

UTILIDAD DE MODELOS RADIÓMICOS EXPLICABLES EN LA TC

URGENTE PARA LA DIFERENCIACIÓN DE ABSCESOS CEREBRALES Y

GLIOBLASTOMAS

Diego Alejandro García Paredes, Diego Galve López, Esther García Casado, José Montilla López, Carmen Campos Ferrer, Inés Pecharromán Las Heras, Agustina

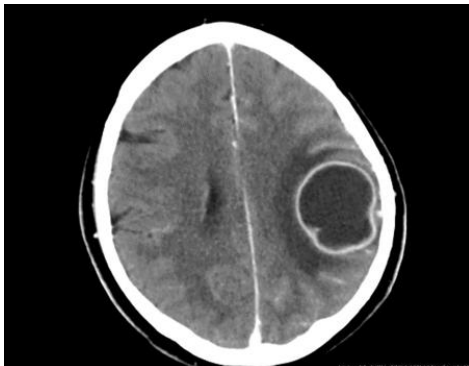
Vicente Bartulos, Ángel Torrado-Carvajal

Hospital Universitario Ramón y Cajal

INTRODUCCIÓN

ABSCESOS CEREBRALES

- Colección de pus alojada en el tejido cerebral
- Pueden generarse por diversas vías
 - Infecciones externas: osteomielitis, sinusitis, mastoiditis
 - Heridas penetrantes
 - Propagación hematógica



GLIOBLASTOMAS

- Tumor del sistema nervioso central
- Formado a partir de tejido glial del encéfalo y la médula espinal
- Extremadamente difícil de tratar debido a su rápido crecimiento



RELEVANCIA CLÍNICA

- Diferenciar las lesiones con realce en anillo en TC —en particular los abscesos y los glioblastomas (GBM)— puede resultar complejo. Un diagnóstico precoz y preciso es fundamental, ya que el tratamiento y el pronóstico difieren de forma significativa.
 - Los abscesos cerebrales requieren tratamiento urgente, antimicrobiano y/o quirúrgico mientras que los glioblastomas precisan tratamiento oncológico multimodal.
 - Un diagnóstico erróneo puede conllevar un aumento significativo de la morbilidad y la mortalidad.
 - La TC es una prueba rápidamente accesible y suele ser el primer estudio de imagen en urgencias, mientras que la RM, considerada el estándar de referencia y utilizada en otros estudios de modelos radiómicos, no siempre está disponible y supone un mayor coste.
-

OBJETIVO

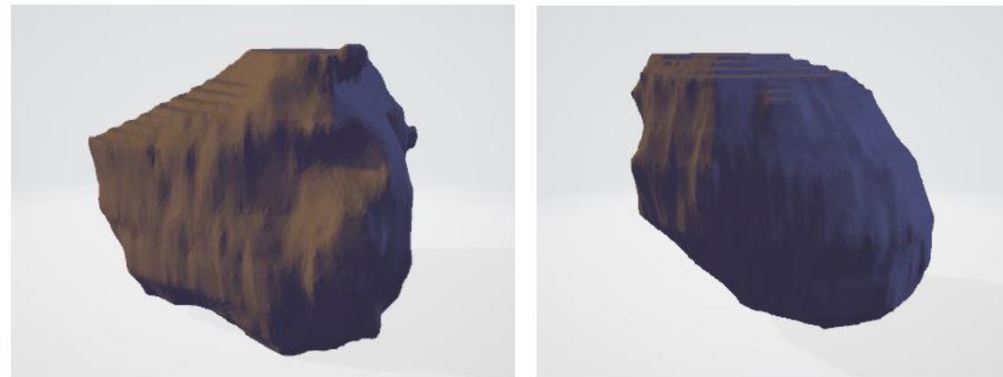
- Diseñar y evaluar un algoritmo radiómico que permita, a partir de imágenes de TC, diferenciar de manera precisa entre abscesos cerebrales y glioblastomas
-

MATERIAL Y MÉTODOS: BASE DE DATOS

- Estudio retrospectivo donde se incluyeron 52 pacientes : 26 con abscesos cerebrales y 26 con glioblastoma
 - Todos los casos habían sido confirmados mediante anatomía patológica
 - Inicialmente hubo un total de 576 casos de glioblastomas que fueron revisados por 5 radiólogos para escoger aquellos que tuviesen las características mas confusas y solapadas con los abscesos cerebrales
-

SEGMENTACIÓN

- Usando un visor 3D se seleccionaba el área de interés manualmente
- Se decidió segmentar exclusivamente la región del tumor o absceso, excluyendo el área de edema y efecto masa adyacente
- Finalmente se exporta en formando STL



Visualización tridimensional de un glioblastoma (izquierda) y un absceso (derecha) segmentados usando Synapse3D

EXTRACCIÓN DE CARACTERÍSTICAS

- Se seleccionaron todas aquellas que pudieran resultar de interés, un total de 107, de los siguientes tipos:
 - Características de forma : 14
 - Características de primer orden : 18
 - GLCM (Gray Level Co-occurrence Matrix) : 24
 - GLDM (Gray Level Dependence Matrix) : 14
 - GLRLM (Gray Level Run Length Matrix) : 16
 - GLSZM (Gray Level Size Zone Matrix) : 16
 - NGTDM (Neighboring Gray Tone Difference Matrix) : 5
-

SELECCIÓN DE CARACTERÍSTICAS

- La selección de características radiómicas reduce la dimensionalidad, mejorando la eficiencia y interpretabilidad del modelo.
 - Se aplicó la correlación de Pearson con un umbral definido para eliminar características redundantes.
 - El algoritmo evalúa de forma iterativa cada par de características y elimina una de ellas cuando la correlación absoluta supera el umbral definido.
 - Este proceso permitió reducir las características de 107 a 71.
-

EXPLICABILIDAD

- La explicabilidad del modelo es clave en el ámbito médico para generar confianza clínica y facilitar la toma de decisiones.
 - Se utilizó el método SHAP, basado en teoría de juegos, para cuantificar la contribución individual de cada característica a la predicción.
 - SHAP permite descomponer las predicciones en aportaciones de cada variable, garantizando transparencia y consistencia en la interpretación.
 - Las visualizaciones SHAP identifican las características más influyentes, proporcionando una comprensión clara del comportamiento global del modelo.
-

CLASIFICACIÓN

- En total se usaron tres distintos modelos radiómicos para analizar las características:
 - Random Forest: Clasificador basado en múltiples árboles, entrenado con validación cruzada estratificada, que mejora la robustez y reduce el sobreajuste.
 - Explainable Boosting Machine (EBM): Modelo interpretable basado en boosting, que combina rendimiento competitivo con análisis claro de la contribución individual de las características.
 - XGBoost: Algoritmo de boosting de gradiente con regularización, altamente robusto y eficaz para el análisis de datos radiómicos complejos.
-

MÉTRICAS DE EVALUACIÓN

- Se seleccionaron las siguientes métricas:
 - Precisión : Proporción de predicciones correctas, clave para evitar falsos positivos innecesarios.
 - Sensibilidad : Capacidad del modelo para detectar casos positivos reales, prioritaria en contexto clínico.
 - F1 Score : Media armónica entre precisión y sensibilidad, que equilibra falsos positivos y falsos negativos.
 - Área bajo la curva ROC (AUC) : Medida global de la capacidad discriminativa del modelo a distintos umbrales, estimada mediante ROC promediada con validación cruzada para mayor robustez.
-

RESULTADOS: SELECCIÓN DE CARACTERÍSTICAS

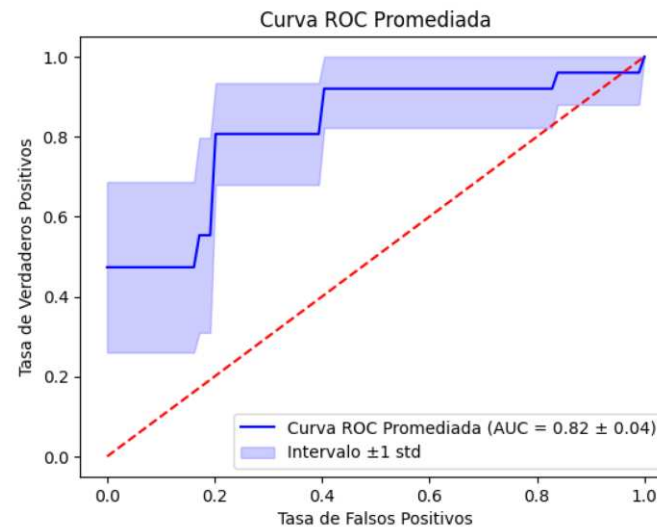
- Tras evaluar distintos clasificadores con varios umbrales de correlación, el valor $r = 0,65$ ofreció el mejor rendimiento global.

r	0,6	0,62	0,65	0,67	0,7	0,72	0,75
Random Forest	0,73	0,717	0,72	0,706	0,74	0,68	0,68
EBM	0,654	0,727	0,769	0,745	0,717	0,717	0,717
XGBoost	0,701	0,703	0,727	0,72	0,745	0,703	0,76

Comparación del F1 score para cada modelo cambiando el índice de correlación

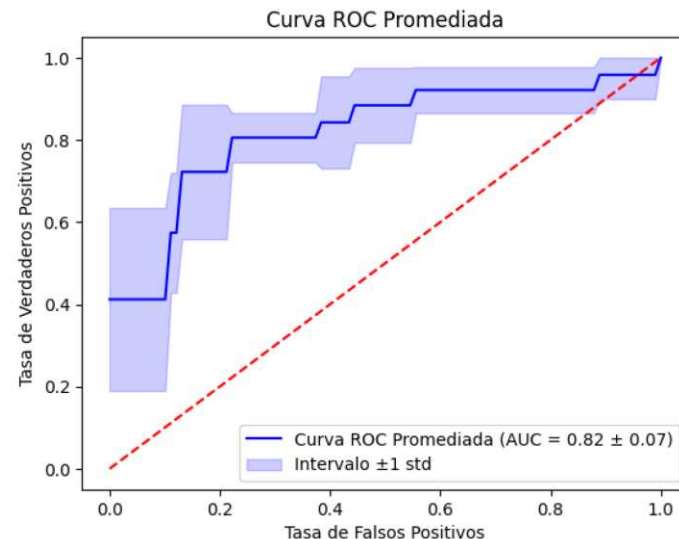
RESULTADOS PARA CADA MODELO : RANDOM FOREST

- El modelo Random Forest obtuvo una precisión media de $0,73 \pm 0,08$, con $AUC = 0,82$, $F1\text{-score} = 0,72$ y sensibilidad = $0,69$, mostrando un buen rendimiento discriminativo con margen de mejora en la detección de positivos.



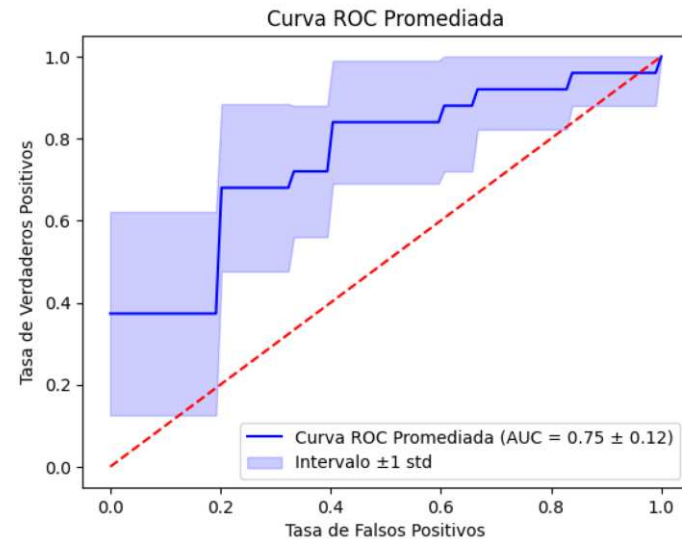
RESULTADOS PARA CADA MODELO : EBM

- El modelo EBM alcanzó una precisión media de $0,77 \pm 0,05$, con $AUC = 0,82$, $F1\text{-score} = 0,77$ y sensibilidad = $0,77$, mostrando un rendimiento equilibrado entre ambas clases.



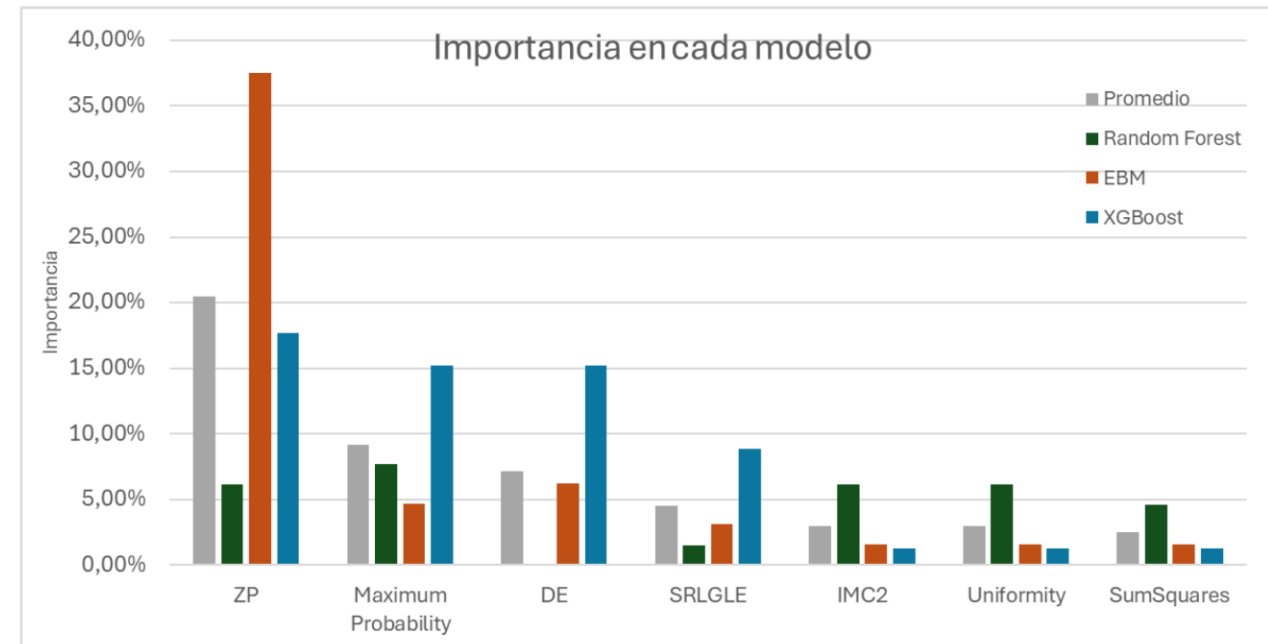
RESULTADOS PARA CADA MODELO : XGBOOST

- El modelo XGBoost obtuvo una precisión media de $0,71 \pm 0,10$, con AUC = 0,75, F1-score = 0,73 y sensibilidad = 0,77, mostrando un rendimiento aceptable con margen de mejora en precisión.



CARACTERÍSTICAS SELECCIONADAS

- Zone Percentage (ZP) fue la característica más relevante en los tres modelos (principal en EBM y XGBoost): valores bajos indican texturas más homogéneas, típicas de abscesos, frente a valores altos y mayor heterogeneidad en glioblastomas.
- Maximum Probability (especialmente relevante en Random Forest): valores bajos reflejan una menor repetición de pares de intensidad, asociada a la heterogeneidad tumoral de los glioblastomas.
- Dependence Entropy (DE) (clave en EBM y XGBoost): valores bajos indican menor complejidad textural, característica de abscesos, mientras que valores altos se asocian a la estructura desorganizada de los glioblastomas.
- Otras características relevantes:
ShortRunLowGrayLevelEmphasis (EBM, XGBoost), *IMC2*, *Uniformity* y *SumSquares* (Random Forest), relacionadas con uniformidad, correlación y variabilidad de intensidad, reforzando la diferenciación entre lesiones homogéneas (abscesos) y heterogéneas (glioblastomas).



Visualización de la importancia de cada característica y su promedio en cada modelo en un grafico de barras

COMPARACIÓN ENTRE MODELOS

- EBM fue el modelo con mejor rendimiento global, mostrando un equilibrio óptimo entre precisión y recall, reflejado en un F1-score elevado.
- Destaca por su alta interpretabilidad, al basar la clasificación en pocas características clave (p. ej., *Zone Percentage*), facilitando la comprensión clínica.
- Presentó mayor estabilidad entre particiones, con AUC medio de $0,82 \pm 0,07$, indicando robustez del modelo.
- Random Forest mostró buen desempeño, pero con sesgo hacia la precisión, y XGBoost presentó mayor variabilidad y menor precisión global.

	Random Forest	EBM	XGBoost
AUC promedio	0,82	0,82	0,75
Precisión	0,73	0,77	0,71
Recall	0,692	0,769	0,769
F1 score	0,72	0,769	0,727

SELECCIÓN ALTERNATIVA DE CARACTERÍSTICAS

- Se modificó el orden de eliminación en la selección por correlación ($r = 0,65$) para conservar la segunda característica de cada par correlado, obteniéndose 82 características y permitiendo evaluar mayor interpretabilidad.
 - Pese al cambio metodológico, la mayoría de las características clave se mantuvieron, confirmando la robustez del proceso; solo algunas (p. ej., *ShortLowRunGrayLevelEmphasis* e *IMC2*) dejaron de ser relevantes.
 - Emergieron nuevas características discriminativas, principalmente *HighGrayLevelZoneEmphasis* y *DependenceVariance* (XGBoost), cuyos valores bajos se asociaron a texturas más homogéneas, típicas de los abscesos.
 - Otras variables complementarias (*Strength* y *Correlation*) reforzaron la diferenciación, apoyando la idea central de que la homogeneidad textural caracteriza a los abscesos, frente a la heterogeneidad de los glioblastomas.
-

DISCUSIÓN

- Los tres modelos (RF, EBM, XGBoost) alcanzaron precisiones del 71–77% y AUC hasta 0,82, resultados prometedores considerando el tamaño muestral limitado y el uso de TC en lugar de RM, reforzando su potencial utilidad en el contexto de urgencias como apoyo al diagnóstico inicial.
 - EBM destacó como el modelo más equilibrado y robusto ($F1 = 0,77$; $AUC = 0,82$), combinando buen rendimiento con alta interpretabilidad, lo que favorece su aceptación clínica frente a modelos más complejos o inestables.
 - Las características ZonePercentage, MaximumProbability y DependenceEntropy fueron consistentemente relevantes y reflejan un patrón común: mayor homogeneidad textural en abscesos frente a la heterogeneidad estructural de los glioblastomas, en concordancia con la literatura previa.
 - Los resultados respaldan la validez de la radiómica en TC, pese a su menor contraste tisular, ofreciendo una herramienta potencialmente útil para estratificación diagnóstica rápida en escenarios donde la RM no está disponible.
 - El uso de SHAP permitió explicar las predicciones y reforzar la confianza clínica, mostrando que modelos interpretables como EBM, basados en pocas variables clave, pueden facilitar la integración en la práctica radiológica diaria.
-

LIMITACIONES

- La principal limitación del estudio es el tamaño reducido de la muestra, especialmente en el grupo de abscesos, lo que puede limitar la generalización del modelo.
 - El uso de características específicas en cohortes pequeñas aumenta el riesgo de sobreajuste, por lo que estudios multicéntricos y con mayor número de casos serían necesarios.
 - No obstante, la consistencia de varias características clave entre los tres modelos y frente a distintos criterios de selección sugiere una generalización interna razonable.
-

CONCLUSIONES

- La radiómica aplicada a TC, una técnica rápida y ampliamente disponible en urgencias permite diferenciar abscesos cerebrales y glioblastomas con un rendimiento comparable al de estudios basados en RM.
 - EBM se posiciona como el modelo más equilibrado e interpretable, combinando buen rendimiento diagnóstico con transparencia, clave para su aceptación clínica.
 - La identificación de patrones texturales consistentes refleja la homogeneidad de los abscesos frente a la heterogeneidad de los glioblastomas, reforzando el valor clínico de la radiómica.
 - Estos hallazgos apoyan el uso de la radiómica en TC como herramienta de apoyo diagnóstico no invasiva, especialmente útil en escenarios donde la RM no está disponible.
-

LINEAS FUTURAS

- Ampliar y diversificar la cohorte, incorporando datos multicéntricos y más casos de abscesos, para mejorar la robustez y generalización del modelo.
 - Integrar información multimodal, combinando distintas secuencias de TC, otras técnicas de imagen y datos clínicos, con el objetivo de aumentar la capacidad discriminativa.
 - Avanzar hacia la automatización e interpretabilidad, mediante segmentación automática, interfaces explicables (SHAP) y la exploración de modelos de deep learning para facilitar la implementación clínica.
-