Interpretation and localization of Thorax diseases using DCNN in Chest X-Ray

Prateek Singhal[1], Pawan Singh[2], Ankit Vidyarthi[3]

[1,2]Department of Computer Science and Engineering, Amity School of Engineering and Technology, AUUP, Lucknow, India
[3]Jaypee Institute of Information Technology University, Noida, India
[1]prateeksinghal2031@gmail.com, [2]pawansingh3@yahoo.com, [3]archietec@gmail.com

Abstract

In recent years, the use of diagnosing images has been increased dramatically. An entry level task of diagnosing and reading Chest X-ray for radiologist but they ought to require a good knowledge and careful observation of anatomical principles, pathology and physiology for this complex reasonings. In many modern hospital's the tremendous number of x-ray images are stored in PACS (Picture Archiving and Communication System). The conditions of plethora been diagnosed by the sustainable number of chest X-Ray. Our aim to predict the thorax disease categories through deep learning using chest x-rays and their first-pass specialist accuracy. In a paper the main application that present a pathology localization framework and multi-label unified weakly supervised image classification that can perceive the occurrence of afterward generation of bounding box around the consistent and multiple pathologies. Due to considering of large image capacity we adapt Deep Convolutional Neural Network (DCNN) architecture for weakly-supervised object localization, different pooling strategies, various multi-label CNN losses and measured against a baseline of softmax regression.

Keywords

Image Processing; Chest Radiography; DCNN; Machine Learning

1. Introduction

Over a past decades the number of x-ray performed has been increased steadily. Of these, a sustainable number are chest x-rays used to diagnose the condition of plethora that's includes such as Pneumonia, Edema, Effusion, Emphysema, Fibrosis, Hernia, Infiltration, Mass, Nodule, Pleural Thickening, Consolidation, Pneumothorax and No finding is also a category for non-diseased patients. Predicting the thorax disease categories through deep convolutional neural network learning in chest x-ray and their metadata. In the recent years a dataset was released by NIH, through image classification we try to improve the f1 score of disease classification [1]. In this dataset there are 30,000 unique patients over a 25,603 gray scale identically sized images corresponding to common thorax diseases types.

Generally, doctors are quite good in diagnosing, mistake can happen, and in-deep details can be left out. In a case study we found that 66% of the time there is a refined opinion from an original diagnose, only 12% of time the diagnose found confirmed, and 21% of time the diagnose was changed completely from original diagnose [2] whenever we take second opinions. In a model, we order to achieve more accurate diagnose there is a reasonable sanity check to predict the diseases based on x-rays.
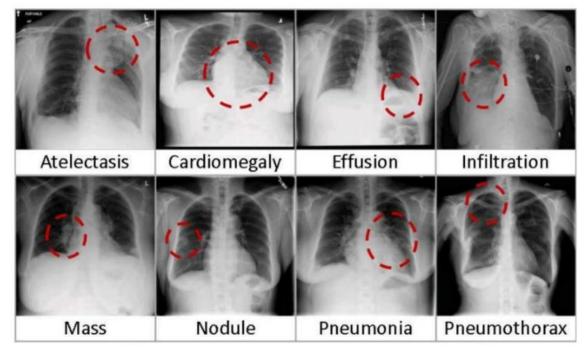

Figure 1. Eight Common Thoracic disease pragmatic in chest X-ray that authenticate inspiring task of fully computerized diagnoses.

Our inputs are 25,603 x-ray images of 1024x1024, as well as metadata on age, gender, and number of visits to the hospital. In modified residual network as well as softmax regression we feed the features to predict the output probability of various thorax diseases, with multi-label classification that range from normal x-ray scan to diagnosis one to many diseases.

2. Related Work

In epoch of deep learning hip computer vision or uses deep neural network [3], various annotated image dataset is built by research efforts with diverse features plays essential role on betterment of forthcoming problems, technological progresses and challenges definitions. The joint learning and relationship of images (chest X-rays) and manuscript (X-ray reports) we basically focus on it and previous generation caption utilizes Flickr8K, MS COCO and Flickr30K to represent images that's hold dataset of 8000,31000 and 123000 respectively. The image is interpreted by the five rulings through Amazon Mechanical Turk.

To address this difficulty, we verify and formulate the disease localization and weakly- supervised multi-label image classification. The VQA technique where all image captioning is depended on ImageNet a pre-trained DCNN model that always perform well in huge number of object and for a great baseline it serves a fine-tuning model. The medical diagnosis domain cannot be applied on this situation. While forming the weakly labelled medical image database we have a knowledge of deep image localization and recognition.

Figure. 2. An example of chest x-ray as input

In this paper, our aim to predict the disease in a multi-label, multi-class, image classification. In earlier only single class diseases were focused in x-ray classification [4] and with specific diseases [5]. The all parts of NIH dataset are being utilized to get the maximum potential of our prediction in multi-label image classification.

3. Problem Statement

The main objective in this paper challenges are: Firstly, the accuracy rate with multi-label classification prediction. Secondly, creation of Deep Convolutional Neural Network (DCNN) architecture (using ImageNet) model to compare the result of softmax regression/ random classifier. Thirdly, Correlation analysis between the patient's traits and thorax diseases.

4. Data

The NIH dataset was released which includes 25,603 gray scale x-ray images of 1024x1024 pixels from 30,000 unique patients, consist the information of patients such as patient age, patient gender and number of follow up visits.

These are the 14 common thorax diseases such as Atelectasis, Edema, Pneumonia, Cardiomegaly, Effusion, Nodule, Fibrosis, Hernia, Pneumothorax, Mass, Emphysema, Pleural Thickening, Consolidation, and Infiltration. No finding is also a category for non-diseased patients.

As an input feature we utilize the x-rays as well as patient's traits with image sized 1024x1024 in gray scale channel, softmax regression can be used directly and in neural net we can predict the downsamples of disease categories. The complex challenge is to categorize a x-ray with multi-label image classification of 14 classes thorax diseases by using the dataset to full potential.

5. Method

The two different models have been used to analyze the x-rays. Firstly, the softmax regression as a baseline which give the probability of the 15 classes dataset of a given image. Secondly, Deep Convolutional Neural Network (DCNN) architecture (using ImageNet [7]) with account of metadata such as patient age, gender, etc.

5.1. Probability of Classification and Accuracy

Before going into a core model, let's discuss our approach with multi-label, multi-class dataset (any data point (x, y), we
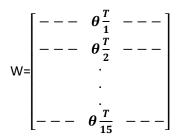
have y $\in$ {0,1}15 and 0 $\leqslant$ $\sum$i yi $\leqslant$ 15, any number of diseases patient could have). After the probability of 15 classes were obtained from both DCNN architecture and softmax regression, categories would be decided by tagged as positive and tagged as negative.

Given a datapoint (x, y) $\in$ R10242 × {0, 1}15, our model predicts yˆ in the following way: let p $\in$ R 15each category is assigned by probabilities for x. yˆ value of one at each index i was given where p[i] top k values of p, where k =∑i yi, and other 15 − k indices is zero. Prediction of our accuracy yˆ for a label y was described to be yˆ · y/k (how many positive categories were identified correctly, divided by the total number of positive categories). It is hard to use this prediction method because for the new patients it is not known a prior that how many diseases they have. Even so, this would be a rigorous accuracy permit to train model well.

The threshold prediction strategy was tested where the probability of all classes is larger than some t marked as 1, other classes marked as 0. The softmax probability gets the opportunity to spread over a multiple class is equally and tagged all of them appropriately, in practice using t = 0.15 led to a similar accuracy to a priori tagging.

5.2. Softmax Regression

A simple softmax regression is been implemented as a baseline. p $\in$ R 15 the matrix calculation was obtained p=Wx where W $\in$ R15×10242. Denoted (W) as:

$$W=\begin{bmatrix} --- & \theta_1^T & --- \\ --- & \theta_2^T & --- \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ --- & \theta_{15}^T & --- \end{bmatrix}$$

Through the optimization of cost function [6] W was calculated:

$$J(W)=\frac{-1}{m}\left[\sum_{i=1}^{m}\sum_{j=1}^{15}\mathbf{1}\left\{y(i)[j]=1\right\}log\frac{\exp(\theta_j^T x(i))}{\sum_{l=1}^{k}\exp(\theta_l^T x(i))}\right]$$

Following gradient θ_j : $\nabla_{\theta_j}J(W) = \frac{-1}{m}\sum_{i=1}^{m}x^{(i)}(1\{y^{(i)}[j]=1\}-p(y^{(i)}=j|x^{(i)};\theta_j))$

5.3. DCNN Unified Framework

The pathology localization and weakly supervised multi-label image classification framework that can perceive various sub-sequential and pathologies bounding boxes around other pathologies. DCNN architecture consider the large image capacity, object localization, various pooling strategies and multi-label CNN.
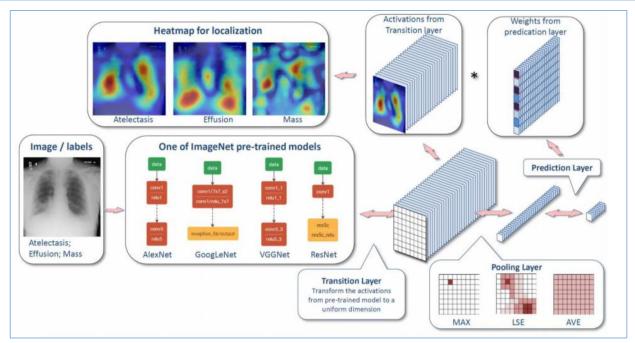
Figure. 3. DCNN unified framework and disease localization

Our priority is to check whether there are one or more than one pathologies is present in each x-ray image then after locating it in network using extracted weights and activation. This challenge can be train by multi-label DCNN classification model. Weakly supervised object localization methods [8, 9, 10, 11] is similar to several DCNN. Network surgery done on various pre-trained models using ImageNet [12, 13] such as GoogleNet [14], RestNet [15], AlexNet [16] and VGGNet-16 [17] through classification layer and fully connected layers. We start inserting transition layer, global layer, pooling layer, prediction layer and loss layer at the termination. To find the plausible spatial location of disease is enabled by the combination of deep activation [9] from transition layer and prediction inner-product layer with weights.

- **Setup of Multi-label:** The numerous choices of image-label has a various option to represent and select the loss function of multi-label classification. Here, 8-dimensional label vector are define y = [y1,...,yc,...,yC],yc $\in$ {0,1},C = 7. yc represent the pathology image presence with respect to it while all-zero vector [0,0,0,0,0,0,0] indicates "Normal" as status (Not any finding). Due to multi-label classification problem there is a loss setting in regression.

- **Transition Layer:** DCNN architectures has been adopt because of huge pre-trained models variety, in this layer there is a uniform dimension of outputs because it transform the activation function from previous layer, S $\times$ S $\times$ D, S $\in$ {7, 14, 28}. D represent spatial location of dimension features (i, j), i, j $\in$ {1, ..., S}, that can be classified by dissimilar model settings, such as D = 1024 for GoogleNet and D = 2048 for ResNet. It supports to allow the weights in a standard form via pre-trained DCNN models, further generation of heatmap in pathology localization step which is perilous by using this activation layer's.

- **Multi-label Classification Loss Layer:** Instead of softmax loss function for traditional multi-label classification such as Euclidean (EL), Hinge loss (HL) and Cross Entropy (CEL), first we use the 3 standard loss functions for regression. However, the instance of positive learning is difficult (pathologies with images) and rather the label of images are spares, means positive "0" or negative "1". Hence, we host the positive/negative balancing factor (βP , βN) for pa-

thologies and normal classes that enforce the positive learning. Such as, weighted CEL (W-CEL) follows:

$$LW-CEL(f(\vec{x}), \vec{y}) =$$

$$\beta_P \sum_{y_c=1} -\ln(f(x_c)) + \beta_N \sum_{y_c=0} -\ln(1-f(x_c)),$$

where β_P is agreed to $\frac{|P|+|N|}{|P|}$ while β_N is agreed to $\frac{|P|+|N|}{|P|}$ the total number of negative or '1's and positive or '0's are |P| and |N| in a group of image labels.

6. Result and Experiment

There is various result has been analyze from the softmax regression and the DCNN (Deep Convolutional Neural Network) model, it is clearly shows that the DCNN performs and give the drastically better result than the doctor diagnosis, matrix regression, softmax regression and random weight.

- Data collection: The unified disease localization and classification framework is evaluated and validated using the ChestX-ray8 database.
- Multi-level setup: There are the various choices of multi-label classification loss function and image-label representation. The 8-dimentional label vector y = [y1, ..., yc, ..., yC], yc∈ {0, 1}, C = 8 for each image is defined. Due to this problem of multi-label classification definition transit into a regression-like loss setting.
- Constructing model: In this stage, some pretrained models like AlexNet, GoogLeNet, VGGNet and ResNet.
- Disease Localization: Due to use of activations from transition and weight from prediction layer we can calculate the heatmap, and also produce the B-Box for apiece pathology candidate.
- Training and Experimentation: The training is being done by the DCNN unified framework which helps to classify the images in Multi-label classification.

Table 1. Multi-label classification using DCNN

Scenery	Atelectasis	Cardiomegaly	Effusion	Infiltration	Mass	Nodule	Pneumonia	Pneumothorax
Diverse pre-trained models initialization								
AlexNet	0.6467	0.6927	0.6645	0.6040	**0.5645**	0.6485	0.5495	0.7427
GoogLeNet	0.6407	0.7057	0.6877	0.6087	0,5365	0.5577	0.5591	0.7825
VGGNet-16	0.6285	0.7085	0.6505	0.5895	0.5105	0.6557	0.5101	0.7515
ResrNet-50	**0.7169**	**0.8140**	**0.7365**	**0.6127**	0.5607	**0.7165**	**0.6335**	**0.7890**
Diverse multi-label loss function								
CEL	0.7065	0.7265	0.7353	0.6085	0.5531	0.6547	0.5165	0.7663
W-CEL	0.7169	0.8140	0.7365	0.6127	0.5607	0.7165	0.6335	0.7890

7. Conclusion

The computerized diagnostic of the radiology image database performance not has been spoken till this work. In many modern hospital's the tremendous number of x-ray images are stored in PACS (Picture Archiving and Communication System). The conditions of plethora been diagnosed by the sustainable number of chest X-Ray. Our shot to create the "human-machine interpreted" which helps to get the comprehensive chest X-ray comparison from the tens of thousands chest x-ray images present in database and became a realistic methodological challenge by using the ImageNet under the DCNN unified Framework. In future we can improve the accuracy validation of the images and made a UI or Android application, so it may be user friendly to everyone.

References

[1] Wang X, Peng Y, Lu L, et al., IEEE Conference on "NIH clinical center provides one of the largest publicly available chest x-ray datasets toscientificcommunity," pp. 2097-2106, Sep 2017.

[2] E. Zimmermann, "Mayo clinic researchers' demon- strate value of second opinions, "Journal of Evaluation of clinical pactice, https://newsnetwork.mayoclinic.org/discussion/ mayo-clinic-researchers-demonstrate-value-of-second-opinions, April 2017.

[3] D. Wang, A. Khosla, R. Gargeya, et al., "Deep learning for identifying metastatic breast cancer," arXiv [q-bio.QM], June 2016.

[4] M. Yann and Y. Tang, "Learning deep convolutional neural networks for X-ray protein crystallization image analysis," Proc. Conf. AAAI Artif. Intell., vol. 30, no. 1, pp. 1373- 1379, Feb 2016.

[5] P. Rajpurkar, J. Irvin, B. Yang, et al., "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning, "arXiv :1711.05225, vol.1, pp.1-7, Dec 2017.

[6] V. Vryniotis "Machine learning tutorial:The logistic regression (softmax regression) ".,http://blog.datumbox.com/machine-learning-tutorial-the-multinomial-logistic-regression-softmax-regression, Nov 2013.

[7] K. He, X. Zhang, S. Ren et al., "Deep residual learning for image recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016.

[8] M. Oquab, L. Bottou, I. Laptev, et al., "Is object localization for free? - Weakly-supervised learning with convolutional neural networks," in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 685–694 Jun 2015.

[9] B. Zhou, A. Khosla, A. Lapedriza, et al., "Learning deep features for discriminative localization,". arXiv:1512.04150 Dec 2015.

[10] T. Durand, M. Cord and N. Thome Weldon: "Weakly supervised learning of deep convolutional neural networks, "Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.4743-4752, June 2016

[11] S .Hwang, H. E. Kim "Self-Transfer Learning for Weakly Supervised Lesion Localization, "In Medical Image Computing and Computer-Assisted Intervention – MICCAI, Lecture Notes in Computer Science. Springer, Cham, Vol 9901, no.2, pp.239– 246, Oct 2016.

[12] J. Deng, W. Dong, R. Sochere et al., "ImageNet: A large-scale hierarchical image database," in IEEE Conference on Computer Vision and Pattern Recognition, pp. 248-255, June 2009.

[13] O. Russakovsky, J. Deng, H. Su, et al., "ImageNet large scale visual recognition challenge," Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, April 2015.

[14] A. Krizhevsky, G. E. Hinton, and I. Sutskever, "Imagenet classification with deep convolutional neural networks, "In Advances in neural information processing systems, pp.1097–1105, 2012.

[15] C. Szegedy, W. Lei,Y. Jia, et al., "Going deeper with convolutions," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1-9, June 2015.

[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556, April 2014.

[17] K. He, X. Zhang, S. Ren, et al., "Deep residual learning for image recognition," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), arXiv preprint arXiv:1512.03385, Dec 2015.