Addressing Chest Radiograph Projection Bias in Deep Classification Models

Author name(s) withheld EMAIL(S) WITHHELD
Address withheld

Editors: Under Review for MIDL 2023

Abstract

Deep learning-based models are widely used for disease classification in chest radiographs. This exam can be performed in one of two projections (posteroanterior or anteroposterior), depending on the direction that the X-ray beam travels through the body. Since projection visibly affects the way anatomical structures appear in the scans, it may introduce bias in classifiers, especially when spurious correlations between a given disease and a projection occur. This paper examines the influence of chest radiograph projection on the performance of deep learning-based chest radiograph classification models and proposes an approach to mitigate projection-induced bias. Results show that a DenseNet-121 model is better at classifying images from the most representative projection in the data set, suggesting that projection is taken into account by the classifier. Moreover, this model can classify chest X-ray projection better than any of the 14 radiological findings considered, without being explicitly trained for that task, putting it at high risk for projection bias. We propose a label-conditional gradient reversal framework to make the model insensitive to projection, by forcing the extracted features to be simultaneously good for disease classification and bad for projection classification, resulting in a framework with reduced projection-induced bias.

Keywords: chest x-ray, deep learning , disparity, anteroposterior, posteroanterior

© 2023 Anonymous.