Introduction

## \* Abstract

Medical images that come from different modalities including X-ray, computed tomography(CT), and magnetic resonance imaging(MRI) performs critical roles in diagnosing and making medical decisions in modern medical caring process. With the arise of artificial intelligence, image processing techniques and the explosion of hash rate, computational analysis of medical images is gaining affinity from researchers. Given that the speed traditional computer image processing methods shown are not satisfied, we developed an image segmentation method based on VNet, whose performance surpass baseline methods(85% acc) with increased speed.

## \* Introduction

### \*\* AF is critical

### \*\* MRI is good

### \*\* GE-MRI is better MRI

### \*\* ML can help segmentation and medical image analysis

Segmentation is a highly relevant task in medical image analysis. Automatic delineation of organs and structures of interest is often necessary to perform tasks such as visual augmentation [10], computer assisted diagnosis [12], interventions [20] and extraction of quantitative indices from images [1]. In particular, since diagnostic and interventional imagery often consists of 3D images, being able to perform volumetric segmentations by taking into account the whole volume con- tent at once, has a particular relevance. In this work, we aim to segment prostate MRI volumes. This is a challenging task due to the wide range of appearance the prostate can assume in different scans due to deformations and variations of the intensity distribution. Moreover, MRI volumes are often affected by artefacts and distortions due to field inhomogeneity. Prostate segmentation is neverthe- less an important task having clinical relevance both during diagnosis, where the volume of the prostate needs to be assessed [13], and during treatment planning, where the estimate of the anatomical boundary needs to be accurate [4,20].

### \*\* benchmark reach 93% dice(2020)

### \*\* issue with multi-modality, lightweight, and robusty yet to be addressed

### \*\* related work:

#### \*\*\* CNNs

Recent research in computer vision and pattern recognition has highlighted the capabilities of Convolutional Neural Networks (CNNs) to solve challenging tasks such as classification, segmentation and object detection, achieving state-of-the- art performances. This success has been attributed to the ability of CNNs to learn a hierarchical representation of raw input data, without relying on hand- crafted features. As the inputs are processed through the network layers, the level of abstraction of the resulting features increases. Shallower layers grasp local information while deeper layers use filters whose receptive fields are much broader that therefore capture global information

CNNs have been recently used for medical image segmentation. Early ap-

proaches obtain anatomy delineation in images or volumes by performing patch- wise image classification. Such segmentations are obtained by only considering local context and therefore are prone to failure, especially in challenging modalities such as ultrasound, where a high number of mis-classified voxel are to be expected. Post-processing approaches such as connected components analysis normally yield no improvement and therefore, more recent works, propose to use the network predictions in combination with Markov random fields [6], vot- ing strategies [9] or more traditional approaches such as level-sets [2]. Patch-wise approaches also suffer from efficiency issues. When densely extracted patches are processed in a CNN, a high number of computations is redundant and therefore the total algorithm runtime is high. In this case, more efficient computational schemes can be adopted.

In the past decade, deep learning techniques, in particular Convolutional

Neural Networks (CNNs), have achieved great progress in various computer vi- sion tasks, and rapidly become a methodology of choice for analyzing medical images [7]. Ciresan et al. [3] firstly introduced CNNs to medical image segmen- tation by predicting a pixel’s label based on the raw pixel values in a square window centered it. But this method is quite slow because the network must run separately for every pixel within every single image and there is a lot of redundancy due to overlapping windows actually

#### \*\*\* VGGNet

VGGNet (Simonyan and Zisserman 2014) has been widely used for developing full convolutional networks (Long, Shelhamer and Darrell 2015, Noh, Hong and Han 2015) for semantic segmentation due to its simplicity, and adaptations of superior architectures, such as ResNet and Inception (Szegedy et al. 2017), are currently the state-of-the-art in the field

#### \*\*\* U-Net

1. Net [1], which consists of a contracting path to capture context and a symmetric expanding path that enables precise localization and can be trained end-to-end from very few images built upon the famous Fully Convolutional Net- work (FCN)

Later on, Ronneberger et al. proposed U-Net [13], which consists of a contracting path to capture context and a symmetric expanding path that enables precise localization and can be trained end-to-end from very few images built upon the famous Fully Convolutional Net- work (FCN) [8]

Then, Cicek et al. [2] replaced the convolution operations in 2D U-Net with 3D counterparts and proposed 3D U-Net for volumetric segmenta- tion. Furthermore, Milletari et al. [10]

#### \*\*\* V-Net

learn a residual function in- spired by [6] which ensures convergence in less training time and achieves good segmentation accuracy

sb proposed V-Net, wherein they introduce a novel loss function based on Dice coefficient and learn a residual function in- spired by [6] which ensures convergence in less training time and achieves good segmentation accuracy

#### \*\*\* ResNet

In this paper, residual block is introduced into the convolution layer of U-net, and the idea of residual learning is used for image feature extraction. The multilayer convolution is used to simulate a residual mapping ??(??), which is easier to optimize than the approximate identity mapping of the direct learning object. Assuming that the input variable is ??, residual learning is to pass the variable ?? to the output directly through the "shortcut connection" as the initial result, and the output result is: ??(??) = ??(??) + ?? .

When ??(??) = 0, ??(??) = ?? , the so-called identity mapping, can avoid the gradient disappear caused by the deepening of the network and make the network better optimized. Therefore, the learning goal of residual is no longer a complete output, but the difference between the goal value ?? (??) and ??, that is, ??(??) = ??(??) − ?? .

#### \*\*\* SE-block

Many studies consider improving network performance from

the spatial dimension level, such as embedding multi-scale in- formation in the Inception structure and aggregating the char- acteristics of a variety of different receptive fields to obtain per- formance improvements. Chudzik et al. [13] applied this struc- ture to the segmentation of retinal exudate and achieved good results. Chen et al. [14] proposed that multiple sets of medical data were segmented by embedding an improved Inception struc- ture in DRINet and the effect was significant. Bell et al. [15] pro- posed the Inside-Outside network, which uses spatial recurrent neural networks to integrate contextual information in the space to further improve accuracy. None of the above networks consider the relationship between feature channels, and the network of- ten learns many features that are irrelevant to the segmentation task during the learning process, resulting in a decrease in the accuracy of the segmentation result. To improve the accuracy of automatic segmentation of 3D medical images, this paper proposes a 3D medical image segmentation method based on the SE-VNet network.

### \*\* our architecture

SE-VNet

In this paper, we develop an automatic 3D atrial segmentation framework

using volumetric fully convolutional networks for 2018 Atrial Segmentation Chal- lenge. The overall pipeline of our method is shown in Fig. 1, it consists of two main stages: 1) in the first stage, we use a segmentation based localization strat- egy to estimate a fixed size target region that covers the whole atria, and leave out pixels outside this region to cut down memory consumption; 2) in the second stage, we train a fine segmentation network based on the cropped target region obtained in the first stage, and transform the predicted masks in target region to the original size volume. The segmentation networks in these two stages are both adapted from V-Net, which can be trained end-to-end and used to segment the atrial cavity fully-automatically.

1. Considering the characteristic channels of the network, a new three-dimensional medical image segmentation net- work SE-VNet is proposed. This method can achieve rapid and accurate segmentation of 3D images without human intervention.

(2) The SE-Net module is embedded after the features ex- tracted from each layer of the encoder, which improves the segmentation accuracy and at the same time proves the ef- fectiveness of the channel attention mechanism in 3D image segmentation

(3) Feature weights are learned by network training. Increase the weight of the key features of the target area, reduce the weight of the invalid or ineffective background area features

1. Extensive experimental results show that the method proposed in this paper has good robustness, and the seg- mentation performance is better than other methods in the literature.

## \* METHODS

### \*\* Datasets and Preprocessing

Out Network used LAScar2022 Atrial Segmentation Challenge dataset. It is consist of 130 3D GE-MRIs as training set, with each containing raw MRI scan and the corresponding gound truth labels of the LA(left atrial) cavity. Original resolution of them varies from 0.625x0.625x1mm and 1x1x2.5mm, and xx of them are with xxx, xx of them are with xxx, xx of them are with xxx and xx of them are with xxx. So we consequently resampled them to a same resolution and size of volumes turned to x x x and x. Actually, only small part of the whole volume represents the LA cavity and applying neural networks to straight-forward segmentation from such high-resolution volumes could be resource-expensive. Accordingly, an adaptive cropping process is introduced to roughly detect and select a smaller patch of volume containing LA cavity. Afterwards, the cropped MRI data was enhanced using a 3D version of contrast limited histogram localization (CLAHE) and was applied to sample-wise normalization, wherein each volume is subtracted with the mean value of intensity and divided by the deviation of intensity.

### \*\* Network architecture

We adapted V-Net to our segmentation network. It is a fully-convolutional network, where features extraction and resolution reduction was accomplished through convolution operations with different cores and appropriate strides. An SE(squeeze-excitationGenerally, the left part of the network

### \*\* Test

### \*\* Train

### \*\* Test

## \* References

[1] Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedi- cal image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)