Diseases Detection On Chest X-Rays With Deep Learning

Guohui Huang [1] **Haowen Zhu** [1] **Meijie Wang** [1] **Mengdi Li** [1]

Abstract

In this paper, we compared several classic neural networks that can detect 14 different thoracic diseases from chest X-rays as well as tried to improve their performance by adding squeeze and excitation blocks which won the first place in 2017 ImageNet competition. We re-implemented the ChexNet on ChestX-ray14 dataset to validate the accuracy published in the pubished paper[1] and employed several classic neural networks with squeeze and excitation blocks which might improve the accuracy. The experiment results indicated that the pre-trained DenseNet121 outperforms other models.

1. Introduction

In the United States, more than 1 million adults are hospitalized for pneumonia, and about 50,000 people die each year from the disease.[5] Chest X-rays are currently the best method for diagnosing pneumonia.[6] However, detecting pneumonia in chest X-rays is a challenging task that relies on the availability of expert radiologists. In 2017, Stanford ML Group published a paper[1] that they could automatically detect chest X-ray pneumonia from levels beyond the practice radiologist with deep learning. Inspired by this, we Implement and train a model(based on DenseNet) and tried to improve the performance by adding squeeze and excitation blocks.

1.1. Research contributions

In this paper, we compared the performances of different classic neural networks including DenseNet 121, DenseNet 169 and ResNet50, and tried to add squeeze and excitation blocks that won the first place in 2017 ImageNet competition into convolutions layers of DenseNet121 and ResNet50 to see weather it can improve the performance. Among these models, the SE-DenseNet121 was implemented by ourselves.

2. Related Work

According to Rajpurkar et al.,[1] the CheXNet is a 121-layer convolutional neural network. The network architecture was not given by this paper, but there are many implementations on github. According to the code provided by arnoweng,[4] we can get the architecture of the CheXNet (shown as Table 1).

Layers	DenseNet -121
Convolution	7*7 conv, stride 2
Pooling	3*3 max pool, stride 2
Dense Block (1)	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \end{pmatrix} * 6$
Transition Layer (1)	1*1 conv
	2*2 average pool
Dense Block (2)	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \end{pmatrix} * 12$
Transition Layer (2)	1*1 conv
	2*2 average pool
Dense Block (3)	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \end{pmatrix} * 24$
Transition Layer (3)	1*1 conv
	2*2 average pool
Dense Block (4)	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \end{pmatrix} * 16$
Classification Layer	7*7 global average pool
	14D fully-connected, elementwise sigmoid

Table 1. Sizes of convolutional kernels for DenseNets121 architectures. Note that each conv layer shown in the table corresponds the sequence BN-ReLU-Conv.

Rajpurkar et al.[1] also mentioned that they pre-trained the DenseNet121 on ImageNet dataset. In order to improve the performance of DenseNet121, we considered of adding the squeeze and excitation module[3] that proposed and won first place in ImageNet competition in 2017 into existing DenseNet 121. Figure 1 shows the schema of the original Inception module (left) and the SE-Inception module (right). The purpose of SE block is to weight the importance of channel of filters so that the network will enhance more useful channels and ignore the less useful ones.

[1]Worcester Polytechnic Institute. Correspondence to: Guohui Huang <ghuang@wpi.edu>, Haowen Zhu <hzhu3@wpi.edu>, Meijie Wang <mwang3@wpi.edu>, Mengdi Li <mli7@wpi.edu>.

CS541 Special Topics Course Final Project, Worcester Polytechnic Institute.

3. Proposed Method

We employed 5 neural networks to solve this problem and CAM(Class activation mapping) to generate heatmap on top of the X-Ray scan to interpret the neural networks' output. Referring to the architecture of ResNet50 and SE-ResNet50 shown in research of Huang et al.[2] We added the SE(squeeze and excitation) modules into the Dense blocks in DenseNet121, the architecture is described as table 2.

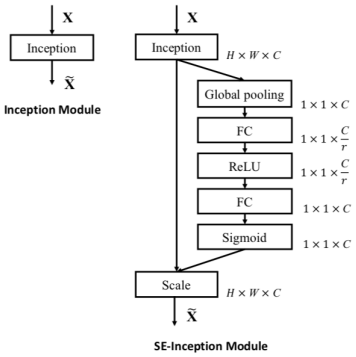

Figure 1. **The schema of the original Inception module (left) and the SE-Inception module (right).**

Layers	DenseNet -121	SE-DenseNet-121
Convolution	7*7 conv, stride 2	
Pooling	3*3 max pool, stride 2	
Dense Block (1)	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \end{pmatrix} * 6$	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \\ fc \end{pmatrix} * 6$
Transition Layer (1)	1*1 conv 2*2 average pool	
Dense Block (2)	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \end{pmatrix} * 12$	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \\ fc \end{pmatrix} * 12$
Transition Layer (2)	1*1 conv 2*2 average pool	
Dense Block (3)	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \end{pmatrix} * 24$	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \\ fc \end{pmatrix} * 24$
Transition Layer (3)	1*1 conv 2*2 average pool	
Dense Block (4)	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \end{pmatrix} * 16$	$\begin{pmatrix} 1*1\ conv \\ 3*3\ conv \\ fc \end{pmatrix} * 16$
Classification Layer	7*7 global average pool 14D fully-connected, elementwise sigmoid	

Table 2. Sizes of convolutional kernels for SE-DenseNets121 architectures. Note that each conv layer shown in the table corresponds the sequence BN-ReLU-Conv.

4. Experiment

4.1. Data Description

We use the ChestX-ray14 dataset released by NIH clinic center which contains 112,120 frontal-view X-ray images of 30,805 unique patients (78468, 22433 and 11219 in training, test and validation set respectively). Each image is with up to 14 different thoracic pathology labels. Figure 2 shows the thoracic pathology labels distribution in the dataset.

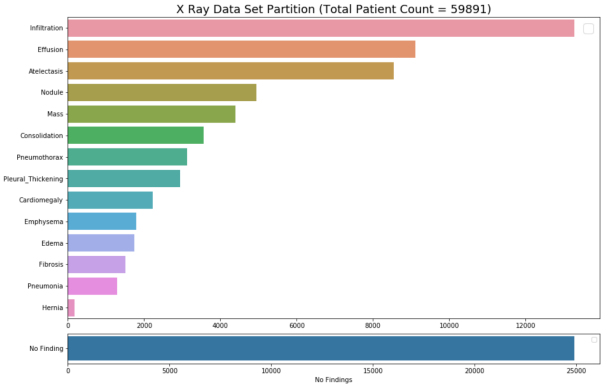

Figure 2. The disease distribution in ChestX-ray14 dataset

Considering of the time of training, we downsize the dataset on condition of dropping duplication with same patient ID and labels. 59,891 X-ray images of 30,805 unique patients were remained(41754, 12151 and 5986 images remained in training, test and validation set respectively). Figure 3 shows the thoracic pathology labels distribution in the downsized dataset.

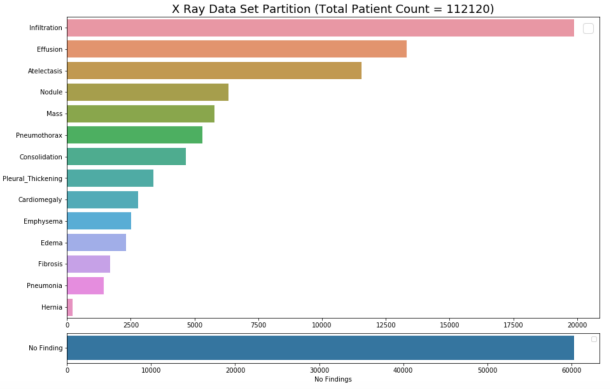

Figure 3. The disease distribution in downsized ChestX-ray14 dataset

4.2. Hyper parameter Settings

According to stanford ML group[1] , the weights of the network are initialized with weights from a model pretrained

on ImageNet. The network is trained end-to-end using Adam with standard parameters (1 = 0.9 and 2 = 0.999). The mini- batches is set as 16. The initial learning rate is 0.0001 (in the Stanford's paper, the initial learning rate is set as 0.001) that is decayed by a factor of 10 each time the validation loss plateaus after an epoch, and pick the model with the lowest validation loss. Fixed the above hyper parameters, we trained 4 pretrained models (DenseNet121, DenseNet169, ResNet50 and SEResNet50) and 2 non-pretrained models (DenseNet121 and SEDenseNet121) on the downsized dataset. The pretrained models were trained on ImageNet dataset.

5. Results and Discussion

Table 3 indicates that our implementation outperforms the published results, which might caused by the difference of initial learning rate. Table 4 gives the results of DenseNet121 trained on our downsized dataset. The performance is slightly worse than the results got by full sized dataset. Therefore, we think it is reasonable to reduce time of training in this way.

Original ChestX-ray14 dataset (pretrained on ImageNet)

Pathology	CheXNet (paper)	DenseNet121 (our implementation)
AUROC Mean	0.841	0.8508
Atelectasis	0.8094	0.8321
Cardiomegaly	0.9248	0.9107
Effusion	0.8638	0.886
Infiltration	0.7345	0.7145
Mass	0.8676	0.8653
Nodule	0.7802	0.8037
Pneumonia	0.7680	0.7655
Pneumothorax	0.8887	0.8857
Consolidation	0.7901	0.8157
Edema	0.8878	0.9017
Emphysema	0.9371	0.9422
Fibrosis	0.8047	0.8523
Pleural Thickening	0.8062	0.7948
Hernia	0.9164	0.9416

Table 3. Our implementation outperforms the CheXNet published results on 14 pathologies' mean AUROC in the original ChestX-ray14 dataset

Table 4 shows that among the pretrained models, DenseNet121 still outperforms other models. Compared with pretrained DenseNet169, the increase of depth did not bring an increase on AUROC. Besides, compared with ResNet50, the performance of SEResNet50 is more close to the performance of DenseNet121. It means the SE blocks did improve the performance of ResNet50. In this case, We considered that weather SE blocks would increase the performance of DenseNet121. Therefore, we added SE block into each Dense block in DenseNet121 like what we described in Table 2. However, there is no existing SEDenseNet121 pretrained on ImageNet. We compared it performance with non pretrained DenseNet121. Table 5 shows the results.

Downsized ChestX-ray14 dataset (pretrained on ImageNet)

Pathology	DenseNet121	DenseNet169	ResNet50	SEResNet50
AUROC Mean	0.8337	0.8308	0.8138	0.8317
Atelectasis	0.8248	0.8156	0.8063	0.8144
Cardiomegaly	0.8982	0.8993	0.9072	0.8981
Effusion	0.8757	0.8792	0.8779	0.8743
Infiltration	0.6738	0.6737	0.6638	0.6756
Mass	0.8469	0.8313	0.8261	0.8342
Nodule	0.7643	0.7491	0.728	0.7583
Pneumonia	0.7781	0.7735	0.7551	0.7608
Pneumothorax	0.8541	0.8602	0.8382	0.8638
Consolidation	0.8107	0.8096	0.8069	0.8092
Edema	0.9033	0.9103	0.9073	0.9081
Emphysema	0.9209	0.91	0.8724	0.9194
Fibrosis	0.8284	0.818	0.7981	0.8291
Pleural Thickening	0.7735	0.7658	0.7431	0.766
Hernia	0.9194	0.9355	0.863	0.9322

Table 4. pre-trained DenseNet121 outperforms the other pre-trained models' results on 14 pathologies' mean AUROC in the downsized ChestX-ray14 dataset

Unfortunately, without pretrained model, DenseNet121 performs much better than our SEDenseNet121. We suppose the reason is that the SE block might only be robust to the ImageNet dataset and images in ChestX-ray14 dataset are in black and white. Therefore, without the pretrianed model, SEDenseNet121 cannot perform better than DenseNet121.

Downsized ChestX-ray14 dataset (didn't pretrain on ImageNet)

Pathology	DenseNet121	SEDenseNet121
AUROC Mean	0.8147	0.7721
Atelectasis	0.8039	0.7756
Cardiomegaly	0.9021	0.8347
Effusion	0.8742	0.8468
Infiltration	0.6701	0.6597
Mass	0.8301	0.7218
Nodule	0.7236	0.6552
Pneumonia	0.7514	0.724
Pneumothorax	0.8461	0.8078
Consolidation	0.81	0.7866
Edema	0.9038	0.8931
Emphysema	0.8738	0.8059
Fibrosis	0.8046	0.7527
Pleural Thickening	0.7497	0.7033
Hernia	0.8626	0.8425

Table 5. DenseNet121 without pre-train outperforms SE-DenseNet121 non pretrained models' results on 14 pathologies' mean AUROC in the downsized ChestX-ray14 dataset

We also generated two heatmap pictures (Figure 4 and 5) with CAM as example to interpreted the outcome of deep learning models.

Figure 4. Patient with congestive heart failure and cardiomegaly (enlarged heart)

Figure 5. Patient with a large right pleural effusion (fluid in the pleural space)

6. Conclusions and Future Work

In this case, we compared the performance of several extisting models (DenseNet121, DenseNet169, ResNet50 and SEResNet50) on solving ChestX-ray detection problem and proposed a new model (SEDenseNet121) to beat the results of DenseNet121. However, the performance of DenseNet121 is still the best.

In the future, if we have enough computational resources, we will pretrain the SEDenseNet121 on ImageNet first and then train it on ChestX-ray14 dataset to validate weather this model can perform better than DenseNet121 in this problem.

References

[1] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... Lungren, M. P. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225.

[2] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[3] Hu, J., Shen, L., Sun, G. (2018). Squeeze-and-excitation networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7132-7141).

[4] A pytorch reimplementation of CheXNet. https://github.com/arnoweng/CheXNet.

[5] Pneumonia Can Be PreventedVaccines Can Help. https://www.cdc.gov/pneumonia/prevention.html

[6] WHO. Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children. 2001.