Sadaf Ashrar, David Robison, Nathan Zencey
DATS 6203 Machine Learning II
Dr. Amir Jafari

Final Project Proposal:

Comparison of Feed Forward Network Architectures for Chest X-ray (CXR) Image Classification

Introduction

Since its invention in 1895 by Wilhelm Conrad Roentgen[1], the x-ray has been synonymous with medical examination. Today, many other radiological exams are available, however, the x-ray remains ubiquitous as one of the most commonly utilized diagnostic tools. In particular, the chest x-ray aides physicians in the diagnosis and management of a variety of lung diseases such as pneumonia and cardiovascular conditions such as heart failure[2].

The interpretation of radiological images often requires the availability of a physician specialist, the radiologist. Recent reports have identified barriers to accessing radiological expertise such as the volume of radiological examinations requiring interpretation. Without accurate interpretation, patients and healthcare professionals struggle to make timely and meaningful care decisions.

As such, radiological interpretation provides a paradigm for a computer vision task. By training a deep learning network for image classification under a supervised learning framework, a computer-aided tool could be developed to assist radiologists. The large-scale implementation of such a tool could ultimately improve access to timely and accurate radiological interpretation[3].

Background

Until recently, the investigation of deep learning frameworks for radiological assistance has been primarily applied to computed tomography (CT) images. The release of the National Institute of Health (NIH) x-ray database, made vastly more labeled data available, reducing research barriers. As such, there have been several recent works on applications of deep learning to x-ray images, which we will use as reference materials.

Wang et al, after using natural language processing to extract labels for x-ray images in the NIH dataset, benchmarked multiple networks using pre-trained Imagenet architectures. They reported that a ResNet architecture had the best overall performance for classification of 8 different conditions[5].

Several research groups have built on the work of Wang et al. One group, using an abnormality detection framework[6], found that GoogLeNet performed best on a subset of 50,000 images. This group performed significant preprocessing such as histogram equalization to increase image contrast, and reported that network performance was not affected by image resizing.

Most recently, a group at Stanford University developed a 121 layer Dense Convolutional Neural Network, achieving superior results for pneumonia detection and F1 scoring across 14 distinct classes[7]. Their model used an Adam optimizer in all layers with parameters pre-initialized from a network trained on ImageNet.

Data

The dataset contains 112,120 x-ray images of 30,805 unique patients, and has been made available by the NIH's Clinical Center. It can be accessed here. The original images are of size 1024 x 1024 x 3. The dataset has been labeled using a novel natural language processing framework to mine 14 disease labels.

The 14 labels are as follows: 1, Atelectasis; 2, Cardiomegaly; 3, Effusion; 4, Infiltration; 5, Mass; 6, Nodule; 7, Pneumonia; 8, Pneumothorax; 9, Consolidation; 10, Edema; 11, Emphysema; 12, Fibrosis; 13, Pleural_Thickening; 14 Hernia

Methods

For our project, we will explore feedforward convolutional neural network architectures for a 14-class classification task using the NIH Chest X-ray database. Our goal will be to investigate the following:

- How is performance affected by changes in network depth?
- How is performance, for a given architecture, affected by changes in distribution and types of transfer function?
- How is performance affected by changes in the distribution and type of pooling layers?
- Is transfer learning using the VGG-16 network developed by Simonyan and Zisserman[8] feasible for x-ray images?

For image processing and network implementation, we will use a combination of tools available in the following Python packages: Keras, Pytorch, sci-kit learn.

During the pre-processing stage, we will reduce the image size from 1024 x 1024 x 3 to 224 x 224 x 3. We will also consider whether to pursue image cropping and flipping as a means of reducing overfitting. Finally, will split the data into train, validation, and test sets with the following ratios: 85%- 10%-5%:

For network training and testing, we will use Pytorch and Keras. We will specifically use Pytorch because of our familiarity with the package, its novel autogradient functionality, and its ability to be modified for use on a GPU. For implementation of a pre-trained network, we will specifically make use of Keras. Keras provides easy implementation of well studied architectures such as the VGG-16 network along with previously initialized network parameters.

We will implement our models in instances hosted by Amazon Web Services (AWS) and Google Cloud, as they allow for easy access to parallelized GPUs.

To assess the performance of our networks, we will use multiple evaluation metrics to understand the behavior of our classifier including:

- Confusion matrices,
- Specificity and sensitivity
- F1 score
- AUC

Learning Outcomes

For this project, we will consider the literature, but also use this as a learning opportunity to become more familiar with several deep learning frameworks. Our goal primary goals are as follows:

- Gain subject matter expertise in a specific type of computer vision classification task, namely x-ray image classification.
- Build intuition for building deep neural network architectures, including implementing data pipelines
- Gain experience evaluating the quality of and trade-offs between different neural network architectures.
- Demonstrate the value of leveraging open source, pre-trained network architectures.

Timeline for Project Completion

Task	***Due Date***
Proposal Submission	4/4/2018
Data preprocessing	4/8/2018
Model prototyping	4/11/2018
Model tuning	4/15/2018
Compiling results	4/20/2018
Submitting report	4/24/2018
Presenting project	4/24/2018

References

1. British Library. Roentgen's discovery of the x-ray. Available at: https://www.bl.uk/learning/cult/bodies/xray/roentgen.html.
2. Mayo Clinic. Chest X-rays. Available at: https://www.mayoclinic.org/tests-procedures/chest-x-rays/about/pac-20393494.
3. Diagnostic Imaging. In Radiology, Turnaround Time is King. Available at: http://www.diagnosticimaging.com/practice-management/radiology-turnaround-time-king.
4. National Institutes of Health. NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community. Available at: https://www.nih.gov/news-events/news-releases/nih-clinical-center-provides-one-largest-publicly-available-chest-x-ray-datasets-scientific-community
5. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, 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, http://openaccess.thecvf.com/content_cvpr_2017/papers/Wang_ChestX-ray8_Hospital-Scale_Chest_CVPR_2017_paper.pdf
6. Tariqul, Mohammad, et al. "Abnormality Detection and Localization in Chest X-Rays Using Deep Convolutional Neural Networks." 27 Sept. 2017, arxiv.org/abs/1705.09850.
7. Rajpurkar, et al. "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning." 25 Dec. 2017, arxiv.org/abs/1711.05225.
8. Simonyan, et al. "Very Deep Convolutional Networks for Large-Scale Image Recognition." 10 Apr. 2015, arxiv.org/abs/1409.1556.