Exploring Augmentations for Contrastive Representation Learning with Chest X-Rays

Image augmentations are quintessential in the training of representations using self-supervised contrastive learning. However, since medical images are vastly different from natural images, it is unknown whether typical augmentation strategies employed in contrastive learning for natural images generalize well to medical images. In this study, we systematically investigate the effect of various augmentations on the quality and robustness of the learned representations, based on abnormality detection in Chest X-Rays. We separately trained image-only Siamese Networks on two large Chest X-ray datasets (ChexPert and MIMIC) and evaluated the learned representations on internal and external validation datasets. We investigated the efficacy of these representations in the context of linear probing, zero-shot transfer learning, fine-tuning, localization efficiency and data efficiency, with both convolutional network and vision transformer backbones. In this paper, we identified a set of essential augmentations and provide further guidelines for hyperparameter tuning.