

Project Deliverable 2

1. Problem statement

I hope to classify a set of chest x-rays into 15 common disease categories using an inception network.

2. Data Preprocessing

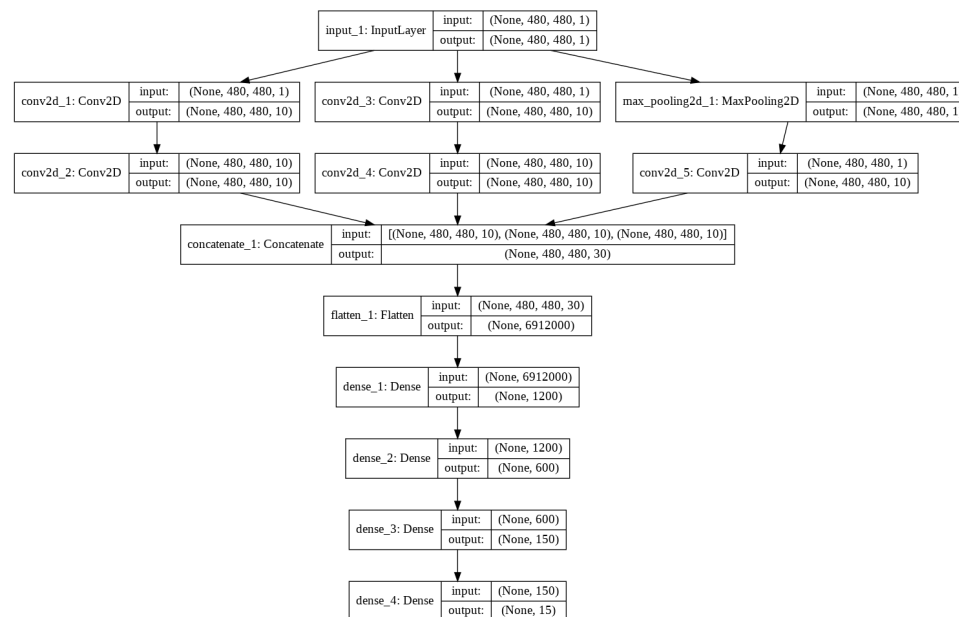
I am currently uploading images from my computer and resizing them manually. However, with an HD dataset, this is suboptimal. In the future, I look to download the images straight from Kaggle, invest in some CGP and use python to preprocess the images. I am also not currently using patient characteristics or the bounding box in this preliminary model. I discussed clustering and relabelling of data points but this has not yet been implemented. I noticed some x-rays have multiple diagnoses. For this model, I am training the model to fit the primary diagnosis but may look to include multiple classifications in the future.

3. Machine learning model

As a preliminary model, I am using a simple inception network to get a feel for what the accuracy is. The model structure is summarized in the following:

4. Preliminary results

Although the model seems to have working structure, I don't have access to enough computing power to fit it to the data and don't yet have preliminary results.



Kusha Sareen

5. Next Steps

In the future, I hope to get access to more computing power so that I can run the model and implement some of the changes listed in the preprocessing section.

Source code:

<https://colab.research.google.com/drive/19r3QtsLvJZw9c8Z68iQeRx4Q55QOgK1m>