Project Proposal

Group 61

Students : Emil Elmarsson and Sabrina Chowdhury

16-04-2020

Project title : Automatic detection of COVID-19 from X-ray images

Brief description :

We will focus our project on detecting COVID-19 cases from X-ray scans (we have chosen not to use CT scans as X-ray scans are more widely available even in more low-income areas). We will utilize convolutional neural networks and will base our experiment on the effects of training our deep network using transfer learning/fine tuning (as the amount of COVID-19 images is so small). The transfer learning will be implemented in the way that we train a model on X-rays from healthy individuals, individuals with pneumonia and individuals diagnosed with SARS (if we find enough data on SARS). The final model parameters will then (using transfer learning) be updated through training with the COVID X-ray dataset. We will experiment with different parameter settings to make both the first model (before transfer learning) and the second model as accurate as possible.

Some papers that inspired us :
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7118364/
https://arxiv.org/pdf/2003.10849

Dataset used for training, validation and testing :

We have found several COVID datasets, but are right now unsure as to whether they contain the same picture, and as we do not want duplicates we will have to look into how to deal with that. We have two options when it comes to the first network we will create. If we find enough SARS images (or even MERS, which we have had no success with so far) we will incorporate SARS images in our model, but if we find that the images we have is not enough we will only utilize the NIH dataset in the first network (before fine tuning).

NIH Chest X-rays: https://www.kaggle.com/nih-chest-xrays/data
Covid-19, pneumonia and normal X-rays:
https://www.kaggle.com/tawsifurrahman/covid19-radiography-database
Covid-19 (majority, also contains SARS but not that many) - X-rays as well as CT-scans
https://www.kaggle.com/bachrr/covid-chest-xray
Covid-19 X-rays: https://www.kaggle.com/andrewmvd/convid19-X-rays

The deep learning software packages we will use :
We will be using Python and Tensorflow.

How much of the software implementation we will write :
We will write all of the software implementation.

The initial set of experiments we will run :
We will initially look at the performance of the first network before fine tuning. We will thus be looking at how well the network distinguishes between X-ray scans of healthy lungs and lungs of patients with pneumonia - we will include a third group/label for SARS if we find enough data.

How we will measure the success of our project :
We will measure the success of our project by looking at how well the final network is able to distinguish between X-ray scans representing healthy lungs, pneumonia without a COVID-19 diagnosis, pneumonia with a COVID-19 diagnosis and pneumonia with a SARS diagnosis (if we find more SARS images).

The skills knowledge each group member aims at acquiring from completing the group project :
Initially we were three people but the third person dropped the course. We hope that we can still manage the scope of this project.
- Sabrina : learning how to implement CNNs and transfer learning, also conducting a scientific deep learning experiment - and contributing to the ongoing COVID-19 research
- Emil : basically what Sabrina said; learning more about CNNs and the finetuning of deeper networks, as well as trying out Tensorflow.

The grade we are aiming for : A