

Unsupervised learning for vascular heterogeneity assessment of glioblastoma based on magnetic resonance imaging: The Hemodynamic Tissue Signature



UNIVERSITAT POLITÈCNICA DE VALÈNCIA
PROGRAMA DE DOCTORADO EN
TECNOLOGÍAS PARA LA SALUD Y EL BIENESTAR

DOCTORAL THESIS

Presented by
Javier Juan Albarracín

Directed by
Dr. Juan M García-Gómez
Dr. Elies Fuster i Garcia

Valencia, Spain
December 2019

Resumen

El futuro de la imagen médica está ligado a la Inteligencia Artificial (IA). El análisis manual de imágenes médicas es hoy en día una tarea ardua, propensa a errores y a menudo inasequible para los humanos, que ha llamado la atención de la comunidad de Aprendizaje Automático (AA). La Imagen por Resonancia Magnética (IRM), que constituye la técnica de imagen estándar para el diagnóstico de muchas enfermedades letales, nos proporciona una amplia y rica variedad de representaciones de la morfología y el comportamiento de lesiones completamente inaccesibles sin una intervención invasiva arriesgada. Sin embargo, explotar la potente pero a menudo latente información contenida en las adquisiciones de IRM es una tarea muy complicada, que requiere técnicas de análisis computacional inteligente.

Los tumores del sistema nervioso central son una de las enfermedades más críticas estudiadas a través de IRM. Específicamente, el glioblastoma representa un gran desafío, ya que, hasta la fecha, continua siendo un cáncer letal que carece de una terapia satisfactoria. De todo el conjunto de características que hacen del glioblastoma un tumor tan agresivo, un aspecto particular que ha sido ampliamente estudiado es su heterogeneidad vascular. La fuerte proliferación vascular de los glioblastomas, así como su robusta angiogénesis y la extensa heterogeneidad de su microvasculatura han sido consideradas responsables de la alta letalidad de esta neoplasia. Por lo tanto, el estudio de estos factores es crucial para entender mejor la agresividad del tumor y diseñar nuevas terapias efectivas que mejoren el pronóstico del paciente.

Esta tesis se centra en la investigación y desarrollo del método Hemodynamic Tissue Signature (HTS): un método de aprendizaje no supervisado para describir la heterogeneidad vascular de los glioblastomas mediante el análisis de perfusión por IRM. El método HTS se basa en el concepto de *hábitats*. Un hábitat se define como una subregión de la lesión con un perfil particular de IRM, que describe un comportamiento fisiológico concreto. El método HTS delinea cuatro hábitats dentro del glioblastoma: el hábitat High Angiogenic Tumor (HAT), como la región más perfundida del tumor con captación de contraste; el hábitat Low Angiogenic Tumor (LAT), como la región del tumor con captación de contraste con un perfil angiogénico más bajo; el hábitat Infiltrated Peripheral Edema (IPE), como la región edematosa sin captación de contraste adyacente al tumor con índices de perfusión elevados; y el hábitat Vasogenic Peripheral Edema (VPE), como el edema restante de la lesión con el perfil de perfusión más bajo. La investigación y desarrollo del método HTS ha originado una serie de contribuciones enmarcadas en esta tesis.

En primer lugar, para verificar que los métodos de aprendizaje no supervisados son

fiables a la hora de extraer patrones de IRM para describir la heterogeneidad de una lesión, se realizó una comparación entre varios métodos de aprendizaje estructurado y no estructurado no supervisados en la tarea de segmentación de gliomas de grado alto. Adicionalmente, se desarrolló un método genérico de postproceso para mapear automáticamente cada etiqueta de una segmentación no supervisada a un tejido sano o patológico del cerebro.

En segundo lugar, se ha propuesto un algoritmo de aprendizaje Bayesiano no supervisado dentro de la familia de los Spatially Varying Finite Mixture Models (SVFMMs). El algoritmo, llamado Non Local Spatially Varying Finite Mixture Model (NLSVFMM), integra con éxito un Gauss-Markov Random Field (MRF) continuo ponderado por la función probabilística Non Local Means (NLM) como densidad a priori del modelo, para codificar la idea de que los píxeles vecinos tienden a pertenecer al mismo objeto semántico. La probabilidad a priori propuesta refuerza simultáneamente la suavidad local en las segmentaciones, a la vez que preserva los bordes y la estructura entre clases.

En tercer lugar, se presenta el método HTS para describir la heterogeneidad vascular de los glioblastomas mediante los hábitats mencionados. El método HTS se ha aplicado a casos reales, tanto en una cohorte local de pacientes de un solo centro, como en una cohorte retrospectiva internacional de más de 180 pacientes de 7 centros europeos. Se llevó a cabo una evaluación exhaustiva del método para medir el potencial pronóstico de los hábitats, así como las capacidades de estratificación de los mismos para identificar poblaciones con pronósticos diferentes. Se encontraron asociaciones estadísticamente significativas entre la mayoría de los hábitats HTS y la supervivencia global de los pacientes, así como diferencias significativas en las tasas de supervivencia de subpoblaciones divididas según mediciones derivadas del HTS.

Finalmente, los métodos y la tecnología desarrollados en esta tesis se han integrado en una plataforma web *online* de acceso público para su uso académico. La plataforma ONCOhabitats se aloja en <https://www.oncohabitats.upv.es>, y ofrece dos servicios principales: 1) segmentación de tejidos de glioblastoma, y 2) evaluación de la heterogeneidad vascular de los glioblastomas mediante el método HTS. Ambos servicios, además de las imágenes preprocesadas y los mapas de segmentación, generan automáticamente un informe radiológico resumiendo los hallazgos del estudio. ONCOhabitats no sólo ofrece a la comunidad científica y médica acceso a algoritmos del estado del arte para el análisis de estos tumores, sino que también permite acceder a su clúster computacional, capaz de procesar cerca de 300 casos al día.

Los resultados de esta tesis han sido publicados en diez contribuciones científicas, incluyendo revistas y conferencias de primer nivel en las áreas de Informática Médica, Estadística y Probabilidad, Radiología y Medicina Nuclear, Aprendizaje Automático y Minería de Datos e Ingeniería Biomédica. También se emitió una patente industrial registrada en España (ES201431289A), Europa (EP3190542A1) y EEUU (US20170287133A1), que representa los esfuerzos de esta tesis para generar activos tangibles además de los méritos académicos obtenidos de las publicaciones de investigación. Finalmente, los métodos, tecnologías e ideas originales concebidas en esta tesis dieron lugar a la creación de ONCOANALYTICS CDX, una empresa enmarcada en el modelo de negocio de los *companion diagnostics* de compuestos farmacéuticos, pensado como vehículo para facilitar la industrialización de la tecnología ONCOhabitats.