Background: 2 sentences

Research question: 1

Methodology:3

Key findings: 3

Implications: 1

Using Multimodal data in medical imaging is a compelling idea in applied AI for medical and diagnostics in which integrating various data types can increase the performances of the learning algorithm. The advantages of Multimodal data in imaging, which each modality can capture different aspects of pathology, might create a more comprehensive view of the anatomy, build up the clinical context as well as cancel out some level of noises and dependences from the image provider. This mini project provides the empirical evaluation of using multimodal data to improve performance in Medical Imaging. The dataset using in this project is from MIMIC CXR (provided by MIT lab) include ChestXray images and the associated clinical text reports. The project consists of 2 parts: feature extraction and binary classification for disease. In feature extraction, Unimodal is using deep learning architecture with AutoEncoder for image modality and Multimodal is combining deep learning for image encoder, attention mechanism for text encoder and neural estimator for fusing method. The classifier will be trained using data embeddings from feature extraction for binary classification for each common pulmonary disease. The evaluation metrics are accuracy for classifier and separability for feature extraction (where embeddings are grouped in positive and negative classes and distance between these classes are measured). In general, for each disease, the accuracy for classifier in MultiModal is from 2 to 5% higher than UniModal and separability is ….. (fill in details here after training completed). For cross disease comparison, CardioMegaly has much higher accuracy than Pneumonia (76.85% compared to 63.29%), but its separabity is lower than Pneumonia’s, around 15% (1.286 compared to 1.508)