

## 1- Resumen

Tradicionalmente, el análisis de imagen médica ha sido principalmente cualitativo. Este tipo de análisis se basa en la experiencia y el conocimiento del médico especialista encargado de realizar el informe. Esto conlleva un elevado coste temporal (los análisis de las imágenes y la creación de un reporte pueden durar varios días) y económico (tanto si hay que subcontratar el análisis como comprar y mantener el equipo especializado).

Gracias a la aparición de las técnicas de análisis por imagen por computador y al avance de los sistemas informáticos, surge el análisis cuantitativo. Al contrario del análisis cualitativo, el análisis cuantitativo tiene como objetivo medir distintas características de una imagen médica (por ejemplo, el tamaño, la textura, la posición de un componente o su evolución temporal) para realizar el informe en lugar de apoyar únicamente en la experiencia del profesional médico. A su vez, el análisis cuantitativo requiere de mediciones lo más precisas posible para poder realizar sus cálculos con mayor exactitud.

Un biomarcador es un programa informático encargado de cuantificar características de comportamientos fisiológicos, químicos, histológicos, anatómicos, físicos o metabólicos en imágenes médicas. Algunos biomarcadores pueden funcionar sin intervención humana, haciendo que los tiempos de entrega de informes médicos puedan reducirse al máximo. Además, al tratarse de software, en algunos casos es posible paralelizar algunas partes de los biomarcadores para conseguir una disminución del tiempo de ejecución (aprovechándose de que los nodos de cómputo dispongan de varios procesadores).

La ejecución de biomarcadores es un proceso que requiere en ocasiones de infraestructuras con gran capacidad de proceso y, en algunos casos, de hardware específico (GPUs) que no está disponible en la mayoría de las instituciones médicas (privadas o públicas). Gracias a los proveedores de servicios Cloud, es posible acceder a hardware específico y potente que se ajuste a las necesidades de la carga de trabajo. Otra ventaja interesante de las plataformas Cloud es el pago por uso, que da la posibilidad de contratar solo los servicios necesarios en cada momento, permitiendo evitar tanto la compra inicial de los recursos como el ahorro de su mantenimiento. En este ámbito definimos elasticidad como la capacidad de las infraestructuras que permite que se ajusten dinámicamente en función de la carga de trabajos y sus necesidades. Existen dos tipos de elasticidad: horizontal y vertical. La elasticidad horizontal permite el aumento o la disminución de nodos en una infraestructura y, la elasticidad vertical, permite el aumento o disminución de los recursos asignados a los nodos de cómputo.

Se considera que cada trabajo sea la ejecución de un determinado biomarcador. Los contenedores Docker [1] se han convertido en la plataforma de encapsulación y distribución de aplicaciones en el Cloud más utilizada actualmente debido a que permiten aprovechar de una manera más eficiente los recursos del nodo (al no tener que virtualizar completamente el SO) y son más ligeros que las máquinas virtuales. Por lo tanto, en este trabajo, los biomarcadores y sus dependencias se encapsulan dentro de contenedores Docker.

Por lo tanto, esta tesis tiene dos objetivos principales (los cuales serán descritos más detalladamente en la sección 5). El primero objetivo es proporcionar un entorno de procesamiento reconfigurable dinámicamente en función de la carga de trabajos junto

con una solución de Continuous Integration. Debido a que el público objetivo de este entorno de procesamiento son entidades médicas y farmacéuticas que pueden demandar que se configure en sus centros, el entorno podrá configurarse en Clouds públicos y privados (on-premise). La solución de Continuous Integration es interesante para que el versionado de los biomarcadores y su distribución sea lo más sencilla y eficiente posible. Cabe destacar que el entorno de Continuous Integration puede ser compartido por varios entornos de procesamiento (por ejemplo, de desarrollo y en producción). El segundo objetivo es paralelizar y encapsular cada biomarcador junto con sus dependencias en contenedores Docker.