Master Thesis Project in Applied Mathematics – Planning Report

1 Preliminary Title

This thesis will concern the design of one or multiple novel cost functions that will be employed in the training of a GAN designed for synthesizing MR images of brains, and an evaluation of the performance of such synthetic datasets when used to train (or augment the training of) current state of the art networks on brain tumor segmentation of (real) MR images of brains. A preliminary title could be "GAN-based synthetic brain MR image synthesization using a [X] cost function" (similiar to the title of the article "GAN-based synthetic brain MR image generation" by Han et al).

2 Problem description

Machine learning applications within neuroimaging suffer from a lack of data; there are few open datasets available with more than 100 subjects, since legal reasons prevent the free distribution of patient data. Naturally, this impedes the training of even the best classifiers. One way to circumvent this issue could be to use GANs to generate synthetic datasets, and in turn use them to train said classifiers without involving new patients or breaching data protection laws.

GANs (generative adversarial networks) were proposed in 2014 as a technique for generating highly realistic data given a training set from the desired distribution, and has since proven useful in a variety of applications. The aim of this thesis project is to examine the usefulness of GANs when applied to the process of synthesizing a dataset of brain MR images together with their corresponding binary image mask (a matrix which labels each pixel in the MR image as part of a tumor or non-tumor area), and to evaluate how much they improve (if at all) the accuracy of a brain tumor segmentation network when trained on this synthetic dataset. This work will emphasize the mathematics of the cost function used to train the GAN, and will involve the comparison of different known cost functions as well as the design of at least one novel cost function which will also be included in this comparison. The following questions will be the focus of this thesis project:

- Can GANs be used to concurrently synthesize realistic MR images and their corresponding brain tumor segmentation masks?
- GANs are normally trained with an adversarial loss function. What is the best loss function for training a GAN that will concurrently synthesize an MR image (where the values could range from 0 - 4095), and a binary mask (where all values are 0 or 1)?
- How well will a segmentation network perform on real data, if it has only been trained on synthetic data?

Figure 1 shows an example of a (real) MR image of a brain, along with its tumor segmentation.

3 Approach

Before any actual work begins, I will spend two weeks familiarizing myself with basic deep learning techniques and practices by reading some material and completing laborations that have been provided to me by my supervisor. I feel that this is a necessary use of the time I have been given, since I am so far inexperienced in the field of deep learning (beyond knowledge of the theoretical fundamentals), and will need some practice to familiarize myself with one of the standard programming libraries that are available in this field.

When this is done, I will begin the literature study, which I expect will require at least three weeks. Each week, I will narrow down the specific aim of the project, hopefully having chiseled out a clear and concrete approach at the end of the third week. So far, I have gathered the set of papers in section 4, and will continually branch out from these and add new items to the reading list as I think of new questions and concepts to investigate. During this third week, I will also begin brainstorming ideas for the cost function(s) I will design, which hopefully will help influence and crystallize the final aim and strategy of this project.

Per my supervisor's suggestion, I will give myself a head start on the implementation phase (de-

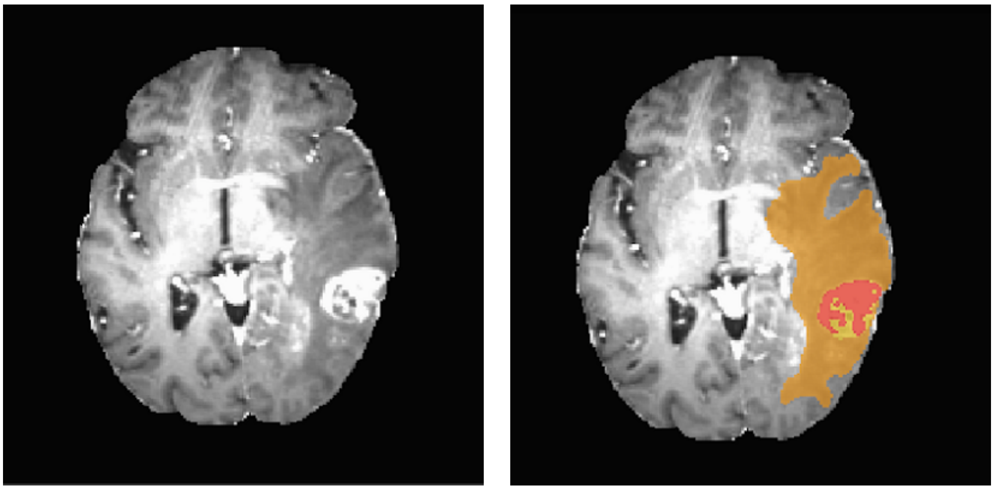

Figure 1: MR image of a brain (left) and its tumor segmentation (right)

tails will follow) by setting up and training the networks (segmentation and/or GAN) during the literature phase. This will streamline the process since I can read the articles while the networks are being trained.

The eleven weeks following the literature study will be fully dedicated towards the implementation and evaluation of the approach that was decided upon in the previous phase. In parallel with the technical implementation, I will work on writing the final report (namely, the literature study and the parts of the implementation that have so far been completed). The reason for this is the fact that the networks that will be used are expected to require hours (or even days) of training, which can be performed simultaneously as I write the report.

Since I haven't begun the literature study yet, it will be difficult for me to outline even a preliminary detailed plan of the specific work that will be done during these eleven weeks; but I will try:

Week 1: I will set up a work environment and a Git repository. I will then import or implement the segmentation network, and in order to confirm that it's implemented correctly, I will train it on a dataset used by the authors of the paper I'm referring to, in an attempt to replicate their test accuracy. When I've succeeded, I will import our own dataset of MR images and train the network on it. Finally, if time permits, I will tune the hyperparameters as best as I can on this dataset by performing a hyperparameter search. I will note the validation accuracy that is achieved in this stage.

Week 2: I will set up a GAN designed for generating images from noise (following instructions and settings from some appropriate paper or code resource) and train it on a known dataset as was done previously, to make sure that everything is implemented correctly and that my outputs look reasonable (since training GANs takes such a long time, this will most likely be done on a heavily downsampled or truncated version of this dataset).

Week 3: When I've convinced myself that the GAN I've set up is alive and healthy, I will try generating images on the MR dataset (since this is expected to require a lot of time, I will start or continue to write my report in parallel). The objective is to generate the MR images and their corresponding tumor masks concurrently. There are at least two ways to do this:

1. Synthesize the MR image and tumor mask simultaneously, as two separate channels in one image tensor.
2. Synthesize the tumor mask first, and from this synthesize the corresponding MR image.

Because it feels like the simpler choice, I will initially proceed with method 1, and return to the second method at a later stage if there is time left over.

Week 4: After hopefully having obtained a synthetic dataset that looks reasonable visually, I will feed it to the segmentation network and observe what happens to the previously obtained validation accuracy. Henceforth, the training will be performed using synthetic images only, and the validation/testing using real images only. If time permits, I may try mixing real and synthetic images in the training phase (while still validating with real images).

Week 5: I will retrain the GAN using the different known cost functions I've decided to evaluate, and redo the process of the previous week. If possible, this process (and similiar ones) will be parallelized by utilizing multiple GPU:s to train multiple instances of the network simultaneously (each using its own cost function).

Week 6: I will implement my own cost function(s), generate new images and redo the training of the segmentation network. I will dedicate the rest of this week towards experimenting and tinkering with this cost function, to see if it can be improved.

Week 7: I will continue experimenting with my cost function(s), and see if any improvements can be made. At the end of this week, I expect all (or at least most) of the code to be finalized and ready for the final evaluation.

Week 8-9: I will perform the final training of the GAN, gather all of the synthetic datasets I have obtained and use them to train different instances of the segmentation network (like in Week 1, but using the different training sets). I will tune the hyperparameters to each specific dataset by performing a hyperparameter search in every training instance.

Week 10-11: I will validate all of my networks on a (previously unused) test set of real images, and gather the final results of every configuration. I will interpret and analyze these final results, and prepare them for the final report.

If there happens to be enough time left (which I'm doubtful will happen, but you never know), I will try Method 2 (as described in Week 3). The following is a preliminary plan for the work involved in doing so:

Bonus Week 1: I will try the second method for generating training data, i.e. by first generating the binary masks and then subsequently "translating" them into MR images. The first step in this process is to import or implement a network that is designed to "translate" between different types of image data (e.g., CycleGAN or pix2pix). This process will be carried out just like the one in Week 2.

Bonus Week 2: I will train the "translation" network to generate MR images from the previously generated image masks, and use it to obtain a dataset of masks + images. This will be done by first training the network on the real dataset (pairs of MR images and corresponding masks), and then feeding it the synthetic image masks that were generated from noise in Method 1 (perhaps, if time is scarce, I will do this using only one of the many generated training sets).

Bonus Weeks 3-4: I will train the segmentation network with these newly obtained datasets like in Week 8-9. Finally, I will validate and analyze the final results on a test set, and prepare them for use in the report, like in Week 10-11.

4 Literature Study

So far, I have gathered the following resources as the basis for the literature study:

Papers:

- U-Net: Convolutional Networks for Biomedical Image Segmentation, Ronneberger et al, 2015
- Generative Adversarial Nets, Goodfellow et al, 2014
- Generative Adversarial Network in Medical Imaging: A Review, Yi et al, 2019
- Progressive Growing of GANs for Improved Quality, Stability, and Variation, Karras et al, 2018
- Feeding the zombies: Synthesizing brain volumes using a 3D progressive growing GAN, Eklund A, 2020
- GAN-based synthetic brain MR image generation, Han et al, 2018
- The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS): A Review, Menze et al, 2015
- Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, Zhu et al, 2018

- Image-to-Image Translation with Conditional Adversarial Networks, Isola et al, 2018

Book(s):

- Deep Learning, Goodfellow et al, 2016

Link(s):

- Hyperparameter Optimization with Keras, Mikko, 2018

5 Time Plan

The table below will outline a preliminary time plan for the project. Weeks 6 to 21 will consist of the activities described in section 3 (with a half-time meeting in week 15). Week 23 will be spent putting the finishing touches on the report and a thorough proof-reading of it. In week 24, I will prepare my presentation, as well as my opposition. In week 24, I will give my presentation and opposition, respectively (perhaps after continuing the work from last week). Week 25 is an empty week that has not yet been allocated to any specific task, and that can be used in case I start to fall behind schedule or if I have completed the tasks I have planned for and want to continue or expand on them in any way (if used, it will most likely be moved back and used in the implementation phase). I will be working full-time on the thesis.

Week #	Activity
6	Deep learning practice
7	Deep learning practice
8	Literature study (& Preliminary implementation)
9	Literature study (& Preliminary implementation)
10	Literature study (& Preliminary implementation)
11	Implementation (& Report writing)
12	Implementation (& Report writing)
13	Implementation (& Report writing)
14	Implementation (& Report writing)
15	Half-time meeting & Implementation (& Report writing)
16	Implementation (& Report writing)
17	Implementation (& Report writing)
18	Final training & Evaluation (& Report writing)
19	Final training & Evaluation (& Report writing)
20	Evaluation & Analysis (& Report writing)
21	Evaluation & Analysis (& Report writing)
22	Final report writing & Proofreading
23	Preparation of presentation and opposition
24	Further preparation & Presentation & Opposition
25	Empty week (to be used if needed)