

En los últimos años, el aprendizaje profundo (DL) se ha convertido en una de las principales áreas de la inteligencia artificial (IA), impulsado principalmente por el avance en la capacidad de procesamiento. Los algoritmos basados en DL han logrado resultados asombrosos en la comprensión y manipulación de diversos tipos de datos, incluyendo imágenes, señales de habla y texto.

La revolución digital del sector sanitario ha permitido la generación de nuevas bases de datos, lo que ha facilitado la implementación de modelos de DL bajo el paradigma de aprendizaje supervisado. La incorporación de estos métodos promete mejorar y automatizar la detección y el diagnóstico de enfermedades, permitiendo pronosticar su evolución y facilitar la aplicación de intervenciones clínicas de manera más efectiva.

Una de las principales limitaciones de la aplicación de algoritmos de DL supervisados es la necesidad de grandes bases de datos anotadas por expertos, lo que supone una barrera importante en el ámbito médico. Para superar este problema, se está abriendo un nuevo campo de desarrollo de estrategias de aprendizaje no supervisado o débilmente supervisado que utilizan los datos disponibles no anotados o débilmente anotados. Estos enfoques permiten aprovechar al máximo los datos existentes y superar las limitaciones de la dependencia de anotaciones precisas.

Para poner de manifiesto que el aprendizaje débilmente supervisado puede ofrecer soluciones óptimas, esta tesis se ha enfocado en el desarrollo de diferentes paradigmas que permiten entrenar modelos con bases de datos débilmente anotadas o anotadas por médicos no expertos. En este sentido, se han utilizado dos modalidades de datos ampliamente empleadas en la literatura para estudiar diversos tipos de cáncer y enfermedades inflamatorias: datos ómicos e imágenes histológicas. En el estudio sobre datos ómicos, se han desarrollado métodos basados en *deep clustering* que permiten lidiar con las altas dimensiones inherentes a este tipo de datos, desarrollando un modelo predictivo sin la necesidad de anotaciones. Al comparar el método propuesto con otros métodos de clustering presentes en la literatura, se ha observado una mejora en los resultados obtenidos.

En cuanto a los estudios con imagen histológica, en esta tesis se ha abordado la detección de diferentes enfermedades, incluyendo cáncer de piel (melanoma spitzoide y neoplasias de células fusocelulares) y colitis ulcerosa. En este contexto, se ha empleado el paradigma de *multiple instance learning* (MIL) como línea base en todos los marcos desarrollados para hacer frente al gran tamaño de las imágenes histológicas. Además, se han implementado diversas metodologías de aprendizaje, adaptadas a los problemas específicos que se abordan. Para la detección de melanoma spitzoide, se ha utilizado un enfoque de aprendizaje inductivo que requiere un menor volumen de anotaciones. Para abordar el diagnóstico de colitis ulcerosa, que implica la identificación de neutrófilos como biomarcadores, se ha utilizado un enfoque de aprendizaje restrictivo. Con este método, el coste de anotación se ha reducido significativamente al tiempo que se han conseguido mejoras sustanciales en los resultados obtenidos. Finalmente, considerando el limitado número de expertos en el campo de las neoplasias de células fusiformes, se ha diseñado y validado un novedoso protocolo de anotación para anotaciones no expertas. En este contexto, se han desarrollado modelos de aprendizaje profundo que trabajan con la incertidumbre asociada a dichas anotaciones.

En conclusión, esta tesis ha desarrollado técnicas de vanguardia para abordar el reto de la necesidad de anotaciones precisas que requiere el sector médico. A partir de datos débilmente

anotados o anotados por no expertos, se han propuesto novedosos paradigmas y metodologías basados en *deep learning* para abordar la detección y diagnóstico de enfermedades utilizando datos ómicos e imágenes histológicas. Estas innovaciones pueden mejorar la eficacia y la automatización en la detección temprana y el seguimiento de enfermedades.