Chest Disease Detection from Human X-Ray Scans Using Deep Learning

Ravi Soni, Rohan Singh, Shadib Shah and Shivang Malik

February 5, 2020

Chest disease detection from human X-ray scans using Deep Learning

Ravi Soni (ravi.soni.cs.2016@miet.ac.in)

Rohan Singh (rohan.gusain.cs.2016@miet.ac.in)

Shadib Shah (shadib.shah.cs.2016@miet.ac.in)

Shivang Malik (shivang.malik.cs.2016@miet.ac.in)

Supervisor : Mrs. Sonika Jalhotra

Department of Computer Science & Engineering

Meerut Institute of Engineering & Technology, Meerut

Chest disease detection from human X-ray scans using Deep Learning

Abstract

Millions of chest X-rays produced world-wide are currently analyzed almost entirely visually on a scan-by-scan basis. This requires a relatively high degree of skill and concentration, and is time-consuming, expensive, prone to operator bias (data distortion or wrong interpretation), and unable to exploit the invaluable informatics contained in such large-scale data.

Errors and delay in these diagnostic methods still contribute to a large number of patient deaths in hospitals, making these errors one of the largest causes of death along with heart disease and cancer.

Moreover, due to the complexity of these scans, it is challenging even for radiologists to differentiate various diseases on them, resulting in the shortage of expert radiologists, particularly in rural areas who are competent to read chest radiographs.

Therefore, it is of utmost significance to design and implement automated algorithms for computer-aided diagnosis of diseases on chest radiography.

Deep learning has transformed healthcare. It’s being used extensively to diagnose cancer, pneumonia, hernia and other diseases. Deep learning is more accurate and faster at diagnosis than real doctors. Automation of X-ray analysis can prevent a lot of mishaps, speed up diagnosis and reveal new patterns thus aiding in medical research.

Hence we get a deep learning model using Deep Convolutional Neural Networks (DCNN) architecture which can predict various chest diseases like Pneumonia, Pneumothorax, Atelectasis, Effusion etc with significant accuracy (>80%) and provide other insights about the analysis performed by generating heat maps and other visualizations. The model will also be able to localize the pathology by generating Class Activation Maps(CAM). We will start first with Pneumothorax owing to the availability of good dataset.

Keywords

DCNN - Deep Convolutional Neural Network, CAM - Class Activation Map,
AUC - Area under the curve
ROC - Receiver Operating Characteristics

1. Introduction:

Healthcare sector has lots of differences from other industries. People have high expectations regarding the service irrelevant of the costs associated. It did not achieve social expectation even though it consumes a huge percentage of the budget. In most of the cases the medical data is analysed by a medical expert.

Because of associated complexity, huge variations among the images, mental fatigue and human nature this interpretation is of limited accuracy and error-prone. Currently deep learning is being successfully applied in various industries to solve real world problems. But now we are seeing its applications in the medical industry too; particularly in medical imaging with good accuracy. It is going to prove itself an important component in future applications in the health sector. Here we will discuss the deep learning architecture and various optimizations used for medical image segmentation and classification. A large number of adults are hospitalized with pneumonia and around 50k die from the disease every year in the United States alone (data obtained from CDC). X-ray scans are currently the best available method for diagnosing various diseases, playing an important role in clinical care and hygiene studies. However, detecting these diseases in chest radiographs is a challenging task that relies on the availability of expert radiologists. In this work, we present a deep learning model that can automatically detect these diseases from chest X-rays at a level at least equal to and hopefully exceeding practicing radiologists

Detecting diseases like node, module or pneumonia in chest x-ray scans is a difficult and time consuming task particularly among amateur radiologists. The appearance of diseases like hernia in radiographs is often indefinite, overlapping with other analysis, and resembles many other gentle abnormalities. These discrepancies cause a significant variance among radiologists in the diagnosis of pneumonia. We assess the performance of both radiologists and our model on the test set for the disease detection task.

It is much more complicated and difficult to share the medical data as compared to real world images. Data privacy is effectively both sociological as well as a technical issue, which must be addressed jointly from both perspectives. Privacy is still a major concern in the healthcare sector. There are legal rights to patients regarding their personally identifiable information and establish obligations for healthcare providers to protect and restrict its use or disclosure.As there is an exponential accumulation of medical data, data analytics researchers are facing big challenges to ensure the privacy of patients. However throwing away information like age, gender etc can make it extremely difficult to catalog the data with a unique individual.

Another major concern is the restriction of the data to an organization because of its business and research needs. These privacy challenges are factors that can lead to situations where, data analytics model likely to impact it negatively from both legal as well as ethical perspective. The main privacy challenges associated with healthcare data analytics, overrunning the privacy concerns of traditional data processing, are as follows: One important issue to address that how to share sensitive data of data while limiting disclosure and limiting its sharing by ensuring the sufficient data utility i.e. Year of birth, Zip code etc.

Supported technologies and algorithms

Our model is a CNN whose input is a chest X-ray scan and predicts the probability of chest disease along with a heat map indicating the areas of the image most indicative of a particular disease. We train the model on the ChestX-ray14 dataset which contains thousands of front facing-view chest X-ray scans, each labeled with up to 14 different chest diseases, including Mass, Nodule etc. Considering a single example in the training set, we can optimize the binary cross entropy loss by using the following formula

$$\begin{aligned} L(X, y) &= -w_{+} \cdot y \log p(Y = 1|X) \\ &\quad -w_{-} \cdot (1 - y) \log p(Y = 0|X), \end{aligned}$$

where p(Y = i|X) is the probability that the network assigns to the label, w+ = |N|/(|P|+|N|), and w− = |P|/(|P|+|N|) with |P| and |N| the number of positive and negative cases of the disease under consideration in the training set respectively.

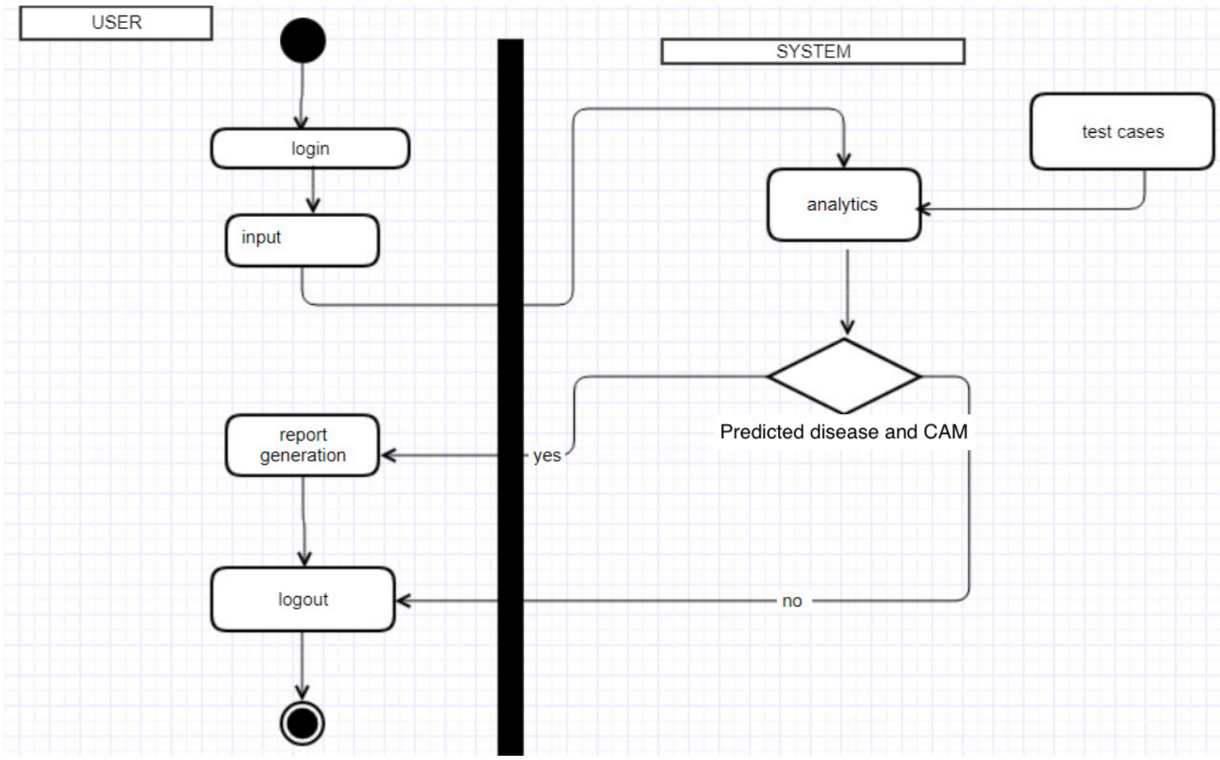
2. Proposed Work Plan:

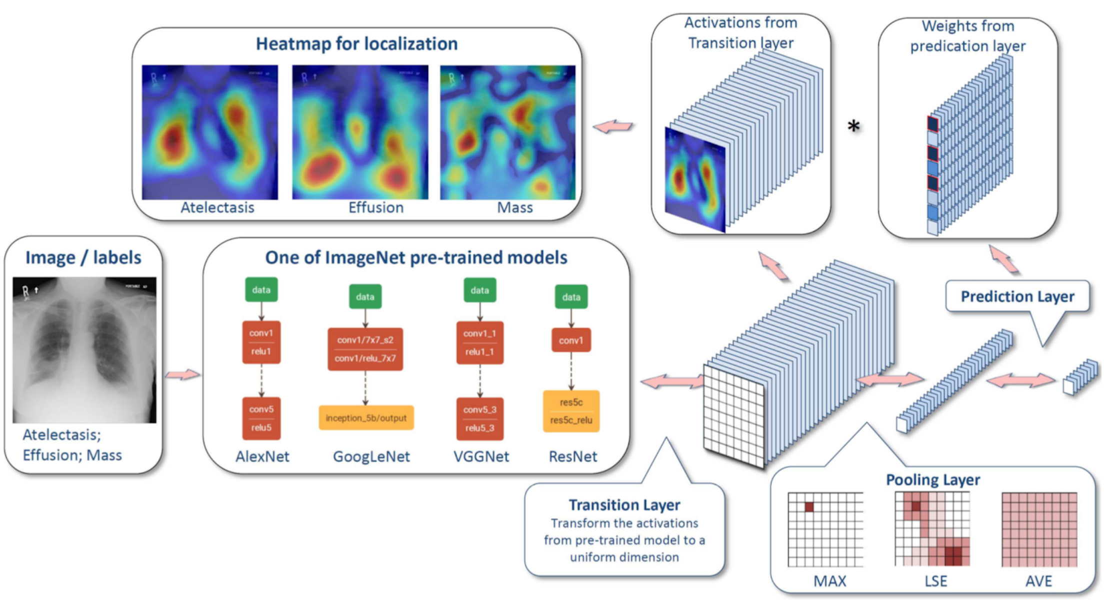
2.1 General architecture of the overall system to be designed.

Artificial neural networks are structurally similar to human biological nervous system which has a biological neuron as its functional unit . Perceptron is one of the earliest neural network that was based on human brain system. A simple perceptron consists of input layer that is directly connected to output layers and is able to classify linearly separable patterns. To solve problems of higher dimensionality, a neural network was created which consists of a layered architecture i.e., input layer, output layer and one or more hidden layers. Neural networks consist of interconnected neurons that take input and perform some processing on the input data, and finally forward the current layer output to the coming layer. Each neuron in the network sums up the input data and applies the activation function to the summed data and finally provides the output that might be propagated to the next layer. Thus adding more hidden layer allows to deal with complex as hidden layer capture nonlinear relationships. These neural networks are known as Deep Neural network. Deep learning provides new cost effective to train DNN were slow in learning the weights. Extra layers in DNN enable composition of features from lower layers to the upper layer by giving the potential of modeling complex data.

Today, in several fields deep learning algorithms are outperforming humansi.e. identifying the markers for cancer in bloodstream and tumors affected areas in MRI scans. This improvement is significantly because of the addition of more hidden layers resulting in detailed analysis and predictions. For several applications the performance is unparalleled speed owing to recent hardware and software advancements i.e. speech recognition, object detection and medical imaging. A deep neural network consists of stacks of multiple layers of neurons, each layer capable of representing a hierarchy. The number of layers can range from 2 to a whopping 1,000! With such an extreme information modeling capacity, a deep network becomes capable of memorizing almost all possible mappings as a result of successful training with a sufficiently large knowledge database. It can then create intelligent predictions for unseen cases. In this way we can see how deep learning is creating a huge impact in computer vision and medical imaging. Similar to its impact in domains like text processing, voice research, etc. There exist various deep learning algorithms and architectures like convolutional neural networks (CNN), Deep Belief Network (DBN), Generative Adversarial Networks (GAN), Variational Autoencoder (VAE), Deep Boltzmann machine (DBM), Recurrent Neural Network (RNN) and its variants The CNN model is of a great interest in computer vision. We will use pre-trained ResNet model to speed up training out of the available CNN

architectures.


USER
SYSTEM
login
input
test cases
analytics
Predicted disease and CAM
yes
report generation
no
logout


Heatmap for localization
Atelectasis
Effusion
Mass
Activations from Transition layer
*
Weights from predication layer
Image / labels
Atelectasis; Effusion; Mass
One of ImageNet pre-trained models
AlexNet
GoogLeNet
VGGNet
ResNet
Prediction Layer
Transition Layer
Transform the activations from pre-trained model to a uniform dimension
Pooling Layer
MAX
LSE
AVE

2.2 Algorithm of main complement of the system.

Above algorithm can be easily extended to perform multiclass classification by making a number of changes. First, instead of computing a single binary label, we compute a vector t of binary labels indicating the absence or presence of the following diseases under consideration: Edema, Effusion, Emphysema, Fibrosis, Hernia, Infiltration, Mass, Atelectasis, Cardiomegaly, Consolidation, Nodule, Pleural Thickening, Pneumonia, and Pneumothorax. Now, we replace the final fully connected layer, with a fully connected layer producing a 14-dimensional output, after which we apply the sigmoid function to introduce non linearity in the model. The final output will give us the predicted probability of the presence of each class of disease. Finally, we modify the loss function to minimize the gross result of unweighted binary cross entropy loss

$$\begin{aligned} L(X, y) &= \sum_{c=1}^{14} [-y_c \log p(Y_c = 1|X) \\ &\qquad -(1 - y_c) \log p(Y_c = 0|X)], \end{aligned}$$

es

where p(Yc = 1|X) is the computed probability that the image represents the disease c and p(Yc = 0|X) is the computed probability that the image does not represent the associated disease. For efficient analysis we randomly split the dataset into training (70%), validation (10%), and test (20%) sets. We also make sure that there is no patient overlap between these splits.

3. Result analysis:

3.1 Description of data set used.

We use the ChestX-ray14 dataset which contains about a hundred thousand front facing X-ray scans of nearly thirty thousand unique patients. Each image is automatically annotated using automatic extraction methods, resulting in 14 different classes of diseases. As an example we annotate images that have pneumonia as one of the labeled classes as positive examples and annotate all other images as negative samples. For the detection task, we randomly split the dataset into training, validation, and test. There is no patient overlap between the sets. Before inputting the images into the network, we downscale the images to 224×224 and normalize based on the mean and standard deviation of images in the training set. To aid training we perform data augmentation using horizontal flipping.

We collected a test set of frontal chest X-ray scans. Annotated data was obtained independently from the work at Stanford University. The radiologists had vast medical experience. Moreover these experts did not have access to any patient information or knowledge of disease most prevalent in the data. In this way we collected the data.

<u>3.2 Calculate the efficiency or accuracy of the designed system according to the parameter used to evaluate the system.</u>

1. TO evaulate our network, we use the **ROC** score. In a ROC curve, we plot the positive predictive value (true positives / (true positives + false positives)) against the negative predictive value (true negatives / (true negatives + false negatives)). We can get different positive and negative prediction values by choosing a different confidence threshold at which we consider a disease to be positive or negative. Area under curve is the metric used to evaluate the quality of the resulting model.

To analyse the model predictions, we also generate heatmaps to visualize the portion of images which are most indicative of a particular disease. This is done by computing class activation mappings (CAMs). To generate the CAMs, we feed an image into the fully trained network and extract the feature maps that are output by the final convolutional layer.

Interpretation of Model's output using Class Activation Mapping (CAM)

1. Class activation mapping generates "heat maps" indicating the regions of the image to which our model "attends" in the final layers.
2. Extract the appropriate feature map from an input image.
3. Compute the CAM for a few sample images. To visualize the image regions, we'll need to retrofit our CAM to our image size.

4. Conclusion:

Early diagnosis and treatment of diseases is critical to preventing complications including death. With billions of procedures per year, chest X-ray scans are one of the most common and important diagnosis tools used in practice. They are used for diagnosing and screening a variety of diseases like pneumonia. Statistics by the World Health Organization indicate that about two thirds of the global population lacks access to proper X-ray diagnostics. Even in the presence of proper imaging equipment the mortality rate is significantly high because of the shortage of experts who can correctly interpret these scans. We develop a public facing platform which is capable of detecting medical anomalies and predictions at levels exceeding medical experts and radiologists. We tested this on the ChestX-ray14 dataset which is one of the largest publicly available chest X-ray dataset. With increasing levels of automation which are simulating the diagnosis done by experts, we hope to achieve an improved healthcare diagnostic access to those parts of the world where there is a shortage of skilled radiologists.

References:

1. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases [http://openaccess.thecvf.com/content_cvpr_2017/papers/Wang_ChestX-ray8_Hospital-Scale_Chest_CVPR_2017_paper.pdf]
2. Zhe Li, Chong Wang, Mei Han, Yuan Xue, Wei Wei, Li-Jia Li, Li Fei-Fei ., et al Thoracic Disease Identification and Localization with Limited Supervision. [https://arxiv.org/abs/1711.06373]
3. Li Yao, Eric Poblenz, Dmitry Dagunts, Ben Covington, Devon Bernard, Kevin Lyman., et al Learning to diagnose from scratch by exploiting dependencies among labels. [https://arxiv.org/abs/1710.10501].
4. Aydogdu, M, Ozyilmaz, E, Aksoy, Handan, Gursel, G, and Ekim, Numan. Mortality prediction in community-acquired pneumonia requiring mechanical ventilation; values of pneumonia and intensive care unit severity scores. Tuberk Toraks, 58(1):25– 34, 2010.
5. Berbaum, K, Franken Jr, EA, and Smith, WL. The effect of comparison films upon resident interpretation of pediatric chest radiograph.
6. Zhou, Bolei, Khosla, Aditya, Lapedriza, Agata, Oliva, Aude, and Torralba, Antonio. Learning deep features for discriminative localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2921–2929, 2016.
7. Neuman, Mark I, Lee, Edward Y, Bixby, Sarah, Diperna, Stephanie, Hellinger, Jeffrey, Markowitz, Richard, Servaes, Sabah, Monuteaux, Michael C, and Shah, Samir S. Variability in the interpretation of chest radiographs for the diagnosis of pneumonia in children. Journal of hospital medicine, 7(4): 294–298, 2012.
8. Islam, Mohammad Tariqul, Aowal, Md Abdul, Minhaz, Ahmed Tahseen, and Ashraf, Khalid. Abnormality detection and localization in chest x-rays using deep convolutional neural networks. arXiv preprint arXiv:1705.09850, 2017.
9. Huang, Gao, Liu, Zhuang, Weinberger, Kilian Q, and van der Maaten, Laurens. Densely connected convolutional networks. arXiv preprint arXiv:1608.06993, 2016.
10. Neuman, Mark I, Lee, Edward Y, Bixby, Sarah, Diperna, Stephanie, Hellinger, Jeffrey, Markowitz, Richard, Servaes, Sabah, Monuteaux, Michael C, and Shah, Samir S. Variability in the interpretation of chest radiographs for the diagnosis of pneumonia in children. Journal of hospital medicine, 7(4): 294–298, 2012.
11. Islam, Mohammad Tariqul, Aowal, Md Abdul, Minhaz, Ahmed Tahseen, and Ashraf, Khalid. Abnormality detection and localization in chest x-rays using deep convolutional neural networks. arXiv preprint arXiv:1705.09850, 2017.
12. Ioffe, Sergey and Szegedy, Christian. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International Conference on Machine Learning, pp. 448–456, 2015.

13. Huang, Peng, Park, Seyoun, Yan, Rongkai, Lee, Junghoon, Chu, Linda C, Lin, Cheng T, Hussien, Amira, Rathmell, Joshua, Thomas, Brett, Chen, Chen, et al. Added value of computer-aided ct image features for early lung cancer diagnosis with small pulmonary nodules: A matched case-control study. Radiology, pp. 162725, 2017.
14. Huang, Gao, Liu, Zhuang, Weinberger, Kilian Q, and van der Maaten, Laurens. Densely connected convolutional networks. arXiv preprint arXiv:1608.06993, 2016.
15. Gulshan, Varun, Peng, Lily, Coram, Marc, Stumpe, Martin C, Wu, Derek, Narayanaswamy, Arunachalam, Venugopalan, Subhashini, Widner, Kasumi, Madams, Tom, Cuadros, Jorge, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama, 316(22):2402–2410, 2016.
16. Grewal, Monika, Srivastava, Muktabh Mayank, Kumar, Pulkit, and Varadarajan, Srikrishna. Radnet: Radiologist level accuracy using deep learning for hemorrhage detection in ct scans. arXiv preprint arXiv:1710.04934, 2017.
17. Goncalves-Pereira, Jo˜ao, Conceicao, Catarina, and P´ovoa, Pedro. Community-acquired pneumonia: identification and evaluation of nonresponders. Therapeutic advances in infectious disease, 1(1):5– 17, 2013.
18. Cherian, Thomas, Mulholland, E Kim, Carlin, John B, Ostensen, Harald, Amin, Ruhul, Campo, Margaret de, Greenberg, David, Lagos, Rosanna, Lucero, Marilla, Madhi, Shabir A, et al. Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies. Bulletin of the World Health Organization, 83(5):353–359, 2005.