como *U-Net++* y *UNet3+*, introducen conexiones más densas entre el encoder y el decoder, facilitando la fusión de características a múltiples escalas y enriqueciendo la representación semántica (Siddique et al., 2020). Entre las propuestas más influyentes derivadas de la arquitectura U-Net se encuentra *nnU-Net* (*no-new U-Net*). Su relevancia no se basa en introducir una arquitectura novedosa, sino en plantear la segmentación automática como un sistema autoregulado. En lugar de requerir un ajuste manual intensivo, *nnU-Net* adapta de forma automática parámetros clave (como la profundidad de la red, el tamaño de entrada o la estrategia de normalización) a las características concretas del conjunto de datos. Asimismo, permite seleccionar el tipo de arquitectura de entrenamiento (2D, 3D o en cascada), lo que facilita su aplicación en diferentes escenarios y con distintos recursos computacionales.

Este enfoque basado en la estandarización y en la configuración adaptativa ha convertido a *nnU-Net* en un punto de referencia en segmentación médica. Ha mostrado un rendimiento competitivo en retos internacionales como el *Medical Segmentation Decathlon* y ha evidenciado capacidad de generalización en diversas modalidades de imagen (TC, RM) y órganos, lo que ha favorecido su uso extendido en investigación y, en ciertos casos, en entornos clínicos (Antonelli et al., 2022; Isensee et al., 2021).

2.7.3 Aplicaciones de la Inteligencia Artificial en neurociencia clínica

Las aplicaciones más consolidadas de la IA en imagen médica pueden agruparse en cuatro grandes áreas funcionales:

- **Segmentación automática de estructuras anatómicas y lesiones.** La segmentación automática consiste en la delineación precisa de estructuras anatómicas y regiones patológicas en imágenes médicas, facilitando el análisis cuantitativo y la planificación clínica. Las CNN, como U-Net (Ronneberger et al., 2015), han demostrado ser altamente efectivas en tareas de segmentación en neuroimagen, gracias a su arquitectura en forma de encoder-decoder con conexiones puente. Más recientemente, *nnU-Net* (Isensee et al., 2021) ha emergido como un marco autoajutable que adapta automáticamente los hiperparámetros de red, el preprocesamiento y la estrategia de entrenamiento a las características del conjunto de datos, alcanzando resultados de referencia en numerosos retos de segmentación. Este enfoque ha mostrado un rendimiento robusto en la segmentación de estructuras cerebrales, tumores y hemorragias (Cai et al., 2020; Isensee et al., 2021; Rusche et al., 2023).
- **Clasificación automatizada de tejidos.** Los modelos de IA también se emplean para la clasificación de tejidos, permitiendo discriminar entre diferentes tipos celulares, entidades tumorales y patrones patológicos. Algoritmos basados en ML clásico (e.g., SVM, Random Forest) y CNN han sido entrenados para diferenciar GBMs de alto grado de gliomas de bajo grado con alta precisión (Bhatele y Bhadauria, 2022; Özcan et al., 2021). En el entorno quirúrgico, la HSI ha permitido avanzar hacia una clasificación tisular en tiempo real. Fabelo et al. (Fabelo, Ortega, Lazcano et al., 2018) demostraron

que el análisis espectral intraoperatorio asistido por IA es capaz de identificar márgenes entre tejido sano y tumoral, con el objetivo de guiar la cirugía y poder optimizar el procedimiento sin comprometer estructuras funcionales.

- **Cuantificación y seguimiento longitudinal.** El análisis longitudinal mediante IA permite cuantificar de forma precisa y reproducible los cambios cerebrales a lo largo del tiempo, siendo clave en el seguimiento de enfermedades neurodegenerativas, lesiones traumáticas o respuesta a terapias. Aunque enfoques clásicos como el registro longitudinal simétrico (Reuter et al., 2012) han sido fundamentales para alinear series temporales de un mismo paciente, desarrollos más recientes han incorporado arquitecturas profundas que modelan explícitamente la dimensión temporal, mejorando la coherencia espacial y reduciendo la variabilidad entre adquisiciones (Iglesias y et al., 2021).

Entre estos avances, destacan las redes siamesas (Chicco, 2021), que permiten emparejar directamente regiones o lesiones a través de diferentes momentos temporales. Estas redes comparan pares de imágenes longitudinales aprendiendo representaciones latentes que capturan similitudes morfológicas y espaciales, incluso ante cambios sutiles o desplazamientos anatómicos. Por ejemplo, Koochaki et al. (Koochaki y Najafizadeh, 2024) implementaron un modelo basado una Red Siamesa Convolutiva para monitorizar lesiones tras traumatismo craneoencefálico leve, con una precisión muy alta en la clasificación de las lesiones, superando las herramientas básicas de registro.

- **Predicción del pronóstico clínico.** La incorporación de la IA en la predicción pronóstica permite desarrollar modelos capaces de anticipar desenlaces clínicos a partir de información cuantitativa extraída directamente de las imágenes médicas. Las características morfométricas (como volúmenes, formas o patrones de distribución anatómica) obtenidas mediante segmentación automática, han mostrado un valor pronóstico relevante en múltiples contextos clínicos (Lagares et al., 2015; Ocaña-Tienda et al., 2024; G. Rao et al., 2023). A ello se suma la radiómica, que permite la extracción sistemática de variables cuantitativas (incluyendo descriptores texturales, estadísticos y geométricos) a partir de regiones de interés segmentadas, capturando patrones complejos y no evidentes al ojo humano. Estudios como el de Lao et al. (Lao et al., 2017) evidenciaron que modelos basados en radiómica obtenida de RM superaban a los modelos clínicos tradicionales en la predicción de supervivencia en GBM. De forma análoga, investigaciones recientes en hemorragias intracraneales han demostrado que los modelos radiómicos aplicados a TC pueden anticipar desenlaces funcionales con una precisión similar o superior a escalas comunmente utilizadas en la práctica clínica (Huang et al., 2022; Pei et al., 2024).

2.8 Casos de estudio: contextualización y relevancia clínica

Los casos de estudio presentados en este capítulo ilustran de forma concreta cómo las tecnologías y metodologías revisadas en el estado del arte pueden aplicarse para abordar desafíos reales en neurociencia clínica. Estos ejemplos responden directamente a la motivación central de esta tesis doctoral (Sección 1.1): desarrollar sistemas automatizados que mejoren la precisión diagnóstica y apoyen la toma de decisiones mediante la integración de técnicas

avanzadas de imagen médica e inteligencia artificial.

Como se expuso en la motivación de este trabajo (Sección 1.1, la práctica clínica actual enfrenta limitaciones significativas relacionadas con la variabilidad interobservador, la carga de trabajo asociada a procesos manuales y la necesidad de herramientas objetivas y reproducibles. En este contexto, los dos casos de estudio seleccionados representan escenarios complementarios que demuestran cómo la automatización y la extracción de información cuantitativa pueden transformar el abordaje diagnóstico y terapéutico:

1. **Clasificación automática de tejidos intraoperatorios mediante imagen hiperespectral:** Este primer caso aborda el desafío de la segmentación manual y la variabilidad en la evaluación visual durante la cirugía cerebral. El proyecto NEMESIS-3D-CM ilustra cómo la HSI, combinada con algoritmos de ML, puede proporcionar información objetiva y en tiempo real que complementa la evaluación visual del neurocirujano, reduciendo la dependencia de la interpretación subjetiva y facilitando una resección más precisa de tumores cerebrales.
2. **Segmentación automática y predicción de pronóstico en hemorragia subaracnoidea aneurismática:** Este segundo caso se centra en el reto de la segmentación manual intensiva en tiempo y la necesidad de biomarcadores cuantitativos para la estratificación pronóstica. Muestra cómo la segmentación automática mediante DL y la extracción de características radiómicas pueden generar información reproducible que supera las limitaciones de las escalas visuales tradicionales, permitiendo una predicción más precisa de desenlaces clínicos.

Ambos casos de estudio comparten un hilo conductor común: la transformación de información visual cualitativa y subjetiva en datos cuantitativos, objetivos y procesables mediante algoritmos de IA. Esta transición desde métodos tradicionales dependientes del observador hacia sistemas automatizados y reproducibles refleja el camino hacia una medicina más precisa y basada en evidencia, tal como se propone en los objetivos de esta tesis (Sección 1.2).

La selección de estos casos no es arbitraria: responden a necesidades clínicas reales identificadas en la práctica neuroquirúrgica y neurocrítica, donde la precisión en la delimitación de lesiones y la anticipación del pronóstico son cruciales para la toma de decisiones terapéuticas. Además, estos ejemplos demuestran cómo diferentes modalidades de imagen (HSI intraoperatoria y TC diagnóstica) pueden integrarse con técnicas de IA para abordar problemas complementarios, desde la guía quirúrgica en tiempo real hasta la estratificación pronóstica en urgencias neurológicas.

A continuación, se presentan ambos casos de estudio en detalle, describiendo el contexto clínico, los desafíos específicos que abordan, las soluciones tecnológicas propuestas y los resultados obtenidos según la literatura disponible. Esta revisión no solo proporciona evidencia del estado actual de estas aplicaciones, sino que también establece el fundamento sobre el cual se desarrollan las contribuciones metodológicas y experimentales de esta tesis doctoral.

2.8.1 Caso de estudio: Clasificación automática de tejidos intraoperatorios (NEMESIS-3D-CM)

El GBM es el tumor cerebral maligno más común y agresivo en adultos. Se trata de una enfermedad especialmente compleja en el ámbito de la neurooncología, tanto por su comportamiento biológico como por las dificultades que plantea su tratamiento. Este tipo de tumor crece de forma rápida, presenta una gran heterogeneidad a nivel molecular y tiende a infiltrarse de manera difusa en el tejido cerebral, lo que dificulta su resección completa y favorece la recaída local (Ostrom et al., 2022).

A pesar de un abordaje terapéutico combinado que incluye cirugía, radioterapia y quimioterapia con temozolomida, el pronóstico sigue siendo muy desfavorable. La supervivencia media apenas alcanza los 12 a 15 meses desde el diagnóstico, y menos del 10 % de los pacientes sobreviven más de cinco años (Stupp et al., 2005). La cirugía juega un papel fundamental en el tratamiento inicial del GBM. No solo busca reducir el tamaño del tumor, sino también mejorar la efectividad de los tratamientos posteriores, como la radioterapia y la quimioterapia. Además, en muchos casos ayuda a aliviar la presión dentro del cráneo o a mejorar síntomas neurológicos causados por la masa tumoral (Brown et al., 2016). Diversos estudios han demostrado que cuanto más completa es la resección del tumor, la posibilidad de supervivencia aumenta, así como la posibilidad de retrasar la progresión de la enfermedad. Sin embargo, conseguir una resección total es complicado porque el GBM crece infiltrándose en el tejido cerebral sano que lo rodea, haciendo difícil distinguir dónde termina el tumor (Neftel et al., 2019). Por eso, si quedan células tumorales en los bordes, es muy probable que el tumor vuelva a aparecer cerca del área operada.

Por esta razón, la identificación precisa de los márgenes tumorales intraoperatorios es uno de los principales retos de la neurocirugía oncológica. La cirugía del GBM exige un equilibrio delicado entre la máxima extensión de resección y la preservación de la función neurológica, especialmente en tumores ubicados en áreas elocuentes. Por ello, se han desarrollado e incorporado múltiples técnicas avanzadas con el objetivo de mejorar la precisión y seguridad de la cirugía.

En este sentido, el diagnóstico temprano es clave para mejorar el pronóstico, pero la identificación exacta del tipo y grado tumoral continúa siendo un desafío. La RM es la técnica más empleada para la detección y caracterización del tumor, ya que permite visualizar diferencias en la composición del tejido cerebral. Sin embargo, durante la cirugía, la RM presenta limitaciones importantes, ya que no puede usarse de forma continua en tiempo real y su resolución espacial y funcional puede ser insuficiente para diferenciar con precisión entre el tejido tumoral infiltrado y el tejido cerebral sano.

La neuronavegación, basada en estas imágenes preoperatorias, es una herramienta fundamental para guiar la resección, facilitando la localización del tumor y sus límites dentro del cerebro. Sin embargo, una vez realizada la craneotomía, se produce el *brain shift*, es decir, el desplazamiento del tejido cerebral tras la apertura del cráneo, la pérdida de líquido cefalorraquídeo o la manipulación quirúrgica. Esto afecta a la precisión de la neuronavegación, lo que hace imprescindible su combinación con otras tecnologías que actualicen la información anatómica en tiempo real.

En este contexto, la fluorescencia con 5-aminolevulínico (5-ALA) ha demostrado ser un avance significativo para mejorar la visualización intraoperatoria del tumor. Al administrarse previamente al paciente, el 5-ALA se metaboliza preferentemente en células tumorales, acumulándose como un compuesto fluorescente que brilla bajo luz ultravioleta o azul (Stummer et al., 2006). Esto permite al cirujano distinguir el tejido tumoral infiltrativo, que puede no ser visible con luz blanca convencional. Numerosos estudios han correlacionado el uso de 5-ALA con un aumento significativo en la tasa de resección completa y, consecuentemente, con una mejor supervivencia (Hadjipanayis y Stummer, 2019; Hadjipanayis et al., 2015). No obstante, el empleo de 5-ALA presenta ciertas limitaciones. Su eficacia se ha demostrado principalmente en gliomas de alto grado, mientras que su utilidad en otros tipos de tumores (como metástasis o gliomas de bajo grado) es más reducida debido a una menor captación de protoporfirina IX (Hefti et al., 2008; Roberts et al., 2011). Además, la intensidad de la fluorescencia puede verse afectada por factores como la vascularización tumoral, el pH tisular o la administración concomitante de determinados fármacos, lo que introduce variabilidad en la interpretación intraoperatoria (Valdés et al., 2011). Asimismo, su implementación requiere un entorno quirúrgico especializado y formación específica del equipo médico para garantizar su efectividad.

En paralelo al 5-ALA, la fluoresceína sódica se ha utilizado como otro agente de fluorescencia intraoperatoria. A diferencia del 5-ALA, su mecanismo de acción se basa en la extravasación a través de la barrera hematoencefálica, lo que permite resaltar áreas de tumor con vascularización anómala o disrupción de dicha barrera (Acerbi et al., 2014). Bajo filtros específicos acoplados al microscopio quirúrgico (habitualmente excitación en el rango de 560–580 nm), la fluoresceína proporciona un contraste inmediato que facilita la delimitación del tejido tumoral frente al parénquima sano. Sus principales ventajas se basan en su bajo coste, amplia disponibilidad y rápida administración intravenosa durante el acto quirúrgico, lo que la convierte en una herramienta práctica incluso en entornos con recursos limitados (Smith et al., 2021).

Sin embargo, la fluoresceína presenta limitaciones relevantes. Su especificidad es inferior a la del 5-ALA, ya que tiende a marcar cualquier área con alteración de la barrera hematoencefálica, incluyendo regiones de edema o inflamación peritumoral (Sutton et al., 2023). Por ello, el neurocirujano debe interpretar con cuidado las áreas fluorescentes para evitar resecciones innecesarias de tejido no patológico. A pesar de esta limitación, estudios recientes sugieren que la fluoresceína puede mejorar el grado de resección en gliomas y metástasis cerebrales, y que su utilidad aumenta cuando se emplea en combinación con técnicas complementarias como la neuronavegación (Catapano et al., 2017).

Por otro lado, la RMi permite obtener imágenes actualizadas durante la cirugía, lo que representa una mejora significativa frente a la neuronavegación basada únicamente en estudios preoperatorios. La RMi ofrece al cirujano información en tiempo real sobre el tumor residual y posibles zonas afectadas que no se aprecian a simple vista (Nimsky et al., 2006). Por lo tanto, puede ser una herramienta que les ayude a decidir si continuar con la resección o dar por finalizada la cirugía, con mayor seguridad y precisión. Sin embargo, su uso conlleva limitaciones. Como se comentaba en la Sección 2.5.2, la RMi requiere quirófanos específicamente diseñados para ello, así como personal altamente entrenado (Senft et al., 2011; Wach et al., 2024). Esto

supone una inversión considerable en infraestructuras y recursos humanos, lo que restringe su disponibilidad a un número reducido de centros hospitalarios.

Entre las tecnologías emergentes aplicadas a la cirugía del GBM, la HSI se ha posicionado como una herramienta prometedora para mejorar la visualización intraoperatoria del tejido tumoral. A diferencia de otras técnicas, la HSI no requiere la administración de contraste ni el contacto directo con el cerebro, lo que la convierte en un método no invasivo, seguro y potencialmente integrable en el flujo quirúrgico sin alterar la dinámica operatoria. El principio de funcionamiento de la HSI se basa en la adquisición de imágenes en un amplio rango del espectro electromagnético, incluyendo longitudes de onda más allá del espectro visible. Esta capacidad permite registrar, en cada píxel de la imagen, una firma espectral única que refleja las propiedades ópticas del tejido cerebral. A través del análisis de esta información, es posible identificar diferencias sutiles en la composición bioquímica y microestructural entre el tejido tumoral y el tejido sano, incluso en zonas donde la infiltración no genera cambios morfológicos evidentes.

Esta sensibilidad espectral abre una vía para detectar márgenes tumorales infiltrativos que, de otro modo, pasarían desapercibidos con técnicas visuales convencionales o incluso con herramientas avanzadas como la neuronavegación basada en RM preoperatoria. En este contexto, la HSI podría aportar información adicional al cirujano, ayudando a decidir en tiempo real hasta dónde es seguro y efectivo resear sin comprometer estructuras funcionales.

Una de sus ventajas más destacadas es su potencial para generar mapas espectrales del campo quirúrgico, que pueden superponerse a imágenes neuronavegadas o integrarse en sistemas de visualización aumentada (Villa et al., 2024). Esto permitiría una resección más precisa, maximizando la eliminación de tejido tumoral y minimizando el daño a zonas funcionalmente críticas.

Aunque la tecnología se encuentra todavía en fase de investigación y validación clínica, los estudios preliminares en cirugía de gliomas han mostrado resultados prometedores. Se ha observado una alta sensibilidad y especificidad para diferenciar tejido tumoral de tejido cerebral sano, lo que refuerza su potencial como herramienta complementaria en la resección de GBM infiltrativos. (Fabelo, Ortega, Lazcano et al., 2018; Leon et al., 2023).

La siguiente tabla resume estudios relevantes publicados desde 2015 que combinan HSI con algoritmos de IA en contexto intraoperatorio para el tratamiento de tumores cerebrales. Se incluyen estudios preclínicos (en modelos animales o tejidos ex vivo) y clínicos, detallando el tipo de algoritmo utilizado, el hardware hiperespectral empleado (incluyendo la técnica de adquisición), las condiciones de adquisición y las métricas de rendimiento disponibles.

Los trabajos analizados muestran una creciente aplicabilidad de la HSI tanto en escenarios in vivo como ex vivo, utilizando diferentes configuraciones de hardware y rangos espectrales. Los estudios de (Fabelo, Ortega, Lazcano et al., 2018) y (Leon et al., 2023) destacan el uso clínico directo de sistemas HSI en quirófano, con cohortes de 22 y 34 pacientes respectivamente. Ambos lograron una alta precisión en la clasificación de tejido tumoral y sano (hasta 97.5% y 91.5%), aplicando algoritmos como SVM, KNN o DNN. El estudio de León también mostró una F1-score más modesta en redes neuronales profundas (70,2%), evidenciando las diferencias de rendimiento entre modelos.

Tabla 2.3: Resumen de estudios sobre el uso de imagen hiperespectral (HSI) en neurocirugía.

Estudio (año)	Hardware y adquisición	Espectro / Bandas	Escenario	Clases / Tumor analizado	Algoritmos de IA	Métricas principales
(Fabelo, Ortega, Lazcano et al., 2018)	Cámara Head-wall Hyperspec (VNIR + NIR), sistema tipo pushbroom, iluminación halógena de 150 W	400–1000 nm (826 bandas) y 900–1700 nm (172 bandas)	In vivo (22 pacientes)	Tejido tumoral, sano, vasos, fondo	PCA (análisis de componentes principales), SVM (máquinas de vectores soporte), KNN (k vecinos más cercanos), K-means	Precisión global $\approx 97.5\%$
(Leon et al., 2023)	Cámara snapshot NIR de 25 bandas	25 bandas (655–975 nm)	In vivo (34 pacientes)	Tumor, sano, vasos, fondo	SVM (lineal y con núcleo RBF), RF, KNN, DNN (red neuronal profunda)	SVM-RBF: 91.5 %, F1 macro DNN: 70,2 %
(Black et al., 2024)	Filtro LCTF + cámara sCMOS, 420–730 nm, con excitación de 5-ALA (fluorescencia de PpIX)	Resolución espectral de 3–5 nm (≈ 80 bandas)	Ex vivo (891 muestras)	Tipo de tumor, grado, márgenes, estado IDH	Unmixing (desmezcla espectral) + RF, MLP (perceptrón multicapa)	Tipo: 84–87 %, Grado: 96 %, Márgenes: 86 %, IDH: 91 %
(Suero Molina et al., 2024)	Filtro LCTF + cámara sCMOS, 420–730 nm, acoplado a microscopio Zeiss con LED 405 nm (5-ALA)	≈ 80 bandas	Ex vivo (muestras neuroquirúrgicas)	Tipo de tumor, grado, márgenes, IDH	Unmixing + MLP, CNN (red neuronal convolucional)	Tipo: 87 %, Grado: 96.1 %, Márgenes: 86 %, IDH: 93 %
(Mühle et al., 2021)	Microscopio con cámara snapshot (550–1000 nm)	VNIR (detalle no especificado)	In vivo + laboratorio	Tejido sano vs tumor	SVM + regresión logística	Relación señal-ruido (SNR) in vivo ≈ 25 dB (laboratorio ≈ 40 dB)
(Giannantonio et al., 2023)	Cámara snapshot acoplada a microscopio quirúrgico	468–787 nm (mosaico tipo snapshot)	In vivo (5 pacientes con resección de gliomas)	Glioma de bajo y alto grado, tejido sano, artefactos, sangre, objetos extraños	CNN para selección de bandas y clasificación	Precisión (accuracy) $\gtrsim 80\%$

Por su parte, los trabajos de (Black et al., 2024) y (Suero Molina et al., 2024) se centraron en muestras ex vivo, combinando desmezcla espectral (*unmixing*) con algoritmos avanzados como perceptrones multicapa y CNN. Ambos presentaron buenos resultados en la clasificación de tipo y grado tumoral, márgenes quirúrgicos e incluso estado mutacional, con precisiones superiores al 90 % en varios casos. Sin embargo, el estudio de (Mühle et al., 2021) exploró el desempeño del sistema en condiciones reales y controladas, evidenciando una relación señal-ruido inferior en entorno quirúrgico (*approx*25 dB) frente al laboratorio (*approx*40 dB), lo que refleja las limitaciones técnicas in vivo. Finalmente, (Giannantonio et al., 2023) incorporaron técnicas de selección de bandas y clasificación basadas en CNN, alcanzando una precisión superior al 80 % en un pequeño grupo de pacientes con gliomas.

En conjunto, estos estudios demuestran que la HSI, combinada con algoritmos de IA, tienen potencial para la discriminación tisular en neurocirugía. No obstante, se evidencian diferencias en las configuraciones técnicas, algoritmos utilizados y métricas alcanzadas.

En este contexto, el proyecto NEMESIS-3D-CM (clasificación intraoperatoria de tumores cerebrales mediante modelos InnerSivos) ha sido una iniciativa de investigación centrada en el uso de imagen hiperespectral (HSI, por sus siglas en inglés) para mejorar las herramientas de navegación intraoperatoria durante la resección de tumores cerebrales. El objetivo principal del proyecto fue facilitar al neurocirujano una visión aumentada del campo quirúrgico, capaz de identificar de forma objetiva y en tiempo real las distintas estructuras tisulares, especialmente diferenciando entre tejido tumoral y tejido sano.

Para ello, se propuso un sistema capaz de adquirir imágenes hiperespectrales en el entorno quirúrgico y generar, a partir de ellas, vídeo en tiempo real con mapas de clasificación tisular. Estos mapas permiten una visualización continua del estado del tejido observado, codificando por colores la probabilidad de pertenencia a distintas clases (por ejemplo, tumor, corteza, sustancia blanca). La clasificación se actualiza dinámicamente a lo largo de la intervención, proporcionando así una herramienta de apoyo a la decisión quirúrgica. Un ejemplo de estos mapas de salida puede verse en la Figura 2.16.

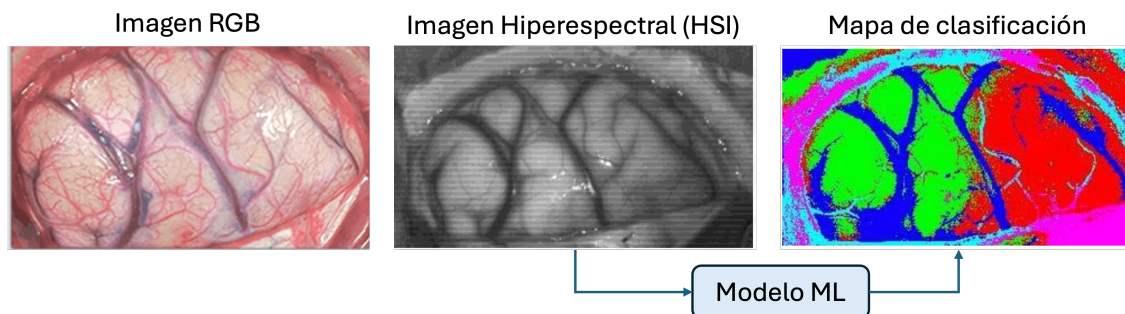


Figura 2.16: Ejemplo de resultado del proyecto NEMESIS-3D-CM. La primera imagen muestra una vista RGB del cerebro obtenida durante la cirugía. La segunda corresponde a la HSI adquirida intraoperatoriamente. Finalmente, se presenta el mapa de clasificación generado a partir de un modelo de ML entrenado para distinguir diferentes tipos de tejido cerebral.

2.8.2 Caso de estudio: Segmentación automática y predicción de pronóstico en pacientes con hemorragia subaracnoidea aneurismática

La Hemorragia Intracraneal (HIC) se refiere a la presencia anómala de sangre dentro del cráneo como resultado de una extravasación vascular, independientemente de su localización o causa. Esta acumulación de sangre puede generar compresión del tejido cerebral, desplazamiento de estructuras, aumento de la presión intracraneal e incluso isquemia secundaria (Caplan, 2015). Se trata, por tanto, de una urgencia neurológica potencialmente letal que requiere atención médica inmediata.

Según la localización anatómica del sangrado, la HIC se clasifica en diferentes tipos (Figura 2.17):

- **Hemorragia intraparenquimatosa (HIP):** Ocurre cuando la sangre se acumula dentro del propio tejido cerebral. Es la forma más común de hemorragia cerebral espontánea y suele estar relacionada con hipertensión arterial crónica, aunque también puede deberse a tumores, fármacos anticoagulantes o malformaciones vasculares (Gebel y Broderick, 2000). Su localización más frecuente es en los ganglios basales, tálamo, tronco encefálico y cerebelo.
- **Hemorragia intraventricular (HIV):** Consiste en la presencia de sangre en el sistema ventricular cerebral. Puede ser primaria (menos frecuente) o secundaria a una HIP o subaracnoidea que se extiende hacia los ventrículos (Hallevi et al., 2008). La HIV puede obstruir la circulación del líquido cefalorraquídeo, provocando una hidrocefalia aguda, que agrava el pronóstico. Clínicamente, puede manifestarse con deterioro del nivel de conciencia, signos de hipertensión intracraneal y alteraciones del sistema nervioso autónomo (Hostettler et al., 2019).
- **Hemorragia subdural (HSD):** Tiene lugar entre la duramadre y la aracnoides. Es más frecuente tras un traumatismo craneal, especialmente en ancianos o pacientes anticoagulados (Miranda et al., 2011). Su evolución puede ser aguda o crónica, y en esta última forma, los síntomas ser alteraciones cognitivas, cefalea o deterioro progresivo del estado general (Feghali et al., 2020).
- **Hemorragia epidural (HED):** Se sitúa entre la duramadre y la zona interna del cráneo, habitualmente causada por la rotura de la arteria menígea media tras un traumatismo craneoencefálico. Tiene un patrón clínico clásico de intervalo lúcido, donde el paciente pierde el conocimiento inicialmente, mejora transitoriamente, y luego se deteriora de forma rápida por la expansión del hematoma (Figuerola y DeVine, 2017).
- **Hemorragia subaracnoidea (HSA):** Se produce en el espacio subaracnoideo, entre la aracnoides y la pía madre, donde circula el líquido cefalorraquídeo. La sangre liberada en este espacio entra en contacto directo con las cisternas basales, los surcos corticales y las raíces de los nervios craneales. A diferencia de otras hemorragias, la HSA no suele formar un hematoma localizado, sino que produce una inflamación difusa que afecta a varias zonas del cerebro al mismo tiempo. Puede ser de origen traumático o espontáneo. Esta última es la más grave desde el punto de vista clínico (Osgood y Singer, 2021;

Van Gijn et al., 2007).

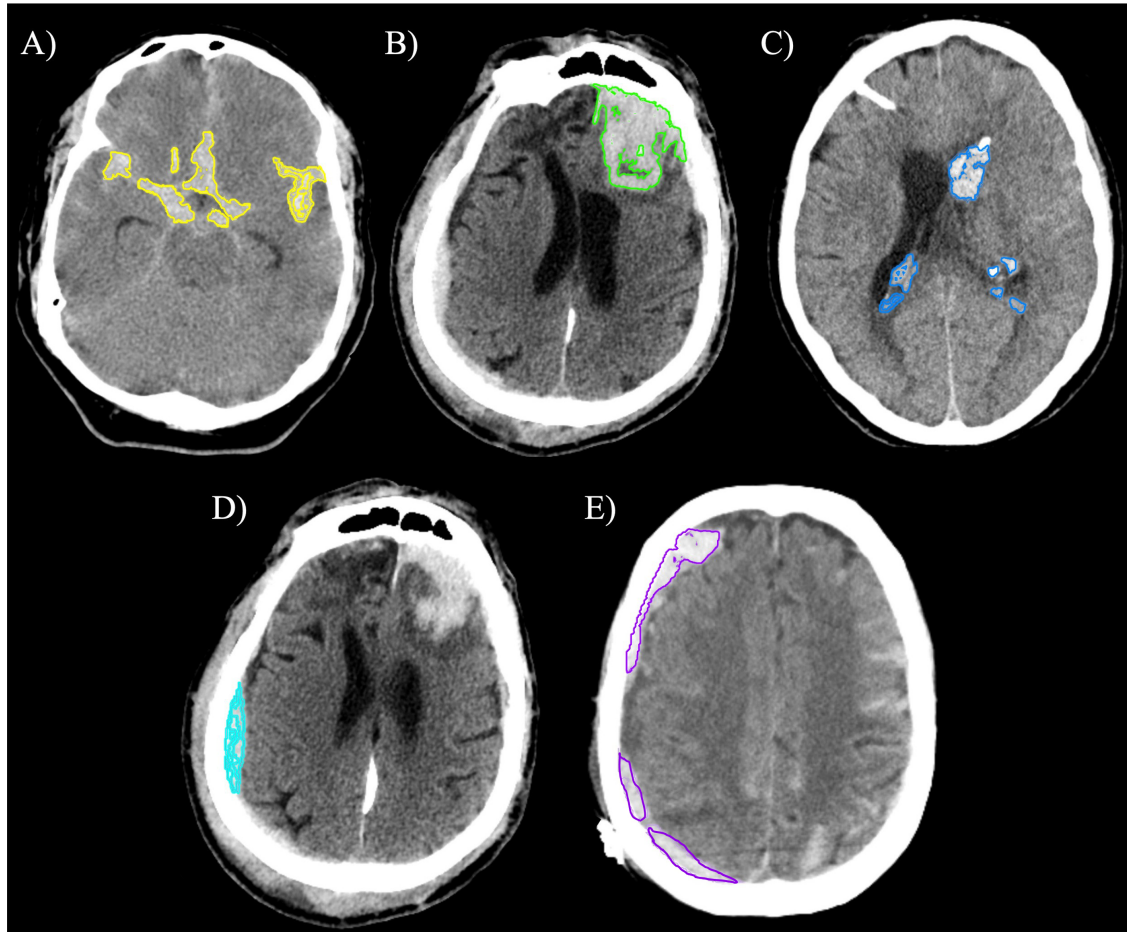


Figura 2.17: Representación esquemática de los principales tipos de HIC según su localización anatómica en TC. A) HSA, B) HIP, C) HIV, D) HED, E) HSD

La hemorragia subaracnoidea aneurismática (HSAa) es la causa más frecuente de HSA espontánea y constituye una forma especialmente grave de ictus hemorrágico. Suele afectar a personas jóvenes y previamente sanas. En aproximadamente el 85% de los casos de HSA no traumática, esta se debe a la rotura espontánea de un aneurisma intracraneal, una dilatación patológica de la pared arterial que, al romperse, provoca la liberación brusca de sangre a presión en el espacio subaracnoideo (Osgood y Singer, 2021; Van Gijn et al., 2007). Desde el punto de vista clínico, la HSA suele aparecer de forma repentina y muy intensa. El síntoma más característico es la cefalea intensa, que muchas personas describen como “el peor dolor de cabeza de su vida”, y que alcanza su máxima intensidad en cuestión de segundos. Este dolor puede acompañarse de pérdida de conciencia, rigidez de nuca, náuseas y vómitos, fotofobia y, en casos graves, de un deterioro rápido del nivel de conciencia o incluso muerte súbita. En algunos pacientes, sin embargo, pueden presentarse cefaleas denominadas centinela en los días previos, causadas por pequeños sangrados iniciales que habitualmente pasan desapercibidos y no reciben atención médica.

La HSA presenta un comportamiento clínico distinto al de otras hemorragias intracraneales,

como la HIV o HSD. Aparte del daño neurológico inicial provocado por la extravasación de sangre al espacio subaracnoideo, la HSA se caracteriza por desencadenar una serie de complicaciones secundarias que pueden comprometer aún más el pronóstico si no se reconocen y tratan a tiempo. Entre las más relevantes se encuentran el resangrado, el vasoespasmo cerebral y la hidrocefalia.

El **resangrado** suele ocurrir en las primeras 24 a 48 horas tras el evento inicial y se asocia con una elevada mortalidad. Por eso, una de las prioridades en el manejo de la HSA es asegurar el tratamiento precoz del aneurisma que lo causa, ya sea mediante cirugía (clipaje) o un procedimiento endovascular (embolización), que permiten evitar nuevas roturas.

El **vasoespasmo** cerebral es una complicación frecuente, que afecta aproximadamente al 30–70% de los pacientes con HSA. Se produce entre el tercer y el decimocuarto día tras la hemorragia. Se trata de una reacción de las arterias cerebrales que, al estar expuestas a los productos derivados de la sangre acumulada, se irritan y comienzan a estrecharse de forma mantenida. Este proceso provoca una vasoconstricción de las arterias cerebrales, que puede reducir significativamente el flujo sanguíneo cerebral y provocar isquemia cerebral, lo que significa que algunas zonas del cerebro dejan de recibir suficiente oxígeno y pueden dañarse. Para su detección, se suelen usar estudios como el Doppler transcraneal, que mide la velocidad del flujo en las arterias del cerebro, o técnicas de imagen como la CTA o DSA (Anderson et al., 2000), que permiten visualizar directamente si las arterias están más estrechas de lo normal. No obstante, su diagnóstico presenta una dificultad añadida: los estándares utilizados son heterogéneos y no existe un consenso universal en cuanto a qué prueba debe considerarse de referencia ni cuáles son los umbrales que determinan la presencia de vasoespasmo. Así, mientras el Doppler es una herramienta no invasiva y fácilmente repetible, su interpretación depende de valores de corte que varían entre instituciones; por otro lado, la CTA y la DSA ofrecen una representación anatómica más directa, pero es una prueba más compleja que conlleva riesgos asociados. Esta falta de homogeneidad en los criterios diagnósticos repercute en la práctica clínica, generando variabilidad en la detección y en la forma de reportar el vasoespasmo en distintos centros y estudios.

Por otro lado, la **hidrocefalia** ocurre cuando la sangre en el espacio subaracnoideo bloquea el flujo normal del líquido cefalorraquídeo o interfiere con su reabsorción (S. Chen et al., 2017). Esto provoca una acumulación anormal de líquido en los ventrículos del cerebro, que se dilatan y aumentan la presión intracraneal. La hidrocefalia se diagnostica TC, donde se observa una dilatación de los ventrículos. Cuando el cuadro es grave o progresivo, puede ser necesario colocar un drenaje ventricular externo de forma temporal, o implantar una válvula (derivación ventrículo-peritoneal) para drenar el exceso de líquido de forma permanente.

A pesar de los avances en diagnóstico y tratamiento, la tasa de mortalidad global de la HSA sigue siendo alta. Se estima que hasta un 50% de los pacientes fallecen, muchos incluso antes de poder recibir atención médica. De los supervivientes, una parte significativa presenta secuelas neurológicas persistentes, que pueden ir desde leves alteraciones cognitivas hasta grados severos de discapacidad física, cognitiva o emocional. En términos de salud pública, esto se traduce en un elevado número de años de vida ajustados por discapacidad, especialmente en poblaciones más jóvenes (Feigin et al., 2017). Esto no solo supone una reducción considerable en la esperanza de vida saludable, sino también un alto costo económico y social, debido al

uso prolongado de recursos médicos, la necesidad de rehabilitación intensiva y la dependencia funcional a largo plazo.

En este contexto, la TC cerebral sin contraste es la prueba diagnóstica fundamental, especialmente en la fase aguda. Gracias a su disponibilidad inmediata, rapidez y capacidad para detectar sangre en el espacio subaracnoideo, se considera la base del diagnóstico inicial. Cuando se realiza en las primeras seis horas desde el inicio de los síntomas, la TC tiene una sensibilidad muy alta, permitiendo confirmar el diagnóstico con gran fiabilidad. A medida que transcurren las horas, la sensibilidad disminuye progresivamente, lo que obliga a considerar pruebas complementarias si la sospecha clínica persiste.

Una vez confirmada la hemorragia, el siguiente paso fundamental es identificar su causa, siendo los aneurismas intracraneales los responsables en la mayoría de los casos. Para ello, se emplean técnicas de angiografía cerebral, entre las que se encuentran la CTA o MRA, que permiten una evaluación vascular rápida y no invasiva. Sin embargo, la DSA se utiliza con mayor frecuencia por su precisión y porque permite intervenir directamente sobre el aneurisma si es necesario.

Desde el punto de vista terapéutico, la HSA requiere atención en unidades especializadas y un enfoque multidisciplinar. Su manejo implica una actuación rápida, precisa y coordinada, que abarca desde la sospecha inicial hasta la rehabilitación neurológica, con el objetivo de reducir la mortalidad y mejorar la calidad de vida de los pacientes que sobreviven.

Escalas clínicas y radiológicas de evaluación inicial

En la práctica clínica habitual, la valoración inicial de los pacientes con HSA se fundamenta en el uso combinado de escalas clínicas y radiológicas, cuyo objetivo principal es estratificar la gravedad neurológica, anticipar la aparición de complicaciones (como el vasoespasmio o la hidrocefalia) (Connolly Jr et al., 2012), y orientar las decisiones terapéuticas desde el momento del ingreso. Estas escalas también permiten establecer una base común para la comunicación entre equipos y para la investigación clínica.

Desde el punto de vista clínico, una de las herramientas más utilizadas es la ***Glasgow Coma Scale (GCS)*** (Teasdale y Jennett, 1974), que evalúa el nivel de conciencia del paciente mediante la respuesta ocular, verbal y motora, con una puntuación total que oscila entre 3 y 15. Aunque no es específica para la HSA, su aplicación sistemática ha demostrado utilidad en la evaluación inicial de la severidad neurológica, especialmente en contextos de urgencia.

Basándose en la GCS, la ***World Federation of Neurosurgical Societies (WFNS)*** propuso una clasificación específica para la HSA que combina el nivel de conciencia con la presencia de déficits neurológicos focales (CG, 1988). Esta escala, ampliamente adoptada, permite categorizar a los pacientes en cinco grados de severidad (WFNS I a V), siendo útil para predecir el pronóstico funcional y la mortalidad. En este sentido, la WFNS se ha validado como un predictor robusto del desenlace funcional a medio y largo plazo, generalmente evaluado mediante escalas como la GOS (Hofmann et al., 2022). Por tanto, su uso es fundamental para anticipar la recuperación neurológica y el grado de discapacidad que el paciente podría presentar tras la fase aguda. No obstante, su aplicabilidad puede verse limitada en situaciones de sedación, intoxicación o intubación precoz, que alteran la valoración del nivel de conciencia.

Desde el punto de vista radiológico, la escala más clásica es la **escala de Fisher**, que clasifica la gravedad del sangrado subaracnoideo según el volumen su distribución en la TC craneal sin contraste (Fisher et al., 1980). Esta escala se basa en una valoración semicuantitativa del sangrado: desde ausencia de hemorragia visible (grado 1), hasta hemorragia densa y extensa con coágulos (grado 4). Sin embargo, una de sus limitaciones más relevantes es que no contempla la HIV, un tipo de hemorragia frecuente en pacientes con HSAa y con importantes implicaciones pronósticas.

Para superar esta deficiencia, se propuso la escala modificada de Fisher, que incorpora la presencia de HIV como un componente adicional, ofreciendo una mejor correlación con la aparición de vasoespasma sintomático e infartos isquémicos (Frontera et al., 2006). Por ello, el Fisher ha demostrado ser un marcador del riesgo de vasoespasma cerebral y complicaciones isquémicas secundarias.

Además de las escalas de Fisher, otras herramientas radiológicas han sido propuestas para caracterizar más detalladamente la carga y distribución del sangrado. La **escala de Hijdra**, por ejemplo, evalúa la presencia de sangre en 10 cisternas basales y 4 ventrículos, asignando una puntuación de 0 a 3 en cada localización según la cantidad de sangre (Hijdra et al., 1990). Esta escala permite una estimación más granular del sangrado subaracnoideo, con buena correlación interobservador, y ha sido utilizada especialmente en estudios clínicos como medida objetiva del volumen total de sangre.

De forma similar, la **escala de LeRoux** ofrece una puntuación global del sangrado en las cisternas, combinando tanto la extensión como la densidad de la hemorragia. A diferencia de escalas más detalladas como la de Hijdra, LeRoux ofrece un enfoque más simple. Esta herramienta ha demostrado tener valor en la predicción de complicaciones neurológicas relevantes, como la aparición de hidrocefalia aguda o crónica (LeRoux et al., 1992).

Junto a estas escalas clásicas, en los últimos años se ha puesto en valor el papel del edema cerebral precoz como marcador pronóstico. En este contexto, se ha desarrollado la **SEBES (Score for the Early Brain Edema Evaluation Scale)** (Ahn et al., 2018), una herramienta simple para cuantificar el grado de edema cerebral en la TC inicial de pacientes con HSAa. La escala SEBES evalúa si los surcos del cerebro están borrados (es decir, menos visibles) en una TC, especialmente a nivel de los ganglios basales y los núcleos talámicos. Se asigna una puntuación de 0 a 4, según si el borramiento es unilateral o bilateral en esas zonas. En algunos estudios, la escala SEBES ha demostrado correlación directa con el pronóstico funcional: puntuaciones altas (3–4) se asocian con mayor riesgo de desenlace desfavorable, independientemente de otras variables clínicas o del volumen del sangrado (Ahn et al., 2018). Además, estudios recientes han propuesto que el edema cerebral temprano reflejado en el SEBES podría ser un marcador indirecto del impacto de la hemorragia sobre la inflamación y la presión intracraneal (Said et al., 2021).

Además del SEBES clásico, se ha descrito una versión ampliada denominada **SEBES-6c** (Eibach et al., 2020). Mientras que la escala original se centra en cuatro niveles de la TC a nivel de ganglios basales, el SEBES-6c incorpora dos cortes adicionales en regiones supraganglionares, con el objetivo de reflejar mejor la extensión del edema cerebral precoz. Los autores señalan que esta variante permite una detección más sensible de la afectación difusa y ofrece una

capacidad pronóstica ligeramente superior para predecir desenlaces funcionales, manteniendo al mismo tiempo la simplicidad de aplicación que caracteriza al SEBES original.

Recientemente, se han desarrollado escalas como la **ictus-WFNS (ictWFNS)**, que a diferencia de la WFNS tradicional, se basa en parámetros extraídos de la TC, específicamente mediante la combinación de las escalas SEBES, Hídra y Le Roux. Los autores afirman que la esta escala permite detectar pacientes que, pese a presentar un mal grado clínico al ingreso, tienen en realidad un potencial de recuperación funcional más favorable. De este modo, la ictWFNS complementa a la escala WFNS y puede mejorar la precisión pronóstica en las primeras horas tras la hemorragia (Hofmann et al., 2022).

Por último, para evaluar la evolución y pronóstico funcional tras la hemorragia subaracnoidea, se utiliza habitualmente la **Escala de Resultados de Glasgow (Glasgow Outcome Scale, GOS)**. Esta herramienta permite clasificar el estado funcional del paciente en cinco categorías que reflejan desde la muerte hasta una recuperación funcional buena. La GOS es fundamental para medir de forma estandarizada el grado de discapacidad o autonomía que presenta el paciente tras el evento y durante el seguimiento. Las categorías de la GOS son:

1. **Muerte.**
2. **Estado vegetativo persistente:** paciente despierto pero sin conciencia ni respuesta al entorno.
3. **Discapacidad severa:** supervivencia con dependencia total o significativa para las actividades básicas de la vida diaria.
4. **Discapacidad moderada:** paciente independiente para la mayoría de las actividades, pero con algunas limitaciones físicas o cognitivas.
5. **Buena Recuperación:** recuperación funcional cercana a la normalidad, con pocas o ninguna secuela significativa.

La evaluación con la GOS tras la HSA ayuda a cuantificar el impacto real de la hemorragia sobre la calidad de vida, orientar el plan rehabilitador y comparar la eficacia de distintos tratamientos o estrategias de manejo clínico.

En los últimos años, diversos estudios han demostrado que el volumen de sangre extravasada en una HIC, cuantificado de forma precisa, guarda una estrecha relación con la evolución clínica del paciente (Jordan et al., 2009; LoPresti et al., 2014). En el caso concreto de la HSAa se ha observado que el volumen de hemorragia presenta una correlación significativa con el pronóstico, especialmente medido con la escala de GOS (Bravo-Garrido et al., 2025; Lagares et al., 2015). A diferencia de las escalas visuales tradicionales, medir con precisión el volumen de sangre permite valorar de forma más continua y objetiva el daño cerebral desde el inicio. Esto no solo ayuda a describir mejor la gravedad del cuadro clínico, sino que también mejora la capacidad para anticipar cómo evolucionará el paciente. En este contexto, estimar el volumen del sangrado se convierte en una herramienta útil para ofrecer un enfoque más personalizado y ajustado a cada caso, apoyado en datos concretos.

Sin embargo, la medición manual del volumen hemorrágico continúa siendo una tarea laboriosa, dependiente del observador y poco escalable en el entorno clínico. Esta limitación ha impulsado

Tabla 2.5: Resumen de las principales escalas clínicas y radiológicas empleadas en la HSAa.

Escala	Dominio	Descripción / Puntuación	Utilidad clínica / Limitaciones
Glasgow Coma Scale (GCS) (Teasdale y Jennett, 1974)	Clínica	Evalúa nivel de conciencia (respuesta ocular, verbal y motora). Rango 3–15.	Amplia utilidad inicial en urgencias; no específica de HSAa. Limitada en pacientes intubados o sedados.
WFNS (CG, 1988)	Clínica	Combina GCS y déficit neurológico focal. Grados I–V.	Predice mortalidad y desenlace funcional. Validada ampliamente, pero limitada por sedación/intubación.
Fisher (Fisher et al., 1980)	Radiológica	Grados 1–4 según volumen y distribución de sangre en TC.	Estima riesgo de vasoespasmos. No contempla hemorragia intraventricular (HIV).
Modificada de Fisher (Frontera et al., 2006)	Radiológica	Incorpora la HIV a la valoración del sangrado subaracnoideo.	Mejora la predicción de vasoespasmos sintomáticos y complicaciones isquémicas.
Hijdra (Hijdra et al., 1990)	Radiológica	Evalúa sangre en 10 cisternas y 4 ventrículos (0–3 puntos cada uno).	Estimación cuantitativa del volumen total de sangre; buena reproducibilidad; usada en investigación.
LeRoux (LeRoux et al., 1992)	Radiológica	Puntuación global de cisternas según extensión y densidad del sangrado.	Más simple que Hijdra; útil en predicción de hidrocefalia y complicaciones neurológicas.
SEBES (Ahn et al., 2018)	Radiológica	Evalúa el borramiento de surcos de la sustancia gris profunda (0–4).	Asociada con edema cerebral temprano y mal pronóstico funcional. Sencilla y rápida en TC inicial.
SEBES-6c (Eibach et al., 2020)	Radiológica	Extiende SEBES al evaluar seis cisternas de la base (0–6).	Aumenta la sensibilidad para detectar edema difuso y mejora la capacidad predictiva frente a SEBES clásico.
ictWFNS (Hofmann et al., 2022)	Radiológica-clínica	Integra SEBES, Hijdra y LeRoux para reclasificar pacientes en grados pronósticos.	Complementa la WFNS clínica. Permite detectar potencial de recuperación en pacientes con mal grado inicial.
Glasgow Outcome Scale (GOS)	Funcional (pronóstico)	Categorías: muerte, estado vegetativo, discapacidad severa, discapacidad moderada, buena recuperación.	Escala estándar para valorar desenlace funcional en el seguimiento.

el desarrollo de herramientas automatizadas capaces de realizar esta cuantificación de forma rápida, precisa y reproducible. En este contexto, la IA, y en particular los modelos de DL como las CNN han demostrado un alto rendimiento en la segmentación de hemorragias intracraneales sobre estudios de TC (Yeo et al., 2021). Estas técnicas permiten delinear automáticamente el sangrado y calcular su volumen, reduciendo la variabilidad interobservador y minimizando el tiempo requerido para el análisis.

La mayoría de los modelos de segmentación disponibles han sido entrenados en cohortes de pacientes con traumatismo craneoencefálico (TCE). Aunque los TCE y la HSA comparten ciertas propiedades estructurales del sangrado, como la alta densidad del coágulo en la TC, su evolución con el tiempo o la importancia de identificar pequeños volúmenes en diferentes áreas del cerebro, también presentan diferencias significativas en cuanto a distribución y morfología. En el TCE, las lesiones hemorrágicas suelen ser multifocales y heterogéneas, y con frecuencia se acompañan de áreas contusivas y edema cerebral. Por el contrario, en la HSAa, el sangrado se concentra principalmente en los espacios subaracnoideos, como las cisternas basales, los surcos corticales o el sistema ventricular, adoptando una forma más lineal y laminar. Estas diferencias en la localización y la fisiopatología dificultan la aplicación directa de modelos de segmentación desarrollados para TCE en pacientes con HSA. Por ello, es necesario crear y validar modelos específicos que consideren tanto las particularidades de la HSA como los aspectos comunes con otras hemorragias intracraneales, con el fin de mejorar la segmentación automática y su utilidad clínica.

En la Tabla 2.6 se resumen los principales algoritmos publicados en los últimos cinco años, incluyendo tanto herramientas de segmentación generalistas para múltiples tipos de hemorragia, como modelos específicamente diseñados para la cuantificación del sangrado en pacientes con HSAa. Para cada estudio se detalla el tipo de hemorragia abordado, la arquitectura utilizada, el número de casos incluidos y su procedencia, las métricas de desempeño (por ejemplo, puntuación de Dice, sensibilidad, explicadas en la sección 3.4.4), la existencia de validación clínica o externa, y la disponibilidad pública del código o herramienta.

Tabla 2.6: Modelos recientes de segmentación automática de hemorragias intracraneales.

Estudio (año)	Tipo de hemorragia	Modelo de segmentación	Cohorte y datos	Métricas de desempeño	Validación clínica / desenlaces	Disponibilidad
(Monteiro et al., 2020)	TCE multiclase: HIP, HIV, extra-axiales (HSA, HSD, HED), edema	DeepMedic (3D CNN, BLAST-CT)	98 TC (manual) + 839 multicéntricos (semi-automático) + 500 externos	Lesiones grandes (>5mL): Dice \approx 0,72 Lesiones pequeñas (\leq 1mL): Dice \approx 0,36–0,49	Validación externa (India); predicción carga lesional y progresión	Código abierto, pesos disponibles.
(Nijiati et al., 2022)	HIC multiclase: HIP, HIV, HED, HSD, HSA	Sym-TransNet (CNN + Transformer)	1.157 pacientes (entrenamiento/test)	Dice global = 0,716; Dice HSA = 0,337; sens \approx 0,98	Evaluado en set balanceado. Sin cohortes externas. No hay predicción de desenlaces.	No disponible públicamente
(Q. Liu et al., 2023)	HIC espontánea	Viola-UNet v2 (3D U-Net + atención) + nnU-Net (ensemble)	INSTANCE 2022 (~600 TC)	Dice global \approx 0,79, Hausdorff Distance \approx 21.56	Validación multicéntrica en INSTANCE challenge. Sin predicción de desenlaces.	Código, Docker y pesos disponibles.
(García-García et al., 2024)	HSAa	Swin-UNETR (Transformer + U-Net)	100 TC HSA (entrenamiento/test); 10 externos	Dice = 0,87; sens. = 0,82	Validación externa; predicción de mortalidad (regresión)	Código abierto, pesos disponibles y herramienta online
(Hu et al., 2023)	HSAa	2D/3D U-Net híbrida	582 TC HSA; 408 entrenamiento, 174 test independiente	Dice \approx 0,842; sens \approx 0,86	Validación independiente; predicción de la aparición de isquemia cerebral diferida, progresión y mortalidad	No disponible públicamente (uso clínico interno)
(Gong et al., 2024)	HIC – hematomas <1ml	nnU-Net	267 TC multicéntricos	Mejor precisión en lesiones pequeñas	Validación interna. Detección temprana de hemorragias.	Código abierto, pesos disponibles
(Rusche et al., 2023)	HIC espontánea	nnU-Net	319 pacientes (entrenamiento/test)	Dice \approx 0,93	Validación interna. Predicción del tiempo de inicio del ictus hemorrágico	Código abierto. Algoritmo de segmentación integrado en TotalSegmentator (Wasserthal et al., 2023).
(Kok et al., 2022)	HIC + edema	Ensemble nnU-Net 3D	1.732 TC (124 centros)	Dice: HIC \approx 0,92; HIV \approx 1,00; edema \approx 0,66	No se analizaron desenlaces clínicos directamente. Validación interna con 174 casos.	Código base open-source; pesos no disponibles
(M. D. Hssayeni et al., 2020)	TCE (IPH, IVH, SAH, EDH, SDH)	U-Net con ventana deslizante	82 TC de pacientes con TC	Slice-level: Dice = 0,31, Jaccard = 0,21, sensibilidad = 0,97	Validación interna. Sin análisis de desenlaces clínicos	Dataset público. Código no disponible públicamente. Pesos no disponibles.

La revisión de modelos recientes para la segmentación automática de hemorragias intracraneales muestra que las arquitecturas basadas en U-Net, especialmente nnU-Net, son las más frecuentemente empleadas. En los estudios analizados (Gong et al., 2024; Kok et al., 2022; Q. Liu et al., 2023; Rusche et al., 2023) esta arquitectura ha demostrado un rendimiento sólido y constante, con coeficientes de Dice en torno a 0,90 o superiores en HIC, tanto en contextos clínicos controlados como en datasets multicéntricos. Además, su disponibilidad como herramienta de código abierto y su capacidad para adaptarse automáticamente a nuevos conjuntos de datos explican su adopción creciente en aplicaciones clínicas y de investigación.

Aunque algunos resultados son prometedores (por ejemplo, un Dice de 0,87 con Swin-UNETR en HSAa (Nijati et al., 2022)), su rendimiento sigue siendo algo variable dependiendo del subtipo de hemorragia (por ejemplo., Dice = 0,337 para HSA en Sym-TransNet), y su validación en distintas cohortes clínicas aún es limitada. Esto sugiere que, si bien representan una línea de desarrollo interesante, parece que estas arquitecturas aún enfrentan desafíos en casos con bordes poco definidos o bajo contraste.

Otra estrategia presente en algunos trabajos es el uso de modelos híbridos o en *ensemble*. Hu et al. (Hu et al., 2023) utilizaron una combinación de redes 2D y 3D tipo U-Net para segmentar HSAa, con buen desempeño global (Dice \approx 0,84). De forma similar, Liu et al. (Q. Liu et al., 2023) propusieron un ensemble entre Viola-UNet v2 y nnU-Net, obteniendo una Dice media cercana a 0,79 en el desafío INSTANCE. Este tipo de combinaciones puede ser útil para mejorar la fiabilidad frente a la variabilidad interpaciente o la heterogeneidad de las imágenes.

Por otro lado, algunos enfoques más clásicos como DeepMedic (CNN 3D multiescala) o U-Net 2D simple con ventana deslizante han mostrado limitaciones. El modelo BLAST-CT (Monteiro et al., 2020), basado en DeepMedic, tuvo un rendimiento aceptable en lesiones mayores de 5 mL (Dice \approx 0,72), pero bajo en lesiones pequeñas (Dice \approx 0,36–0,49). Hssayeni et al. (M. D. Hssayeni et al., 2020) reportaron también un Dice modesto (0,31) con una U-Net 2D, lo que sugiere que las arquitecturas más simples pueden no ser suficientes para segmentar con precisión hemorragias intracraneales, especialmente cuando son pequeñas, irregulares o poco contrastadas.

Un aspecto aún poco explorado, pero relevante, es la conexión entre la segmentación automática y la predicción de desenlaces clínicos. Aunque algunos estudios como los de García et al. (García-García et al., 2024) y Hu et al. (Hu et al., 2023), incorporan modelos que relacionan la segmentación con variables como la mortalidad o la aparición de complicaciones, este tipo de integración sigue siendo minoritario. Sin embargo, representa una dirección clave para que estos modelos tengan un impacto real en la práctica clínica.

Una vez segmentada la hemorragia, la radiómica ofrece una vía adicional para enriquecer en análisis. Estudios han demostrado que el volumen total de hemorragia, su distribución anatómica y la afectación intraventricular han demostrado ser variables relevantes para predecir mortalidad, aparición de vasoespasma y deterioro funcional (Dupont et al., 2009; Lagares et al., 2001). La radiómica por su parte, permite extraer otros descriptores cuantitativos relacionados con la textura, intensidad, forma y relaciones espaciales en las regiones de interés que no son perceptibles visualmente.

Por ello, diversos estudios recientes han explorado el uso de radiómica extraída directamente del sangrado cerebral en imágenes de TC para predecir desenlaces clínicos adversos en pacientes con hemorragias intracraneales. La tabla 2.8 resume investigaciones publicadas en los últimos años que utilizan radiómica obtenida a partir de la segmentación de la hemorragia en la imagen inicial. Estos parámetros alimentan modelos de ML para predecir variables clínicas de interés como mortalidad, resultado funcional (GOS), vasoespasma o hidrocefalia.

Tabla 2.8: Estudios recientes que emplean radiómica del sangrado cerebral para predecir desenlaces clínicos.

Autor (año)	Tipo de hemorragia	Número de casos	Características radiómicas	Modelo/algorithm	Desenlace(s)	Desempeño	Val. ext.
(Huang et al., 2022)	HIC	1098	Hematoma y edema: forma, intensidad, textura; selección automática	Regresión logística multivariable (modelo combinado con variables clínicas)	Desenlace funcional a los 3 meses	AUC = 0,89 (val. externa); buena calibración	Sí
(Pei et al., 2024)	HIC	161	8 características seleccionadas de 1130 (LASSO)	Regresión logística, SVM, RF, KNN, XGBoost (modelos clínicos, radiómicos y combinados)	Desenlace funcional al alta	Entrenamiento AUC = 0,8583 (combinado); validación = 0,9153	Sí
(Dierksen et al., 2024)	HIC	852	Hemorragia y edema perilesional; Distintos métodos de selección de características	Modelos combinados (radiómica + clínica); comparativa de múltiples clasificadores	Desenlace funcional a 3 meses	AUC ≈ 0,85 (modelo combinado)	Sí
(L. Chen et al., 2025)	HSA	377	Textura y localización del sangrado subaracnoideo; selección con LASSO	SVM + nomograma clínico-radiómico	Isquemia cerebral diferida (DCI)	AUC = 0,831; Accuracy = 77.5 %	No
(Shan et al., 2023)	HSA	105	13 características (forma, textura, primer orden)	Comparativa de 7 clasificadores; mejor: SVM	Desenlace funcional a los 3 meses	AUC ≈ 0,82; precisión ≈ 88 % (validación cruzada)	No
(Rusche et al., 2023)	HIC	644 casos tras criterios de selección	Desarrollo de modelo de segmentación automática (nnU-Net); radiómica clásica	Árboles de regresión y CNN supervisada	Tiempo de inicio de la HIC	MAE: 9.96 h (radiómica), 9.77 h (CNN); error clínico: 11–13 h	No

En todos los estudios incluidos, la región hemorrágica fue segmentado de forma manual o automática, y se emplearon algoritmos de ML para entrenar modelos predictivos. La mayoría de los trabajos reportan métricas robustas de desempeño, principalmente AUC, y algunos incorporan validación externa para garantizar generalización. En conjunto, la evidencia respalda el potencial de la radiómica aplicada a la hemorragia cerebral como herramienta pronóstica complementaria en el manejo clínico.

Por otro lado, se ha observado que el daño cerebral inducido por la HSA no se restringe a las áreas donde es visible el sangrado. Estudios recientes han mostrado que existen alteraciones microestructurales subyacentes en el parénquima cerebral, incluso en ausencia de lesiones evidentes en la TC convencional. Estas alteraciones pueden estar asociadas a mecanismos secundarios como vasoespasmo, isquemia difusa, disfunción metabólica o procesos inflamatorios. La sustancia blanca es particularmente vulnerable al daño axonal e hipoxia, mientras que la sustancia gris puede presentar alteraciones funcionales tempranas (Hill et al., 2016; Y. Wang et al., 2016).

En este sentido, existen métodos establecidos para la segmentación automática del parénquima cerebral en TC, que permiten delimitar la sustancia gris, la sustancia blanca y otras regiones anatómicas. Entre ellos, destacan los enfoques basados en atlas deformables y desarrollos específicos para TC como el CTSeg (Brudfors, 2020; Brudfors et al., 2020) basado en el conocido software SPM12 (Ashburner et al., 2014), que permiten registrar automáticamente cada escaneo a un espacio anatómico común y etiquetar regiones cerebrales con alta resolución. Alternativamente, se han empleado modelos de DL capaces de generar mapas multiclase directamente a partir de TC no contrastado (Gupta et al., 2010; Manniesing et al., 2017).

La segmentación estructurada del encéfalo posibilita la extracción de biomarcadores radiómicos a nivel regional, incluyendo medidas de intensidad, textura y heterogeneidad tanto en áreas afectadas como en zonas aparentemente normales. Esto resulta especialmente relevante en el contexto de la HSAa donde los efectos difusos sobre el tejido cerebral pueden pasar desapercibidos visualmente.

En resumen, el contexto actual de la HSAa plantea múltiples desafíos diagnósticos y pronósticos que no son completamente resueltos por las herramientas clínicas tradicionales. La integración de técnicas de segmentación automática y análisis radiómico, aplicadas tanto al sangrado como al tejido cerebral adyacente, podría ser una estrategia prometedora para avanzar hacia una medicina de precisión, apoyada en datos objetivos y reproducibles derivados de imagen.

Los resultados obtenidos en el marco de este proyecto se presentan en el Capítulo 4, en la que se recoge la contribución de esta tesis doctoral al desarrollo e integración del sistema de clasificación intraoperatoria mediante HSI.

Capítulo 3

Materiales y métodos

La metodología de esta tesis doctoral constituye el eje central para alcanzar el **Objetivo Principal** (Sección 1.2): desarrollar y validar un sistema automatizado capaz de mejorar la precisión diagnóstica, predecir la evolución clínica del paciente y aportar información relevante para la toma de decisiones en el ámbito de la neurociencia clínica.

Para ello, se ha diseñado un flujo de trabajo integral que abarca desde el diseño del estudio y la adquisición de datos hasta la implementación final y su evaluación clínica siguiendo un esquema adaptado a las particularidades de cada modalidad de imagen y contexto clínico.

El flujo metodológico se organiza en cuatro grandes bloques (Figura 3.1):

- I. **Diseño del estudio y fuentes de datos**, que define el planteamiento metodológico, detalla los criterios de inclusión/exclusión, los procedimientos de aprobación ética, el proceso de adquisición de imagen y datos clínicos y la definición de las clases objetivo.
- II. **Preparación y anotación de datos**, que comprende la generación de anotaciones de referencia destinadas tanto a la validación como al entrenamiento del modelo (*ground truth*), el preprocesado de imagen y de datos clínicos, y la definición de estrategias de partición que garanticen la comparabilidad de resultados.
- III. **Extracción de características**, centrada en la obtención de características radiómicas a partir de TC y de firmas espectrales a partir de HSI, junto con su normalización y estructuración para el análisis posterior.
- IV. **Modelado y evaluación**, que engloba el desarrollo, validación y análisis de modelos de segmentación, clasificación y predicción, así como la interpretación de resultados y su comparación con escalas clínicas establecidas.

Cada bloque metodológico incluye tareas específicas para las modalidades HSI y TC, integradas en un esquema común que permite analizar ambas modalidades de forma coherente y complementaria. La Figura 3.1 presenta de forma sintética este flujo de trabajo adaptado a ambos casos de uso.

El presente trabajo se estructura sobre un diseño metodológico que integra dos escenarios clínicos y modalidades de imagen complementarias: la resección quirúrgica de GBM asistida

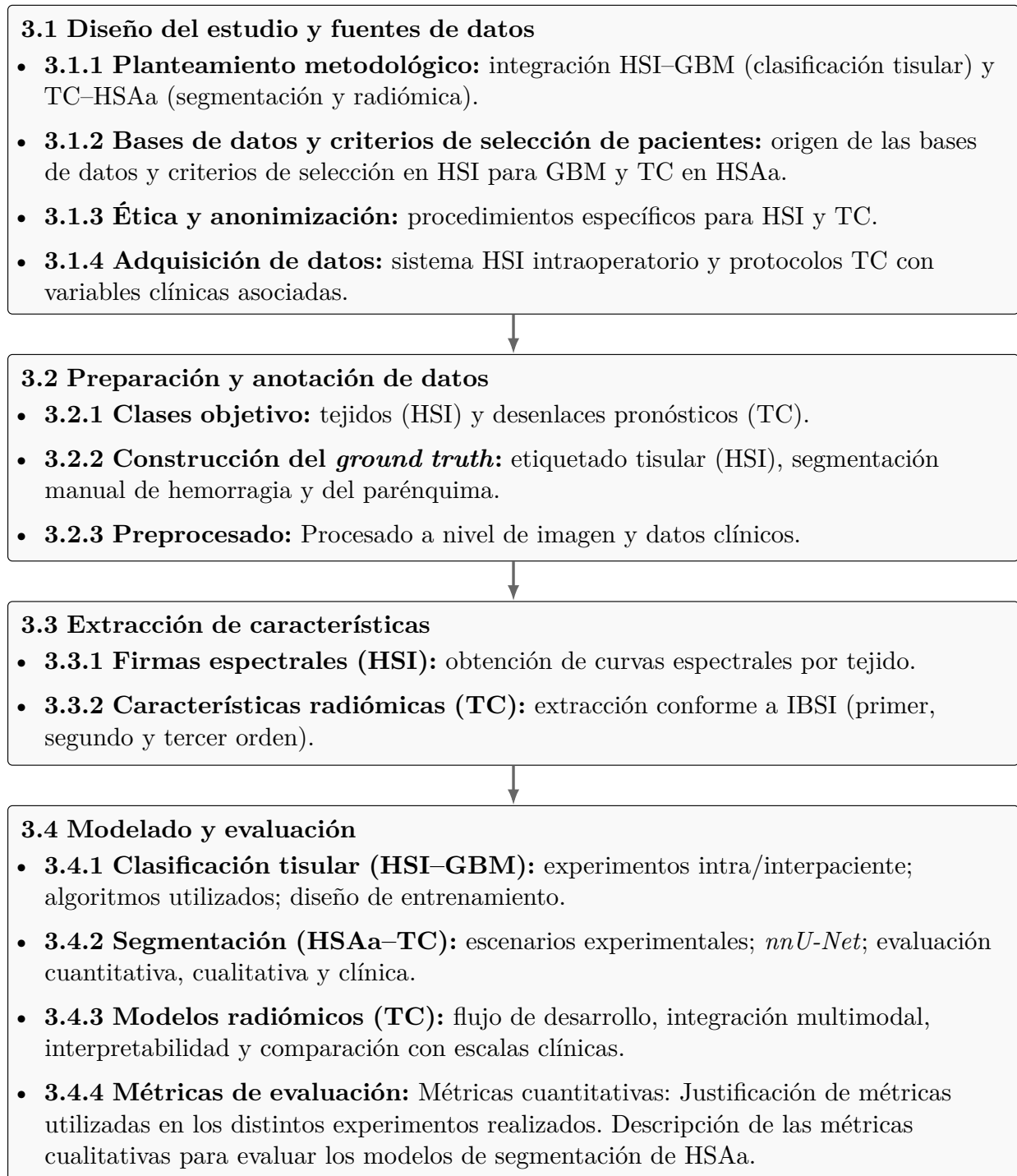


Figura 3.1: Estructura metodológica general del estudio, organizada en cuatro bloques principales (3.1–3.4) con tareas específicas para HSI en GBM y TC en HSAa.

con HSI intraoperatoria y la evaluación de la HSAa mediante TC.

3.1 Diseño del estudio y fuentes de datos

El primer bloque metodológico (Figura 3.2) resume la estructura del diseño del estudio, las fuentes de información y los procedimientos asociados a la adquisición y definición de las cohortes, tanto para el conjunto hiperespectral de glioblastoma (HSI–GBM) como para las imágenes de TC en hemorragia subaracnoidea aneurismática (TC–HSAa).

3.1 Diseño del estudio y fuentes de datos

- **3.1.1 Planteamiento metodológico:** Definición de las decisiones metodológicas para la integración de HSI–GBM (clasificación tisular) y TC–HSAa (segmentación y radiómica).
- **3.1.2 Bases de datos y criterios de selección de pacientes:** origen de las bases de datos y criterios de inclusión/exclusión en HSI para GBM y TC en HSAa.
- **3.1.3 Ética y anonimización:** procedimientos éticos y de anonimización específicos para HSI y TC, así como gestión de los datos.
- **3.1.4 Adquisición de datos:** descripción del sistema HSI intraoperatorio y sus componentes y protocolos TC con variables clínicas asociadas.

Figura 3.2: Bloque 3.1 — Diseño del estudio y fuentes de datos.

3.1.1 Planteamiento metodológico

El planteamiento metodológico se basa en desarrollar un flujo automatizado que, a partir de imagen médica y datos clínicos, permita extraer biomarcadores cuantitativos y construir modelos predictivos de valor clínico.

Este trabajo adopta una perspectiva transversal, abarcando tanto el contexto *intraoperatorio* como el *no intraoperatorio*, e integrando HSI y TC como escenarios complementarios, que permiten probar la robustez de la metodología en dos contextos clínicos distintos:

- **Contexto intraoperatorio (HSI en GBM):** adquisiciones *in vivo* en un entorno quirúrgico variable, con el desafío añadido de la heterogeneidad en la captura. El objetivo es demostrar que las firmas espectrales permiten discriminar tejido tumoral y estructuras adyacentes, aplicando protocolos de adquisición y preprocesado que garanticen mayor reproducibilidad.
- **Contexto no intraoperatorio (TC en HSAa):** adquisiciones volumétricas en un entorno clínico controlado, con el reto de caracterizar lesiones heterogéneas y complejas. El objetivo es superar las limitaciones de la evaluación cualitativa y de las escalas ordinales, integrando segmentación anatómica y radiómica en modelos predictivos de desenlaces clínicos.

Integración de adquisición hiperespectral y análisis espectral para la clasificación de tejidos intraoperatorios en glioblastoma

En el contexto de la cirugía de resección de GBM, se parte de la premisa de que, durante la intervención, la distinción entre tejido tumoral y tejido sano no siempre es evidente a simple vista. Esto puede derivar en una resección excesiva que comprometa áreas funcionales esenciales (como corteza motora o lenguaje) con el consiguiente deterioro neurológico y pérdida de calidad de vida; o, en una resección incompleta que deje restos tumorales y favorezca la recurrencia precoz, limitando tanto la supervivencia global como el beneficio terapéutico de la intervención (Grasso, 2016; Hardigan et al., 2023). En estas circunstancias, los avances en neurociencia intraoperatoria requieren herramientas capaces de proporcionar información objetiva y reproducible, que asista al cirujano en la toma de decisiones rápidas sin comprometer la seguridad del paciente. Entre las modalidades emergentes, la HSI ofrece la posibilidad de registrar, *in vivo* y sin contacto, la firma espectral de los tejidos expuestos, incorporando información que trasciende el rango visible y que potencialmente refleja diferencias fisiológicas y estructurales entre las distintas regiones cerebrales y vasculares (Fabelo et al., 2016; Ruiz et al., 2020).

En este escenario, la hipótesis inicial es que las firmas espectrales obtenidas mediante HSI durante la neurocirugía de GBM contienen información discriminativa suficiente para diferenciar, de forma reproducible, el tejido tumoral y las principales estructuras anatómicas y vasculares presentes en el campo quirúrgico. Asimismo, se plantea que la incorporación de un flujo estandarizado de adquisición, preprocesado y extracción de características espectrales reduciría la variabilidad técnica entre capturas y pacientes, favoreciendo la generalización de los modelos de clasificación resultantes (Mandal et al., 2021).

Por tanto, la primera decisión metodológica es definir un protocolo de adquisición intraoperatoria que garantice la calidad y la comparabilidad de las imágenes. La posición de la cámara, las condiciones de iluminación y la distancia al tejido se estandarizan y registran sistemáticamente, asegurando la coherencia de las capturas a lo largo de toda la cohorte. Paralelamente, se establece un procedimiento de preprocesado destinado a corregir artefactos ópticos, normalizar la señal y preservar la integridad espectral de cada píxel.

La segunda decisión clave es desarrollar un esquema robusto de construcción de conjuntos de datos anotados, basado en la identificación y etiquetado manual de regiones de interés por especialistas. Este paso resulta esencial para disponer de un *ground truth* fiable, sobre el que evaluar y entrenar los algoritmos de clasificación.

A partir de estas anotaciones, las firmas espectrales se transforman en características cuantitativas capaces de capturar no solo la magnitud de la reflectancia, sino también su variación a lo largo del espectro. De forma análoga a la radiómica en imágenes médicas convencionales, estas características constituyen un espacio de variables reproducible, menos dependiente de la interpretación subjetiva y más sensible a variaciones sutiles en el tejido.

En este contexto, la hipótesis de trabajo es que la combinación de un protocolo estandarizado de adquisición con una extracción sistemática de características espectrales permite desarrollar modelos de clasificación con un rendimiento técnico competitivo y con potencial de utilidad clínica en la diferenciación intraoperatoria de tejidos. La integración de estos modelos en

quirófano podría, a futuro, mejorar la precisión en la delimitación de márgenes y optimizar las decisiones quirúrgicas.

Operativamente, el proceso se estructura en fases encadenadas: (i) adquisición intraoperatoria de imágenes hiperespectrales; (ii) preprocesado para la corrección y normalización de la señal; (iii) anotación de regiones de interés por expertos; (iv) extracción de firmas espectrales; (v) organización y preparación de los datos para el modelado; y (vi) desarrollo, evaluación e interpretación de los modelos de clasificación.

El objetivo de este trabajo es, por tanto, el desarrollo integral de un sistema intraoperatorio basado en HSI, abarcando desde el diseño y la implementación de una estructura de adquisición adaptada al entorno quirúrgico hasta la creación de una base de datos propia de capturas intraoperatorias y el desarrollo de modelos basados en IA para la clasificación automática de tejidos. El propósito último de esta línea de investigación es proporcionar al neurocirujano una herramienta de apoyo objetiva y en tiempo real, capaz de distinguir entre tejido tumoral y tejido sano durante la intervención, optimizando así la precisión de la resección y contribuyendo a mejorar los resultados clínicos y funcionales del paciente.

Integración de segmentación automática y radiómica para la caracterización y predicción de desenlaces en Hemorragia Subaracnoidea Aneurismática

En este trabajo se parte de la hipótesis de que describir la hemorragia de forma cuantitativa y reproducible, más allá de la evaluación cualitativa tradicional, puede facilitar su integración en modelos que relacionen de manera más consistente los hallazgos de imagen con desenlaces clínicos de interés.

La HSAA es una enfermedad compleja que presenta una gran variabilidad en su extensión y localización, y cada patrón anatómico puede asociarse a riesgos y complicaciones diferentes. La afectación de las cisternas basales se asocia con un mayor riesgo de vasoespasmó (Dupont et al., 2009; Lagares et al., 2001); la extensión intraventricular puede bloquear la circulación del LCR y provocar hidrocefalia (Lagares et al., 2015); y la afectación parenquimatosa (ya sea por sangrado directo o por lesiones isquémicas) se asocia a un peor pronóstico funcional y mayor mortalidad (Jiménez-Roldán et al., 2013).

Por tanto, la primera decisión metodológica es desarrollar un sistema de segmentación automática capaz de identificar la HSAA y, al mismo tiempo, delimitar el parénquima cerebral para analizar cambios volumétricos y estructurales de forma global. Antes de iniciar el desarrollo, se lleva a cabo una revisión bibliográfica centrada en estas dos áreas: la segmentación de parénquima cerebral y la segmentación de sangre en el contexto de HSAA. Esta revisión permite identificar que, aunque existen trabajos previos en ambos ámbitos, los resultados obtenidos para segmentación de hemorragia en esta enfermedad son mejorables (García-García et al., 2024; Hu et al., 2023). Esto refuerza la necesidad de diseñar un algoritmo de segmentación específico que supere las limitaciones de las propuestas anteriores en HSAA. Para lograrlo, es necesario contar con segmentaciones manuales de las hemorragias realizadas por clínicos expertos, que sirvan como referencia (*ground truth*) para entrenar y validar el modelo. Basándose en el buen rendimiento demostrado en el estado del arte (Sección 2.8.2), se adopta el marco *nnU-Net* para su desarrollo (Isensee et al., 2021).

Una vez desarrollado el modelo de segmentación de hemorragias y, de forma paralela, la segmentación automática del parénquima cerebral, dichas segmentaciones pueden emplearse como base para el análisis cuantitativo. De esta manera, la segmentación deja de ser un fin en sí mismo para convertirse en el paso inicial de un flujo de trabajo más amplio, orientado a la extracción de biomarcadores cuantitativos.

La segunda decisión clave es, por tanto, ampliar el alcance de la caracterización más allá de las medidas volumétricas convencionales. Si bien estas métricas aportan información global de la carga hemorrágica o del volumen parenquimatoso, pueden resultar insuficientes para capturar la complejidad morfológica y la heterogeneidad interna de la HSAa. Las escalas radiológicas y clínicas habitualmente empleadas (como la WFNS, el SEBES o la escala de Fisher modificada) han demostrado valor pronóstico (Ahn et al., 2018; CG, 1988; Frontera et al., 2006), pero su naturaleza ordinal y su dependencia de la interpretación subjetiva introducen variabilidad y reducen la sensibilidad para detectar cambios sutiles o patrones no evidentes a simple vista.

Bajo esta premisa, se plantea que una descripción más rica y cuantitativa puede revelar asociaciones clínicas que permanecen ocultas cuando la evaluación se limita a variables categóricas o semicuantitativas. La radiómica ofrece precisamente esa capacidad: a partir de las regiones segmentadas, es posible extraer variables numéricas que capturan no solo el tamaño, sino también la forma, la distribución espacial, la intensidad y la textura tanto del parénquima cerebral como de la hemorragia. Estas características, al ser reproducibles y libres de subjetividad, permiten construir un espacio de variables más amplio y sensible a variaciones locales o patrones heterogéneos.

En este escenario, la hipótesis es que la integración de radiómica con segmentación anatómica detallada genera un conjunto de biomarcadores informativos, que pueden servir como entrada a modelos estadísticos y de aprendizaje automático, y además ofrecer una representación más fiel del estado inicial del paciente.

La inclusión de variables clínicas y de laboratorio se plantea como un complemento esencial. Si bien la radiómica puede revelar patrones invisibles a la inspección visual en imagen médica, la predicción de desenlaces clínicos en la HSAa también depende de las variables clínicas del paciente. De esta forma, la hipótesis de trabajo es que esta integración multimodal puede mejorar la robustez y la capacidad discriminativa de los modelos predictivos.

Operativamente, el proceso se estructura en fases encadenadas: (i) análisis del conjunto de datos disponible (imágenes de TC y variables clínicas); (ii) segmentación de las regiones de interés (sustancia blanca, sustancia gris y áreas de hemorragia); (iii) extracción de características radiómicas a partir de las segmentaciones; y (iv) desarrollo, evaluación e interpretación de los modelos predictivos.

En conjunto, este planteamiento metodológico busca transformar el conocimiento clínico y fisiopatológico de la enfermedad en un flujo de trabajo automatizado y cuantitativo, que va desde la segmentación precisa de las estructuras implicadas hasta la combinación de información procedente de distintas fuentes en modelos predictivos. El objetivo final es que las herramientas desarrolladas no solo ofrezcan un rendimiento técnico competitivo, sino que también generen biomarcadores interpretables, con potencial para integrarse en la práctica clínica habitual y contribuir de forma directa a la toma de decisiones en pacientes con HSAa.

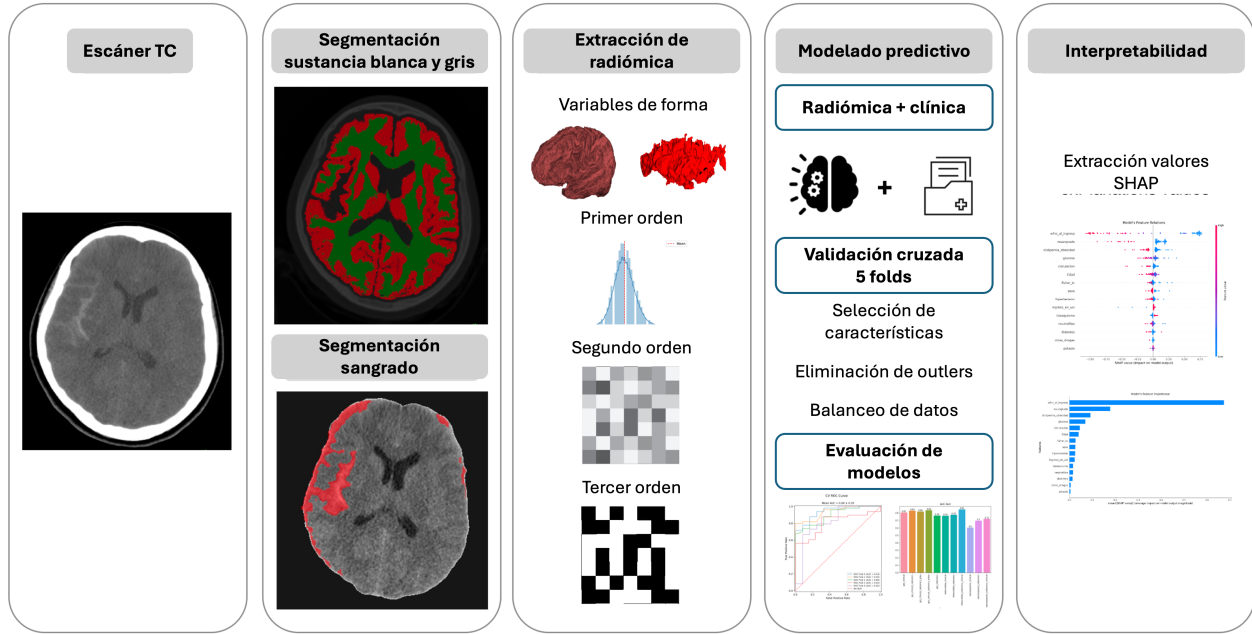


Figura 3.3: Flujo de trabajo de la metodología: a partir de una TC, se segmentan la parénquima y las lesiones. Se extraen las características radiómicas y se combinan con los datos clínicos para desarrollar modelos de IA (Urbanos et al., 2025).

3.1.2 Bases de datos y criterios de selección de pacientes

Este trabajo se apoya en bases de datos retrospectivas recopiladas íntegramente en el Hospital Universitario 12 de Octubre (H12O), en Madrid, España. Este centro cuenta con servicios de Neurocirugía, Neurorradiología intervencionista y Oncología neurológica de alta especialización, lo que garantiza una atención integral y centralizada para pacientes con patología neurovascular aguda, así como para aquellos con tumores cerebrales de alto grado, que requieren abordaje quirúrgico especializado.

Su papel como hospital de referencia implica que, además de atender a la población de su propia área sanitaria, recibe pacientes remitidos desde otros centros ante la sospecha de HSAa o para el tratamiento quirúrgico de tumores cerebrales complejos, incluyendo resecciones de GBM.

Esta concentración de casos se traduce en un volumen elevado y heterogéneo de pacientes, que abarca desde presentaciones clínicas leves hasta casos de gran complejidad técnica y pronóstica. Este espectro amplio y representativo resulta especialmente valioso para el desarrollo de modelos predictivos con potencial de aplicación en distintos escenarios clínicos, tanto en el ámbito vascular como en el oncológico neuroquirúrgico.

La selección de pacientes se llevó a cabo siguiendo criterios predefinidos, establecidos para preservar la homogeneidad clínica de cada cohorte y garantizar que las variables de imagen y los desenlaces se registraran de forma consistente y comparable en todos los casos.

Criterios de inclusión y exclusión

Ambos estudios contaron con criterios de inclusión y exclusión ajustados a las características de cada cohorte y modalidad de imagen.

Criterios para la cohorte de glioblastoma (Imagen hiperespectral)

- **Criterios de inclusión**

- Diagnóstico clínico de GBM en el paciente.
- Disponibilidad de capturas HSI intraoperatorias, con ficheros de calibración y parámetros técnicos estandarizados.
- Campo de visión que incluya tanto el área tumoral expuesta como tejido cerebral adyacente macroscópicamente sano.
- Documentación completa de metadatos quirúrgicos y técnicos (distancia cámara-superficie, ángulo de incidencia, condiciones de iluminación).

- **Criterios de exclusión**

- Tumores intracraneales distintos a GBM (p. ej., astrocitomas, oligodendrogliomas, metástasis).
- Capturas HSI con artefactos graves (desenfoque, saturación, pérdida de señal) que impidan la segmentación o el análisis espectral fiable.
- Capturas incompletas, ausencia de calibración o metadatos técnicos esenciales.
- Procedimientos quirúrgicos en los que no se expuso el tumor de forma directa al campo óptico del sistema HSI.

Criterios para la cohorte de Hemorragia Subaracnoidea Aneurismática (TC)

- **Criterios de inclusión**

- Diagnóstico de HSAa confirmado mediante TC craneal sin contraste realizada en el momento de la valoración inicial (al ingreso), antes de cualquier intervención terapéutica que pudiera alterar la distribución del sangrado o la morfología de las estructuras intracraneales. Este criterio asegura la homogeneidad temporal de la adquisición y maximiza la comparabilidad entre casos, dado que todos los estudios corresponden a la fase aguda de la enfermedad.
- Identificación del aneurisma responsable del sangrado mediante angiografía por CTA y/o DSA durante la hospitalización. La exigencia de esta confirmación angiográfica se justifica en la necesidad de excluir otras causas no aneurismáticas de HSA.

- **Criterios de exclusión**

- HSA secundaria a traumatismo craneoencefálico o a causas no aneurismáticas (por ejemplo: malformaciones arteriovenosas, vasculitis o coagulopatías graves), ya que la fisiopatología, la distribución del sangrado y la evolución clínica difieren

sustancialmente de las formas aneurismáticas, comprometiendo la comparabilidad de los datos.

- Ausencia de seguimiento clínico a los seis meses tras el evento índice. Este criterio se definió para asegurar que todos los pacientes incluidos contaran con información completa y fiable sobre los desenlaces, ya que la predicción del resultado funcional a medio plazo es uno de los objetivos clave del modelado.

3.1.3 Procedimientos éticos y anonimización

En ambos estudios se aplicaron procedimientos de aprobación ética y anonimización de datos, adaptados a las particularidades de cada cohorte y modalidad de imagen, garantizando el cumplimiento de la normativa vigente en materia de investigación biomédica y protección de datos.

Procedimientos éticos y de anonimización en glioblastoma (Imagen Hiperespectral)

Este estudio prospectivo ha sido revisado y aprobado por el comité ético del H12O, cumpliendo con la legislación nacional y el RGDP de la Unión Europea. Todas las adquisiciones hiperespectrales y los metadatos asociados se gestionan siguiendo protocolos estandarizados para garantizar la confidencialidad y la trazabilidad de la información.

Previo a su incorporación al entorno analítico, todos los casos se pseudonimizan mediante la asignación de un identificador ordinal consecutivo (*Patient ID*) del 1 al N , utilizado de manera consistente en todas las fases del estudio. Esta estrategia de gestión de la información asegura que se conserve únicamente la información técnica necesaria para el análisis, manteniendo la integridad de los datos y minimizando el riesgo de identificación de pacientes.

Los datos hiperespectrales se almacenaron en un repositorio interno seguro de la UPM, configurado con control de versiones y acceso restringido. La estructura de almacenamiento siguió una jerarquía *paciente* \rightarrow *estudio* \rightarrow *captura*, lo que garantiza la coherencia entre las distintas adquisiciones y preserva la relación entre los datos brutos, las calibraciones y los metadatos. Esta organización asegura la trazabilidad completa de cada adquisición y reduce el riesgo de errores o pérdidas de información.

Cada captura hiperespectral se acompaña de los elementos necesarios para reconstruir las condiciones de adquisición y permitir un preprocesado coherente: el cubo hiperespectral bruto, los ficheros de calibración (blanco y negro) y un conjunto estandarizado de metadatos técnicos y contextuales. La descripción detallada del contenido de cada adquisición y de la estructura de archivos se incluye en el Anexo A, Tabla A.1.

Procedimientos éticos y de anonimización en Hemorragia Subaracnoidea Aneurismática (TC)

En el estudio de HSAA con imágenes de TC todos los procedimientos utilizados para este trabajo forman parte de la atención clínica habitual y no implican intervenciones adicionales fuera de la práctica asistencial estándar. El estudio ha sido revisado y aprobado por el Comité de Ética de la Investigación con medicamentos (CEIm) del H12O (*código de aprobación: CEIm-19/078*), que determina que, dado su carácter retrospectivo y el uso exclusivo de

datos previamente anonimizados, no es necesaria la obtención de consentimiento informado individual. Esta decisión se basa en la ausencia de riesgo añadido para los pacientes y en el estricto cumplimiento de la normativa vigente en materia de protección de datos, incluido el Reglamento General de Protección de Datos (RGPD) de la Unión Europea.

Antes de su utilización en el análisis, todos los datos se someten a un proceso de pseudonimización. Al igual que en el caso de uso anterior (HSI en GBM), a cada paciente se le asigna un identificador interno (*Patient ID*) compuesto por un número ordinal consecutivo del 1 al N , que sirve como clave para vincular los datos clínicos y las imágenes asociadas. Este identificador se utiliza en el momento de la subida de los estudios a la plataforma *QUIBIM Precision*[®] V3.0,3 (Quibim, Valencia, España), la cual proporciona un entorno seguro y conforme a normativa para la gestión y almacenamiento de imágenes y datos médicos, cumpliendo con estándares internacionales de interoperabilidad y con la legislación europea en materia de protección de datos.

En paralelo, el registro de variables clínicas se pseudonimiza siguiendo el mismo esquema de numeración, de modo que la correspondencia entre datos clínicos e imágenes se mantiene de forma unívoca, pero sin que los investigadores tengan acceso a identificadores personales.

La plataforma *QUIBIM Precision*[®], configurada específicamente para este proyecto. Entre sus funcionalidades clave se incluyen:

- **Pseudonimización automática** de los estudios en el momento de la importación, eliminando o codificando los campos sensibles de la cabecera (p.ej. identificadores del paciente, fechas de nacimiento, números de serie).
- **Validación de integridad** para detectar series incompletas o inconsistencias en los metadatos.
- **Control de versiones** de las imágenes subidas y de las segmentaciones generadas, manteniendo un registro histórico de todas las modificaciones realizadas.

Asimismo, el sistema permite la exportación de los estudios en distintos formatos compatibles con entornos clínicos y de investigación, incluyendo la recuperación en cualquier momento de las series originales.

Gracias a esta infraestructura es posible mantener un control riguroso de la información, reducir riesgos asociados a la manipulación manual de archivos y asegurar la reproducibilidad de los análisis.

3.1.4 Proceso de adquisición de datos

El proceso de adquisición de datos en este estudio abarcó tanto imágenes médicas como información clínica y técnica asociada, adaptándose a las particularidades de cada modalidad y contexto clínico. Para el GBM, las adquisiciones se basaron en HSI intraoperatoria, empleando un sistema de captura optimizado para el entorno quirúrgico y acompañado de un registro exhaustivo de metadatos técnicos. En el caso de HSAA, se utilizaron imágenes de TC obtenidas bajo protocolos estandarizados para patología neurovascular aguda, junto con variables clínicas extraídas de la historia clínica electrónica (HCE). En ambos casos, la homogeneidad en

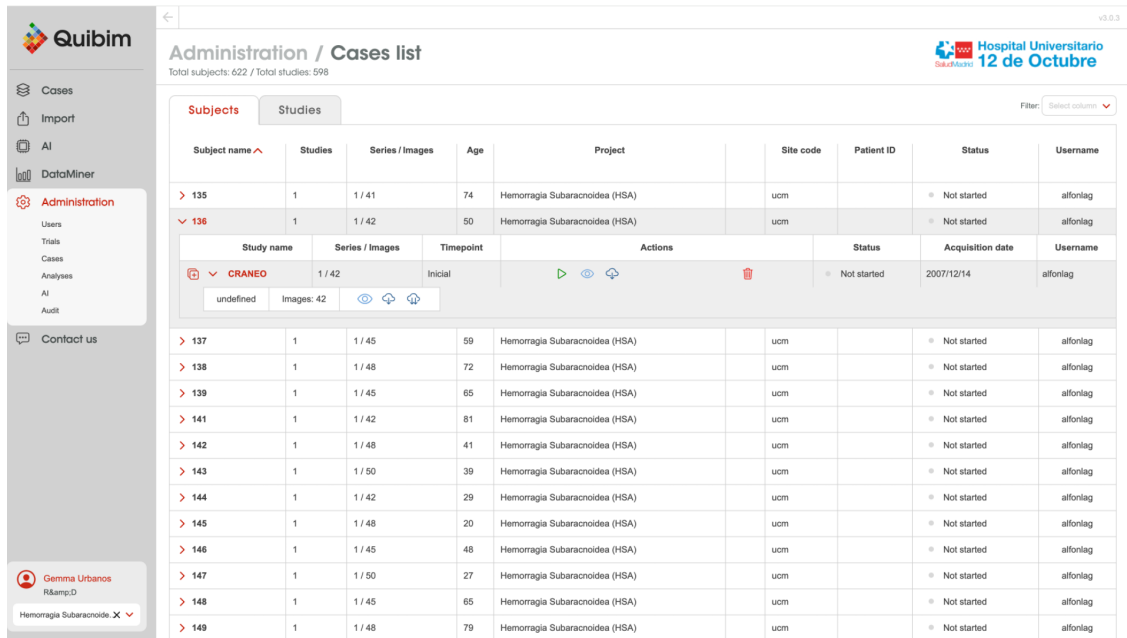


Figura 3.4: Interfaz general de la plataforma QUIBIM Precision® utilizada para la gestión de estudios. Se muestra el listado de casos almacenados con identificadores pseudonimizados y la estructura jerárquica de organización de los datos.

los protocolos de adquisición y la documentación sistemática de parámetros garantizan la trazabilidad, reproducibilidad y comparabilidad de los datos a lo largo de todo el flujo metodológico.

Imagen y datos espectrales en glioblastoma mediante imagen hiperespectral

Esta sección describe el sistema empleado para la adquisición intraoperatoria de HSI en casos de GBM, junto con la organización general del flujo de captura, la estructura de metadatos y las estrategias adoptadas para asegurar la calidad y la coherencia de las imágenes obtenidas.

Sistema de captura de imagen hiperespectral intraoperatorio

Las HSI se adquirieron *in vivo* durante procedimientos de neurocirugía en pacientes diagnosticados de GBM durante 2019-2022. Para ello se empleó un sistema de captura de tipo *snapshot* diseñado específicamente para el entorno quirúrgico. Esta modalidad permite registrar en una única adquisición la totalidad del cubo espectral, evitando movimientos o artefactos asociados a sistemas de escaneo secuencial y reduciendo el impacto sobre la dinámica operatoria. El diseño compacto y ligero del sistema favoreció su integración en el campo quirúrgico sin interferir con la esterilización ni prolongar el tiempo de exposición del paciente.

Unidad de adquisición.

El sistema central de captura se basa en una cámara hiperespectral instantánea con sensor multibanda capaz de registrar 25 bandas espectrales de manera uniforme en el rango visible e infrarrojo cercano (VNIR, 655–975 nm). Este intervalo cubre tanto las longitudes de onda

sensibles a la reflectancia superficial del tejido como aquellas asociadas a la oxigenación y al contenido hídrico, lo que resulta de particular relevancia para la diferenciación entre tejido tumoral y sano (Lu y Fei, 2014). El empleo de una cámara de tipo matricial elimina la necesidad de movimiento mecánico durante la adquisición, mejorando la repetibilidad de las capturas y la estabilidad espectral. Las especificaciones completas del dispositivo (incluyendo modelo, dimensiones, resolución y características del sensor) se detallan en el Anexo B, Tabla A.2.

Óptica y filtrado.

La cámara se equipa con una lente VIS/NIR de baja distorsión y gran apertura, seleccionada para mantener la nitidez en todo el rango espectral. A la lente se le acopla un filtro de corte en torno a 650 nm para suprimir componentes no deseadas del espectro y optimizar la respuesta del sensor en el rango VNIR. Con el fin de asegurar la comparabilidad intercaso, se mantienen **parámetros homogéneos de captura**, empleando valores fijos de exposición y apertura para todas las tomas (valores típicos: 70 ms y $f/4$; véase el Anexo A, Tabla A.2). Esta estandarización permite reducir la variabilidad instrumental y garantizar una calibración espectral coherente entre intervenciones.

Iluminación.

La escena quirúrgica se ilumina mediante una fuente halógena continua acoplada a dos guías de fibra óptica orientables situadas a ambos lados de la cámara. Esta disposición permite dirigir la luz de forma precisa y minimizar sombras y reflejos especulares, asegurando una iluminación homogénea del campo de interés. El espectro de la fuente cubre completamente el rango de sensibilidad del sensor (VNIR), evitando pérdidas de señal en los extremos del espectro. Los parámetros de potencia, geometría y disposición se incluyen igualmente en el Anexo A, Tabla A.2.

Geometría de adquisición.

La posición de la cámara y de las fuentes de iluminación se ajusta individualmente para cada intervención, adaptándose a la anatomía expuesta y a la ergonomía del cirujano. Se han definido rangos preestablecidos de distancia y ángulo de incidencia que garantizaban una cobertura uniforme del campo y reducen artefactos de iluminación. Los valores de distancia y ángulo de las adquisiciones se detallan en el Anexo A, Tabla A.2.

Control y registro de datos. El flujo de adquisición se gestionó inicialmente mediante el software proporcionado por el fabricante (*XiSpec01* (XIMEA GmbH, 2017)) y, posteriormente, a través de una herramienta propia desarrollada para este proyecto. Esta aplicación permite la captura automática del cubo HSI, su almacenamiento inmediato en un entorno seguro y la asociación de un conjunto estandarizado de metadatos técnicos y contextuales (exposición, apertura, distancia, ángulo, tipo de tejido, fase quirúrgica, operador y estado de calibración). El registro estructurado de estos metadatos facilita la validación posterior de la información espectral y la integración con las anotaciones clínicas y quirúrgicas.

Ventajas operativas.

El uso de una cámara *snapshot* de reducidas dimensiones, junto con la homogeneización de

los parámetros de captura y un control exhaustivo de las condiciones de iluminación, permite obtener conjuntos de imágenes con elevada consistencia técnica y mínima interferencia en la dinámica operatoria. La ausencia de partes móviles reduce artefactos por microvibraciones o desplazamientos, mientras que la captura instantánea favoreció la obtención de datos en momentos críticos de la cirugía. En conjunto, el protocolo desarrollado proporciona cubos hiperespectrales estables, comparables y adecuados para su análisis cuantitativo y para el entrenamiento de modelos de clasificación de tejidos basados en IA.

Imagen y datos clínicos en Hemorragia Subaracnoidea Aneurismática mediante TC

Esta sección describe las fuentes y características de la información empleada para el análisis de HSAa, incluyendo las imágenes obtenidas mediante TC y los datos clínicos asociados.

Protocolos de adquisición de imagen

Las imágenes se han adquirido en equipos de TC multidetectores pertenecientes al Servicio de Radiodiagnóstico del Hospital 12 de Octubre, siguiendo protocolos estandarizados para la valoración de patología neurovascular aguda. Posteriormente, se exportan en formato *Digital Imaging and Communications in Medicine* (DICOM), el estándar internacional para el almacenamiento y transmisión de imágenes médicas. Este formato integra no solo la información de los voxels de la imagen, sino también un conjunto amplio de metadatos relativos al paciente, los parámetros técnicos de adquisición y la calibración del escáner, lo que garantiza la preservación de la calidad original, la trazabilidad del estudio y su interoperabilidad con sistemas PACS hospitalarios.

Para facilitar su manejo en entornos de análisis cuantitativo y de neuroimagen, los volúmenes DICOM se convierten a formato *Neuroimaging Informatics Technology Initiative* (NIfTI). A diferencia del DICOM, que organiza la información en múltiples ficheros correspondientes a cortes bidimensionales, NIfTI almacena el volumen tridimensional completo en un único fichero acompañado de una cabecera compacta que define dimensiones, resolución y orientación espacial. Esta estructura permite un acceso más eficiente a los datos y simplifica la integración con librerías especializadas en segmentación, extracción de características radiómicas y modelado predictivo.

Fuentes de datos para las variables clínicas

Además de las imágenes, se han recopilado variables clínicas relevantes obtenidas directamente de la HCE del hospital. La selección de estas variables se centra en información clínica y de laboratorio de interés para el estudio. La extracción se realiza por neurocirujanos siguiendo un protocolo predefinido, lo que asegura que los datos sean completos, coherentes y homogéneos entre pacientes. Finalmente, todos los datos se almacenan y gestionan en la plataforma REDCap (Research Electronic Data Capture), una herramienta ampliamente utilizada en investigación biomédica para la captura y gestión de datos clínicos. REDCap ofrece una infraestructura flexible basada en formularios electrónicos personalizables, lo que facilita la estandarización de las variables y la reducción de errores en la entrada de datos.

Desde el punto de vista de seguridad, REDCap cumple con estándares internacionales de

protección de datos, incluyendo control de accesos mediante autenticación, encriptación de la información y registro de auditorías para el seguimiento de modificaciones. Además, permite la exportación de los datos en formatos compatibles con software estadístico habitual (como Excel o CSV), lo que facilita su integración en los flujos de análisis (Harris et al., 2009, 2019).

3.2 Preparación y anotación de datos

El segundo bloque metodológico (Figura 3.5) detalla las etapas de preparación, anotación y preprocesado de los datos. Incluye la generación de anotaciones de referencia para imagen (*ground truth*) en cada modalidad, así como las fases de preprocesado específicas para imagen y datos clínicos.

3.2 Preparación y anotación de datos

- **3.2.1 Clases objetivo:** tejidos (HSI) y desenlaces pronósticos (TC) definidos como *ground truth* objetivo.
- **3.2.2 Construcción del *ground truth*:** flujo de etiquetado tisular (HSI), segmentación manual de hemorragia y del parénquima cerebral (TC).
- **3.2.2 Preprocesado de datos:** Preprocesado de imagen: HSI en GBM, TC en HSAa para la segmentación automática y modelado radiómico. Procesado de datos clínicos en la base de datos HSAa.

Figura 3.5: Bloque 3.2. Preparación y anotación de datos: generación del *ground truth*, preprocesado de imagen y variables clínicas.

3.2.1 Definición de las clases objetivo

En este apartado se describen y justifican las clases objetivo empleadas en el proceso de modelado, especificando los criterios utilizados para su definición y su relevancia para ser usadas dentro del problema de predicción abordado por los modelos de IA.

Clases tisulares en glioblastoma (Imagen Hiperespectral)

El objetivo principal de la clasificación en este estudio es la diferenciación, a nivel de píxel, de las principales clases tisulares y estructurales presentes en el campo quirúrgico durante la resección de GBM. Estas clases se definieron en colaboración con neurocirujanos, considerando su representatividad en la práctica intraoperatoria y la plausibilidad de que presenten firmas espectrales diferenciables dentro del rango de adquisición (VNIR). El conjunto de clases (representado en una imagen real de quirófano en la Figura 3.6) establecido es el siguiente:

1. **Tejido sano cortical.** Incluye tanto la corteza cerebral como la sustancia blanca o gris visible tras la apertura de la duramadre y exposición del campo. Este tejido presenta patrones de reflectancia asociados a su contenido en mielina, agua y vascularización superficial. Su identificación es fundamental para preservar áreas funcionales y evitar resecciones excesivas.

2. **Tejido tumoral (GBM).** Corresponde a las áreas identificadas intraoperatoriamente como tumorales según el criterio del neurocirujano, basándose en la textura, color, infiltración y consistencia del tejido. El GBM presenta una alta heterogeneidad espectral debido a su variabilidad histológica (zonas hipercelulares, necrosis, proliferación vascular). Su discriminación respecto al tejido sano es el objetivo principal del sistema, ya que permite guiar la extensión de la resección.
3. **Vaso arterial.** Vasos de alto flujo y presión, ricos en oxihemoglobina, con paredes más tensas y un tono más claro que los venosos. Presentan picos y valles de absorción característicos en longitudes de onda asociadas a la hemoglobina oxigenada. Su preservación es crítica para mantener la perfusión cerebral en áreas funcionales.
4. **Vaso venoso.** Estructuras vasculares de retorno venoso, de menor presión y mayor contenido de hemoglobina reducida (desoxihemoglobina). Espectralmente, los vasos venosos muestran una absorción más marcada en bandas sensibles a la hemoglobina desoxigenada, con menor reflectancia global que el tejido circundante. Su identificación es importante para evitar lesiones que comprometan el drenaje venoso cerebral.
5. **Duramadre.** Tejido meníngeo fibroso que cubre el encéfalo, visible en fases tempranas de la exposición o en áreas no incluidas en la craneotomía. Su aspecto blanquecino y su elevada reflectancia en la región visible lo diferencian del tejido cerebral y vascular. Aunque no suele ser objetivo de resección en este contexto, su identificación correcta se incluye para el entrenamiento de los algoritmos, evitando que se confunda con otras clases.

Desenlaces pronósticos en Hemorragia Subaracnoidea Aneurismática (TC)

La elección de los desenlaces se realizó considerando su relevancia clínica (véase Sección 2.8.2), su disponibilidad en la cohorte y su capacidad para ser tratados como variables binarias en modelos de aprendizaje automático.

Desenlace principal: GOS a 6 meses

El resultado principal evaluado es la GOS a los seis meses del diagnóstico, obtenida mediante seguimiento clínico y verificación en la HCE. Se ha seleccionado este punto temporal (6 meses tras diagnóstico) porque proporciona una estimación estable del estado funcional una vez superada la fase aguda.

A partir de la GOS se plantearon dos tareas de clasificación independientes, diseñadas para capturar aspectos complementarios del pronóstico:

- **Predicción del resultado funcional:** Se sigue una división ampliamente utilizada en la literatura (Hofmann et al., 2022; Lagares et al., 2015) que permite distinguir entre pacientes que recuperan un grado aceptable de autonomía funcional y aquellos con dependencia grave o fallecimiento.
 - Buen resultado: GOS 4–5 (discapacidad moderada o buena recuperación).
 - Mal resultado: GOS 1–3 (discapacidad grave, estado vegetativo o muerte).

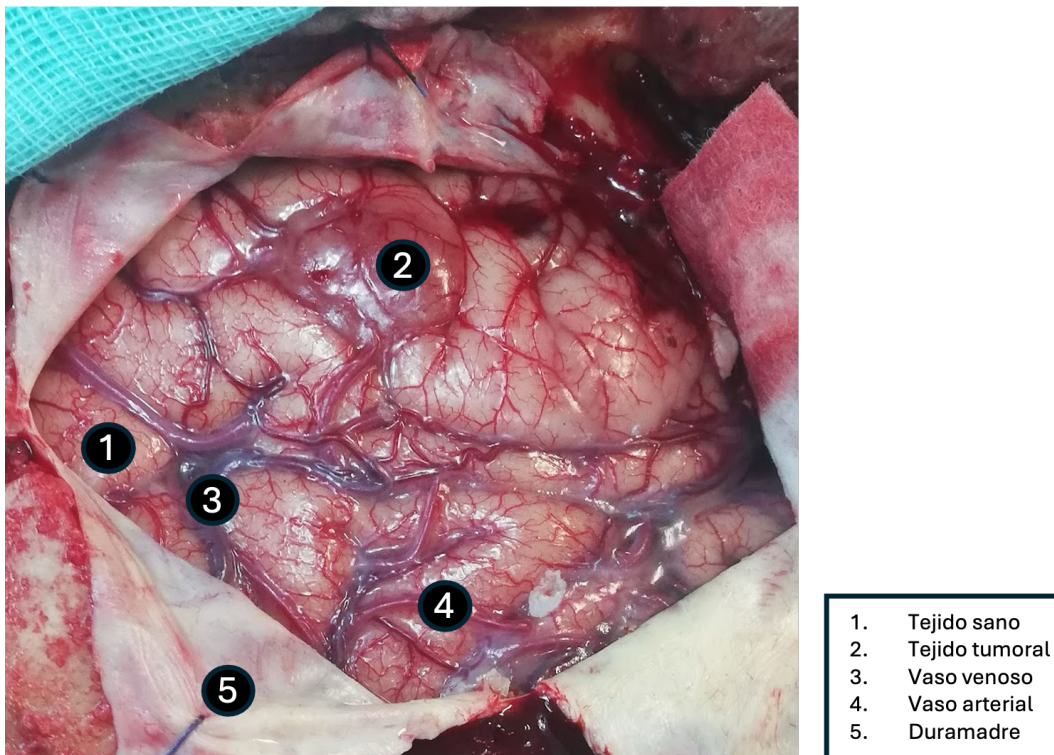


Figura 3.6: Ejemplo de imagen real intraoperatoria con identificación numérica de las cinco clases tisulares utilizadas. (1) tejido sano, (2) tejido tumoral, (3) vaso venoso, (4) vaso arterial, (5) duramadre.

- **Predicción de mortalidad:** Tratar la mortalidad como desenlace independiente permite evaluar el rendimiento del modelo en un resultado crítico y no lineal, evitando que la señal se diluya al agruparla con otros grados de discapacidad.
 - Supervivencia: $GOS \geq 2$.
 - Muerte: $GOS = 1$.

Resultados secundarios

Se incluyen dos resultados binarios adicionales, seleccionadas por su frecuencia en la evolución de la HSAa y su impacto potencial en el pronóstico (descripción clínica resumida en Sección 2.8.2). Estos desenlaces secundarios son el vasoespasma y la hidrocefalia, medidos también a los seis meses tras el diagnóstico. Sin embargo, con el fin de evitar sesgos en la estimación de la evolución clínica y en la incidencia de complicaciones, se han excluido los pacientes que fallecieron antes de los siete días posteriores al ingreso, ya que no permiten una observación suficiente del desarrollo de complicaciones, así como aquellos casos en los que no consta registro del desenlace correspondiente.

- **Vasoespasma (sí/no):** Variable dicotómica definida por la presencia de criterios clínicos, como nuevos déficits neurológicos focales o disminución del nivel de conciencia no atribuibles a resangrado, hidrocefalia o crisis, y criterios radiológicos, consistentes en estenosis arterial identificada en angio-TC o angiografía digital y confirmada por un neurorradiólogo especializado.
- **Hidrocefalia (sí/no):** Codificada cuando el agrandamiento ventricular sintomático requiere una derivación definitiva de LCR, incluyendo aquellos casos en los que no es posible retirar el drenaje ventricular externo durante el ingreso o en los seis meses posteriores.

La combinación de un desenlace principal, formulado en dos aproximaciones complementarias (pronóstico funcional y mortalidad), junto a desenlaces secundarios, permite obtener una visión completa tanto de la evolución global del paciente como de la aparición de complicaciones específicas. La elección de codificar todas las variables de manera binaria permite facilitar la comparación directa de métricas de rendimiento entre tareas y con estudios previos y simplificar el entrenamiento y validación de modelos supervisados de clasificación.

3.2.2 Construcción del *ground truth*

La obtención de un *ground truth* fiable resulta esencial para el entrenamiento y validación de los algoritmos de segmentación y clasificación. En ambos escenarios, las anotaciones se han realizado mediante la construcción de máscaras manuales adaptadas a cada modalidad: en HSI (GBM) se han delimitado las principales clases tisulares de interés, mientras que en TC (HSAa) se han segmentado la hemorragia y el parénquima cerebral.

Etiquetado tisular en Imagen Hiperespectral de glioblastoma

Para la construcción del *ground truth* en HSI se ha empleado una herramienta de anotación basada en el trabajo de (Fabelo, Ortega, Ravi et al., 2018). Se trata de una herramienta

Tabla 3.1: Resumen de desenlaces utilizados en el estudio, con su codificación binaria, fuente de datos y criterios temporales de evaluación.

Desenlace	Codificación binaria	Fuente de datos	Criterios temporales
GOS – Resultado funcional	4–5 (buen resultado) / 1–3 (mal resultado)	Seguimiento clínico + HCE	6 meses post-HSAa
GOS – Mortalidad	≥ 2 (supervivencia) / 1 (fallecimiento)	Seguimiento clínico + HCE	6 meses post-HSAa
Vasoespasmio	Sí / No	HCE + informes radiológicos	Evaluado hasta 6 meses; excluidos fallecidos antes de 7 días
Hidrocefalia	Sí / No	HCE + informes de neurocirugía	Evaluado hasta 6 meses; excluidos fallecidos antes de 7 días

interactiva diseñada para identificar y clasificar tejidos de forma precisa y reproducible en el contexto intraoperatorio.

Flujo de trabajo de etiquetado

Para facilitar la inspección visual y la anotación experta, a partir de cada cubo hiperespectral preprocesado (véase Sección 3.2.3) se genera una imagen RGB sintética en *falso color*. La composición se construye mapeando tres bandas representativas de regiones baja, media y alta del rango VNIR a los canales R, G y B, respectivamente, con el objetivo de favorecer el contraste de estructuras vasculares, parénquima y tejido tumoral. Más información sobre los valores utilizados para cada canal se detallan en el Anexo A, Sección A.2.

El proceso de etiquetado consiste en: (i) selección por parte del usuario de un píxel de referencia o vecindad dentro de la imagen RGB sintética; (ii) segmentación inicial mediante similitud espectral (principalmente el *Spectral Angle Mapper* (SAM) (Kruse et al., 1993)) ; (iii) refinado manual poligonal para aislar la región anatómica de interés; y (iv) asignación de la etiqueta tisular correspondiente según los criterios del protocolo. La anotación se realiza por el neurocirujano responsable de la intervención, siguiendo un conjunto de clases predefinidas (Sección 3.2.1). Las definiciones operativas y ejemplos visuales de la herramienta de anotación utilizada se muestran en el Anexo A, Figura A.1.

Métrica de similitud (SAM).

El *Spectral Angle Mapper* mide el ángulo α entre el vector espectral del píxel de referencia \mathbf{r} y el de un píxel objetivo \mathbf{t} :

$$\alpha = \arccos \left(\frac{\mathbf{t} \cdot \mathbf{r}}{\|\mathbf{t}\| \|\mathbf{r}\|} \right)$$

Cuanto menor es α , mayor es la similitud entre ambos. Dado que α está acotado entre 0 y $\pi/2$ radianes, es habitual normalizarlo como:

$$\text{SAM} = \frac{2\alpha}{\pi}$$

Esta métrica es insensible a cambios de iluminación o de ganancia, ya que se basa en la forma del espectro más que en su magnitud.

Segmentación manual de hemorragia

Para la anotación manual de las imágenes se ha empleado la plataforma QUIBIM Precision[®]. El proceso de segmentación es llevado a cabo íntegramente por neurocirujanos con experiencia en el manejo clínico y quirúrgico de la HSAa, lo que garantiza la coherencia anatómica y la validez clínica de las etiquetas generadas.

En cada estudio, las áreas afectadas por los distintos patrones de sangrado intracraneal secundarios a la HSAa se han segmentado de forma diferenciada, asignando a cada una una etiqueta independiente con su identificador y color propios. Esta clasificación permite distinguir las tres localizaciones de sangrado más relevantes en este contexto: la HSA, HIV y HIP. Las máscaras resultantes se exportaron en formato estándar NIFTI para su integración en el flujo de análisis. Una descripción más detallada de las modalidades de segmentación y del entorno de trabajo empleado se incluye en el Anexo B, B.1.

Segmentación automática del parénquima cerebral

La segmentación del parénquima pretende cuantificar información global de la sustancia gris y sustancia blanca, y por tanto, evaluar posibles cambios estructurales del cerebro asociados a la hemorragia o a procesos secundarios como el vasoespasma o el edema.

Para segmentar el parénquima se utiliza **CT-Seg**, un algoritmo desarrollado para la segmentación automatizada de TC craneales (Adduru et al., 2020). Este software está construido sobre SPM12 (Ashburner et al., 2014) y emplea un enfoque basado en atlas, que combina registro espacial y clasificación probabilística de tejidos.

El flujo de trabajo de **CT-Seg** se puede desglosar en las siguientes etapas:

1. **Registro al espacio estándar:** cada volumen de TC es alineado al espacio anatómico de referencia MNI152 (Fonov et al., 2011) mediante un registro afín inicial seguido de un ajuste no lineal, corrigiendo así diferencias en tamaño, orientación y morfología craneal.
2. **Corrección de intensidad:** se aplica un ajuste de histograma para homogeneizar la distribución de valores Hounsfield y reducir variabilidad entre exploraciones.
3. **Clasificación tisular:** el algoritmo asigna a cada vóxel una probabilidad de pertenencia a seis clases: sustancia gris, sustancia blanca, LCR, cráneo, tejido blando extracraneal y fondo.
4. **Umbralización y generación de máscaras binarias:** se aplica un umbral de probabilidad para cada clase, generando máscaras binarias que delimitan las regiones