X-Ray Image Analysis Using Convolutional Neural Networks: A Udacity Capstone Project Proposal

By: Justin Merkel

Domain Background

X-Rays are a medical analysis technique in which x-ray waves are sent through the body in order to get an image of what the inside of a person's body looks like. In the past, these images had to be analyzed by a person in order to evaluate their meaning. With the advances in image recognition, computers are able to recognize more and more patterns in human medical data. Image recognition using deep neural networks has recently been applied to medical image data, and these models have already seen some success. [1]

Problem Statement

Doctors are still having to manually evaluate every single x-ray that is taken. This takes not only time but also costs hospitals money. A solid first step towards solving this problem would be a mechanism for flagging x-rays that are very likely healthy vs. those that need a doctor's attention. This would allow doctors to focus more time on the 'unhealthy' x-rays that require more of their attention.

Datasets and Inputs

In order to facilitate researchers in creating more effective machine learning models for x-ray analysis, the U.S. National Institute of Health has recently released a dataset of anonymous chest x-ray images. This dataset provides over 100,000 images of chest x-rays along with some relevant information on the patient in the x-ray such as sex, age, and visit number. There are eight common pathological diseases presented in this dataset: atelectasis, cardiomegaly, effusion, infiltration, mass, nodule, pneumonia, and pneumothorax.

Each image either contains one to many of these diseases or marked as 'no finding'. [2] Any images labeled as 'no finding' will be considered 'healthy', while the rest will be considered 'unhealthy'. The distribution of dataset based on this classification is as follows…

Total number of images: 112,120
Total number of healthy images: 60,361

---

[1] Machine Learning in Medical Imaging. Miles Wernick-Yongyi Yang-Jovan Brankov-Grigori Yourganov-Stephen Strother - https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4220564/

[2] NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community. https://www.nih.gov/news-events/news-releases/nih-clinical-center-provides-one-largest-publicly-available-chest-x-ray-datasets-scientific-community

Total number of unhealthy images: 51,759

Benchmark Model

X-ray image analysis is currently a human problem; and no standard model currently exists for analyzing these images that does not require human involvement. In actuality, however, the benchmark model can be thought of as always guessing that an image is ‘unhealthy’, because that is what hospitals do currently when they treat every x-ray as an image that requires human analysis. For this reason, the benchmark model for this project will be the model that always guesses an image is ‘unhealthy’.

Solution Statement

My proposed solution is to use these images to train a convolutional neural network (CNN) to return a probability that an x-ray image is ‘abnormal’ or having signs of one or more of the eight diseases this dataset contains.

Evaluation Metrics

A percentage of the images will not be trained on the neural network and instead be used as a testing set. After the model is trained, predictions will be compared against the actual labels from the dataset in order to evaluate the performance of the model.

It is important to note that precision should be taken into account but recall should be given much more weight when evaluating models because in a problem like this avoiding false negatives is much more important than avoiding false positives (Note that a ‘positive’ prediction would be a prediction that marks an x-ray as abnormal, or having one of these known diseases). Allowing a false positive through means the doctor has to look at an x-ray that ends up being healthy. A false negative, however, means that we have marked an ‘unhealthy’ x-ray as not important; leading the doctor to ignore a patient that.

To measure recall and precision will be measured with f-beta score, with a beta value much higher than one in order to favor recall more than precision.

Project Design

My goal is to build a convolutional neural network that returns the probability that the x-ray image presents signs of one of the eight diseases the dataset contains. Transfer learning will likely be used as a way to develop a good model without having to start from scratch, but the details of the neural network will be worked out by me during development of the model. As noted in the previous section, this model will be designed to classify an image as either ‘healthy’ or ‘unhealthy’. Success will be evaluated based on the performance of the predictions made on the testing set.

Citations

1. Machine Learning in Medical Imaging. Miles Wernick-Yongyi Yang-Jovan Brankov-Grigori Yourganov-Stephen Strother - https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4220564/
2. NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community. https://www.nih.gov/news-events/news-releases/nih-clinical-center-provides-one-largest-publicly-available-chest-x-ray-datasets-scientific-community