Thrifty Neural Architecture Search for Medical Image Segmentation (Student Abstract)

Ruibin Chen, [1] **Miao Zhang,** [2] **Xin Zheng,** [1] **Shirui Pan** [1]

[1] Monash University, Melbourne, Australia
[2] Aalborg University, Aalborg, Denmark
rche0046@student.monash.edu, miaoz@cs.aau.dk, xin.zheng@monash.edu, shirui.pan@monash.edu

Abstract

Convolutional neural network (CNN) based image segmentation has been widely used in analyzing medical images and benefited many real-world disease diagnosis applications. However, existing advanced CNN-based medical image segmentation models usually contain numerous parameters that require massive computation and memory, limiting the applicability of these models in the data-constrained or hardware-constrained environments. By leveraging the recently proposed neural architecture search (NAS), this paper presents a novel approach, dubbed Thrifty NAS, to design computation and memory-efficient models for medical image segmentation automatically. The searched models by Thrifty NAS are with much fewer parameters while retaining competitive performance. More specifically, we design a micro level space for cell structure search and a macro level cell path for better network structure modeling. Extensive experimental results in different medical image datasets verify the effectiveness of the proposed method with competitive segmentation performance, especially with minuscule neural architecture model size, i.e., 0.61M that is superior to U-Net (7.76 M) and UNet++ (9.04 M).

Introduction

Medical image segmentation plays a crucial role in diagnosing disease, surgical planning, and prognostic assessment. Recently, convolutional neural network (CNN) based methods have been the mainstream for medical image segmentation. These methods are widely used in analyzing different types of medical images, including X-ray, microscopic, ultrasound, magnetic resonance imaging (MRI), and computerized axial tomography (CAT). However, existing CNN based medical image segmentation models have two following potential problems:

Lack of universality. First, most of the existing CNN-based models are designed for specific datasets and tasks for medical image segmentation. Hence, once an artificially designed CNN-based model needs to be applied to other datasets, it would take too much time to manually tune the model's hyperparameters and structures, resulting in limited universality.

Hardware limitations and overfitting. Second, considering that medical images are characterized by high-resolution image sizes, training a large-scale CNN-based model with massive parameters often consumes amounts of GPU memory and computation. Due to such hardware limitations, existing medical image segmentation models have to split images into several patches to process them separately. This strategy could lead to information loss and does not generalize well to all images. Moreover, the medical images are not easily obtainable in the real world, causing insufficient data sources. Hence, using a limited-scale dataset to train a large-scale CNN model often leads to inevitable overfitting.

By leveraging the recently-proposed neural architecture search (NAS), we propose a novel method, called Thrifty NAS, to solve the above two problems. Taking the advantage that NAS can automatically tune the neural architecture, we can design suitable architectures for different datasets. In addition, the network searched by our Thrifty NAS has fewer parameters to alleviate hardware limitations and overfitting. Specifically, at the macro level, we redesign the backbone architecture with dense connections to reuse feature maps and thus reduce the number of parameters. At the micro level, we design a search space of the cell structure to filter and keep partial feature maps only, reducing the GPU memory consumption. As a result, we can obtain high-performance models with fewer parameters, which allows us to use computational resources more efficiently, increase batch size, and reduce training time for medical image segmentation.

Proposed Approach

Macro level. Thrifty NAS takes inspiration from UNet++ (Zhou et al. 2018) and applies it to the design of our macro architecture. UNet++ is a state-of-the-art manually designed network for medical image segmentation. We redesigned its dense skip pathways to reuse feature maps more efficiently. Figure 1 shows our proposed search space, where the whole network is composed of several cells. We define two different cell structures: downsampling (DN) and upsampling (UP) cells. These DN and UP cells are densely connected in the search space, with data flowing from left to right. The dimensions of input features are halved after the calculation of DN cells and doubled after UP cells. Each cell has two inputs, in_0 and in_1. For $X^{0,j}$, a DN cell, in_0 and in_1 are
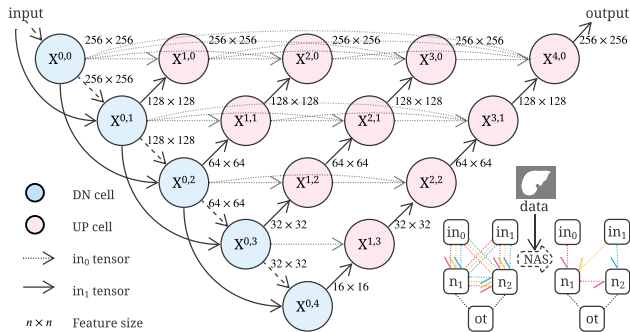
Figure 1: Overview of the proposed network architecture. Each block represents a cell structure, and each edge represents an input-output tensor. Use the relaxation method to obtain the best cell structure.

the outputs of its two predecessors $X^{0,j-2}$ and $X^{0,j-1}$. For an UP cell $X^{k,j}(k \geq 1)$, in_0 is a concatenation of the outputs of $X^{m,j}(m \in 0,...,k-1)$ while in_1 is the output of $X^{k-1,j+1}$. Cells of the same output size are densely connected so that we can reuse features more efficiently.

Our model differs from UNet++ in three ways. First, UNet++ consists of several convolutional blocks designed by data scientists, while Thrifty NAS stacks several cells that the cell structure is found by NAS. Secondly, each block of UNet++ is single-input and single-output, while Thrifty NAS has two inputs per cell, where each deals with features of different sizes. Finally, UNet++ extracts features by doubling the number of convolutional kernels in the encoder path. However, Thrifty NAS maintains the same number of convolutional kernels per cell. The intuition behind this design is distillation, and only the most critical features need to be extracted. As a result, the model parameters are significantly reduced with getting better performance.

Micro level. For the cell structure, we proposed a new search space to filter features by reducing redundant channels. Each cell consists of N nodes, and in Figure 1, $N = 2$. The red, yellow, and blue lines that connect these nodes represent candidate operations, including convolution, pooling, skip connection, or even none. We search for these candidate operations using PC-DARTS based search strategy (Xu et al. 2020), which is memory efficient and stable. We search the weights of candidate operations by relaxation and gradient descent, and the K operations with the highest weights are retained to build the final cell. Inside the cell, most existing methods choose to concatenate the outputs of N nodes to form a tensor with $N \times C$ channels. Instead, we use a 1×1 convolution to filter the channels and output a tensor with only C channels. It significantly reduces the number of channels and minimizes GPU memory consumption.

Experiments

We evaluate our model on two different benchmark datasets. CHAOS (Kavur et al. 2019) is a Combined (CT-MR) Healthy Abdominal Organ Segmentation dataset, from which we use computed tomography (CT) data.

		Dataset	
Architecture	Params	CHAOS	PROMISE12
U-Net	7.76M	96.12	82.85
UNet++	9.04M	95.89	85.56
Thrifty NAS	0.61M	**96.21**	**87.69**

Table 1: Segmentation results (DSC: %) for U-Net, UNet++ and our searched architecture.

PROMISE12 (Litjens et al. 2014) is a dataset of magnetic resonance imaging (MRI) of the prostate. For the loss function, we used a combination of cross-entropy and dice loss. Dice similarity coefficient (DSC) is used to measure the performance of our model. We first search neural architectures on both datasets respectively and then train the models from scratch for evaluation. To evaluate our model, we compare the performance of Thrifty NAS with U-Net and UNet++, which are state-of-the-art models for medical image segmentation. The hyperparameters and data augmentation are set identically for all models. As shown in Table 1, Thrifty NAS achieves better performance than the baseline models on both CT and MR images. More importantly, our model size is only 6.7% of UNet++.

Conclusions

In this paper, we propose a novel approach, dubbed Thrifty NAS, to design more efficient and lightweight neural architectures for medical image segmentation. We design a novel backbone architecture involving both macro and micro level search spaces and use a memory-efficient search strategy for architecture search. Unlike most existing image segmentation models that double the number of kernels per convolutional block, our method requires only the same number of kernels per cell. Experimental results in different medical image datasets demonstrate the superior performance of the proposed Thrifty NAS while the model parameters are significantly cut down with $6 \sim 8$ % compared to advanced U-Net and UNet++ models.

References

Kavur, A. E.; Selver, M. A.; Dicle, O.; Barış, M.; and Gezer, N. S. 2019. CHAOS - Combined (CT-MR) Healthy Abdominal Organ Segmentation Challenge Data. In *Zenodo*.

Litjens, G.; Toth, R.; van de Ven, W.; Hoeks, C.; Kerkstra, S.; van Ginneken, B.; Vincent, G.; Guillard, G.; Birbeck, N.; Zhang, J.; et al. 2014. Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge. *Medical image analysis*, 18(2): 359–373.

Xu, Y.; Xie, L.; Zhang, X.; Chen, X.; Qi, G.-J.; Tian, Q.; and Xiong, H. 2020. PC-DARTS: Partial channel connections for memory-efficient architecture search. In *ICLR*.

Zhou, Z.; Siddiquee, M. M. R.; Tajbakhsh, N.; and Liang, J. 2018. Unet++: A nested u-net architecture for medical image segmentation. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, 3–11. Springer.