
Curriculum learning from patch to entire image for screening pulmonary abnormal patterns in chest-PA X-ray: intra- and extra-validations on multi-center datasets

Beomhee Park¹ Joon Beom Seo^{1,3} Sang Min Lee^{1,3} Kyunghee Lee^{2,3} Eun Sol Lee¹
Young Hoon Cho^{1,3} Yongwon Cho¹ Namkug Kim^{1,4*}

¹University of Ulsan College of Medicine, Asan Medical Center, Seoul, South Korea

²Seoul National University Bundang Hospital, Seongnam, South Korea

³Department of Radiology

⁴Department of Convergence Medicine

{beomheep, joonbeom.seo, sangmin.lee.md, kyung8404, eslee0214, yhcho84, dragon1won, namkugkim}@gmail.com

Abstract

In applying the deep-learning method to medical images, there is always the lack of data, and high dimensionality and complexity of medical images make this problem even more serious. Although chest X-ray image is two-dimensional data, accurately detecting abnormal patterns is a very difficult task due to its intrinsic limitations. Therefore, we proposed a computer-aided detection (CAD) for detecting 5 kinds of pulmonary abnormalities in chest-posterioranterior (PA) X-ray images with a curriculum learning strategy to train complex problem after training relatively easy problem to guide the CAD toward better local minima. In addition, extra-validation using multi-center datasets was performed to demonstrate the accuracy and robustness of this strategy.

1 Introduction

Chest X-ray is considered as one of the easiest accessible radiological examinations to screen and diagnose pulmonary problems and secondary prevention. For this reason, there have been several studies to detect the pulmonary disease in chest X-ray using deep-learning methods based on ChestX-ray14 database [1]. However, such studies require huge amounts of data to train weak supervision. As a part of solutions to these problems, Yoshua et al. have suggested curriculum learning to train gradually from simple to more complex concepts [2]. Using this strategy, we proposed a curriculum which fine-tunes complicated entire images after training lesion-specified image patches. We hypothesized that this curriculum can guide the network toward better local minima by preceding patch-based learning. Therefore, the purpose of this study is to evaluate the efficacy of the curriculum which trains with two steps for detecting pulmonary abnormalities in chestPA X-ray image.

2 Materials and Methods

Chest-PA X-rays collected from two hospitals, which consisted of 10137 healthy subjects and 3244 patients including 944, 550, 280, 1364, and 331 patients with nodule (ND), consolidation (CS), interstitial opacity (IO), pleural effusion (PE) and pneumothorax (PT) from Asan Medical Center (AMC), and 1035 healthy subjects and 4404 patients including 1189, 853, 1009, 998 and 944

patients with ND, CS, IO, PE and PT from Seoul National University Bundang Hospital (SNUBH), respectively. These datasets were composed of adults PA images considering the homogeneity and complexity of the data. For accurate experiments, 60% and 20% of AMC dataset were used for training and validation, respectively. 20% of AMC and SNUBH datasets were used for test. Every abnormality lesion was manually drawn by expert thoracic radiologists with more than 10 years' experiences.

The overall procedure for curriculum learning consisted of two steps. First, an initial learning with image patches was performed to specifically train the regional patterns of abnormalities. In addition, entire images were used to fine-tune the network subsequently. Resnet-50 architecture [3] was selected to train weak supervisions and modified for multi-label problem for detecting independent disease patterns. Finally, class activation mapping (CAM) [4] was employed to localize the trained abnormal patterns.

2.1 Training Architecture

In the weakly supervised learning, Resnet has been the most widely used CNN architecture since it has proven its performance in ImageNet Large Scale Visual Recognition Competition (ILSVRC) [5]. As the layer becomes deeper, high performance could be expected. However, since there is a trade-off with accuracy, training time and overfitting, a 50-layer architecture (Resnet-50) which showed appropriate training time and performance was used. Resnet-50 designed in the ILSVRC employs a softmax function as a classifier for multiclass problem. In contrast, the problem of detecting various disease patterns in chest X-ray should be regarded as multi-label problem because various diseases in a image could exist independently. Therefore, we modified the classifier with multiple sigmoid functions. Accordingly, the loss was calculated with the sum of binary cross-entropies.

2.2 Curriculum Learning Strategy

In chest-PA X-ray images, training directly with the entire image could lead to the wrong local minima, since patterns overlapping by organs, tissues, and bones make the problem more difficult. Thus, we employed a simple curriculum learning strategy consisting of two steps to train the complex disease patterns. First, the Resnet-50 network with pre-trained on the ILSVRC dataset was trained using lesion-specified image patches. These patches were extracted around the abnormal lesion, and their size was defined to be half of its original size to contain the disease patterns sufficiently and various patterns surrounding lesion. Subsequently, the network was fine-tuned using entire images, since there is difference in distribution between patches and entire images.

2.3 Experimental Results

To assess the effectiveness of our approach, we compared accuracies of our model with or without the curriculum strategy using six-metrics including area under the curve (AUC), accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). Both models were sufficiently trained and converged on the validation set as shown in Figure 1. Since the curriculum learning-based model preceded training on the patches, the convergence rate was highly fast and showed to converge better local minima. With curriculum learning, the AUC in AMC test set was 93.2,

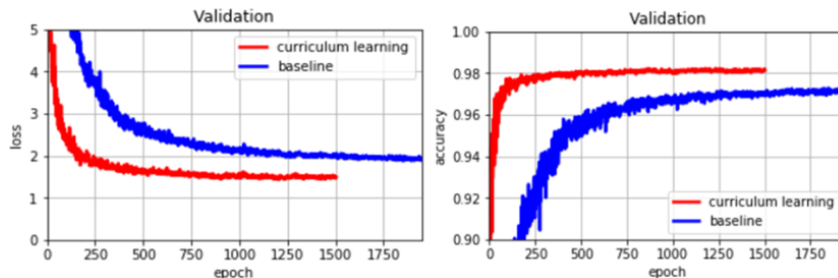


Figure 1: Training curve for loss (left) and accuracy (right) on validation set (red: curriculum learning, blue: baseline)

Table 1: The classification results, compared with or without the curriculum strategy
 ((a): AMC dataset, (b): SNUBH dataset)

	Baseline results (%)						Curriculum learning (%)						Difference in PPV
	AUC	Acc	Sen	Spe	PPV	NPV	AUC	Acc	Sen	Spe	PPV	NPV	
(a) Nodule	91.6	93.7	70.2	95.5	54.5	97.7	93.2	96.1	75.4	97.7	71.3	98.1	+16.8
Consolidation	89.6	96.6	59.6	98.2	60.2	98.2	88.6	96.9	57.9	98.6	65.3	98.1	+5.1
Interstitial Opacity	98.2	99.4	93.0	99.5	76.9	99.9	97.7	99.4	90.7	99.6	78.0	99.8	+1.1
Pleural Effusion	99.4	98.6	95.0	99.0	91.8	99.4	99.5	98.8	94.7	99.3	94.3	99.4	+2.5
Pneumothorax	96.1	98.6	69.1	99.2	64.4	99.4	96.6	99.6	81.8	99.9	95.7	99.6	+31.3
Average	95.0	97.4	77.4	98.3	69.6	98.9	95.1	98.2	80.1	99.0	80.9	99.0	+11.3
(b) Average	89.4	87.7	69.2	92.0	70.9	93.0	91.3	89.8	71.5	94.0	75.5	93.6	+2.0

88.6, 97.7, 99.5, 96.6% for ND, CS, IO, PE, and PT, respectively, as shown in Table 1. Compared to the model without this strategy, performance improvement was achieved in all metrics, and PPV showed the largest improvement of 11.3% and 4.6% for AMC and SNUBH, respectively.

3 Discussion and Conclusion

The proposed curriculum learning strategy successfully showed the outperforming results compared with baseline through multi-center validation. The quantitative accuracies of our method were higher for all classes, especially in case of pneumothorax with the smallest number of data, which might indicate that the curriculum strategy is more effective as the problem becomes more complex and the number of dataset is small.

Acknowledgments

This work was supported by the Industrial Strategic technology development program (10072064) funded by the Ministry of Trade Industry and Energy (MI, Korea) and by Kakao and Kakao Brain corporations.

References

- [1] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M. Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. *CoRR*, abs/1705.02315, 2017.
- [2] Yoshua Bengio, Jerome Louradour, Ronan Collobert, and Jason Weston. Curriculum learning. In *Proceedings of the 26th annual international conference on machine learning*.
- [3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015.
- [4] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. Learning deep features for discriminative localization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- [5] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. A large-scale hierarchical image database. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.