Resource Optimized Neural Architecture Search for 3D Medical Image Segmentation

Woong Bae 1 , Seungho Lee 1 , Yeha Lee, Beomhee Park, Minki Chung, and Kyu-Hwan Jung ∗

VUNO Inc., Seoul, South Korea
{iorism, bleaf.lee, yeha.lee, beomheep, brekkanegg, khwan.jung}@vuno.co

Presenter: Nyanye

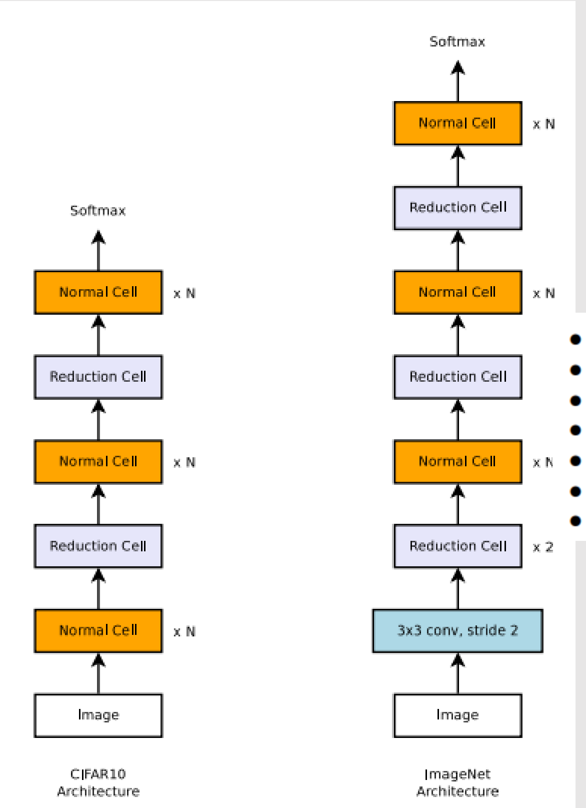
Introduction

- Research using deep neural networks for 3D medical image segmentation has exploded producing excellent methods such as U-Net
- Performance of methods is influenced by manual tasks : post-processing, hyperparameter tuning, and designing an optimal architecture.

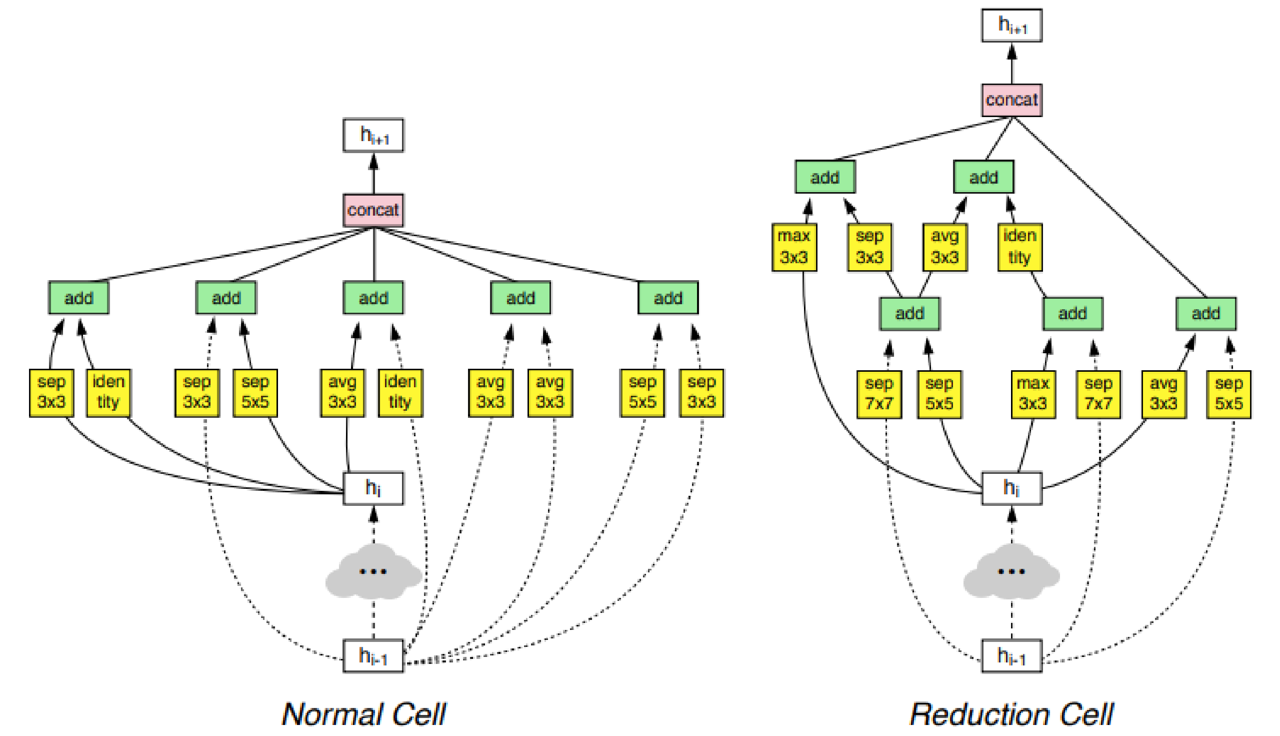
⇒AutoML, Neural Architecture Search

⇒Despite their success in natural image processing, these methods are difficult to apply to the segmentation of 3D medical image

Related Works - What is NAS



- identity
- 1x7 then 7x1 convolution
- 3x3 average pooling
- 5x5 max pooling
- 1x1 convolution
- 3x3 depthwise-separable conv
- 7x7 depthwise-separable conv
- 1x3 then 3x1 convolution
- 3x3 dilated convolution
- 3x3 max pooling
- 7x7 max pooling
- 3x3 convolution
- 5x5 depthwise-seperable conv



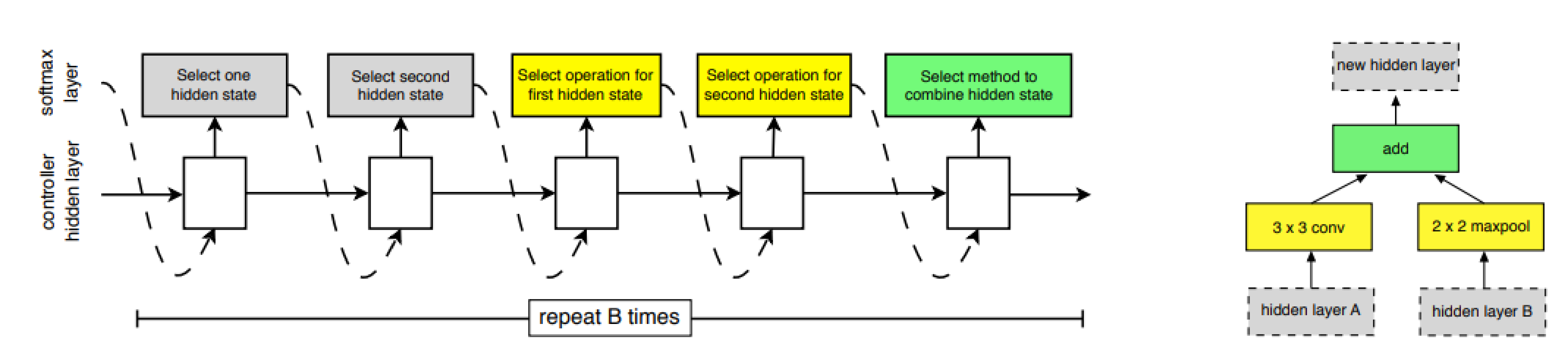
Baseline architecture

Building blocks

Convolution Cells

Learning Transferable Architectures for Scalable Image Recognition @ Google Brain

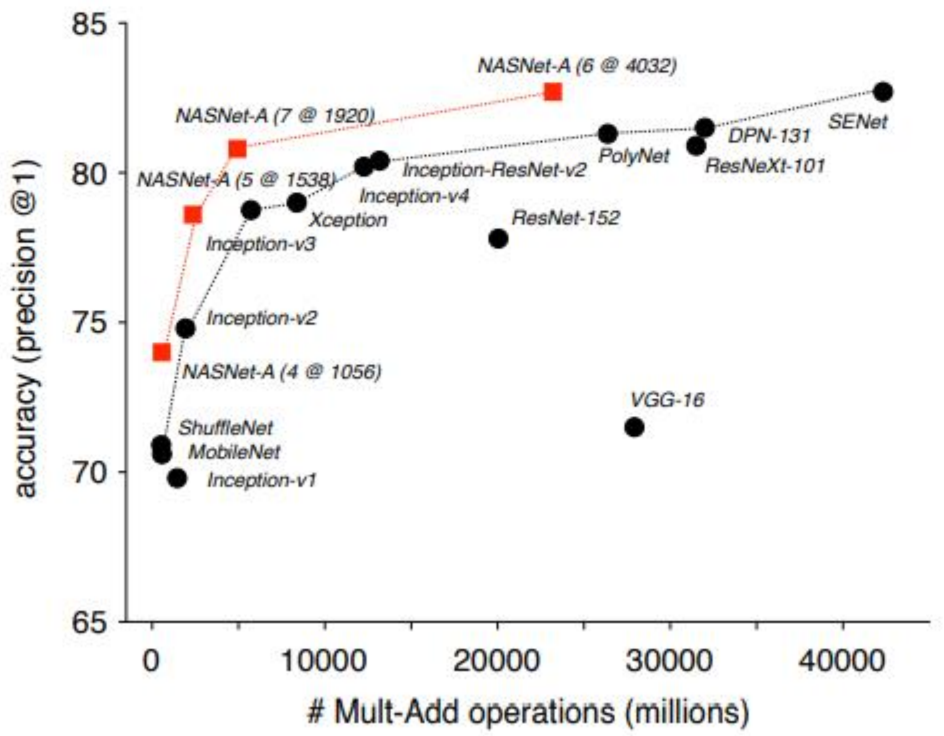
Related Works - What is NAS

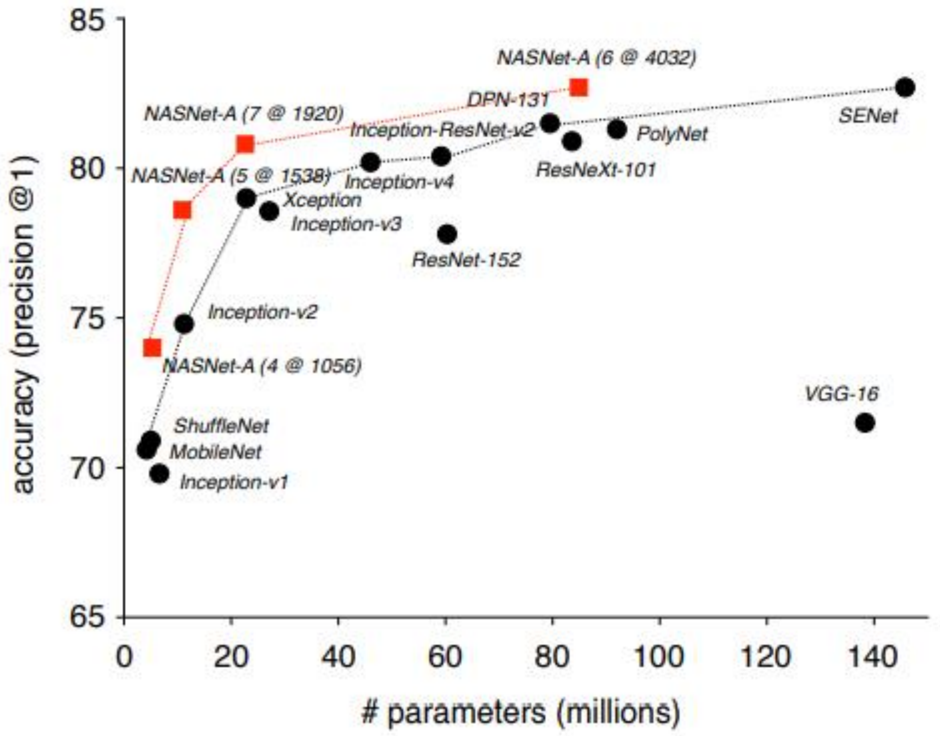


Controller architecture

Learning Transferable Architectures for Scalable Image Recognition @ Google Brain

Related Works - What is NAS





Better than human

Learning Transferable Architectures for Scalable Image Recognition @ Google Brain

Related Works - What is NAS

Requirements & Limitations

- Large amount of computational costs (FLOPS)
- Large amount of memory

  The controller RNN was trained using Proximal Policy Optimization (PPO) [51] by employing a global workqueue system for generating a pool of child networks controlled by the RNN. In our experiments, the pool of workers in the workqueue consisted of 500 GPUs.

Related Works - What is NAS

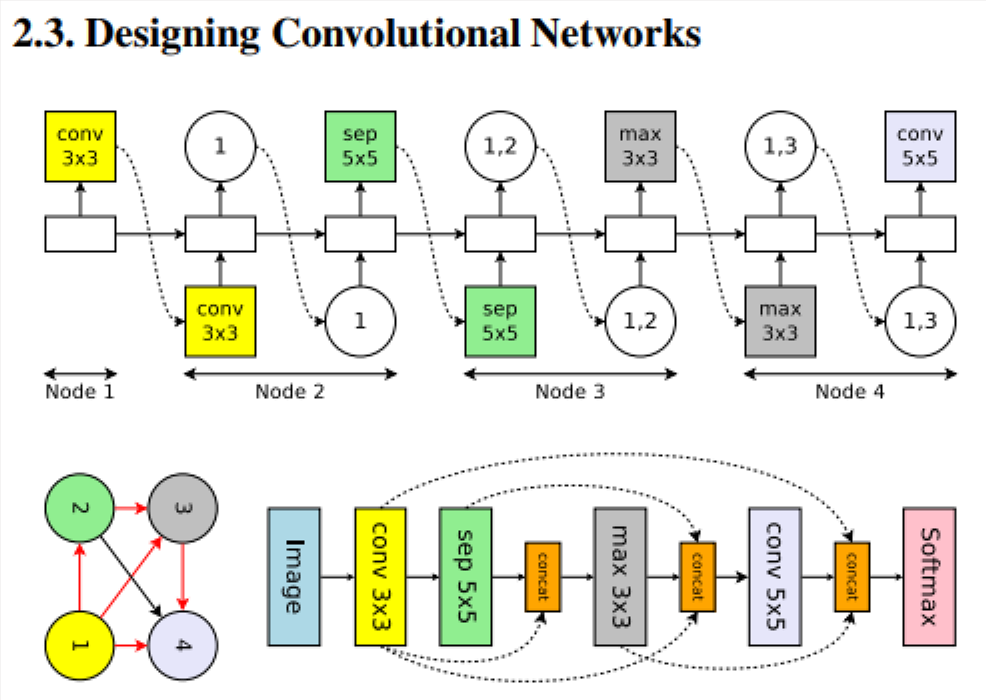
Main Tasks

- Image Classification (2D, ImageNet, CIFAR10)
- Image Segmentation (2D, PASCAL VOC)
- Text Classification (Penn Treebank)

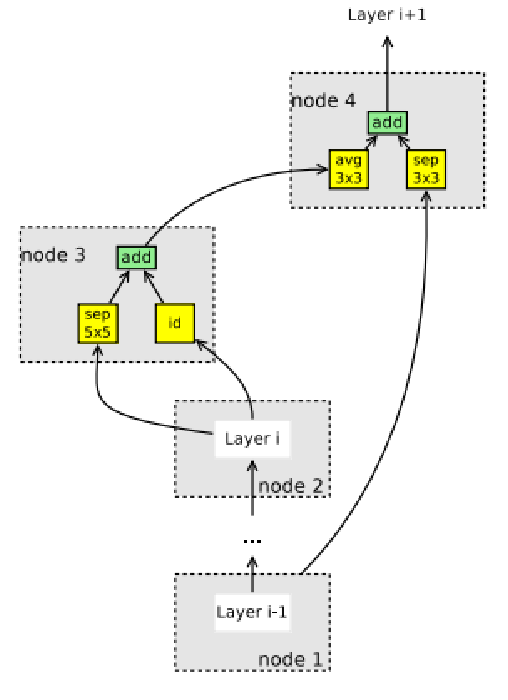
RONASMIS

Resource optimized neural architecture search for 3D medical image segmentation

- Short training time (1.39 days for 1GB dataset)
- Small amount of gpu computational power (1 RTX 2080Ti with 10.8GB)
- Focus on micro search space



macro search space



micro search space

RONASMIS

Resource optimized neural architecture search for 3D medical image segmentation

- Avoid retraining (continuously training child network during architecture search process)
- Addition-based skip-connection
- only includes elements that significantly impact the final performance of 3D medical image segmentation

RONASMIS

Resource optimized neural architecture search for 3D medical image segmentation

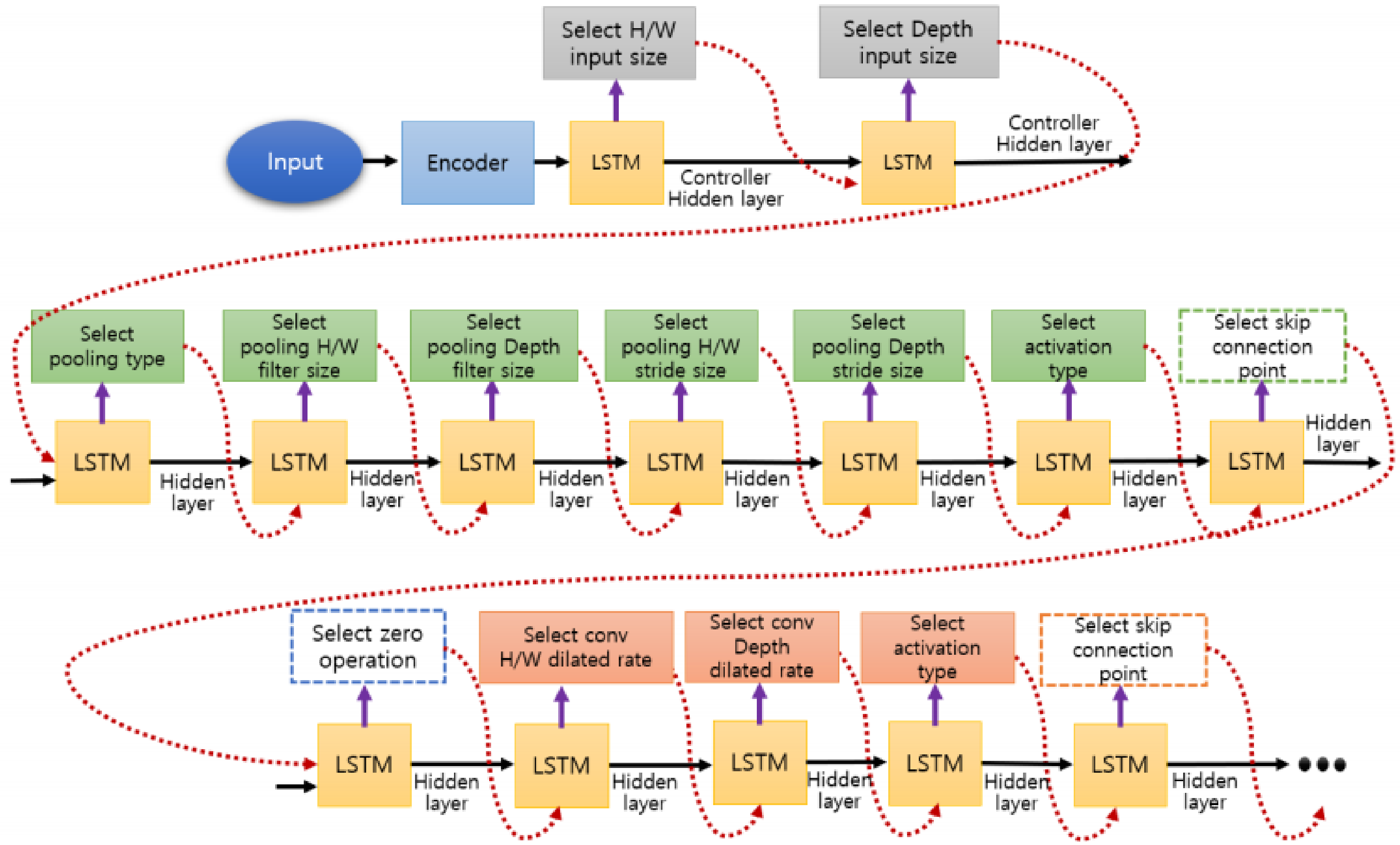


Table 1. Proposed search space

	Search space
Input Patch Size H/W	Eq. 1
Input Patch Size Depth	Eq. 2
3D Pooling Type	Max, Avg
3D Pooling H/W Filter	1, 3, 5
3D Pooling Depth Filter	1, 3, 5
3D Pooling H/W Stride	1, 2
3D Pooling Depth Stride	1, 2
Activation Type	ReLU, Leaky ReLU, ELU
Zero Operation	0, 1
3D Conv H/W Dilate rate	1, 2, 3, 4
3D Conv Depth Dilate rate	1, 2, 3, 4
Skip Connection Point	0 ~10

Fig. 1. The proposed RNN based controller and search space. The left figure shows the structure of the proposed controller with dotted boxes indicating that there is a point not applicable in certain sections. Table 1. shows the proposed search space.

RONASMIS

Resource optimized neural architecture search for 3D medical image segmentation

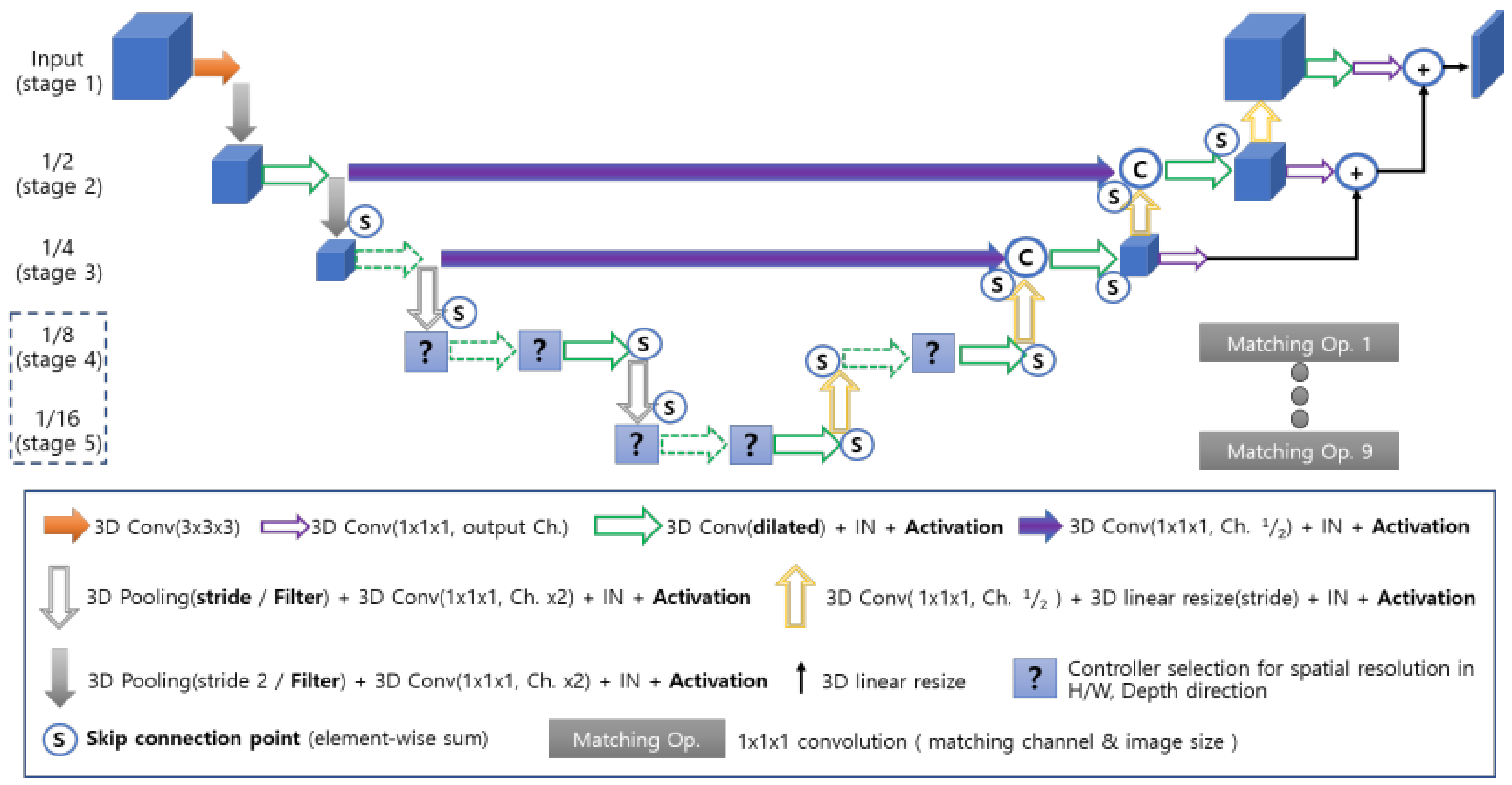


Fig. 2. The proposed base architecture.

The Proposed Base Architecture : We use an architecture modified from UNet[9] for our base architecture as shown in Figure 2. The architecture combines the decoder 1 × 1 × 1 convolution skip connections in DeepLabV3+[2] and the deep supervision scheme proposed in [3]. Batch normalization is replaced with instance normalization to account for GPU memory.

Training

Resource optimized neural architecture search for 3D medical image segmentation

- Use a parameter sharing based reinforcement learning proposed by ENAS
- Controller creates 20 child networks and observe validation patient-wise dice scores of each network

Evaluation

Resource optimized neural architecture search for 3D medical image segmentation

Datasets

- Medicaldecathlon (brain, heart, prostate)

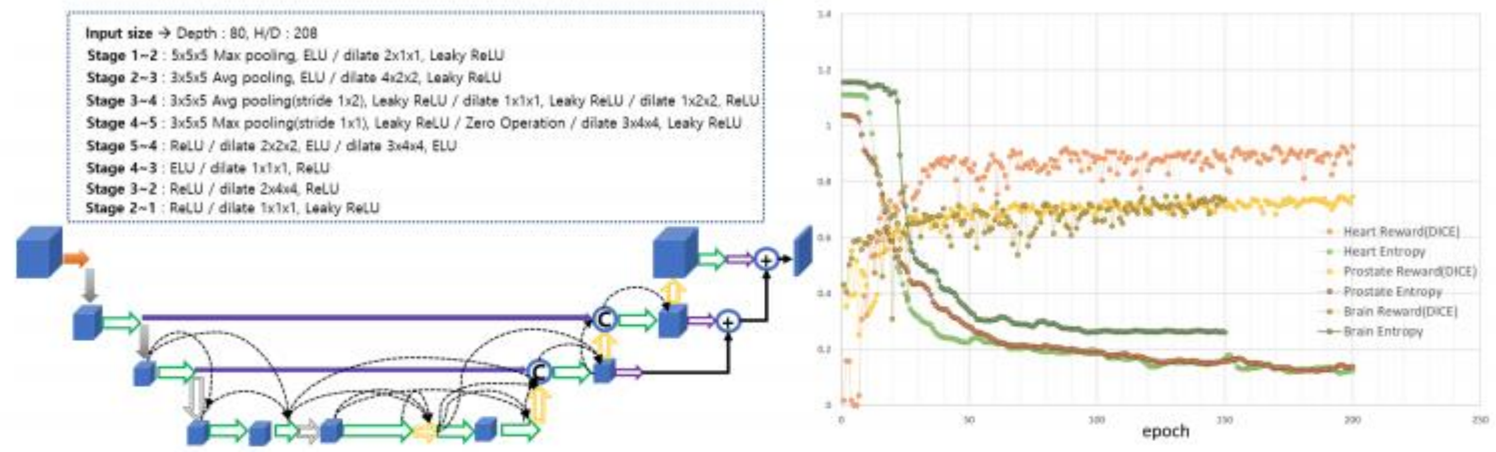


Fig. 3. **Left** : One of the optimal architectures chosen for the validation set of heart task. **Right** : The entropy of the controller and the reward of the selected architecture for each task.

Evaluation

Resource optimized neural architecture search for 3D medical image segmentation

Datasets

- Medicaldecathlon (brain, heart, prostate)

Table 2. Mean Dice score for Brain tumor, Heart, and Prostate 3D segmentation tasks. V.D.A, Ensemble, T.T.A, and P.P indicate whether Various Data Augmentation, Ensembling, Test Time Augmentation, and Post-Processing were used to obtain the final result.

	3D U-ResNet[10]	SCNAS[10]	SCNAS (transfer)	nnUnet[5]	RONASMIS (non fine-tunning)
Brain Tumor	71.61	72.04	-	74.00	**74.14**
Heart	89.60	89.99	90.47	92.70	**92.72**
Prostate	63.77	65.30	67.92	74.54	**75.71**
V.D.A, Ensemble, T.T.A, P.P.	No	No	No	Yes	**No**
Training GPU	Tesla V100	Tesla V100	Tesla V100	-	**One RTX 2080Ti**
Inference of network	Overlapped patch-wise	Overlapped patch-wise	Overlapped patch-wise	Weighted overlapped patch-wise	**One-shot**

Conclusion

Resource optimized neural architecture search for 3D medical image segmentation

- Apply NAS to 3D medical imaging segmentation tasks
- Resource-optimized NAS framework outperforms state-of-the-art results obtained by manual design in the 3D medical image segmentation challenge
- Achieves excellent performance without using various data augmentation, ensembling, T.T.A, and post-processing.