CHAOS Challenge: Liver segmentation from CT images

Shruti A. Mali
s.a.mali@student.tue.nl

25th June 2020

1 Introduction

The proposed system is an adaptation of self-supervised models called *Models Genesis* from [2] making it more generic across different variants in multi-modal images. The goal is to perform volumetric segmentation using a 3D-UNet architecture [1], so that significant 3D anatomical information is retained while fine-tuning the network for Task 2 of the challenge.

2 Methods

2.1 Models Genesis framework

With an intent to make most use of unannotated data, *Models Genesis* [2] uses a self-supervised learning technique to pre-train huge amount of data for 3D imaging target tasks, preserving the 3D information which would have been otherwise lost if reproduced into 2D images [1]. Hence, the performance evaluation in paper [2] proves to top transfer learning from 2D approaches (even ImageNet), stressing on the significance of *Models Genesis* for 3D medical imaging tasks.

Figure 1 shows the encoder-decoder structure for the purpose of pre-training. This 3D-UNet architecture is build using base architecture from [1]. The main goal of the encoder-decoder structure is to perform image restoration task after the unannotated 3D images undergo transformations as seen in Figure 1. To elaborate, 3D patches $X = x_1, x_2, x_3, ..., x_n$, where n is the number of slices in depth of the image, are cropped randomly at random locations from unannotated 3D image, which further undergo transformations creating a set of transformed patches $\widetilde{X} = \widetilde{x_1}, \widetilde{x_2}, \widetilde{x_3}, ..., \widetilde{x_n}$. X is considered the ground truth and the encoder-decoder structure tries to restore the transformed patch back to X indicating self-supervised pre-training to produce effective target models. The restored patch is X' in the Figure 1. After pre-training, main segmentation process is carried out by fine-tuning the same 3D-UNet architecture as mentioned before for target tasks. For Task 2, pre-training is performed on CT train sets only. These pre-trained weights are further loaded for fine-tuning the network for final liver segmentation task from CT test data.

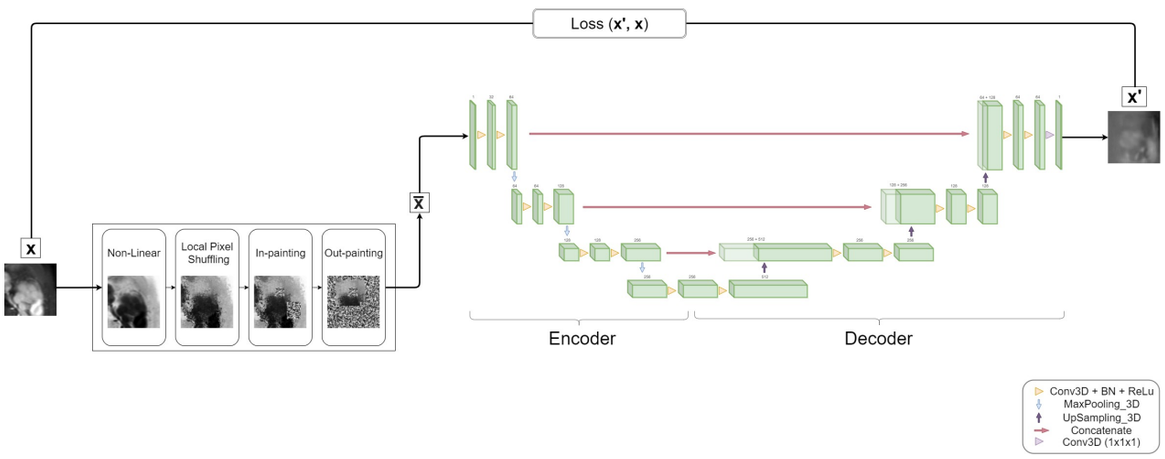

Figure 1: Network architecture

The purpose of this task is to segment liver from CT test data and this is achieved by utilizing all possible liver annotations coming from CT as well as MRI train sets while fine-tuning.

2.2 Pre-training details

The network is trained for the purpose of reconstruction and randomly extracted patches of size [64, 64, 32] were used from CT plus MRI train set. Transformations such as non-linear transformation (Bezier curve), local pixel shuffling and painting-based methods were applied on the patches [2]. Mean squared error is used as loss function. For optimization, SGD is used with initial learning rate of 1e-2, which further decreases if validation loss doesn't decrease after a patience of 10 epochs. The network is trained with a batch size of 12. The pre-training alone took a whole day to execute.

2.3 Fine-tuning details for liver Segmentation

The network is fine-tuned using patch-based approach again. This time patches of size [128, 128, 48] were used with zero overlap. The labels are encoded using one hot code vectors. Note that here any input size can be used as long as it is divisible by $16(= 2^4)$, because in the 3D-UNet architecture four down-sampling layers are used. The loss function is binary crossentropy. For optimization, Adam is used with initial learning rate as 1e-2, with patience of 6. The network is trained with batch size of 2. The fine-tuning takes more than 24 hours to execute.

2.4 Evaluation

The training set is split into 80-20% respectively for train and validation set. Evaluation metrics such as Dice score and Mean-IoU is computed for validation sets where Dice score of 0.91 and mean-IoU of 0.89 is obtained. The best model is saved with the least validation loss of 0.06.

2.5 Future work

Same implementation is under process for Task 4 to perform multi-modal learning for multi-organ segmentation.

Thesis Project taken under Professor Mitko Veta and Dr. Samaneh Abbasi Sureshjani at TU Eindhoven, The Netherlands.

References

[1] Ahmed Abdulkadir, Soeren S Lienkamp, and Olaf Ronneberger. 3D U-Net : Learning Dense Volumetric Segmentation from Sparse Annotation.

[2] Z Zhou, V Sodha, M M Rahman Siddiquee, R Feng, N Tajbakhsh, M B Gotway, Zongwei Zhou, Vatsal Sodha, and Mahfuzur Rahman Siddiquee. This is the full version of our MICCAI-2019 paper on Models Genesis with the whole Supplementary Materials This paper was awarded a Young Scientist Award at MICCAI 2019 Models Genesis : Generic Autodidactic Models for 3D Medical Image Analysis. pages 1–27, 2019.