Deep learning applications in lung cancer imaging

Summary

This thesis explored imaging-based applications of deep learning methods in diagnosing and treating non-small cell lung cancer (NSCLC) patients. It comprises a mixture of perspective review articles as well as experimental studies on routine patient imaging data.

PART 1: Artificial Intelligence in Cancer Imaging

Part 1 provided an overarching introduction into the topic through a collection of three perspective articles. **Chapter 2** reviewed a selection of AI applications in oncology from the lens of a patient moving through clinical touch points along the cancer care path. It also mapped the challenges faced in clinical translation across clinical validity, usability, and utility. **Chapter 3** established a general understanding of AI methods particularly pertaining to image-based tasks in oncology. It also explored how up-and-coming AI methods will impact multiple radiograph-based practices within oncology. Finally, it discussed the challenges and hurdles facing the clinical implementation of these methods. **Chapter 4** shifted to discussing AI applications in cancer therapeutics, namely radiotherapy. It provided an overview of the potential for AI to transform radiotherapy by walking through each step of the workflow. It also highlighted examples where AI may increase efficiency, accuracy and quality of radiotherapy, thereby enhancing value-based cancer care delivery in today's resource-limited healthcare environment.

PART 2: Prognostic and Therapeutic Deep Learning Applications

Part 2 explored the development and validation of deep learning applications for the prognostication and treatment of NSCLC patients. **Chapter 5** provided evidence that deep learning networks may be used for mortality risk stratification based on standard-of-care CT images of NSCLC patients. This evidence motivates future research into better deciphering the clinical and biological basis of deep learning networks as well as validation in prospective data. **Chapter 6** investigated the utility of convolutional neural networks (CNN) in non-invasively predicting histology in early-stage NSCLC patients, using routinely acquired noninvasive radiologic images. The association of CNN-derived quantitative radiographic image feature maps with histologic phenotype in this cohort was also assessed. **Chapter 7** demonstrated that deep learning can analyse CT imaging scans at multiple time-points to improve clinical cancer outcome prediction, namely progression, distant metastases and local-regional recurrence. This highlights the impact AI-based non-invasive biomarkers can have in the clinic, given their low cost and minimal requirements for human input. **Chapter 8** validated DL models for localizing and segmenting primary NSCLC tumors and involved lymph nodes in CT images for RT targeting. Beyond establishing inter and intraobserver benchmarks, we performed multi-tiered validation on external datasets including clinical trial and diagnostic radiology data. We also carried out additional dosimetric and metabolic validation,

and measured the models' stability and accuracy. Finally, we conducted human subject experiments to measure clinical utility and physician acceptance.

PART 3: AI Methods and Best Practices

Part 3 highlighted best practices in conducting experimental studies, both on the data science and computational methodology fronts. **Chapter 9** investigated two radiomics methodologies for the prediction of response to cancer therapy: handcrafted feature-based and deep learning-based. It also underscored the importance of model interpretability efforts in understanding the biology of the cancer-normal tissue interface, and ultimately targeting localised cancer therapies such as RT and surgery. **Chapter 10** explored building labels that highlight the key ingredients in a dataset such as meta-data as well as unique or anomalous features regarding distributions, missing data, and comparisons to other 'ground truth' datasets. Such a label may afford data specialists a better and more efficient process of data interrogation, which will produce higher quality AI models. **Chapter 11** underscored the importance of transparency in AI research through the sharing of reproducible computational methods as well as data whenever possible. Such practices may increase the impact of published AI algorithms and accelerate their translation into clinical settings.

PART 4: Beyond Cancer Imaging

Part 4 and **Chapter 12** identified unique challenges in the global health system that may be addressed through AI applications, while assisting in reaching the United Nations' Sustainable Development Goals.