Semi-supervised Learning for Multi-label Classification

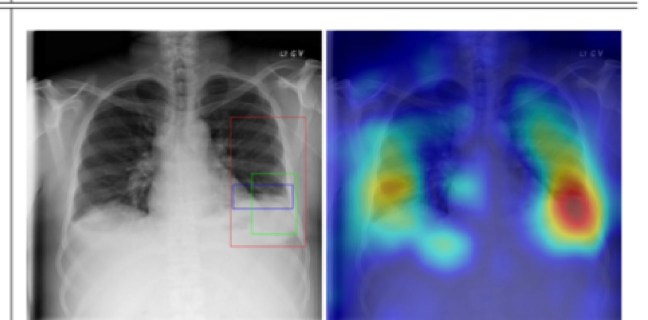
Liyue Shen, Ruiyang Song

Stanford University

Introduction

- Semi-supervised learning: A small portion of the training data are labeled. Prediction by exploiting information from both labeled and unlabeled data.
- Medical Imaging: Data annotation requires expertise. Labeled data are much more expensive than unlabeled data.

Data Overview

Radiology report	Keyword	Localization Result
findings include: 1. left basilar atelectasis/consolidation. 2. prominent hilum (mediastinal adenopathy). 3. left pic catheter (tip in atriocaval junction). 4. stable, normal appearing cardiomediastinal silhouette. impression: small right pleural effusion otherwise stable abnormal study including left basilar infiltrate/atelectasis, prominent hilum, and position of left pic catheter (tip atriocaval junction).	Effusion; Infiltration; Atelectasis	

- 112,120 frontal-view X-ray images of 32,717 unique patients with 14 common thoracic disease categories.
- 60,412 healthy/negative images and 51,708 positive samples.
- 78,484 images (70%) used for training, 11,212 images (10%) for validation, 22,000 images (20%) for test.

Supervised Learning

Method	Benchmark [1]	DenseNet-LSTM [2]	CheXNet [3]	ResNet-18	ResNet-50
Atelectasis	0.7158	0.772	0.8209	0.8190	**0.8276**
CM	0.8065	0.904	**0.9048**	0.8998	0.9013
Effusion	0.7843	0.859	0.8831	0.8881	**0.8903**
Infiltration	0.6089	0.695	0.7204	0.7165	**0.7229**
Mass	0.7057	0.792	0.8618	0.8534	**0.8690**
Nodule	0.6706	0.717	0.7766	0.7738	**0.7884**
Pneumonia	0.6326	0.713	**0.7632**	0.7593	0.7588
PTX	0.8055	0.841	0.8932	0.8934	**0.9033**
CONSOL	0.7078	0.788	0.7939	0.8116	**0.8178**
Edema	0.8345	0.882	0.8932	0.9061	**0.9106**
Emphysema	0.8149	0.829	**0.926**	0.9083	0.9198
Fibrosis	0.7688	0.767	0.8044	0.8149	**0.8197**
PT	0.7082	0.765	**0.8138**	0.8007	0.8048
Hernia	0.7667	0.914	**0.9387**	0.8822	0.8700
NoFinding	-	0.762	-	0.7229	**0.7894**
Average	0.7379	0.798	0.8424	0.8377	**0.8432**

- Per-class AUC scores of different supervised methods.
- Training loss function:

$$L(Y,T) = -\sum_{i=1}^{|\mathcal{C}|}\left[t_i \log\frac{1}{1+e^{-y_i}} + (1-t_i)\log\frac{e^{-y_i}}{1+e^{-y_i}}\right]. \tag{1}$$

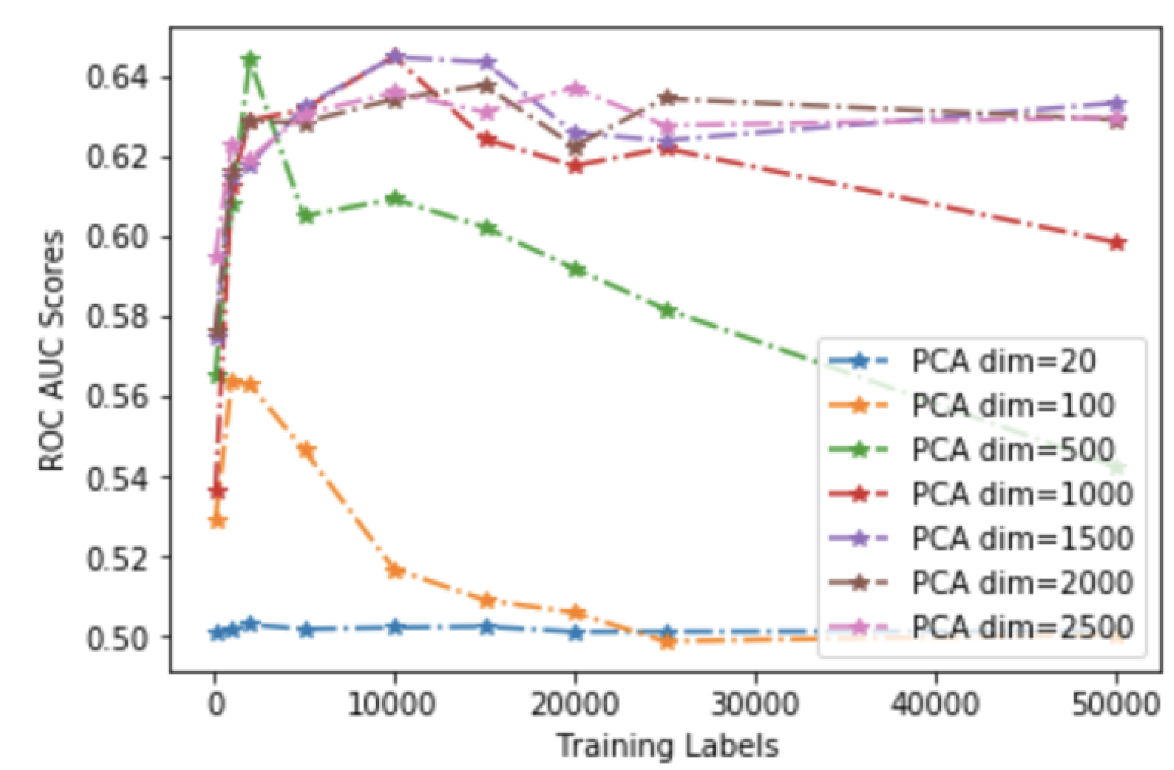
Joint project with CS 331B.

Semi-supervised Learning

Method	PCASVM	ResNet18	ResNet50	ResNet18SP	ResNet50SP	CNNLadder
Atelectasis	0.619	0.7273	0.7084	**0.7312**	0.7120	-
CM	0.689	0.7501	0.7334	0.7453	0.7404	**0.7939**
Effusion	0.674	0.8104	0.8118	**0.8123**	0.8043	0.7395
Infiltration	0.598	0.6593	0.6500	**0.6624**	0.6552	0.5803
Mass	0.558	**0.6758**	0.6729	0.6687	0.6646	0.6108
Nodule	0.539	**0.6511**	0.6386	0.6364	0.6436	0.5312
PM	0.590	0.6792	0.6711	0.6823	**0.6861**	-
PTX	0.616	0.7803	0.7735	0.7755	**0.7821**	0.6630
CONSOL	0.672	**0.7699**	0.7598	0.7616	0.7645	0.6711
Edema	0.761	0.8573	0.8559	0.8594	**0.8659**	0.8170
EP	0.596	0.7570	0.7789	**0.7806**	0.7677	0.6879
Fibrosis	0.618	0.7263	0.7285	0.7307	**0.7377**	-
PT	0.584	0.6971	0.6982	**0.7088**	0.6964	-
Hernia	0.661	0.8179	0.8145	0.8219	**0.8497**	-
Average	0.6268	0.7399	0.7354	**0.7412**	0.7407	-

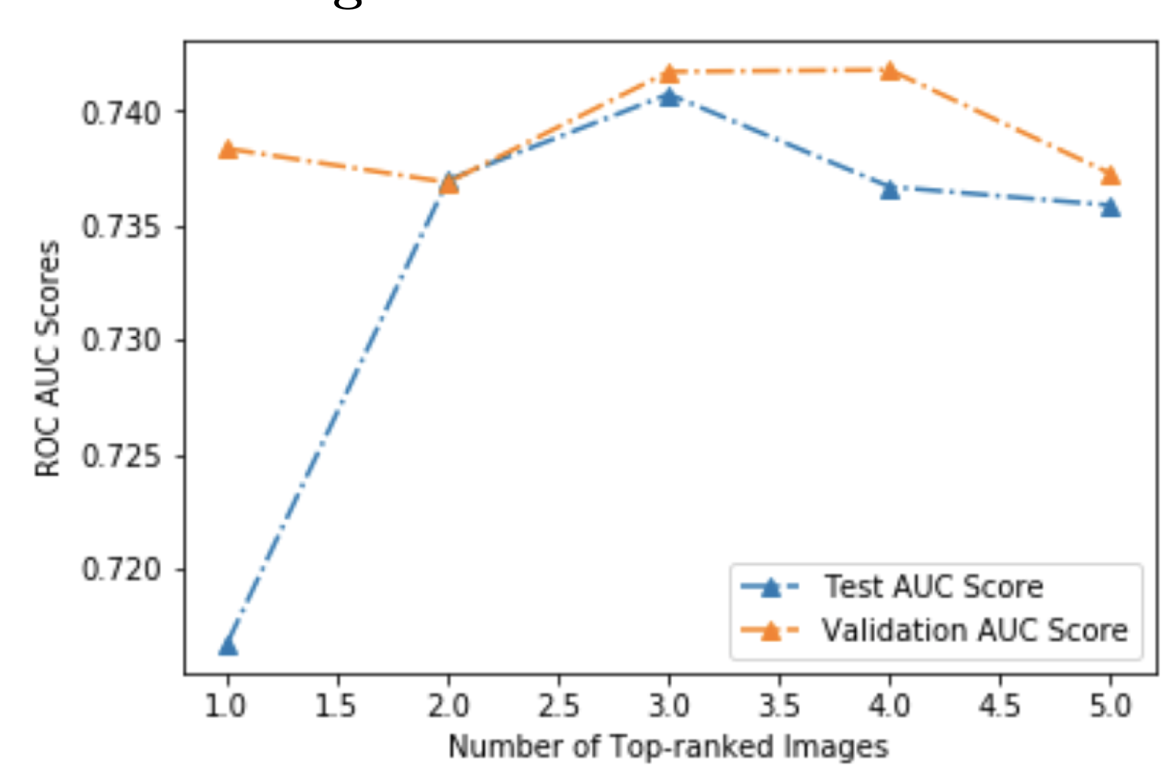
- Per-class AUC scores of different semi-supervised methods with 2000 training images.

PCA-SVM



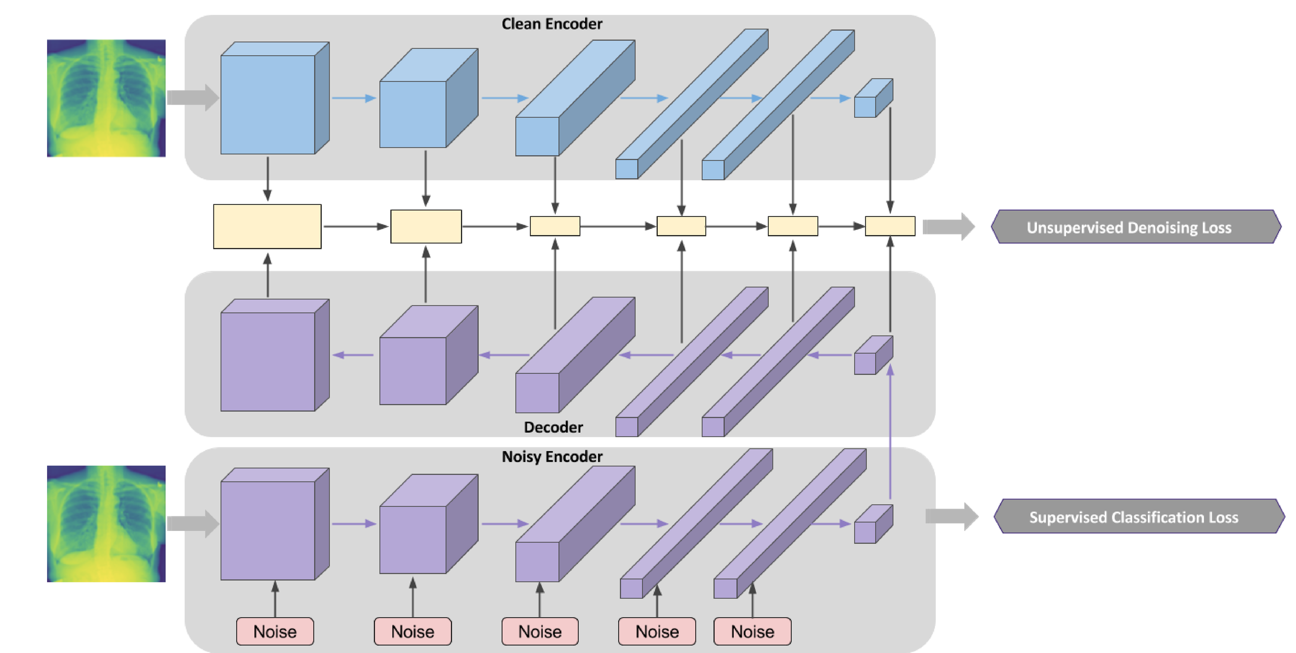
- Performance of PCA-SVM versus number of training labels.

Self-paced Learning



- Validation and test performances of self-paced learning.

Ladder Network



- CNN-based ladder network [4] with clean encoder path, noisy encoder path and denoising decoder.
- Classification loss:

$$L_s(Y,T) = -\frac{1}{N}\sum_{i=1}^{N}\log P(Y=T(n)|X(n)) \tag{2}$$

- Reconstruction loss:

$$L_u(\hat{z},z) = \sum_{l=1}^{L}\frac{\lambda_l}{Mm_l}\sum_{n=1}^{M}\|\hat{z}^{(l)}(n) - z^{(l)}(n)\|^2 \tag{3}$$

Conclusion

- Fine-tuning ResNet with pre-trained weights achieves state-of-the-art performance on the ChestXRay14 dataset.
- Label-efficient semi-supervised approaches enable learning well-performed classifier with less than 2000 training examples with self-pace learning.
- CNN-based ladder network structure exploits information and connection among labeled and unlabeled data.

Reference

[1] Le Lu Zhiyong Lu M. Bagheri Ronald Summer Xiaosong Wang, Yifan Peng. Chestx-ray8 hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *CVPR*, 2017.

[2] Poblenz Eric Dagunts Dmitry Covington Ben Bernard Devon Yao, Li and Kevin Lyman. Learning to diagnose from scratch by exploiting dependen- cies among labels. *arXiv preprint*, 2017.

[3] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, Matthew P. Lungren, and Andrew Y. Ng. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *CoRR*, abs/1711.05225, 2017.

[4] Antti Rasmus, Harri Valpola, Mikko Honkala, Mathias Berglund, and Tapani Raiko. Semi-supervised learning with ladder network. *CoRR*, abs1507.02672, 2015.