Diagnosing Chest X-ray Diseases with Deep Learning

Jiawei Li(jiaweili), Zixi Xu (zixixu), Yue Zhang(yzhang16)

Predicting

The diagnosis of pneumonia and its complications (such as effusion and infiltration) remains a challenging task that relies on the availability of expert radiologists. Based on the previous successes of CheXNet [1], we focus on investigating the correlations among different thoracic diseases:

- **Inputs:** Normalized chest X-ray RGB images and corresponding ground truth labels
- **Model:** DenseNet121 as an encoder to interpret essential features of the input images, followed by LSTM or GRU as a decoder to exploit the correlations among different diseases
- **Outputs:** Predicted label of all 14 thoracic diseases.

We compare our DenseNet121-GRU model against the published results and other kinds of models such as DenseNet121-LSTM.

Data

The dataset is CheXNet14[2] published online by NIH. It contains 112,120 chest X-ray 1024 × 1024 gray-scale images which are labeled by 14 kinds of thoracic diseases. We converted the images into normalized 224 × 224 RGB images and fed them in the model.

Features

We have 15 features in this task. The first feature is the probability that the patient is healthy. The other 14 features is the probability that the patient has the disease for each 14 diseases.
We generate a category vector of size 15 through the tables in the data set and set it as the ground truth label.

Models

- ***Two loss functions:***
  1. $L_1(X,y) = \sum_{i=1}^{14} (-y_i \log p(Y_i=1|X) - (1-y_i)\log p(Y_i=0|X))$
  2. $L_2(X,y) = y_0 \sum_{i=1}^{14} (-y_i \log p(Y_i=1|X) - (1-y_i)\log p(Y_i=0|X)) + (-y_0 \log p(Y_0=1|X) - (1-y_0)\log p(Y_0=0|X))$
- **Weight decay for L2 regularization:** 5e-5
- **Metric:** AUC ROC scores

DenseNet121 model

- **Fine tuning:** Add a fully connected layer and a Sigmoid function at the end of the pretrained model
- **Learning Rate:** 5e-5

CNN-RNN model

- **CNN encoder:** DenseNet121
- **RNN decoder:** *Bi-directional* LSTM or GRU
- **Number of layers:** 2
- **Learning Rate:** 1e-4

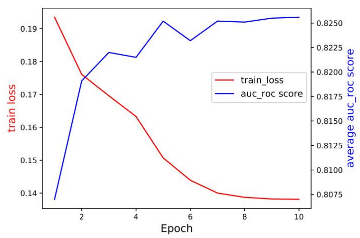

Figure 1: DenseNet 121 ChexNet Results

Results

Results are compared with models in literatures. The highlighted numbers are better predictions by current models.

Table 1: DenseNet121

Pathology	Ng et al.	Loss 1	Loss 2
Atelectasis	0.809	**0.811**	0.781
Cardiomegaly	0.925	0.882	0.872
Effusion	0.863	**0.884**	**0.868**
Infiltration	0.734	0.714	0.647
Mass	0.867	0.846	0.836
Nodule	0.780	0.770	0.748
Pneumonia	0.768	0.745	0.727
Pneumothorax	0.888	**0.889**	0.873
Consolidation	0.790	**0.802**	**0.792**
Edema	0.887	**0.899**	**0.895**
Emphysema	0.937	0.915	0.912
Fibrosis	0.804	**0.812**	0.776
Pleural Thickening	0.806	**0.807**	0.792
Hernia	0.916	0.831	0.874
Average	0.841	0.829	0.814

Table 2: CNN-RNN

Pathology	Yao et al.	LSTM	GRU
Atelectasis	0.772	0.768	0.771
Cardiomegaly	0.904	0.797	0.854
Effusion	0.859	**0.863**	**0.877**
Infiltration	0.695	0.557	0.617
Mass	0.792	**0.816**	**0.816**
Nodule	0.717	0.698	0.699
Pneumonia	0.713	0.640	0.667
Pneumothorax	0.841	**0.849**	**0.844**
Consolidation	0.788	0.777	0.785
Edema	0.882	0.861	0.878
Emphysema	0.829	**0.878**	**0.882**
Fibrosis	0.767	0.625	0.731
Pleural Thickening	0.765	0.693	0.741
Hernia	0.914	0.754	0.78
Average	0.798	0.760	0.782

Discussion

1. Loss function 1 is defined the same as in literature[1], this is used to validate our code. By tuning the hyperparameters, we get similar results as in [1].
2. Loss function 2 attempts to examine the existence of disease before the disease classification, which better imitates the diagnostic process. We expect it to perform better especially when incorporating more information like gender, age, body temperature, diseases history etc.
3. From Table 2, using Bi-directional LSTM helps to increase the ROC AUC scores of Mass and Emphysema. This is possibly because these two have some correlation so Bi-LSTM performs better.
4. According to Table 2, GRU performs better than LSTM in both loss functions. GRU contains less parameters thus it is less prone to over-fitting.

Future

We will focus on extending the current RNN decoder to a more sophisticated attention model (that is add an one-directional RNN on top of the current bi-directional layers), so as to better capture the correlations among the diseases.

References

[1] A. Y. N. et al.
Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning.
CoRR, abs/1711.05225, 2017.

[2] X. W. et al.
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.