Capstone proposal

Niels Ota

November 19, 2021

1 ‖ Introduction

Large medical datasets allow deep neural nets to achieve state-of-the-art performance on computer vision tasks. In this capstone project, I would like to deploy a vision-transformer model in PyTorch, and deploy it to AWS in the form of an API. In theory, medical professionals could use this API to quickly diagnose medical conditions, without a radiologist having to examine the image. Work by Lian et al., 2021 investigated a more complex variant of this problem, where not only diseases were classified, but also localized, and further investigated co-occurrence of thoracic diseases. The latter points are a step forwards compared to work by Rajpurkar et al., 2017, who solely focused on classification. In that paper, ChestXNet outperformed Stanford radiologists and displayed the potential of deep learning models in medical imaging classification. Since this is the end of the MLE nanodegree I have several specific learning goals (**and perhaps receive feedback on during reviews**)

1. Implementing transformers.
2. Making the implementation of a much deeper neural network tractable by using techniques such as batch normalization, and feasible by leveraging cloud computing.
3. Deploying a trained model using an API.

I expect to need to take much more care whilst optimizing a deep network compared to more shallow networks that I have so far trained locally.

2 ‖ Data

The ChestX-Det consists of 3578 images from NIH ChestX-14, annotated by three board-certified radiologists with 13 common categories of diseases or abnormalities. The goal of this research is to develop a model that can classify images correctly.

3 ‖ Model

3.1 ‖ Model

The model used for classification will be the vision transformer. Although transformers were originally used for NLP tasks, they can also work for computer vision tasks.

3.2 ‖ Metrics

The percentage of correctly classified images is an obvious metric. Depending on class imbalance, it may be appropriate to compare models based on F1 score. Since Rajpurkar et al., 2017 compare using F1 score, to make results comparable, I will also calculate the F1 score on the different diseases using the vision transformer. I will compare my results to that of achieved in Rajpurkar et al., 2017. I do not expect to outperform CheXNet, but I do expect to

3.3 ‖ Outline

To go from data exploration to a deployed model, I will use a various notebooks in SageMaker, each with their own task

1. Data downloading and exploration/engineering: downloading the data from an external source automatically, and engineering it into a format suitable for a deep learning model to use it as input. Furthermore, this notebook's goal is to get a feel for what the dataset looks like. Finally, the data will be uploaded to S3 when it is ready to be used by a model.

2. Implementing a transformer model in PyTorch. Since I have never used a transformer model before, I will write a class implementing the vision transformer and train it on a small portion of the data to ensure everything works properly, before attempting to deploy it

3. Training and deploying a vision transformer model using the AWS managed PyTorch container.

4. (*not a notebook*) Creating an AWS Lambda function and API. The endpoint resulting from the previous step is where the lambda function will send the image to, after first doing any required preprocessing.

References

Lian, J., J. Liu, S. Zhang, K. Gao, X. Liu, D. Zhang, and Y. Yu (2021): *A Structure-Aware Relation Network for Thoracic Diseases Detection and Segmentation.*

Rajpurkar, P., J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng (2017): *CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning.*

Appendix