Deep Learning Final Project Proposal

Team Members

Ambuj Arora, Nathan Marshak, and Jordan Saethre

Project Description

What is the problem? The world is facing an unprecedented crisis due to the COVID-19 pandemic. Being able to test patients for COVID rapidly is crucial to all containment efforts. However, not all countries and municipalities have the resources to rapidly test all patients, especially when the healthcare system is overwhelmed with patients, e.g., during a significant spike in coronavirus cases.

Therefore, imaging has been proposed as a way to complement existing methods for COVID testing, especially in "low-resource" situations [3]. Ultrasound, in particular, is appealing because ultrasound scanners are cheap, portable, readily available, and do not expose the patient to radiation the way that X-ray or CT imaging does [3]. Therefore, there is a demand for automated classification of ultrasound images for the purpose of diagnosing COVID-19.

The goal of this project is to accurately classify ultrasound lungs images into three categories, namely, COVID-19 infected, pneumonia infected or healthy lungs.

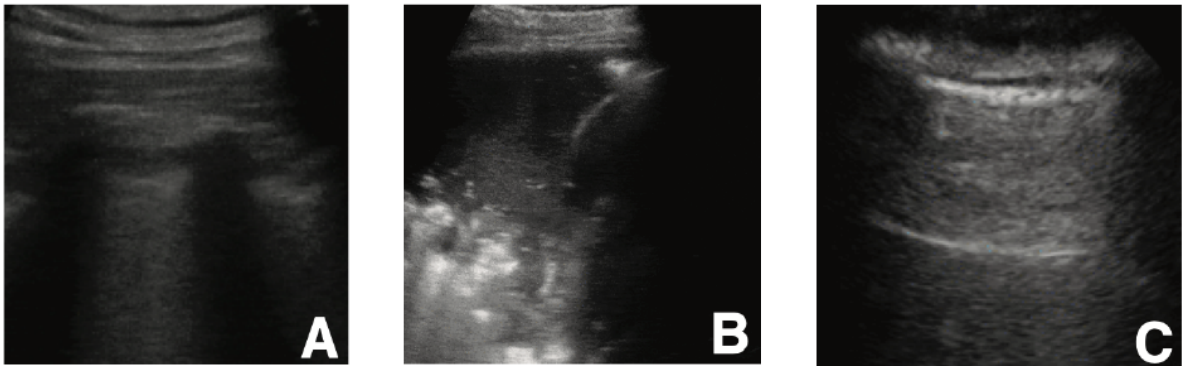

Figure 1: "Example lung ultrasound images. (A): A typical COVID-19 infected lung, showing small subpleural consolidation and pleural irregularities. (B): A pneumonia infected lung, with dynamic air bronchograms surrounded by alveolar consolidation. (C) Healthy lung. The lung is normally aerated with horizontal A-lines." Image and caption from Born et. al. 2020 [1]

Why would deep learning solve this? Deep learning has emerged in recent years as one of the best ways to classify images, as long as there is enough training data. Fortunately (or unfortunately, depending on your perspective), the number of training images for COVID is currently growing as the pandemic continues to spread. Therefore, a deep learning based approach not only has the potential to be effective but would also become more promising in terms of classification performance as time goes on since more training data will become available.

Network Architectures

- The original model was built in TensorFlow and uses the convolutional part of VGG-16 with an additional hidden layer of 64 neurons with ReLU activation, drop out of 0.5 and batch normalization. The output layer used softmax activation. Only the weights of the last three layers were optimized during the training. The authors named this model as "POCOVID-Net" [1].
- We propose to reimplement POCOVID-Net in Pytorch and try other pretrained base models such as (but not limited to):
  - ResNet

  - SqueezeNet
  - GoogLeNet
- For all the base models that we try, we would make some modifications for the models to learn better on the COVID-19 ultrasound dataset. These modifications can be (but not limited to) adding some additional layers, or hyper-tuning the parameters, or using different activation functions.
- In a follow-up paper [2], additional base models were compared to VGG-16. One such model was "NasNet Mobile" which has a fraction of the parameters of VGG-16 and was intended to be a "light weight" version for use on mobile devices. This model underperformed as compared to the VGG-16 base but laid the foundation for further research in coming up with a well performing light weight model. To take this further, we may try to use MobileNet as a base model to achieve better results.

Dataset

Source of the dataset

We will be using the dataset provided by the authors of the original POCOVID-Net paper. This dataset is freely available on GitHub [3].

Description of the dataset

- The dataset consists of images (obtained from videos) which are labelled into the following three categories:
  - COVID-19 infected
  - Pneumonia infected
  - Healthy
- The images and videos are recorded by an **ultrasound** transducer, also called a **probe**, which is a device that produces sound waves that bounce off body tissues and make echoes.
- The linear probe is a higher frequency probe yielding more superficial images.
- Depending on the type of probe, there are two types of data (*paraphrased from the authors [3]*)
  1. Convex Probe
     - 162 videos (46x COVID, 49x bacterial pneumonia, 64x healthy, 3x viral pneumonia)
     - 20 videos from the Butterfly dataset (18 COVID, 2 healthy, see below how to use the provided scripts to process the data)
     - 53 images (18x COVID, 20x bacterial pneumonia, 15x healthy)
  2. Linear Probe
     - 20 videos (6x COVID, 2x bacterial pneumonia, 9x healthy, 3x viral pneumonia)
     - 6 images (4x COVID, 2x bacterial pneumonia)
     - 45 videos of *possible* COVID-19 patients collected in Piacenza at the peak of the crisis in Italy; there were not enough PCR tests available, so the label is not clear. For more information and comments by medical experts, see the metadata sheet or metadata csv.

Train/Validation/Test Split

The data set will be split into 5 folds for cross validation. Frames taken from videos will be included in only one fold per video to ensure that training and testing sets are completely disjoint.

CNN Visualization

In a very recent follow-up paper, the POCOVID-Net authors implemented "Class Activation Maps" a.k.a. CAMs (these are basically saliency maps) for visualizing which parts of an ultrasound image are most important for classification of the lung as either "COVID-19", "Pneumonia", or "Healthy" [2]. We can take this a step further by implementing other visualization techniques, such as:

- Creating "maximal activation" images using gradient ascent.
  - E.g. what lung "looks like" the most COVID-infected lung possible, at least from the perspective of the neural network?
- Optimize in some latent space (e.g. of one of the fully-connected layers) instead of pixel space in order to get "maximal activation" images that are members of the lower-dimensional manifold represented by that latent space.
- Visualize the space of features at the last layer using a dimensionality reduction technique like t-SNE.
- Do feature inversion.

Comparison of CAMs Futhermore, the original the authors do not present a comparison of CAMs for different network architectures. Comparison of the differences in saliency maps between architectures (i.e. are different neural networks looking in different) could be an interesting exercise, *if time allows*.

References

1. Born, J., Brändle, G., Cossio, M., Disdier, M., Goulet, J., Roulin, J. and Wiedemann, N., 2020. POCOVID-Net: automatic detection of COVID-19 from a new lung ultrasound imaging dataset (POCUS). *arXiv preprint arXiv:2004.12084*. https://arxiv.org/abs/2004.12084
2. Born, J., Wiedemann, N., Brändle, G., Buhre, C., Rieck, B. and Borgwardt, K., 2020. Accelerating COVID-19 Differential Diagnosis with Explainable Ultrasound Image Analysis. *arXiv preprint arXiv:2009.06116*. https://arxiv.org/abs/2009.06116
3. https://github.com/jannisborn/covid19_pocus_ultrasound