Imperial College
London

IMPERIAL COLLEGE LONDON

DEPARTMENT OF COMPUTING

Background and Progress Report

Author:
Tianyang Sun

Supervisor:
Name of supervisor

September 2015

Contents

Chapter 1

Task description

1.1 Motivation

We here describe the problem setup for this project. The project aims to explore several few shot segmentation methods in medical domain, specifically 3D Lung computed tomography (CT) images. One of the target for this method is to provide solutions to the few shot segmentation challenge proposed in the COVID19 segmentation benchmark[1]. Given this is a relatively small dataset and might not present good evaluation to the result, we also plan to evaluate the performance on Strucseg dataset.

1.2 Feasibility

1.3 Challenge

In non-medical domain, both few shot classification and few shot segmentation has been well explored and provide good performance on some of the segmentation benchmarks. In medical application, few shot learning mainly focus on classification tasks while dense segmentation does not provide promising results as far as we know.

We so far beleive the task is doable based on the previous analysis. We aim to implement a method that provide comparable performance compared to U-net, which is a good model on small sample medical image segmentation.

---

[1]http://medicalsegmentation.com/covid19/

Chapter 2

Background

2.1 Dataset

2.2 Machine Learning Preliminaries

Since the

2.2.1 Machine learning methods in medical domain

such as Unet

2.3 Data augmentation

In medical domain, huge dataset consists of large numbers of carefully labelled samples is rarely available due to heavy workload for annotations, rarity of disease, ethic issues of data acquire process and data privacy. Furthermore, different data acquire protocols (i.e. CT machines used by different hospitals) brings difficulties to clinical practice for good accuracy of existing pre-trained models. As a result, few shot learning and/or few shot segmentation has been explored in recent years. In this section, we focus on a few approaches that has been used in current literature which aim to explore the potential of existing training samples through various augmentation methods to alleviate the insufficient training samples in medical imaging. We discuss data augmentation method as well as the amount of data used in each work. Table **??** provide an overview of each method.

2.3.1 Traditional Data Augmentation

Traditional Data augmentation method in imaging domain refers to the process that does not require such training data to learn a transformation.

[1] investigated data augmentation methods under 3D medical domain of MR and ultrasound images. The data augmentation process consists of a sequence of traditional transformation techniques. The paper argued that sharpness in medical images during training process limits model generalization thus applying gaussian filter to images take noise into consideration. Brightness and contrast difference caused by variations in scanning protocols brings potential domain shift thus a sequenced random shift followed gamma correction and random linear transform in intensity are reasonable data augmentation methods. Finally spatial transformations including rotation, scaling and deformation is added to the augmentation process. The source domain in this method is Prostate dataset [1] consists 48 4D volumes. We argue here that the stacked transformation is physical transformation process independent to the size of dataset because no learning or training process is required in this augmentation method thus might bring benefits to our task. However we doubt the improvement due to the difference between CT and MR images.

Another method "mixup" by [2] based on generic vicinal distribution, which generates new samples through interpolation between two existing data. The author argued that this method works as a regularizer which encourages linear behavior between training samples. In terms of imaging, the augmentation is applied to CIFAR 10 (2D non medical) Dataset. The work was originally implemented on training GANs. Later in medical domain, [3] further investigated mixup method on Knee MR images on OIA database [2] that consists of 88 3D MRI scans. The work showed mixup improves generalization under their experiment setup while having risk of slight underfitting due to to the strong regularization. The author further mentioned not using weight decay in the experiment solve the underfitting issue. [4] summarized mixup augmentation method gives "soft labels". Variations of the mixup method utilize asymmetric is further explored in [5] trained on Brain MR images, and the method reports huge gain under their experiment setup. We cannot guarantee the effectiveness of this augmentation method due to the domain difference between MRI and CT scans.

2.3.2   Learning for augmentation

Following the traditional augmentation methods, we here want to discuss a few medical image augmentation based on deep learning. Specifically we focus on those who used one or a few data during the augmentation which is close to the problem we are facing.

Adversarial defense was deployed in [6] for the augmentation of small samples in Brain MR segmentation. Adversarial samples generated though Fast Gradient Sign Method[] and then added as training data to improve the robustness, which according the study in [] without decreasing the performance. So far we believe this method worth trying because the experiment was trained on 7 brain volumes, which

---

[1]http://medicaldecathlon.com/index.html#tasks
[2]http://www.oai.ucsf/.edu/

is close to our task.

Paper	Method	Dataset	Number of samples
[1]	Stacked traditional transform	Prostate dataset	48 4D volumes
[2]	mixup	CIFAR 10	Huge
[3]	mixup	Knee MR images	88 3D MRI
[5]	Asymmetric mixup	BraTS ?	
[6]	Adversarial defense	MRBrainS18 challenge	7 train, 14 test

2.4   Few shot learning in non-medical domain

Recent work in non-medical few shot segmentation have explored several network structures that tend to provide correct segmentation for few labelled samples from dense to sparse annotations.

Although few-shot learning for classification has been wide explored, and segmentation, in some sense, can be viewed as classification task, a direct extension from those networks might not give good results.

Experiment of extending Siamese Network in meta learning classification to pixelwise segmentation failed to scale well enough, but this idea inspire the work in [7] that proposed a two-branch network for few-shot semantic segmentation. The work assumes that semantic class labels for training and testing have no overlap, that is $L_{train} \cap L_{test} = \varnothing$. Large dataset updates a pre-trained VGG network on ImageNet in the "conditioning branch", and a query sample updates the "segmentation branch" including a FCN-32 and a conv-fc7. The authors further utilize masking and weight hashing to combine the two branches. The experiment reported promising results in non medical dataset (PASCAL-5) the work further removed the effect of ImageNet and still show good performance. However, in our CT segmentation task, we might want to take 3D features into consideration.

Another work in [8] relax the restriction of requiring dense segmentation masks to sparse annotations with similar backbone of VGG net. They utilize the guidance network that provide guiding inference to extract visual features through an encoder structure. The work is evaluated on both 2D images and video sets while the structure used in video sets.

2.5   Few-shot learning in medical application

2.5.1   Transfer learning method

In medical domain, few shot learning mainly focus on transfer learning from pre-trained networks that leverage both medical and non-medical datasets.

In Lung CT segmentation area, Sports-1M dataset has been used as source domain to train a multi-task learning model for nodule malignancy prediction and rating [9]. The author reported significant improvement in the prediction accuracy, however, did not mention the proportion of data used for transfer training.

People tend to choose datasets from closer domain for transfer learning. It is reasonable to consider methods that transfer across disease in the same structure under the same modality. In our case, we might want to investigate transfer learning from NSCLC Dataset to Covid segmentation set given that both of them are lung CT scans.

Recent work explored several across disease transfer learning training techniques under MRI domain [10]. The paper evaluated three transfer learning methods trained on 3D U-Net by Fine tuning the last three layers, Fine tuning the decoder and Fine tuning all model parameters. The Source Dataset: Multiple Sclerosis Dataset consists 3630 MRI volumes and used Brain Tumor Dataset as Target dataset including 210 high-grade glioma (HGG) and 75 low-grade glioma (LGG) Brain MRI scans. The training target is a decaying weighted categorical cross entropy loss weighted by relative voxel. Their best validation performance of pre-trained network achieved validation performance AUC 0.77. Experiment result on 20, 50, 100 and 150 samples during Fine tuning respectively showed that Fine tuning all parameters out performed the rest methods in most cases.

One potential drawback is that compared to the our task, the target training set is relatively larger, the performance is expected to be less ideal when using "fine tune all" method using 4 or less volumes in our case.

Paper [6] trained segmentation of Brain MR image on 7 brain volumes after Adversarial defense augmentation. The work first split the segmentation from easy to hard into 2 individual classifiers then joint learning with dense pixel segmentation. The author reported that the result outperformed Unet and Vnet method trained from scratch. We argue that the augmentation provide good result in the segmentation accuracy while the segmentation part is not well explained in the paper. We have emailed the author for further details.

2.5.2   Special network design

Transfer learning methods usually require small samples to update millions of parameters that take the risk of overfitting [7]. The design of multiple branch network, usually includes conditioner arm and segmentation arms inspired the deep learning in medical domain to go beyond transfer learning while encourage a stronger between-arms interaction to compensate the lacking in pre-trained model [11]. The proposed method perform 3D volume segmentation at test time while use 2D images during training. However the method requires start and end slice to be indicated for each query volume and still not achieving good dice score.

Chapter 3

Progress Section

3.1 Progress so far

3.1.1 Brief plan

3.1.2 robustness

Bibliography

[1] L. Zhang, X. Wang, D. Yang, T. Sanford, S. Harmon, B. Turkbey, H. Roth, A. Myronenko, D. Xu, and Z. Xu, "When unseen domain generalization is unnecessary? rethinking data augmentation." [Online]. Available: http://arxiv.org/abs/1906.03347 pages

[2] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond empirical risk minimization." [Online]. Available: http://arxiv.org/abs/1710.09412 pages

[3] E. Panfilov, A. Tiulpin, S. Klein, M. T. Nieminen, and S. Saarakkala, "Improving robustness of deep learning based knee MRI segmentation: Mixup and adversarial domain adaptation," in *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. IEEE, pp. 450–459. [Online]. Available: https://ieeexplore.ieee.org/document/9022164/ pages

[4] N. Tajbakhsh, L. Jeyaseelan, Q. Li, J. N. Chiang, Z. Wu, and X. Ding, "Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation," vol. 63, p. 101693. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S136184152030058X pages

[5] Z. Li, K. Kamnitsas, and B. Glocker, "Overfitting of neural nets under class imbalance: Analysis and improvements for segmentation." [Online]. Available: http://arxiv.org/abs/1907.10982 pages

[6] X. Ren, L. Zhang, D. Wei, D. Shen, and Q. Wang, "Brain MR image segmentation in small dataset with adversarial defense and task reorganization," in *Machine Learning in Medical Imaging*, H.-I. Suk, M. Liu, P. Yan, and C. Lian, Eds. Springer International Publishing, vol. 11861, pp. 1–8, series Title: Lecture Notes in Computer Science. [Online]. Available: http://link.springer.com/10.1007/978-3-030-32692-0_1 pages

[7] A. Shaban, S. Bansal, Z. Liu, I. Essa, and B. Boots, "One-shot learning for semantic segmentation," in *Procdings of the British Machine Vision Conference 2017*. British Machine Vision Association, p. 167. [Online]. Available: http://www.bmva.org/bmvc/2017/papers/paper167/index.html pages

[8] K. Rakelly, E. Shelhamer, T. Darrell, A. Efros, and S. Levine, "Few-shot segmentation propagation with guided networks," p. 10. pages

[9] S. Hussein, K. Cao, Q. Song, and U. Bagci, "Risk stratification of lung nodules using 3d CNN-based multi-task learning," vol. 10265, pp. 249–260. [Online]. Available: http://arxiv.org/abs/1704.08797 pages

[10] B. Kaur, P. Lematre, R. Mehta, N. M. Sepahvand, D. Precup, D. Arnold, and T. Arbel, "Improving pathological structure segmentation via transfer learning across diseases," in *Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data*, Q. Wang, F. Milletari, H. V. Nguyen, S. Albarqouni, M. J. Cardoso, N. Rieke, Z. Xu, K. Kamnitsas, V. Patel, B. Roysam, S. Jiang, K. Zhou, K. Luu, and N. Le, Eds.   Springer International Publishing, vol. 11795, pp. 90–98, series Title: Lecture Notes in Computer Science. [Online]. Available: http://link.springer.com/10.1007/978-3-030-33391-1_11 pages

[11] A. G. Roy, S. Siddiqui, S. Plsterl, N. Navab, and C. Wachinger, "'squeeze & excite' guided few-shot segmentation of volumetric images." [Online]. Available: http://arxiv.org/abs/1902.01314 pages