Project proposal for Degree Project in Computer Science and Communication

Tao Wang
taowan@kth.se

December 14, 2017

1 PRELIMINARY THESIS TITLE

Adversarial Domain Adaptation for Screening Mammograms

2 BACKGROUND/CONDITIONS

In Sweden, women of age between of 40 and 74 go through regular screening of their breasts every 18-24 months. The screening mainly involves obtaining a mammogram and having radiologists analyze them to detect any sign of breast cancer. This screening is vital and in one study has reduced mortality by 30%. In Sweden and generally world-wide, we have shortage of radiologists for screening in such a scale. Thus, a computer-assisted analysis using machine learning tools is highly relevant not only to reduce the burden on the radiologists but also increase the sensitivity of the analysis.

However, we are facing a big challenge while training a machine learning model, which is the appearance of the mammograms coming from different hospitals. Different hospitals have different imaging equipment and are manually tuned with different settings. So it is hard to learn a single model that can make use of images coming from different places.

3 RESEARCH QUESTION

3.1 The Question

In this project we plan to use adversarial domain adaptation to transfer mammograms from different datasets into a standard format, so that mammograms from different sources could be classified by one classifier.

3.2 The Research Area

From the perspective of application view, this research will be mainly used in the medical area. And from the perspective of academical view, this research focus on domain adaption area.

3.3 Connection to Research/Development

As we have discussed before, the mammograms recognition is facing a pretty big problem. It is a waste to training a recognition model for each mammograms source. And sharing a recognition model between different sources could not get a ideal result. So what we want to do in this project is to apply

adversarial domain adaptation into the recognition model. So that we could use a domain adaptation network to transfer mammograms from different datasets into a standard format.

3.4 Hypothesis

The hypothesis of this thesis is that it is possible to find a adversarial network which is suitable for all kind of mammograms and the classifier's accuracy after transfer still keeps at a high level.

3.5 Evaluation

The model we proposed will be tested by several individual datasets which do not get involved in training. And also I will run these datasets on some other models. These classified and marked results will then be analyzed and a conclusion is drawn based on these. And we will define several evaluation metrics, such as, mean intersection-over-union (mIoU), frequency weighted intersection-over-union (fwIoU), pixel accuracy, etc. The result will be quantified by these metrics.

4 BACKGROUND OF THE DEGREE PROJECT STUDENT

I have taken a lot of courses about machine learning and artificial intelligence: Artificial Intelligence (DD2380), Machine Learning (DD2431), Artificial Neural Networks and Other Learning Systems (DD2432), Statistical Methods in Applied Computer Science (DD2447), Deep Learning in Data Science (DD2424), Machine Learning, Advanced Course (DD2434). Also I have taken Data-Intensive Computing (ID2221) to get some experience about dealing with big data. What's more, since this thesis project is based on generative adversarial networks(GAN) and in Deep Learning in Data Science (DD2424) course, my course project is about implementing a Deep Convolutional Generative Adversarial Networks(DCGAN), which gave me a basic understand of adversarial network.

5 LIMITS/RESOURCES

5.1 Resources

Here is the list of the resources expected to be needed to solve the problem.

- Hardware, a computer equipped with a graphics card
- Software, matlab, deep learning frameworks, such as tensorflow, theano, etc
- Data, mammography datasets, such as CBIS-DDSM, INBREAST, dream pilot dataset, etc
- Perseverance, because it's tough to write a thesis

6 ELIGIBILITY AND STUDY PLANNING

6.1 Eligibility

I assure that I have got a bachelor's degree at Zhejiang University and I have already got 67.0hp which includes a course in scientific theory and method(Introduction to the Philosophy of Science and Research Methodology for Computer Scientists (DA2210)).

6.2 Study planning

I will have 126.5hp before applying for the master's degree. So I don't need to choose other courses.