CNN based techniques for diagnosing multiple conditions in Chest X-Rays

Isaac Godfried
Data Analyst
Eastern Maine Medical Center
Bangor, ME 04401
imckillengodfried@emhs.org

Abstract

Researchers have successfully used Convolutional Neural networks to both automatically detect nodules in CT scans as well as classify them as cancerous or benign. However, the research and results with respect to Chest X-Rays has been comparatively more limited. Chest X-rays are the most commonly performed and least invasive radiography exam. As such a variety of conditions can appear in X-rays ranging from pneumonia, to cardiomegaly, to nodules. However, few papers have focused on detecting and/or diagnosing all of the possible conditions present in chest X-rays. Recently the National Institute of Health (NIH) released a dataset of over 100k chest X-rays with a variety of medical conditions. In this paper we hope to empirically compare two possible methods of diagnosis 1. Performing multi-label classification with Inception3 with a modified loss function 2. Using Faster RCNN supported by either an Inception3 or Resnet 50 base network.

1 Introduction and related work

Convolutional neural networks such as InceptionNet and ResNet have achieved high results on image classification tasks such as ImageNet and CIFAR-10. Convolutional models have also been successfully applied to medical imaging. The majority of these attempts have focused on the binary classification of X-Ray images or CT scans. This is most commonly found either in the form of detecting normal versus abnormal or benign versus malignant. A much more difficult problem is detecting all the conditions present in an image and where they occur.

Classifying multiple objects within an image has been a goal of the larger computer vision research community for quite sometime. Various methods have been developed some involving object detection (and thus requiring annotated bounding boxes) while others perform multi-label classification by merely modifying the loss function. From a medical imaging standpoint there are advantages and disadvantages to both approaches. Region proposal networks like Faster RCNN generally perform better at both detecting and classifying multiple objects [5]. Also it is useful for doctors be able to see what regions the network identified with a specific classification. However, the main drawback of RCNN compared to modifying the loss function or employing the type of architecture described in paper by Y Wei et al. [7] is that RCNN requires labeled bounding boxes. They are also generally more difficult to train and tweak. In contrast multi-label image classification with an existing single image classification net is relatively easy to setup as usually it only require modifying the loss function. It may also benefit from the larger training data available. For instance, only a small part of the NIH dataset contains bounding box information (approximately 1000 of the 100,000 images) [8] and many other X-Ray imaging datasets such as the Indiana University [10] one do not contain any.

There has been some work that has described diagnosing multiple conditions, however it has generally been limited in scope. “Deep learning with non-medical training used for chest pathology identification,” by Bar et al. [1] did attempt multi-label classification, however it used the convolutional neural network solely as a feature extractor then trained separate SVM classifiers to perform classification on each class [1]. In a more recent article by Cicero et al. used GoogleNet to classify cardiomegaly, pulmonary edema, consolidation, pleural effusion, and pneumothorax [2]. This approach was fairly successful at diagnosing "normal" X-rays but ran into difficulty classifying pleural effusion with cardiomegaly. Finally the authors of the NIH dataset performed benchmarking the ChestX-Ray8 datasets (a smaller sub-sample containing only eight classes) using AlexNet, GoogleNet, VGG-16, and Resnet-50 and achieved AUC values generally between .61-.81 [8].

Despite the widespread success of the R-CNN "family" of models on the Pascal-VOC dataset they have not achieved widespread use in medical imaging applications. This can likely be attributed to the previously mentioned reasons including the complexity of training RCNN (particularly from scratch) and the limited availability of bounding box data for medical data. The winning team, "grt123" in the 2017 Kaggle Data Science demonstrated the success of using (a 3D) F-RCNN at detecting nodules in CT scans [4]. They then fed the region proposal to a separate network which determined malignancy.

Dataset

We use the recently released dataset by the NIH for training and (most of) our testing purposes [5]. The dataset is publicly available on the NIH website. It contains more than 112,120 images from more than 30k patients. The authors of the dataset (Wang et al.) have already performed a number of experiments and created benchmarks on a smaller version of the dataset titled Chest X-Ray8. On the Chest X-Ray14 (the full dataset) they tested a Resnet 50 and achieved AUCS ranging from .6326 to 0.865 [8].

Methodology

Chest X-Rays pose numerous challenges due to their non-uniform size, differences in lighting, and inconsistent number of pixels. In both cases we preprocess the images with scikit-image with all images undergoing histogram equalization. Histogram equalization is a fairly common technique in the processing of X-Ray images which allows better contrast between bone/fluid and air in the lungs. For part one images are then resized to 299x299 for Inception3. We use the Inception3 implementation provided by PyTorch. As of now we are training completely from scratch, but may experiment with using a pre-trained Inception3 in a manner similar to Bar et al in the future [1]. Images are then converted to RGB format in order to have the proper dimensions for Inception3 network. The network architecture for the Inception3 is kept the same with the only change coming in the loss function. We are currently experimenting with several different loss functions. Right now we are using Multi-Label Soft Margin Loss function (sometimes referred to as Sigmoid Cross-entropy loss [3]) since it is already implemented in PyTorch, but we hope to implement the custom loss function discussed in Wang et al soon [8]. This function is popular for multi-label classification problems [3, 8]. We chose Adam as the optimization function due to its ability to minimize the training loss.

In the second case images are left in their original size for F-RCNN. We process the bounding boxes from the separate bounding boxes file. Since this dataset is much more limited we are considering using a pre-trained F-RCNN then fine tuning and possibly using image augmentation. Region proposals are identified using F-RCNN then classified using the Inception3 base network. Since no existing implementations of F-RCNN support Inception3 out of the box we are currently working on modifying the Pytorch implementation[1] made by Ruotian Lou to support it. ResNet-101 is supported, so only minor modifications should be necessary for ResNet50.

Results

As a work in progress we do not have meaningful results at this point. For the results we are planning on using cross validation for training while also maintaining a random holdout of around 1000 images for our test set. We also hope to test parts of the Indiana University [10] and JSRT [9] datasets to determine how well the models perform on other data sources.

1 https://github.com/ruotianluo/pytorch-faster-rcnn

References

[1] Y. Bar, I. Diamant, L. Wolf, and H. Greenspan, “Deep learning with non-medical training used for chest pathology identification,” in SPIE Medical Imaging, pp. 94140V–94140V, International Society for Optics and Photonics, 2015.

[2] M. Cicero , A. Bilbily, Colak E, T. Dowdell, B. Gray, K. Perampaladas et al.: Training and Validating a Deep Convolutional Neural Network for Computer-Aided Detection and Classification of Abnormalities on Frontal Chest Radiographs. Invest Radiol 52:281–287, 2017

[3] Z. Maurer, T. Thapliyal, S. Deseai. Land Cover Classification in the Amazon http://cs231n.stanford.edu/reports/2017/pdfs/917.pdf

[4] grt123 Winning Submission https://github.com/lfz/DSB2017/blob/master/solution-grt123-team.pdf

[5] S. Ren, K. He, R. Girshick, and J. Sun. Faster r-cnn: Towards real-time object detection with region proposal net- works. arXiv preprint arXiv:1506.01497, 2015.

[6] Wang, Jiang, et al. “Cnn-rnn: A unified framework for multi-label image classification.“ Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016. https://www.ics.uci.edu/ yyang8/research/cnn- rnn/cnn-rnn-cvpr2016.pdf

[7] Y. Wei, W. Xia, J. Huang, B. Ni, J. Dong, Y. Zhao, and S. Yan. CNN: single-label to multi-label. arXiv, 2014. https://arxiv.org/pdf/1406.5726.pdf

[8] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, Summers RM. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR 2017,

[9] JSRT http://www.jsrt.or.jp/data/

[10] OpenI University of Indiana Dataset https://openi.nlm.nih.gov/gridquery.php?q=&it=xg&coll=cxr