Chest X-ray Disease Diagnosis with Deep Convolutional Neural Networks

Christine Herlihy, Charity Hilton, Kausar Mukadam
Georgia Institute of Technology, Atlanta, GA

Abstract

This project uses deep convolutional neural networks (CNN) to: (1) detect and (2) localize the 14 thoracic pathologies present in the NIH Chest X-ray dataset. Our approach is to develop a multi-label CNN using the network architectures from previous works as baselines, to detect and visualize the selected diseases. We also use a simpler feed forward network to incorporate patient information and compare results against the CNN.

Our project presentation slides are available here.

Our video presentation is available here.

The code is available on Github here

1 Introduction

Many common thoracic pathologies, including pneumonia, pulmonary disease, and edema, are typically diagnosed through chest radiographs (X-rays). Radiographic diagnosis can be life-saving, but it requires radiologists with extensive domain knowledge to interpret X-ray results. As such, radiographic diagnosis does not scale well to points of care where such professionals are in short supply, including rural areas and developing countries. In addition, even highly trained radiologists may fail to correctly diagnose, and/or distinguish between various pathologies with overlapping physical and/or radiographic features[6].

Due to these limitations, research into computer-aided diagnosis of these images through rule-based methods, machine learning, and most recently, deep learning, has garnered attention[4]. Computational approaches have the potential to improve clinical outcomes by increasing diagnostic accuracy and speed; they may also allow make it easier to deliver diagnostic care in underserved communities.

2 Problem Description

Since the release of the NIH chest X-ray dataset[3], research into computer aided pathology detection through X-ray analysis is now easier. As with previous research, our aim is to use the available images and socio-demographic patient information, to enable the accurate prediction of 14 disease labels. We also attempt to visualize the relationships learned through the convolutional neural networks through heat maps for better interpretability.

3 Related Work

While X-rays have been used to diagnose a variety of pathologies, including chronic obstructive pulmonary disease[5], pneumonia[6,10,20], lung cancer[8], and tuberculosis[15], for more than a century, advancements in computer-aided diagnosis have until recently been hindered by the relative scarcity of training data, given the cost and PHI concerns associated with assembling a labelled data set of X-ray scans large enough for use in deep learning. In 2017, the National Institute of Health (NIH) released one of the largest publicly available chest X-ray datasets to date, which includes demographic information and images from more than 30,000 patients[9].

Many researchers have leveraged this dataset, as well as others like it, including the Indiana University Chest X-ray Collection[12], the Japanese Society of Radiological Technology (JSRT) database[13], and the Shenzhen dataset[14], to develop deep learning models capable of detecting and distinguishing between various lung pathologies, including in instances where a single patient may present with more than one pathology, as well as localizing these pathologies within the image files[3,6,11].

While introducing the NIH dataset and its labelling technique, Wang et al.[3] provide a DCNN based multi-label weakly supervised method for disease classification, setting benchmark results for future work on the dataset. Pre-trained networks on ImageNet[17] including AlexNet, GoogLeNet, VGGNet-16, and ResNet-50 were used without the final classification layers. Some additional layers were added to the architecture to adapt it to the classification task, including a pooling layer, transition layer and weighted Cross Entropy Loss layer. For localization, the pooling filters information to be passed down, and the weights of the prediction layer can then be used in spatial maps, producing weighted maps localizing active X-ray areas for each disease class.

Islam et al.[11] apply deep convolutional neural networks (DCN) to the Indiana, JSRT, and Shenzhen datasets in an effort to detect a range of abnormalities, with a particular focus on binary classification and localization of cardiomegaly and pulmonary atelectasis. They demonstrate improvement in classification accuracy relative to rule-based methods, and also find that ensemble learners consisting of multiple DCNs, in which the final labels are obtained by averaging the probabilities output by each individual model, resulted in increased area under the ROC curve relative to individual models. To derive information about localization, they occlude small portions of each X-ray scan, and note which region(s) have the largest impact on the softmax score output by the model when occluded[11].

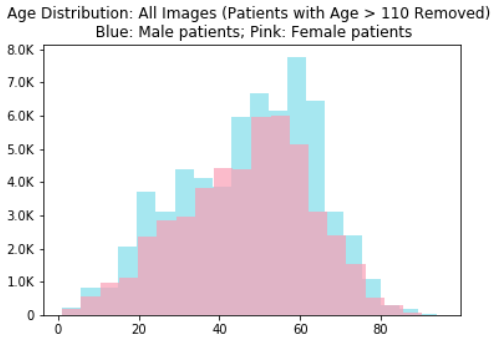
Rajpurkar et al.[6] produced state of the art results for predicting pneumonia through a 121-layer convolutional neural network based on the NIH dataset. The model was based on DenseNet[16] pre-trained on the ImageNet[17] dataset using the Adam optimizer with default parameters. In comparison to the average accuracy achieved by four independent, practicing radiologists when diagnosing pneumonia, the algorithm performed significantly better, achieving an AUC of 0.7680. Class Activation Maps were used to localize the cause of diagnosis in the image. To generate these maps, the feature maps generated by the final convolutional layer of the trained network are extracted and their weighted sum, based on associated weights in the network, is used to extract the most salient features.

4 Experimental Setup

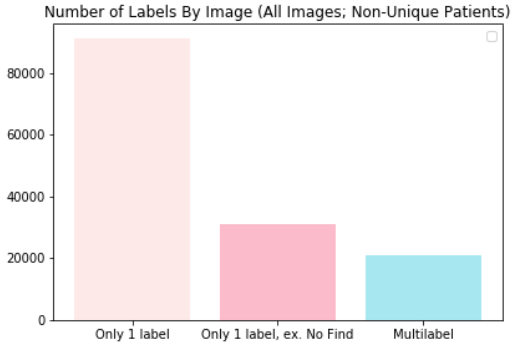
4.1 Data

The project is based on the data provided by the NIH Chest X-ray dataset[3], which comprises $112,120$ 1024×1024 pixel frontal-view X-ray images with 15 labels, including 14 different disease labels and a "No Finding" label, obtained using text mining. Limited patient-level socio-demographic information is included with each X-ray image; from this, we are able to compute summary statistics about the patient population represented in this dataset.

The $112,120$ images are associated with $30,805$ unique patients; $14,175$ (54%) of these patients are men, and $16,630$ (46%) are women. There are 16 patients in the dataset whose ages are listed (presumably erroneously) as > 110. When these patients are excluded, the average age of patients in this dataset is 46.9 years, with a min of 1.0 years, a max of 95.0 years, and a standard deviation of 16.6 years.



(a) Age Distribution by Gender



(b) Number of Labels Per Image

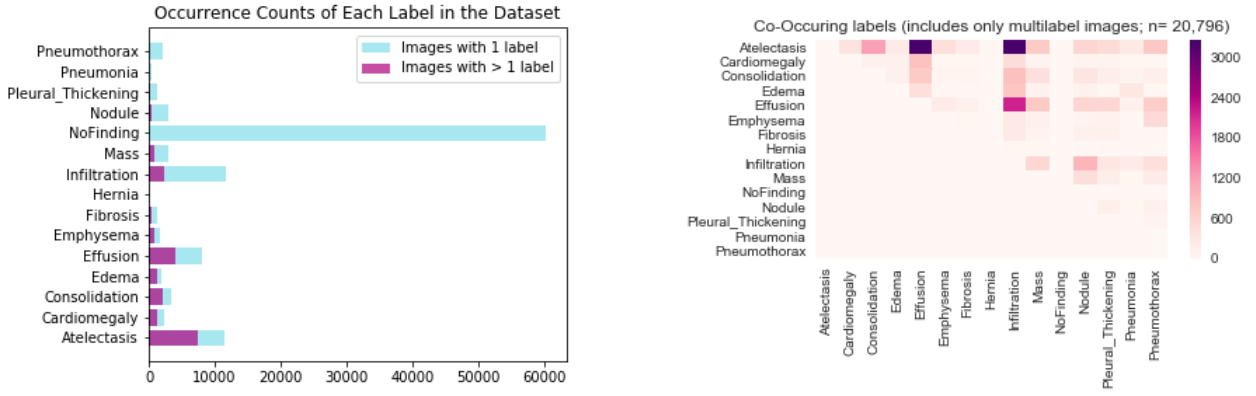
Figure 1: Visualizing the NIH Chest X-Ray Dataset

Each image is mapped to a patient ID and a follow-up visit number; from this information, one can infer that some patients have multiple images in the dataset that may be associated with multiple visits. There is not a precise visit date associated with each image, and patients' ages are recorded only as integers; thus, it is difficult to determine with certainty whether a set of diagnosis labels co-occur when they are mapped to a single patient ID-age value pair, but different images with different follow-up visit numbers.

When all images are considered independently of patient ID, $91,324$ images are associated with only one label that may include "No Finding." Of these, $30,963$ images are associated with only one label that is one of the 14 diseases detailed below; the remaining $20,796$ images are associated with more than one label. By definition, the list of diagnoses associated with multilabel images excludes "No Finding."

The 14 diseases labeled in this dataset include atelectasis, consolidation, infiltration, pneumothorax, edema, emphysema, fibro-

sis, effusion, pneumonia, pleural thickening, cardiomegaly, nodule, mass and hernia. It should be noted that a "No Finding" label does not necessarily mean that the patient in question is healthy, as he/she may be suffering from condition(s) other than those enumerated above[3]. Since multi-labelling of images is possible, many images (and, as mentioned above, sets of images), may represent the co-occurrence of diseases, which can make identifying the exact cause for disease diagnoses substantially difficult. The plots below illustrate the distribution and co-occurrence of labels within the dataset. Note that the co-occurrence heatmap only considers single images associated with multiple labels, and does not consider a set of labels associated with a unique patient ID to be co-occurring if they are split across multiple images with different follow-up visit numbers.



(a) Label Frequency (Single & Multilabel Images) **(b)** Heatmap of Co-Occuring Labels for Multilabel Images

Figure 2: Label Composition of the NIH Chest X-Ray Dataset

4.2 Software Stack

We used Apache Spark 2.3.0 for our ETL pipeline; specifically, we use Python and PySpark to ingest the data, including images and patient-level demographic data, and compute summary statistics for initial exploration. In addition, we used Seaborn and Matplotlib to develop high-level data visualizations to represent the composition of the patient population. For modeling the networks, we have relied primarily on PyTorch. For visualizing the heatmaps and ground-truth bounding boxes super-imposed over the original X-ray images, we used PyTorch and Pillow (a Python-based imaging library).

4.3 Hardware Specifications

Since training CNNs requires substantial hardware resources including GPUs, we used the computing resources at our disposal: ICEHAMMER and PACE. PACE is a current GaTech environment for high performance computing, consisting of 9 nodes, 2 with 7 Tesla C2050 Nvidia GPUs, 1 with a Tesla C2050 GPU, 5 with AMD GTX 580 GPUs, and 1 node with a Nvidia TeslaT10 GPU. ICEHAMMER[18] is a private resource available to GTRI researchers (all members of this team are currently employed there). ICEHAMMER is a cluster environment with 1200 cores, 5TB of memory, 3 IBM POWER8 with Nvidia GPUs, 22 Kepler GPUs, 4 Maxwell GPUs, and 9 Pascal GPUs. In addition, it is equipped with JupyterHub and Slurm.

We successfully implemented a PyTorch environment, including torchvision, OpenCV, and Pillow on the ICEHAMMER cluster using a single NVIDIA GeForce GPU.

5 Approach

This project uses deep convolutional neural networks to: (1) detect and (2) visualize the 14 thoracic pathologies present in the NIH Chest X-ray dataset. We note that it is of particular importance to achieve high levels of predictive performance for the detection of atelectasis, infiltration and effusion, as these are the predominant classes available in the dataset with 16057, 25366, and 18974 images respectively.

5.1 CNN for Pathology Detection Using X-Ray Images

Since the release of the NIH dataset, many convolutional networks have been proposed to solve this task. Our approach was to develop a multi-label CNN using the network architectures from previous works as baselines, to detect the selected diseases. Given that we intended to develop and build on a variety of networks by tuning both network architecture and hyper parameter settings of previous work, we attempt to compare the performance of each network to one another.

The initial approach we used was to use a pre-trained network as a baseline, as suggested by Wang et al.[3]. The first network builds on the ResNet-50[20] pre-trained network in PyTorch, following the modifications used by the previous authors. A 2048 x 2048 transition layer, a global pooling layer, and a loss layer using cross entropy loss were added to the existing model, to adapt it for the 14 class dataset; a sigmoid activation layer was used for prediction.

Later, we deployed PyTorch implementations of the CheXNet models[22,23] which use a 121-layer DenseNet convolutional neural network. The first model is the same as the standard DenseNet architecture with an additional sigmoid function applied to produce independent probability estimates for each class (i.e. diagnosis labels). The second model was similarly trained, using experiments from DenseNet-121, DenseNet-169 and DenseNet-201.

5.2 FFN for Pathology Detection Using Patient-Level Sociodemographic Features and Diagnostic History

In addition to leveraging convolutional neural networks on the X-ray images, research aimed at using temporal information to improve patient-level prediction models inspired us to investigate whether feed-forward neural networks could be used to predict diagnostic labels for each image, based on the limited patient-level sociodemographic features included in the dataset (i.e., age and gender), as well as each patient's diagnostic history (i.e., the cumulative sum of all labels seen up to that point on previous visits, or the multi-hot vector representing the diagnostic label(s) from the patient's most recent visit)[25].

To implement this approach, we grouped the NIH metadata by `Patient ID`, and then considered which feature(s) could be reliably used to determine temporal ordering of each patient's associated image(s). This was challenging for three primary reasons: (1) the dataset does not include a visit date or timestamp; (2) patient ages are listed as integers, and in some cases, patients may have multiple images associated with the same age, such that the temporal proximity of these images is unclear; and (3) when the dataset is sorted by `Patient ID` and `Patient Age`, the `Follow-up #` field does not sequentially increase in the way we might expect; the same is true if we sort by `Patient ID` and `Follow-up #` for `Patient Age`.

Thus, we decided to use `Patient Age` to sort each unique patient's image(s). For patients with more than 1 image, we computed a vector to represent the set of diagnoses associated with previous image(s), by lagging each set of diagnoses and computing the cumulative sum of the multi-hot lagged vectors available for a given patient at each time step. Our model is paramaterized such that either the vector associated with the most recent previous visit can be used in isolation, or the cumulative sum vector, representing the patient's full diagnosis history up to but excluding the current visit, can be used.

We train a feed-forward neural network with one 128-node hidden layer, a cross-entropy loss layer, and a sigmoid activation function, to produce independent probability estimates for each diagnosis label. As part of cross-validation, we perform 5 different splits of the data into train and test sets, with test set size $= 0.3$ of the full dataset. We then compute average loss per epoch during the training and test periods (averaged over all CV splits), as well as average test-set AUC for each diagnosis label.

5.3 Pathology Localization and Visualization Using Heat Maps

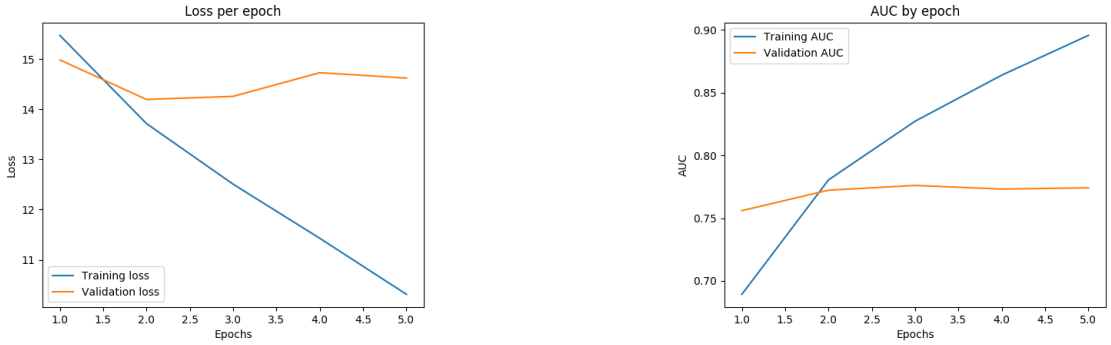
The results from the DenseNet model, since it was the better performing architecture, were visualized using heat maps and superimposed with the bounding boxes provided in the data set. The heat maps were generated using the weights obtained in the final layer of the CNN, mapped to image coordinates. This helped us visualize specific thoracic conditions and analyze image areas that led to certain disease diagnoses.

6 Experimental Results

Our approach was to tune and test a variety of networks, leading to architecture and hyper parameter changes from previous work, and compare results across baselines to measure success. The architectures are compared based on AUC and the results are given below in Table 2.

The ResNet-50 network[24] was trained on the complete available data, with 78484 images used for training, 11214 used for

validation and the remaining for testing. The CNN was trained for 5 epochs, on an ICEHAMMER GPU with 45GB of memory. While initially the model was to be trained for a larger batch size (24), until convergence, we eventually used a smaller batch size (12) to ensure in-memory execution. During each epoch, the model is first trained and then validated using the hold out validation set. From these generated models (one during each epoch), the best performing model is chosen as the final model. This model was then evaluated using the hold out test set, and the performance is reported below. The training and validation loss changes during epochs, along with the AUC changes during training, are depicted in Figure 3.



(a) Training and Validation Loss by epoch

(b) AUC changes with epochs

Figure 3: Training statistics for ResNet implementation

The first DenseNet model[22] was trained on all 112,120 images using a 121-layer Dense Convolutional Network (DenseNet) with PyTorch and torchvision. This model is the same as the standard DenseNet architecture with an additional sigmoid function applied to produce independent probability estimates for each class (i.e. diagnosis labels).

The second model[23] was similarly trained, using experiments from DenseNet-121, DenseNet-169 and DenseNet-201. The models were evaluated on the ICEHAMMER cluster using Slurm on an NVIDIA GeForce GTX 1080 Ti GPU. The AUC results are slightly higher than Rajpurkar et al.[6]

Table 1: Comparison of Results for Convolutional and Feed Forward Neural Networks

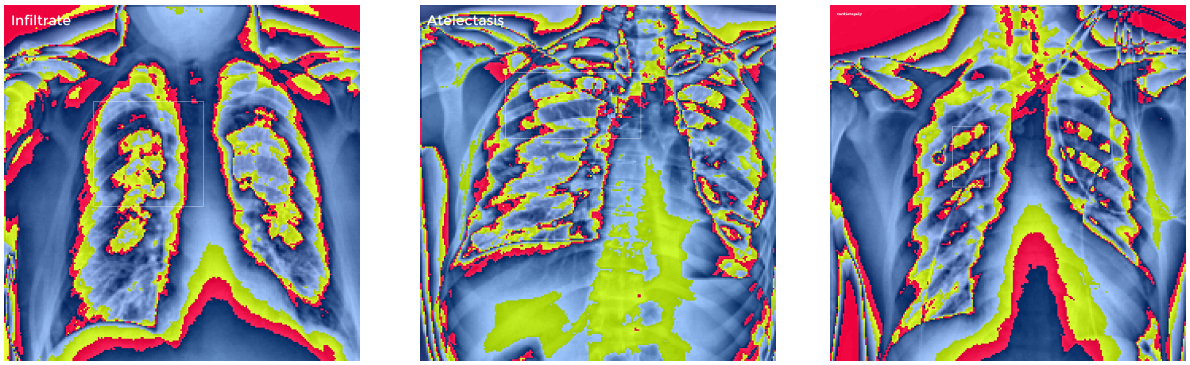
	Average AUC	Atelectasis	Cardiomegaly	Effusion	Infiltration	Mass	Nodule	Pneumonia
ResNet-50	**0.7742**	0.7393	0.8507	0.8367	0.7004	0.7077	0.6438	0.6795
DenseNet-121+Sigmoid	**0.843**	0.8294	0.9165	0.887	0.7143	0.8597	0.7873	0.7745
DenseNet-[121,169,201]	**0.8508**	0.8321	0.9107	0.886	0.7145	0.8653	0.8037	0.7655
FFN-CumulativeSumMultiplier	**0.4980**	0.5004	0.5391	0.4932	0.4766	0.4334	0.4879	0.5048
FFN-CumulativeSumNoMultiplier	**0.5037**	0.5000	0.5516	0.5014	0.4998	0.5000	0.4786	0.4999

	Pneumothorax	Consolidation	Edema	Emphysema	Fibrosis	Pleural Thickening	Hernia
ResNet-50	0.8526	0.7595	0.8948	0.8638	0.7509	0.7493	0.8094
DenseNet-121+Sigmoid	0.8727	0.8142	0.8933	0.9254	0.8304	0.7831	0.9104
DenseNet-[121,169,201]	0.8857	0.8157	0.9017	0.9422	0.8523	0.7948	0.9416
FFN-CumulativeSumMultiplier	0.5757	0.4982	0.4962	0.5083	0.4872	0.4858	0.4862
FFN-CumulativeSumNoMultiplier	0.5189	0.5005	0.5000	0.5015	0.4999	0.5000	0.5000

As illustrated by the results in the table above, the performance of the FFN with respect to average AUC per diagnosis label was quite weak, and in many cases, worse than random. To attempt to improve model performance, we implemented several tweaks, including discretizing patient age into a series of buckets, and/or applying a multiplier to positive diagnosis labels when generating each patient's diagnostic history vector, such that Type II errors would be more heavily penalized. While these tweaks did improve AUC performance for a subset of diagnosis labels, they typically resulted in worse performance on other labels, and also resulted in an increased number of false positives.

Visualizing the final layer weights for different images, with diverse predicted labels allowed us to understand the relationships

learned by the convolutional neural net and increase model interpretability. Figure 4 contains some sample heat maps generated for the chest X-ray images, superimposed with the bounding boxes from the dataset, along with disease labels.



(a) Infiltration **(b)** Atelectasis **(c)** Cardiomegaly

Figure 4: Heat map visualization of CNN weights superimposed with bounding boxes for localization

7 Evaluation

As seen above, the DenseNet based implementations perform significantly better than the ResNet-50 network.[3,6]. This is not surprising, given that the DenseNet network is substantially deeper, providing a better approximation of the relationship between the chest X-ray images and disease labels.

The AUC results for the ResNet architecture are better than the baseline results from previous research[3]. Similarly, the AUC results from the DenseNet architectures are slightly higher than Rajpurkar et al.[6] As mentioned above, our focus was to improve the results for atelectasis, infiltration and effusion, as these are the predominant classes available in the data set, but contrary to our expectations the best results were for Hernia, Emphysema, Cardiomegaly and Edema. While there is a lesser number of images available for these better performing classes, their identification could be a clearer, less complicated endeavour and hence be represented better through the CNN.

The FFN results were generally disappointing. While this is undoubtedly due in part to the naive nature of our implementation, we also hypothesize that since patient age and follow-up visit number did not move in parallel as expected, our use of patient age to assign patient history diagnoses vectors might have resulted in the corruption and/or obfuscation of temporal information related to disease progression. In addition, it is possible that model performance might improve if this model were applied on a richer dataset with additional sociodemographic information and/or information about each patient's other comorbid condition(s).

8 Future Work

The CNN architectures used in this project are influenced by previous research. Our project aims at additionally using the socio-demographic information on patients to inform the disease diagnoses, especially in cases of recurring patient ids. Going forward, these two approaches can be combined to provide a single, more accurate prediction. This method would hence combine a deep convolutional neural network for image based diagnosis with a simpler feed-forward network capturing static patient features, like gender and average age and/or a recurrent neural network capturing the patient's age and vector of previous diagnoses. The hope is that with such a network we would be able to leverage all the information contained in the dataset and hence get better performance.

The method and pipeline used in this project could potentially be used on additional X-ray datasets, and/or on multimodal patient-level datasets consisting of time-varying patient-level features and medical imaging results. In particular, we are interested in continuing to work on this topic within the context of explainable AI, such that models could be used to provide patient-level predictions, while other types of inputs, including imaging results and/or clinical notes, could be used to enhance predictions and "point" to specific aspects of the patient's medical history to provide context for physicians as they receive and interpret the model's output.

In addition, it would be interesting to explore whether we could achieve better performance with respect to average ROC AUC if we trained separate models for each diagnosis, rather than a single model with a cross-entropy loss layer for all 14 diagnosis labels. Finally, while we have generated heat maps based on CNN weights and superimposed them with bounding boxes,

our analysis is limited due to the small number of annotated images provided with the input data set. Additional information regarding bounding boxes from physicians and radiologists would help us understand the accuracy of our localization better. We could also improve our method to generate bounding boxes based on image weights and/or a separate neural network trained on these ground truth boxes and images.

9 Conclusion

In this project we have tried to leverage all the information provided in the NIH Chest X-ray data set i.e. images, patient information like gender and age, and patient history, to predict the occurrence of 14 disease labels. Through the feed forward network, we have depicted a method to incorporate patient information alongside images. In comparison to the previous research that guided our methods, our approach has produced slightly better performance in terms of AUC.

References

1. Pryor TA, Gardner RM, Clayton RD, Warner HR. The HELP system. J Med Sys. 1983;7:87-101.
2. Gardner RM, Golubjatnikov OK, Laub RM, Jacobson JT, Evans RS. Computer-critiqued blood ordering using the HELP system. Comput Biomed Res 1990;23:514-28.
3. Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, Ronald M. Summers. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR, pp. 3462-3471, 2017.
4. B.V. Ginneken. Fifty years of computer analysis in chest imaging: rule-based, machine learning, deep learning. Radiol. Phys. Technol., 2017, pp. 23-32.
5. Elizabeth Pudney, Martin Doherty. Plain chest x-ray (CXR) in the diagnosis of chronic obstructive pulmonary disease (COPD). European Respiratory Journal Sep 2016, 48 (suppl 60) PA3936; DOI: 10.1183/13993003.congress-2016. PA3936.
6. Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul, Curtis Langlotz, Katie Shpanskaya, Matthew P. Lungren, Andrew Y. Ng. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. arXiv:1711.05225.
7. Antin, Benjamin, Kravitz, Joshua, Martayan, Emil. Detecting Pneumonia in Chest X-Rays with Supervised Learning. Available from: http://cs229.stanford.edu/proj2017/final-reports/5231221.pdf.
8. Yu.Gordienko, Peng Gang, Jiang Hui, Wei Zeng, Yu.Kochura, O.Alienin, O. Rokovyi, and S. Stirenko. Deep Learning with Lung Segmentation and Bone Shadow Exclusion Techniques for Chest X-Ray Analysis of Lung Cancer.
9. NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community [Internet]. National Institutes of Health. U.S. Department of Health and Human Services; 2017 [cited 2018Mar10]. Available from: https://www.nih.gov/news-events/news-releases/nih-clinical-center-provides-one-largest-publicly-available-chest-x-ray-datasets-scientific-community.
10. Huang, Gao, Liu, Zhuang, Weinberger, Kilian Q, and van der Maaten, Laurens. Densely connected convolutional networks. arXiv:1608.06993, 2016.
11. Islam, Mohammed Tariqul, Aowal, Md Abdul, Minhaz, Ahmed Tahseen, and Ashraf, Khalid. Abnormality detection and localization in chest x-rays using deep convolutional neural networks. arXiv:1705.09850, 2017.
12. Demner-Fushman D, Kohli MD, Rosenman MB, Shooshan SE, Rodriguez L, Antani S et al. Preparing a collection of radiology examinations for distribution and retrieval. Journal of the American Medical Informatics Association. 2016 Mar 1;23(2):304-310. Available from: DOI: 10.1093/jamia/ocv080.
13. JSRT Database, Japanese Society of Radiological Technology. Available from: http://db.jsrt.or.jp/eng.php.
14. Jaeger, Stefan, Candemir, Sema, Antani, Sameer, Wng, Y-Xing J., Lu, Pu-Xuan, and Thoma, George. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. Quantitative Imaging in Medicine and Surgery. 2014; 4(6):475-477. DOI:10.3978/j.issn.2223-4292.2014.11.20.
15. Gordienko Y., Kochura Y., Alienin O., Rokovyi O, Stirenko S., Gang P., Hui J., Zeng W. Dimensionality Reduction in Deep Learning for Chest X-Ray Analysis of Lung Cancer. arXiv preprint arXiv:1801.06495.
16. Huang, Gao, Liu, Zhuang, Weinberger, Kilian Q, and van der Maaten, Laurens. Densely connected convolutional networks. arXiv preprint arXiv:1608.06993, 2016.
17. Deng, Jia, Dong, Wei, Socher, Richard, Li, Li-Jia, Li, Kai, and Fei-Fei, Li. Imagenet: A large-scale hierarchical image database. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pp. 248255. IEEE, 2009.
18. ICEHAMMER User Guide, https://github.gatech.edu/ICEHAMMER/UserGuide.
19. Spark Deep Learning Pipelines, https://github.com/databricks/spark-deep-learning.
20. Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun. Deep Residual Learning for Image Recognition
21. Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich. Going Deeper with Convolutions. Computer Vision and Pattern Recognition (CVPR) 2015.
22. CheXNet: A pytorch reimplementation of CheXNet, https://github.com/arnoweng/CheXNet.
23. chexnet: Implementation of the CheXNet network (PyTorch), https://github.com/zoogzog/chexnet.
24. Chest X-Ray: Deep learning practice on CXR dataset, https://github.com/TRKuan/cxr8
25. Choi, Edward, Taha M, Schuetz, Andy, Stewart, et al. Doctor AI: Predicting Clinical Events via Recurrent Neural Networks [Internet]. [1511.05942] Doctor AI: Predicting Clinical Events via Recurrent Neural Networks. 2016 [cited 2018Apr24]. Available from: https://arxiv.org/abs/1511.05942