Computational Challenges

- Accurate diagnosis would require frontal and lateral radiographs. However, most large scale datasets only contain frontal view.
- The patient history is not provided whereas during a general diagnosis, the radiologist has access to the diagnosis history
- Representing a medical image in a semantic space that captures the essence of the image is a key challenge. It involves determining an appropriate image representation and appropriate matching tools suitable for catego-ization and retrieval. The representation needs to be general enough to accommodate multiple modalities yet robust enough to handle the large variability of the data.
- Obtaining large quality labelled dataset for training - form of labelling images can have error
- Balanced Dataset/ Class imbalance. Rare pathologies are hard for the model to learn from.
- Challenge of accessing high fidelity labeled training data within siloed hospital information technology systems
- Dealing with the high resolution with which images are acquired for diagnostic purposes. These images, usually greater than 2000 pixels in each dimension, are much larger than what is acceptable in other domains CNNs have had success in and this leads to increased computational requirements.
- Many medical problems have turned into multi-label classification with a large number of targets, many of which may be subtle or poorly defined and are likely to be inconsistently labeled.
- A thorax disease usually happens in (small) localized areas which are disease specific. Training CNNs using global image may be affected by the (excessive) irrelevant noisy areas
- Due to the poor alignment of some CXR images, the existence of irregular borders hinders the network performance