

COMPUTER GENERATIVE METHOD ON BRAIN TUMOR SEGMENTATION IN MRI IMAGES

By

Yanye Li

Senior Thesis in Electrical Engineering

University of Illinois at Urbana-Champaign

Advisor: Zhi-Pei Liang

May 2020

Abstract

Computer generative method has been used for long time in brain tumor segmentation tasks on magnetic resonance images. The popularity of machine learning also prompts people to explore the use of generative methods to better train their segmentation models. At the early stage, brain tumor segmentation competitions like BraTS 2012 used computer synthetic MR images with tumor to solve the lack of enough data in the training set, and now, with the rise of computer generative models in deep learning, more researchers have started to work on this track to find a better solution for the task. This thesis addresses the implementation and analysis of some existing methods, specifically a tumor synthetic tool called TumorSim and a competition winning deep learning model that incorporates variational auto-encoder as a generative model. This thesis also reports on an experiment that uses imperfect segmented tumors from simple models as the input to a generative adversarial network to generate a better result.

Subject Keywords: magnetic resonance imaging; biomedical imaging; signal processing; machine learning; generative model

Contents

1. Introduction.....	1
2. Literature Review	3
2.1 TumorSim	4
2.2 Variational Auto-encoder	6
2.3 cGAN to Correct Imperfect Result.....	8
3. Description of Research Results	10
3.1 TumorSim	10
3.2 Variational Auto-encoder	12
3.3 cGAN to Correct Imperfect Result.....	14
4. Conclusion and Future Work.....	16
References.....	18

1. Introduction

Magnetic resonance imaging (MRI) is a noninvasive medical imaging technique that has a wide clinical use as it generates images with high spatial resolution and high contrast between different tissues. Because of these advantages, most of the research on brain tumors uses MR images [1]. MRI itself also has different modalities such as T1-weighted, T2 weighted images and Fluid-Attenuated Inversion Recovery (FLAIR). Different modalities also contain different information, which could be used in combination during the research process.

Brain tumor segmentation is a process of separating the pathological tumor from normal brain tissues. It is largely performed on the outputs of magnetic resonance imaging, and a good segmentation map can be very helpful in offering diagnosis and pathological information of the patients. Despite the evolution of MRI techniques and the improvement of MR image quality, brain tumor segmentation remains challenging because of the irregular forms of tumors and confusing boundaries between tumors and normal tissue cells [1]. Another major problem to this task is that human segmentation process often takes a large amount of time. The completion of a thorough brain segmentation map may take a radiologist from dozens of minutes to an hour or two, depending on the skill of the radiologist.

Scientists have been exploring automated brain segmentation with computers for long time. From normal tissue segmentation to tumor segmentation, these researches are aimed to automate this time-consuming process efficiently and to achieve accurate and unbiased results. At the early stage, edge and corner-based segmentation is very dominant but the performance of these methods suffers from the artifacts and noises, which are very common in MR images. Gradually, some model-based segmentation methods that exploit statistical representation of data have emerged, but they still fail to generate satisfying result for complex structures [2]. Recently, deep learning has shown its strong performance in

instance segmentation and more people start to use it on brain tumor segmentation. To gauge the state-of-the-art methods for the segmentation of brain tumors, Multi-modal Brain Tumor Segmentation Challenge (BraTS) is organized in conjunction with the MICCAI conference [3, 4]. Generative modeling is one of the trends in this competition. In general, generative methods include all the strategies that facilitate learning by generating data. Even though different scientists aim to improve this process by different ways, for example some use generative method to create new data cases while others use generative method to capture the distribution of existing cases, the fundamental idea behind these methods is the same: discriminative method is insufficient to recover all the information needed in a perfect segmentation job [5]. This thesis will present three generative methods, including two that have gained public acknowledgement and one that the author has been working on. These methods are: TumorSim by Marcel Prastawa and Guido Gerig from the University of Utah and Elizabeth Bullitt from UNC at Chapel Hill [6], 3D MRI brain tumor segmentation using autoencoder regularization by Andriy Myronenko from NVIDIA [7], and using GAN to correct imperfect segmentation map. While TumorSim helps to solve the lack of training data by generating synthetic MR images with tumors, the following two aim to learn the distribution of data by incorporating different generative models into neural networks.

2. Literature Review

Computer generative methods can be applied in many aspects to improve the efficiency of machine learning in biomedical imaging problems. On the one hand, generative methods are used to perform data augmentation, which often can be a vital factor in the success of learning, because the relative short of training data is common in this area. Previous research on using generative adversarial network (GAN) based model to synthesize medical images as training data has proved successful in liver lesion classification [8]. Brain tumor segmentation competition BraTS also uses synthetic MRI images with tumors starting in 2012. On the other hand, generative methods can be incorporated into a neural network to improve accuracy. Many works on this track need the use of realistic generative models, as figure 1 shows, to catch the data distribution. Here I want to present two methods, variational auto-encoder and GAN, and both are widely recognized in instance segmentation task and now people start to land them on brain tumor segmentation.

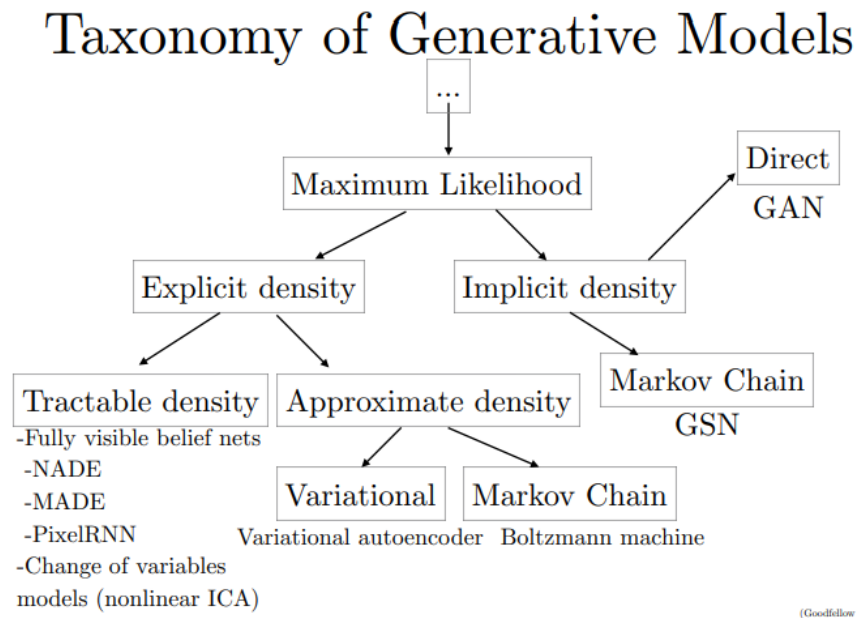


Figure 1: A summary of generative models by Ian Goodfellow [9]

2.1 TumorSim

BraTS 2012 offers an image dataset for competitors to have experiments on. There are 130 cases of the brain with a tumor in total and 65 of them are synthetic. These images were generated using the TumorSim software, a cross-platform simulation tool that combines physical tumor models and statistical models to generate the synthetic MR images [10].

The process of generating an MR image with a tumor in TumorSim takes two major steps: forming a tumor and synthesizing the MR image.

The tumor is formed through four sequential processes. First TumorSim will simulate the deformation that is due to tumor mass effect. The author of the software claims that his work provides an example for tumor mass effect that likely represents metastatic lesions or small glioblastomas, which tend to be ring-enhancing [11]. Second some modification of diffusion tensor MRI is needed following the finite strain reorientation strategy proposed by Daniel Alexander [12]. This is to get the properties of white matter fibers and thus simulating tumor and edema infiltration. Then TumorSim will simulate the proposed tumor's infiltration process using a reaction-diffusion model guided by Clatz's model of tumor growth [13]. Finally, it will compute the displacement that is due to the infiltration of brain tissue and the mass effect of edema. The whole process is summarized as in figure 2.

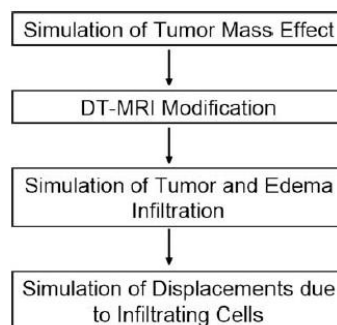


Figure 2. Construction of a tumor [11]

After the construction of the tumor, there are two more processes that carry out a synthetic MR image: contrast agent accumulation and texture synthesis. Contrast agent accumulation happens in real MR scans where biological processes such as blood flow create inconsistencies in contrast. In order to simulate this inconsistency, the software studies the reaction-diffusion equation of the contrast agent and applies this pattern to synthetic images [11]. Texture synthesis is to generate an intensity pattern for different anatomical structures through drawing from a probability distribution created by real MR images. This part follows a training that creates the probability distribution of each region being white matter, gray matter, or different parts of the tumor and brain. Figure 3 indicates the flow diagram of the whole process.

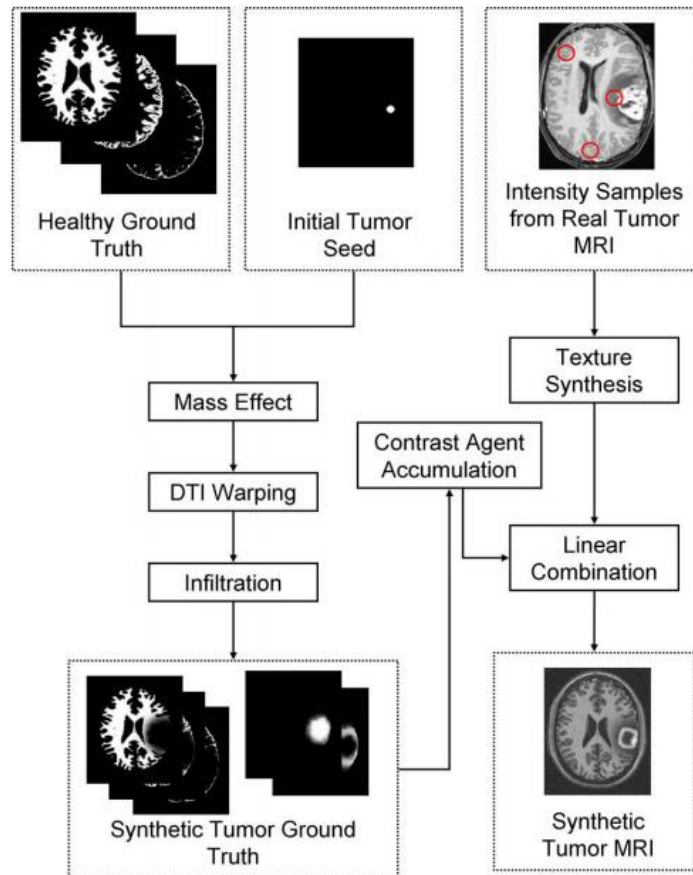


Figure 3: Flow Diagram of TumorSim [11]

TumorSim’s aim is to generate a database that has similar challenges to brain tumor segmentation of real brain scans and is not to create a database that is indistinguishable from real brain scans. So, the focus of it is to simulate the difficulty and have certain reproducibility.

2.2 Variational Auto-encoder

Variational auto-encoder (VAE) is a powerful generative model that is widely used in synthesizing images and music. The idea behind VAE is that when generating outputs, we not only want the output similar to our training data but also want to explore the variations on these data. VAE provides guidance that lets the model explore in a specific direction instead of randomly walking around [14]. It can be defined as an autoencoder that ensures a distribution over latent space in the training process has good properties [15]. This is achieved by outputting a vector of means and a vector of standard deviation in addition to an encoded feature vector at the end of the encoder, as in figure 4. Thus, we could have the distribution of sampled data.

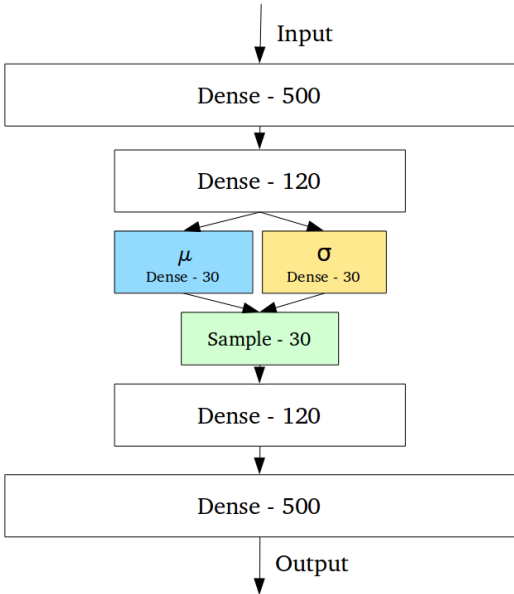


Figure 4: VAE [14]

Andriy Myronenko from NVIDIA combines a standard encoder-decoder based neural network and a VAE to bring up a new segmentation approach. He added a VAE branch at the end of the encoder to reconstruct the original image and thus providing additional regularization to the encoder part and better grouping the features of the encoder. The VAE branch will create a Gaussian distribution and a sample is drawn and reconstructed into the original input [7]. Figure 5 shows the architecture of the model.

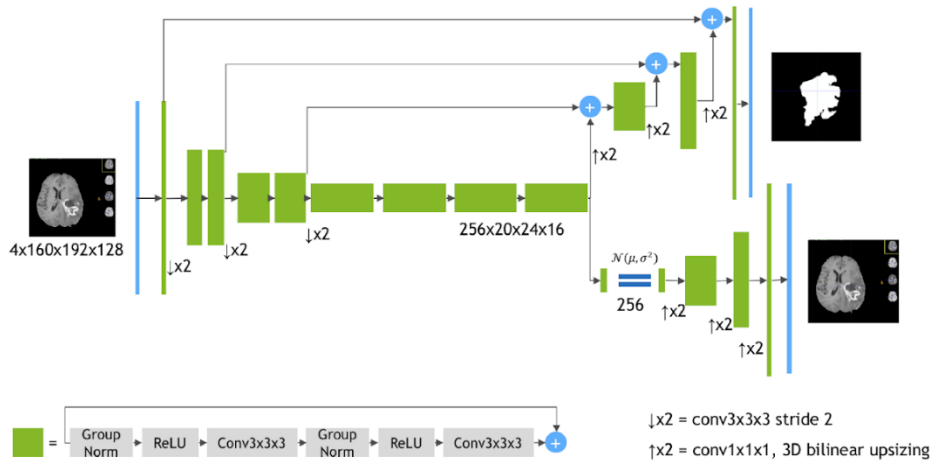


Figure 5: Scheme of the network that incorporates VAE [16]

The network also requires a loss function that includes a VAE penalty. Specifically, the total loss function is:

$$L = L_{dice} + 0.1 * L_{L2} + 0.1 * L_{KL} \quad (2.1)$$

Where L_{dice} is a loss that applied to prediction P_{pred} and ground truth P_{truth} :

$$L_{dice} = \frac{2 * \sum p_{truth} * p_{pred}}{\sum p_{truth}^2 + \sum p_{pred}^2 + \epsilon} \quad (2.2)$$

And L_{L2} is a loss applied to VAE output I_{pred} and original input I_{input} :

$$L_{L2} = \left\| I_{input} - I_{pred} \right\|_2^2 \quad (2.3)$$

And L_{KL} is KL divergence defined as:

$$L_{KL} = \frac{\sum u^2 + \sigma^2 - \log \sigma^2 - 1}{N} \quad (2.4)$$

The author wins BraTS 2018 competition with an ensemble of 10 of VAE models. His winning dice scores of 0.9100 on whole tumor, 0.8668 on tumor core, and 0.8233 on enhanced tumor core [16].

The dice score is defined as:

$$D = \frac{2 * (P_{truth} \cap P_{pred})}{|P_{truth}| + |P_{pred}|} \quad (2.5)$$

2.3 cGAN to Correct Imperfect Result

Generative adversarial nets (GAN), proposed by Ian Goodfellow, is a framework that simultaneously trains two models, a discriminative and a generative model, to generate images [16]. Conditional GAN (cGAN) is an extension of GAN that allows the user to assign a specific condition or characteristic to generate specific fake samples [17]. Inspired by cGAN, we assume that any prediction segmentation from computer models are imperfect. Since we already have the relatively perfect segmentation mask from radiologists, we can use a generative model to receive the imperfect results as features and output a more accurate mask. Pix2pix is a type of cGAN model that has a variety of applications and is mainly used for image translation [18]. We are making efforts to implement it in this scenario. Figure 6 should give you some ideas of what pix2pix can be applied to.

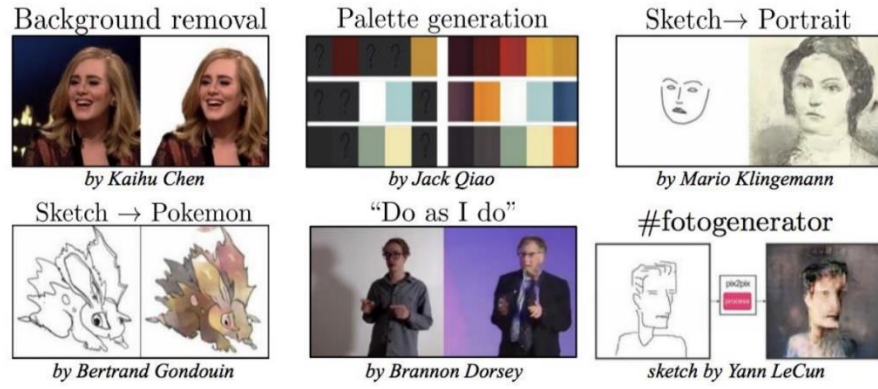


Figure 6: Applications of pix2pix [19]

3. Implementation and Experiment

3.1 TumorSim

We downloaded the TumorSim software and tried to create some synthetic images with a tumor on our own.

TumorSim requires an input of an integer map and four float maps. The integer map indicates a segmented normal tissue brain that includes background, white matter, gray matter, CSF and falx cerebri. Four float model voxel maps of CSF, gray matter, white matter and vessels indicate the proportion of each type present in that voxel and their ranges are 0 to 1 [11]. Besides these inputs, the software will also need a diffusion tensor MR image. TumorSim's developer offers 5 sample packed inputs (essentially 5 brains) and the user can also get these data from online brain databases such as Brainweb [20], another tool that synthesizes normal brain MR image.

We also need to give the software an initial seed of the tumor. This seed is simply created by labeling pixels true in a zeros array that has the same size as the input maps. The size of the seed will also affect the growth of the tumor. For example, in our first experiment we set up the initial seed to be a pixel and the result after a standard number of iterations only shows edema. An initial seed with a size of 10x10x10 or larger will be good enough to produce realistic tumors. Each simulation takes around 2-4 hours and will generate a 3D MR image sequence. Some sample experiment results are displayed in figure 7-9.

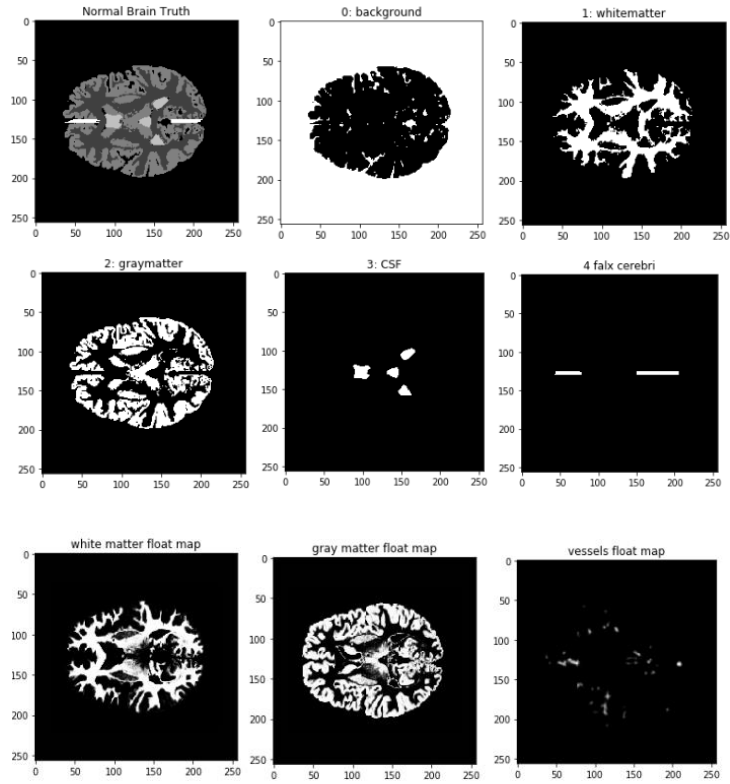


Figure 7: Inputs to the TumorSim. The first two rows show the integer segmentation maps and the last row shows float maps.

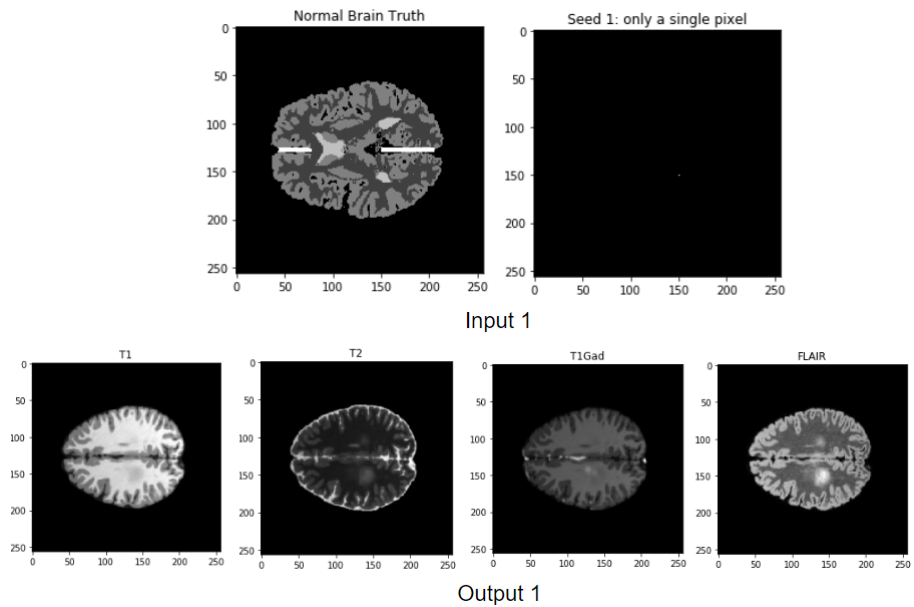


Figure 8: Output of TumorSim when given a single pixel seed

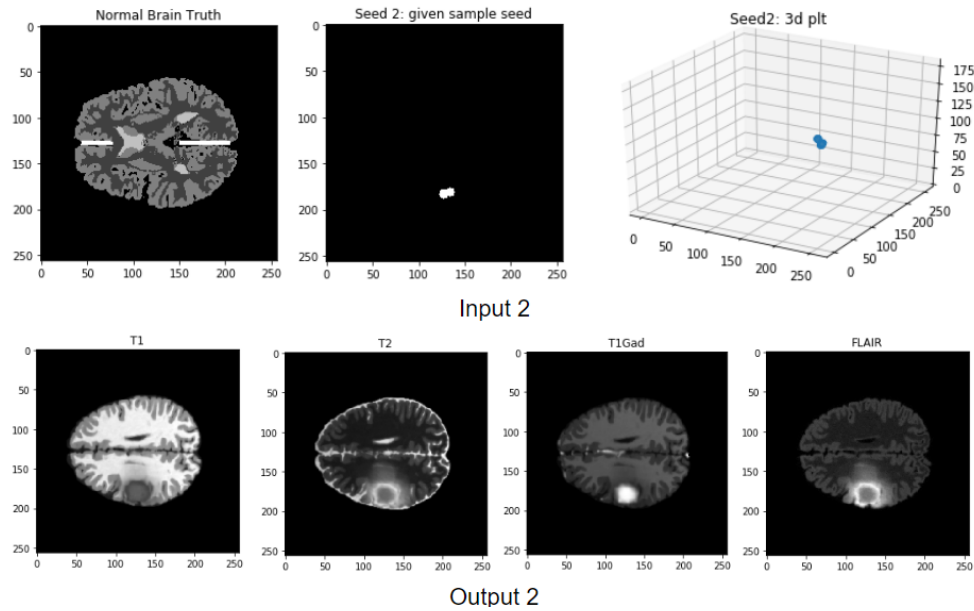


Figure 9: Output of TumorSim when given a more appropriate seed

3.2 Variational Auto-encoder

We also experimented on the encoder-decoder model with VAE on BraTS 2019 dataset. Due to hardware limitations, we resize the input to 80x112x112 instead of the size of 160x192x128 used by the paper's author [16].

Despite the compromise in data, the model still shows very strong performance with a dice score of over 0.83 on whole tumor after 50 epochs. As a reference, the author reached a dice score of 0.91 after 200 epochs. Training 50 epochs takes around 12 hours on Colab. The result of a batch (batch size is 1) is shown in figure 10.

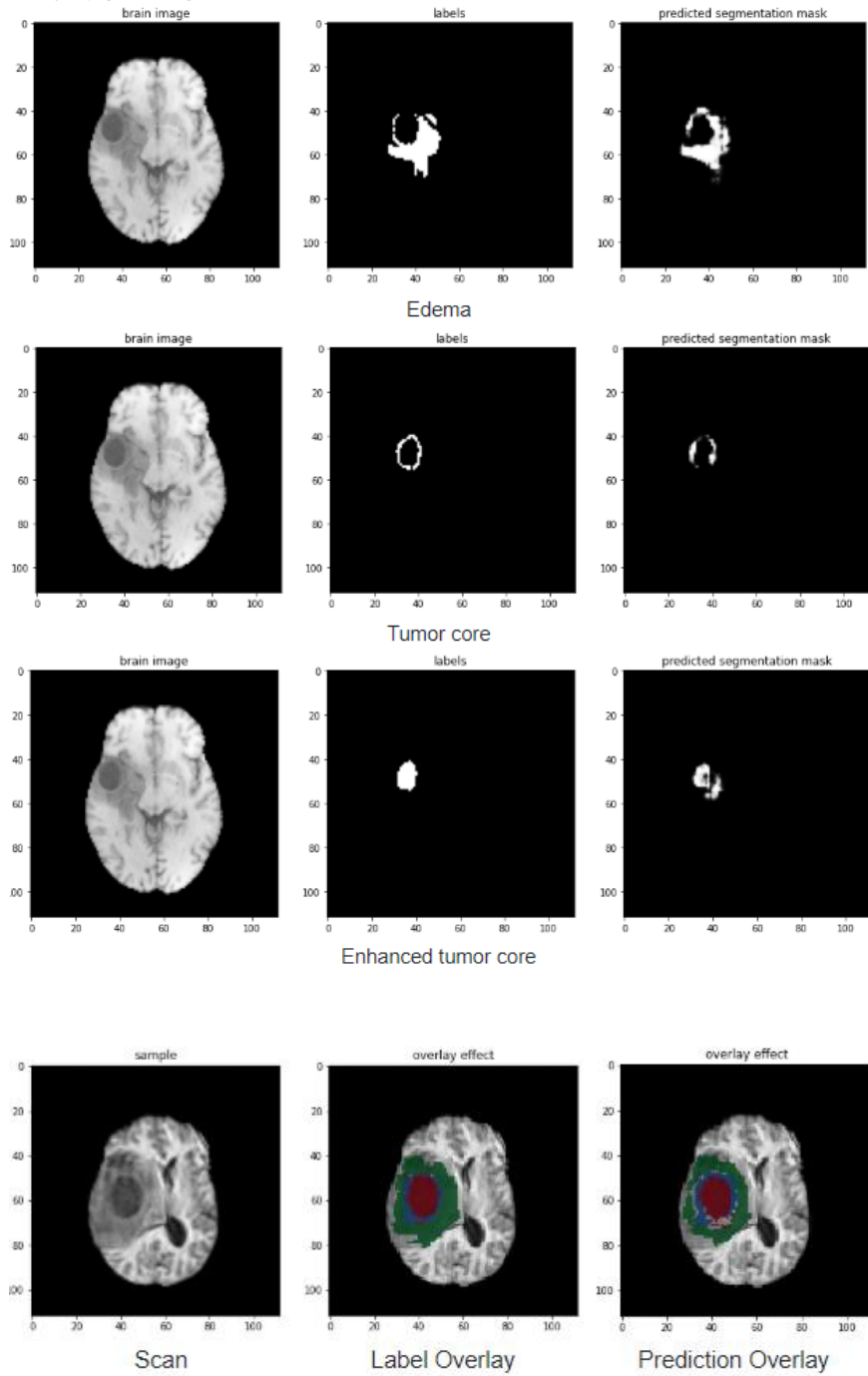


Figure 10: Prediction result of VAE model. The first three rows show segmentation truth and prediction for each part of the tumor and the last row shows an overlay of them on a real scan.

3.3 cGAN to Correct Imperfect Result

We experimented on this idea based on Tensorflow's official implementation of pix2pix model [21]. The original model is used to translate the building facades to real buildings. We keep the skeleton of it and change the data loading, data preprocessing, and the parameter of the optimizer to run on our dataset.

This pix2pix takes 2D input of size 256x256 that has three channels. We adjust our data to be aligned with this form and see if there is any improvement in prediction. We manually select 330 2D slices of predicted segmentation maps from the VAE model's result. These slices all contain a certain amount of tumor tissue because we do not want any training data that does not have a positive sample, in this case, tumor, that will potentially hurt the learning. In addition, we use T1 images, T2 images and FLAIR images as three channels.

The generator uses a U-net that contains 8 up-sampling layers and 8 down-sampling layers, with skip-connections between the encoder and decoder. The discriminator uses a patchGAN that takes input image, target image and generated image and classifies the latter two as real or fake [21, 22].

Training of 280 images over 300 epochs takes around one hour, and the result is very variant. On some of the images that the shapes of their tumors are caught in the training data (slices from the same brain, especially those that are adjacent in the original 3D array, tend to share a similar shape of tumors), we can see an apparent improvement of the segmentation prediction. But on some of the cases that the shapes of tumors never show up in the training (tumor shapes are very different in each patient so this kind of sample is very common in clinical practice), there is hardly any substantial improvement in generated prediction. A mathematical criterion for the result is hard to find but the visual effect of these outputs is very effective to judge. Figure 11 and 12 are respectively some of the good and bad predictions.

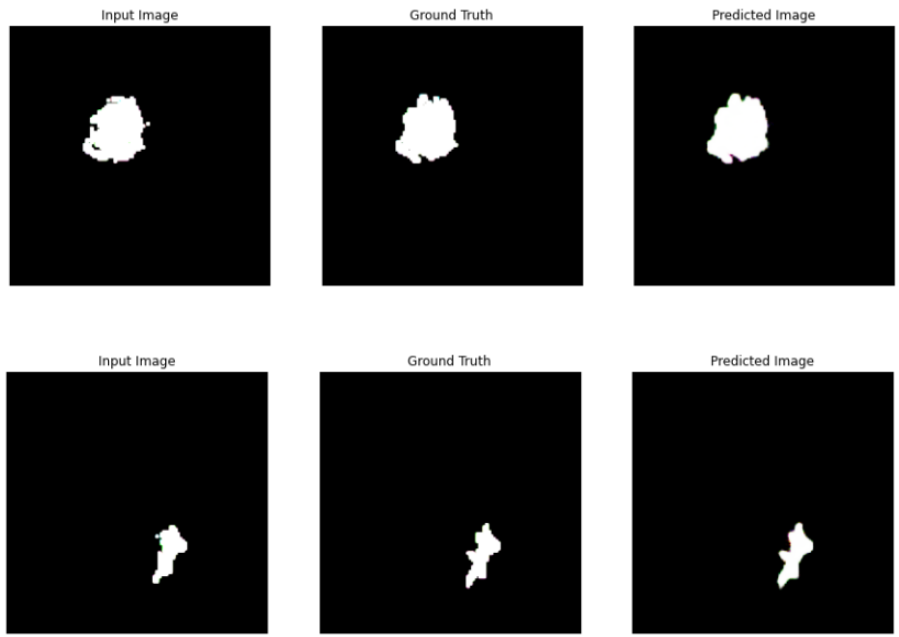


Figure 11: Samples that show good correction result

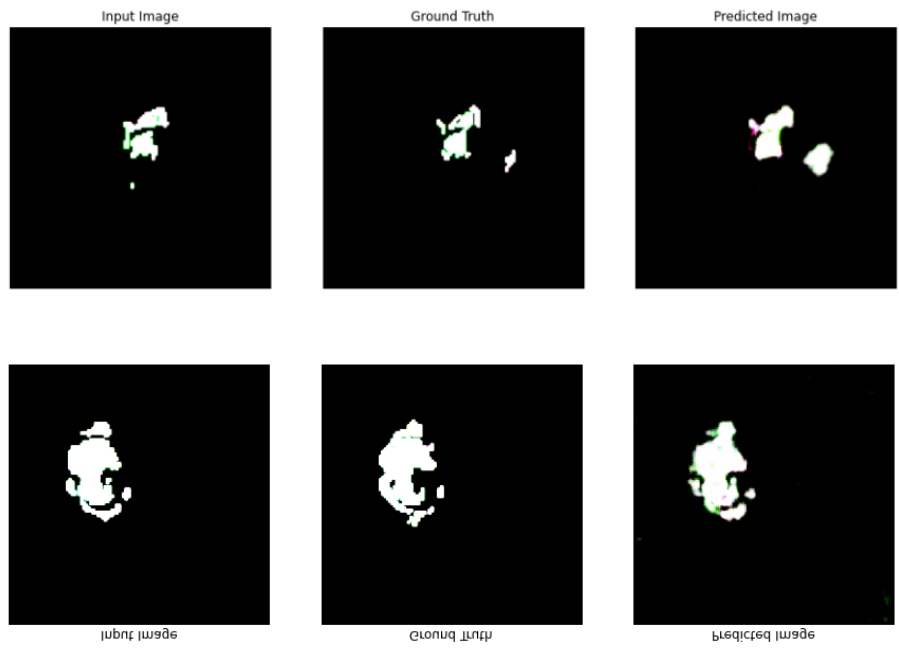


Figure 12: Samples that do not show good correction result

4. Conclusion and Future Work

To sum up, a trend of using generative methods of brain tumor segmentation in magnetic resonance images is on fire in recent years. The use of synthetic data in BraTS competition as well as some other researches proves that generative methods are authentic in data augmentation of a biomedical imaging problem. We need to be careful about the procedure of generating data because if they are not “real” or standard enough, they may contaminate the training. The winning on BraTS 2018 with a variational auto-encoder in the neural network proves that generative models can be very powerful in segmentation tasks. We also bring up a new idea that uses a conditional generative adversarial network to correct imperfect segmentation prediction, but it is still under development and there is a lot of space for improvement.

Toward “cGan to correct imperfect result”, there are some points that we would like to address in the future. First, the original pix2pix is used for RGB channels image, and we directly transplant the model onto our data. To address the issue of variant results described in the last section, we might need to redesign this architecture into the one we need, such as a one-channel model that only takes T1 images, and it may also include a revision on the loss function. Second, we would like to change it into a 3D model for that 2D slices could lose information on the third dimension, and manually selecting 2D slices that contain tumors is a very not intelligent way to choose inputs. Third, we would like to use more imperfect predictions from different models as features to train our cGAN. The final expected diagram should be indicated in figure 13.

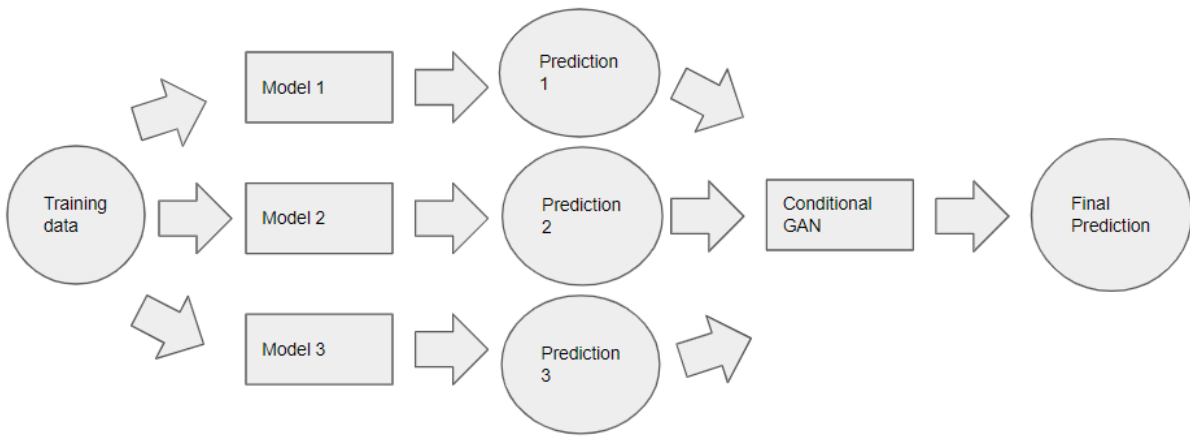


Figure 13: Expected final diagram

References

- [1] Bousselham, A., Bouattane, O., Youssfi, M. and Raihani, A., 2019. Towards Reinforced Brain Tumor Segmentation on MRI Images Based on Temperature Changes on Pathologic Area. *International Journal of Biomedical Imaging*, 2019, pp.1-18.
- [2] Sharma, N., Ray, A., Shukla, K., Sharma, S., Pradhan, S., Srivastva, A. and Aggarwal, L., 2010. Automated medical image segmentation techniques. *Journal of Medical Physics*, 35(1), p.3.
- [3] Challenge.kitware.com. *Multi-Modal Brain Tumor Segmentation Challenge (Brats 2012)*. [online] Available at: <<https://challenge.kitware.com/#challenge/54e4baeacad3a55533264cce>> [Accessed 23 April 2020].
- [4] Med.upenn.edu. *Multimodal Brain Tumor Segmentation Challenge 2019 | CBICA | Perelman School Of Medicine At The University Of Pennsylvania*. [online] Available at: <<https://www.med.upenn.edu/cbica/brats2019.html>> [Accessed 23 April 2020].
- [5] Medium. *Generative Deep Learning: Let'S Seek How AI Extending, Not Replacing Creative Process*. [online] Available at: <<https://towardsdatascience.com/generative-deep-learning-lets-see-how-ai-extending-not-replacing-creative-process-fded15b0561b>> [Accessed 23 April 2020].
- [6] Nitrc.org. *NITRC: Tumorsim: Tool/Resource Info*. [online] Available at: <<https://www.nitrc.org/projects/tumorsim/>> [Accessed 23 April 2020].
- [7] Myronenko, A. *3D MRI Brain Tumor Segmentation Using Autoencoder Regularization*. arXiv preprint arXiv: 1810.11654, 2018.

- [8] Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J. and Greenspan, H., 2018. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*, 321, pp.321-331.
- [9] Goodfellow, I., 2016. *Generative Adversarial Networks (Gans)*. [online] iangoodfellow.com. Available at: <<https://www.iangoodfellow.com/slides/2016-12-04-NIPS.pdf>> [Accessed 23 April 2020].
- [10] B. H. Menze *et al.*, "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)," in *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993-2024, Oct. 2015.
- [11] Prastawa, Marcel et al. "Simulation of brain tumors in MR images for evaluation of segmentation efficacy." *Medical image analysis* 13 2 (2009): 297-311.
- [12] Alexander, D., Pierpaoli, C., Basser, P. and Gee, J., 2001. Spatial transformations of diffusion tensor magnetic resonance images. *IEEE Transactions on Medical Imaging*, 20(11), pp.1131-1139.
- [13] Clatz, O., Sermesant, M., Bondiau, P., Delingette, H., Warfield, S., Malandain, G. and Ayache, N., 2005. Realistic simulation of the 3-D growth of brain tumors in MR images coupling diffusion with biomechanical deformation. *IEEE Transactions on Medical Imaging*, 24(10), pp.1334-1346.
- [14] Medium. 2020. *Intuitively Understanding Variational Autoencoders*. [online] Available at: <<https://towardsdatascience.com/intuitively-understanding-variational-autoencoders-1bfe67eb5daf>> [Accessed 23 April 2020].
- [15] Medium. 2020. *Understanding Variational Autoencoders (Vaes)*. [online] Available at: <<https://towardsdatascience.com/understanding-variational-autoencoders-vaes-f70510919f73>> [Accessed 23 April 2020].

- [16] Goodfellow, Ian J., Pouget-Abadie, Jean, Mirza, Mehdi, Xu, Bing, Warde-Farley, David, Ozair, Sherjil, Courville, Aaron C., and Bengio, Yoshua. Generative adversarial nets. NIPS, 2014.
- [17] M. Mirza and S. Osindero. Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, 2014.
- [18] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In CVPR, 2017
- [19] Phillipi.github.io. *Image-To-Image Translation With Conditional Adversarial Networks*. [online] Available at: <<https://phillipi.github.io/pix2pix/>> [Accessed 23 April 2020].
- [20] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan, A.C. Evans : "*BrainWeb: Online Interface to a 3D MRI Simulated Brain Database*"
NeuroImage, vol.5, no.4, part 2/4, S425, 1997 -- Proceedings of 3-rd International Conference on Functional Mapping of the Human Brain, Copenhagen, May 1997.
- [21] TensorFlow. n.d. *Pix2pix | Tensorflow Core*. [online] Available at: <<https://www.tensorflow.org/tutorials/generative/pix2pix>> [Accessed 23 April 2020].
- [22] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In MICCAI, pages 234–241. Springer, 2015.