Artificial Intelligence for diagnosis and image synthesis in breast cancer

8

Summary

In this thesis, we develop various AI models aimed at benefiting the healthcare of breast cancer patients.

PART I — Feature extraction and prediction

Breast cancer is the most common cancer in women, and the leading cause of cancer mortality in women. With the continuous development of computer science, AI-based methods have received more and more attention in the medical field and have been widely explored in the field of breast cancer. In **Chapter 2**, we provided a summary of the application of existing AI-based methods to breast cancer. First we conducted a quick survey on computer-aided diagnosis and treatment, AI and breast screening, and pointed out that It is particularly important to analyze what is possible beside the commonly discussed early breast cancer detection in screening examinations. Then we summarized the application of AI in the field of breast cancer, including breast cancer risk prediction, breast lesion classification, histological characteristics of breast cancer, molecular subtypes recognition of breast cancer, prediction of response to neoadjuvant chemotherapy, prediction of lymph node status, and Radiomics prediction of breast cancer recurrence. Finally, we discussed the challenges of AI and its application prospects in the field of breast cancer.

Digital health data for various patient care and oncology research continues to accumulate exponentially, however, most medical information, and particularly radiology results, are stored in free text format. Therefore, the data are inaccessible for computer analysis and their potential remains unused. In **Chapter 3**, we proposed a large language healthcare model for processing electronic health records and decision-making in breast disease. The model can automatically extract valuable features from unstructured electronic health records and evaluate human health characteristics and predict pathological outcomes of breast diseases through transfer learning.

Breast cancers can be divided into molecular subtypes based upon the expression levels of ER, PR, HER2 and Ki-67, resulting in Luminal A, Luminal B, HER2-enriched and Triple-negative breast cancer. These molecular subtypes are an important prognostic factor and can guide pre- and postoperative systemic therapy, because these therapies typically target these receptors. Accurately determining the molecular subtypes of breast cancer is important for the prognosis of breast cancer patients and can guide treatment selection. In **Chapter 4**, multimodal deep learning with intra- and inter-modality attention modules was proposed for predicting molecular subtypes of breast cancer. Our model caould potentially be used to predict the molecular subtypes and discriminate Luminal disease from Non-Luminal disease of breast cancer, while being a completely non-invasive, cheap and widely available effective method. This supports the idea that combining multi-modal medical imaging may indeed provide relevant imaging biomarkers for predicting therapy response in breast cancer, thereby potentially guiding treatment selection for breast cancer patients.

PART II — Image synthesis

Multi-parameter MRI information has been shown to improve radiologist performance in lesion classification, as well as improving the performance of artificial intelligence models in various tasks. However, obtaining multi-parameter MRI makes the examination costly in both financial and time perspectives, thus making acquisition of the full spectrum of MRI sequences less durable. In **Chapters 5 and 6**, we proposed a series of models to synthesize absent MRI images.

The provided technical innovations for medical image synthesis can potentially benefit healthcare in the use scenarios when certain MRI sequences are absent and it may also be applied for other modality-image synthesis.

Conclusion

In summary, this thesis investigates the progress of AI in the field of breast cancer and develops large language, predictive and generative models to benefit the healthcare of breast patients. In the future, AI-based radiomics methods may potentially be incorporated into clinical use for accurately diagnosing breast cancer, combining multi-modal imaging information with other data from the patient records to optimally select treatments. The combination of large language model, predictive model and generative model can potentially promote the application and development of AI in the field of breast cancer. Although there are challenges, AI models based on multimodal radiomics will be an important direction in the future of breast cancer research.