



DISEASE TYPE PREDICTION

Zhen Yan¹, Yumin Sun¹, Yuchang Zhang¹, and Xiaojing Li¹

¹Technical University of Munich



Introduction

Chest X-ray is a widely used medical imaging procedures. A large hospital typically produces over 40,000 chest X-rays per year. However, lacking qualified radiologists to review these X-rays is a major challenge. Reviewing chest X-rays heavily depends on the experience of radiologists and many images are difficult to read when the lesions are in low contrast or overlap with large pulmonary vessels. Since CNN shows an ability to extract high level features, we explore the possibility of designing computer aided diagnosis for chest X-rays using deep learning approach. We work on 18,000+ images provided from Hackerearth Deep Learning Challenge Cup(HDLCC) to predict single-disease. At last we reach 0.41 weighted f1 score on final off-line test set, with which we locate in the 20th place in the HDLCC (rank 20 / 151 solved / 7079 participants).

Methodology

Generalization Results

Content in your block.

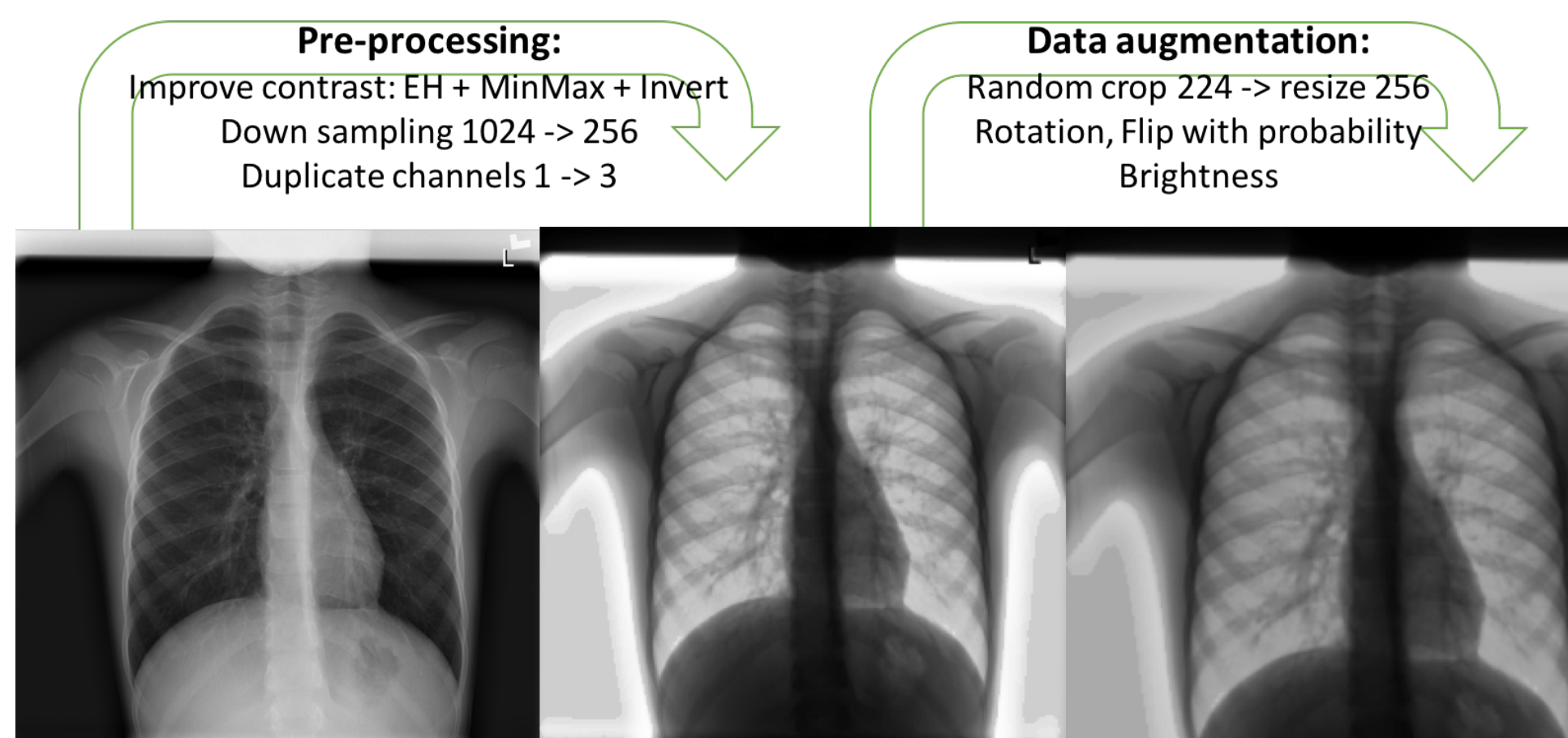
Data pre-processing and augmentation

•Dataset

The dataset includes 18,000+ chest X-ray images of size $1024 * 1024 * 1$. There are in total 14 diseases and the dataset is strongly unbalanced with 5000+ images from most likely diseases and 50+ images from most unlikely diseases.

•Pre-processing and augmentation

Viewing X-ray images, they usually has low contrast in details like blood vessels. We first improve contrast using equalized histogram and then saturate the darkest and lightest of 0.5% gray values. We also did data augmentation to avoid over-fitting. And we find out even rotation and flips also help to avoid over-fitting, though intuitively those augmented images could never happen in nature.



Conclusion

Content in your block.