

## Letter

### Differentiable Automatic Data Augmentation by Proximal Update for Medical Image Segmentation

Wenxuan He, Min Liu, Yi Tang, Qinghao Liu, and Yaonan Wang

Dear editor,

This letter presents an automatic data augmentation algorithm for medical image segmentation. To increase the scale and diversity of medical images, we propose a differentiable automatic data augmentation algorithm based on proximal update by finding an optimal augmentation policy. Specifically, on the one hand, a dedicated search space is designed for the medical image segmentation task. On the other hand, we introduce a proximal differentiable gradient descent strategy to update the data augmentation policy, which would increase the searching efficiency. Results of the experiments indicate that the proposed algorithm significantly outperforms state-of-the-art methods, and search speed is 10 times faster than state-of-the-art methods.

Deep neural networks have made great progress in medical image segmentation and contributed to the rapid development of intelligent healthcare in recent years [1]. It is conducive to the surgical planning, pathological analysis, and disease diagnosis of patients. A variety of medical segmentation models have been proposed based on obtained training data [2], [3], the performance of which depends heavily on large-scale labeled data, while medical images are extremely difficult to obtain. This is because the increasing awareness of patients' privacy protection makes it more difficult to obtain patient case data, and many professional physicians need to spend a lot of time and effort labeling medical images for deep learning-based methods [4]. Moreover, multimodal data of diseases are characterized by a large span and low density due to the variety of diseases and various inspection methods. Therefore, it still has many challenges in medical image segmentation.

Data augmentation is an effective way for enlarging data size, which has been commonly employed in computer vision tasks with remarkable results [1], [2], [4]–[7]. More specifically, data augmentation is regularly utilized to address the problem of insufficient medical image data. Typical data augmentation methods in medical image segmentation include rotation, vertical flip, and random scaling [1], [4]. Since large differences exist in the samples for different types of diseases, resulting in different optimal data augmentation policies. And some experiments in [8] demonstrate that selecting an inappropriate data augmentation policy will reduce the segmentation accuracy. In other words, we need to redesign a reasonable and effective data augmentation policy for a new medical segmentation task, which requires professional experience and a variety of time to manually adjust the probability and the magnitude of the data augmentation operations. Additionally, the model [9] is also used to generate extra data, but it may not be optimal for the specific task because of the large variation in organs, tissues, and lesions and the connection of spatial contexts in medical images. Therefore, it is necessary to develop an automatic data augmentation

Corresponding author: Min Liu.

Citation: W. X. He, M. Liu, Y. Tang, Q. H. Liu, and Y. N. Wang, "Differentiable automatic data augmentation by proximal update for medical image segmentation," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 7, pp. 1315–1318, Jul. 2022.

The authors are with the College of Electrical and Information Engineering, Hunan University, Changsha 410082, and also with the National Engineering Research Center of Robot Visual Perception and Control Technology, Changsha 410082, China (e-mail: hwx@hnu.edu.cn; liu\_min@hnu.edu.cn; tyhnu@hnu.edu.cn; liuqinghao@hnu.edu.cn; yaonan@hnu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JAS.2022.105701

algorithm for medical image segmentation. In this letter, we propose an automatic data augmentation algorithm to find the optimal data augmentation policy. Firstly, a dedicated search space is designed for the medical image segmentation task. Then an optimization strategy is proposed to transform the automatic data augmentation problem into a single-step optimization problem, which is resolved by a proximal differential update. The experimental results on the publicly available liver tumor dataset, the publicly available pneumothorax dataset, and our private liver tumor dataset demonstrate that the proposed algorithm achieves state-of-the-art performance with the basic network architecture. Besides, the efficiency of the search policy of the proposed algorithm is improved by at least one order of magnitude compared to existing algorithms.

**Related work:** Deep neural networks have made significant progress in medical image segmentation. Currently researchers generally utilize UNet [2] and its variants nnUNet [1], UNet++ [4], etc. for medical image segmentation tasks.

Data augmentation is critical to the generalization performance and robustness properties of the network with only a relatively small number of training samples available. A combination of expert hand-designed augmentation strategies is commonly used for medical image segmentation. Valanarasu *et al.* [10] apply the horizontal flip, vertical flip, and addNoise to the task of brain anatomy segmentation, and Ronneberger *et al.* [2] apply shift, rotation, and elastic deformations to the microscopical images during preprocessing. However, these combinations of transformations are manually designed, which consumes the massive efforts of experts, and segmentation accuracy is hard to improve. Different from these methods, the first automatic data augmentation algorithm in [11] is proposed to find augmentation policies automatically based on reinforcement learning. However, this method takes 768 GPU hours by searching the probability of the augmentation policy. In [12], network weights and parameters can be optimized simultaneously with less time than in [11], but the time taken to find a reasonable policy is still hard to meet the requirements of the researchers [12]. In contrast, differentiable automatic data augmentation (DADA) [13] greatly reduces the time cost of searching policy in a differentiable way. Nevertheless, it also has the problem of the inaccurate searched policy and being not robust [13]. Moreover, medical images differ from natural images, which contain spatial contextual information, a smaller dataset size, and diverse morphologies of lesions and tissues. Thus, the transformations in natural images are hard to be adapted to medical images. To this end, we develop a novel and huge search space. Moreover, an efficient and accurate search algorithm is proposed. The proposed algorithm flow is as follows: The original training set is firstly divided into a training set and a validation set. Then a data augmentation sub-policy is sampled from the search space and applied to the training set. Meanwhile, in the search stage, the network weights are updated and used for forward propagation, when the validation set loss is used to update the data augmentation policy parameters by proximal differentiable gradient descent. The above process is repeated until the loss of the validation set is converged. Finally, in the retraining stage, we select  $N = 25$  sub-policies with the highest probability for expanding the scale of the original training set for network retraining, as shown in Fig. 1.

**Problem statement:** The automatic data augmentation (DA) algorithm finds the optimal DA policy for model training. The DA policy is composed of  $N = 25$  sub-policies  $s$  sampled from search space  $S$ , which has a vector parameter  $a$  (the probability of applying the sub-policies). A sub-policy contains two operations sampled from search space  $S$ , which are applied in sequence. Every operation has two parameters:  $b$  (the probability of applying the operation) and  $v$  (the magnitude of the operation).  $O$  is a set of image processing operations. The original training set is divided into a training set  $\Phi_{\text{train}}$  and a validation set  $\Phi_{\text{val}}$ . The optimization objective is to find the optimal parameters  $= \{a, b, v\}$  by minimizing the validation loss

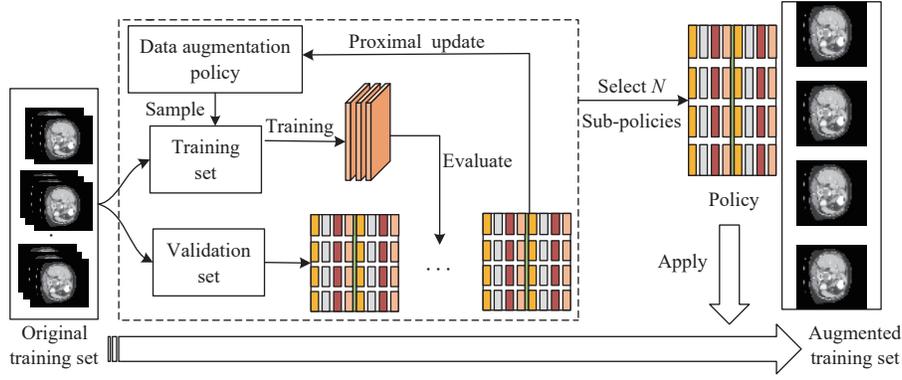


Fig. 1. The structure of the proposed algorithm.

$L_{\text{val}}$  when training loss  $L_{\text{train}}$  is converged via updating network weight  $w$ . Then the data augmentation policy optimization framework is represented as

$$\begin{aligned} \min L_{\text{val}}(w^*, \Phi_{\text{val}}) \\ \text{s.t. } w^* = \arg \min_w E(L_{\text{train}}(w, \varphi, \Phi_{\text{train}})). \end{aligned} \quad (1)$$

**Search space:** Since the automatic data augmentation algorithm needs to be employed for medical image segmentation, a dedicated search space is designed for diverse morphologies of tissues and lesions in medical images. We select 17 operations containing pixel-level and spatial-level transformations from albumentations [14], which are suitable for the variety of diseases and the uniqueness of the lesions in medical image, including contrast, randomgamma, brightness, clahe, gaussnoise, horizontalflip, verticalflip, elasticsearchtransform, opticaldistortion, gridfistortion, randomscale, rotate, shiftscalerotate, translate  $X/Y$ , shear  $X/Y$ . A data augmentation policy consisting of 136 sub-policies is sampled from the search space  $S$ . To increase the diversity of policies, we set the magnitude of operations in a continuous range. Therefore, we need to determine the scope of the interval. Besides the magnitude of the operation, the probabilities of applying these operations are also to be searched. The range of magnitudes and possibilities of the seventeen operations are illustrated in Table 1.

**Proposed search algorithm:** Based on (1), since the gradient of  $\varphi$  is not available directly,  $\varphi$  is difficult to be updated via gradient descent. To this end, we use Gumbel-Softmax [13], [15] to approximate the discrete parameter  $\varphi$  to continuous values, and then validation loss  $L_{\text{val}}$  is differentiable to  $\varphi$ . Finally, an efficient optimization method is introduced to update  $\varphi$  with discrete constraints and  $w$  by proximal gradient descent.

Firstly, the sub-policy selection and operation application are sampled from Categorical and Bernoulli distributions, respectively [13]. In order to determine which sub-policy is selected, probability  $h$  is sampled from the Categorical distribution  $H$ .

$$\bar{s}(x) = \sum_{s \in S} h_s s(x) \quad (2)$$

$$H_s = \frac{\exp(a_s)}{\sum_{s' \in S} \exp(a_{s'})} \quad (3)$$

where  $x$  denotes image.

To estimate the gradient of the DA policy parameters  $\{a, b\}$ , the Gumbel-Softmax reparameterization trick is utilized to make the parameters  $\{a, b\}$  differentiable. Next, we describe in detail the differentiable of relaxation of sub-policy selection  $a$  and operation application  $b$ . With the Gumbel-Softmax reparameterization, (2) could be represented as

$$\bar{s}(x) = \sum_{s \in S} k_s s(x) \quad (4)$$

$$k \in \text{Distribution}(a, \eta) = \frac{\exp((\log(a_s) + f_s)/\eta)}{\sum_{s' \in S} \exp((\log(a_{s'}) + f_{s'})/\eta)} \quad (5)$$

where  $f = -\log(-\log(\mu))$  with  $\mu \sim \text{Uniform}(0, 1)$ , and  $\eta$  is the

Table 1. The Designed Search Space for Medical Image Segmentation

Operation	Magnitude range	Probability range
Contrast	[0.1, 1.9]	[0, 1]
RandomGamma	[80, 120]	[0, 1]
Brightness	[0.1, 1.9]	[0, 1]
CLAHE	[1, 7]	[0, 1]
GaussNoise	[10, 50]	[0, 1]
HorizontalFlip	–	[0, 1]
VerticalFlip	–	[0, 1]
ElasticTransform	[0, 10000]	[0, 1]
OpticalDistortion	[-2, 2]	[0, 1]
GridDistortion	[-0.3, 0.3]	[0, 1]
RandomScale	[0.5, 1.5]	[0, 1]
Rotate	[-45, 45]	[0, 1]
Shift Scale Rotate	[-0.0625, 0.0625]	[0, 1]
Translate $X$	[-0.15, 0.15]	[0, 1]
Translate $Y$	[-0.05, 0.05]	[0, 1]
Shear $X$	[-0.15, 0.15]	[0, 1]
Shear $Y$	[-0.1, 0.1]	[0, 1]

temperature coefficient. Then, the reparameterization is applied to the Bernoulli distribution [13]

$$z \in \text{Distribution}(\eta, b) = \sigma\left(\log \frac{b}{1-b} + \log \frac{\mu}{1-\mu}\right)/\eta. \quad (6)$$

The magnitude parameter  $v$  is optimized by approximating the gradient. Since the magnitude parameter is not differentiable, we apply a gradient estimator [13], [16] to optimize it for an image  $\bar{x}$ ,

$$\frac{\partial L}{\partial v} = \sum_{i,j} \frac{\partial L}{\partial \bar{x}_{i,j}}. \quad (7)$$

With the reparameterization trick, we keep  $\{a, b\}$  continuous, which is optimized by gradient descent, but constrains the selections of  $\{a, b\}$  to be discrete. Therefore, we generalize (4) as below:

$$\bar{s}(x) = \sum_{s \in S} k_s s(x), \quad \text{s.t. } k \in M \quad (8)$$

where  $M = \{k \mid \|k\|_0 = 1, \text{ and } 0 \leq k_s \leq 1\}$ . While  $k$  is continuous to be differentiable,  $k$  is a one-hot vector when sampled. We formulate the applications of the combination of operations with the same discrete constraint. Then, the selection of operations  $z$  is sampled.

The proximal algorithm could quickly solve such convex optimization problems: A part of the objective function may be non-differentiable, but it can be split into the sum of a differentiable convex function and a non-differentiable convex function [17]. The proximal algorithm is difficult to obtain reasonable solutions with discrete constraints. And proximal algorithm cannot be applied with the additional constraint.

To solve the optimization problem as  $\min L_{\text{val}}(w^*, \Phi_{\text{val}})$ , following [17], a similar proximal algorithms is adopted. We optimize  $\varphi$  as a

continuous variable subject to the constraint  $Q_2$ . The discrete  $\bar{\varphi}$  is constrained during the iteration by the constraint  $Q_1$  derived from  $\varphi$ . We denote  $Q = Q_1 \cap Q_2$ , where  $Q_1 = \{\varphi \| k_s \|_0 = 1 \text{ and } 0 \leq \|z_i\|_0 \leq 2\}$  and  $Q_2 = \{\varphi | 0 \leq k_s, z_i \leq 1\}$ . Then, the proximal update is given by

$$\text{prox}_Q(\varphi) = m^* = \arg \min_m \frac{1}{2} \|\varphi - m\|_2^2, \varphi \in Q. \quad (9)$$

The constraint  $Q_1$  is meant to imply that  $m^*$  could be expressed as  $jd_i$ , where  $j$  is the parameter to be determined and  $d_i$  is a vector that contains the value 0 or 1. Let  $\varphi$  be an  $n$ -dimensional vector, and (9) is divided into  $n$  problems to be resolved

$$j_i^* = \arg \min_j \frac{1}{2} \|\varphi - jd_i\|_2^2, \text{ for } i = 1, \dots, n. \quad (10)$$

Let  $r_i = (1/2)(\varphi_i - j_i^*)^2$  and  $R = [r_1, \dots, r_n]$ , which contains  $n$  solutions for (9). To get the minimum, we should find  $i^* = \arg \max_i r_i$ , and  $m^* = j_{i^*}^* d_{i^*}$ . Then, it could be represented as

$$\text{prox}_Q(\varphi) = \text{prox}_{Q_1}(\text{prox}_{Q_2}(\varphi)). \quad (11)$$

The procedure of the proposed algorithm is given in Algorithm 1.

#### Algorithm 1 The Proposed Algorithm

- 1: **Initialize**  $\varphi_0 = \{a_0, b_0, v_0\}$ , create mixture sub-policy  $\bar{s}$  and operation  $\bar{O}$  with reparameterization.
- 2: **while**  $L_{\text{val}}$  not converged **do**
- 3:   Sample sub-policy  $s$ :  $k_t = \text{prox}_{Q_1}(k_t)$ ,  $\bar{z}_t = \text{prox}_{Q_2}(z_t)$ ;
- 4:   Update  $\varphi_{t+1} = \text{prox}_{Q_2}(\varphi_t - \nabla_{\bar{\varphi}_t} L_{\text{val}}(w_t, \bar{\varphi}_t))$ ;
- 5:   Update  $w_t$  by  $\nabla_{w_t} L_{\text{train}}(w_t, \varphi_{t+1})$  with backpropagation;
- 6: **end while**
- 7: Get a data augmentation policy consisting of top- $N$  sub-policies.

**Dataset:** We conducted experiments on three datasets including the publicly available LiTs dataset, the publicly available pneumothorax dataset, and our private liver tumor dataset.

The LiTs dataset is a publicly available liver tumor dataset containing liver and tumor labels. The dataset is provided by the 2017 Liver Tumor Segmentation Challenge organized by Medical Image Computing and Computer Assisted Intervention and IEEE International Symposium on Biomedical Imaging. The training set and the test set contain 130 CT scans and 70 CT scans, respectively.

The pneumothorax dataset is also a publicly available dataset containing only pneumothorax labels. The dataset is provided by Society for Imaging Informatics in Medicine-American College of Radiology (SIIM-ACR) pneumothorax segmentation. The training set contains 10 712 images and the test set has 1377 images.

The liver tumor dataset is an unpublished dataset containing the liver tumor label. The dataset is provided by a tertiary care hospital in Changsha, China, which contains 107 liver tumor CT scans of patients. This dataset is divided randomly into a training set and a test set in the ratio of 4:1.

**Evaluation metrics:** For the segmentation task, the common evaluation metrics are used [1], [4]: Dice coefficient and IoU.

$$\text{Dice} = \frac{2|X \cap Y|}{|X| + |Y|}. \quad (12)$$

$$\text{IoU} = \frac{|X \cap Y|}{|X \cup Y|}. \quad (13)$$

**Implementation detail:** As the test set labels of the LiTs dataset are unpublished, we need to submit the predictions of the test set. Since the number of submissions is limited, we compare only two well-known methods: UNet and nnUNet on the LiTs datasets. The Dice coefficient is the main evaluation metric on the LiTs and pneumothorax datasets.

Our experiments are divided into two parts: the data augmentation policy search stage and the network retraining stage. We divide the training set into two parts, half for optimizing the policy parameters and the remaining for updating the network weights in the policy search stage. Firstly, the top-25 sub-policies are obtained, when the validation loss is converged. Then the searched sub-policies are applied to the model for retraining. The model is trained using a cross-entropy loss function with a standard stochastic gradient

descent (SGD) optimizer and a learning rate of 0.001. The data augmentation parameters are optimized using an Adam optimizer with a learning rate of 0.005 and a weight decay of 0. All experiments are implemented based on PyTorch and the model is trained on an NVIDIA RTX 2080Ti.

**Comparisons with state-of-the-art methods:** We aim to search for a group of data augmentation sub-policies suitable for multiple medical image segmentation tasks, and then the sub-policies obtained from the search stage are applied to common medical segmentation networks, such as UNet [2], UNet++ [4], DenseUNet, MANet [18], nnUNet [1], and FPN [19]. DenseNet [20] achieves better performance with fewer parameters and computational costs by linking features on channels to achieve feature reuse. Densenet161 is chosen as the encoder for UNet, which is a variant of UNet, called DenseUNet. We conducted comparison experiments on liver tumor, LiTs, and pneumothorax datasets, respectively.

Comparative results of liver tumor segmentation and pneumothorax segmentation are shown in Tables 2 and 3. The results in Tables 2 and 3 show that the searched data augmentation policy could improve segmentation accuracy when applied to networks. And the best segmentation results we achieve outperform current methods, especially nnUNet [1], which is considered to be the best segmentation framework in medical image segmentation. It is worth mentioning that our algorithm is used for the basic network architecture, and still achieves the best segmentation accuracy. This demonstrates the effectiveness of our algorithm. Moreover, as shown in Table 4, [11] takes 768 hours, and [12] takes nearly 100 hours. And \* represents the results from the paper in Table 4. By contrast, our algorithm takes approximately 5 hours, which is at least one order of magnitude faster than [11] and [12]. Thus, the proposed algorithm has an advantage in speed.

Table 2. Comparative Results of Liver Tumor Segmentation (Liver Tumor and LiTs Dataset)

Dataset	Method	Dice (%)	IoU (%)
Liver tumor	UNet [2]	51.46	39.87
	UNet++ [4]	51.47	39.88
	MANet [18]	49.65	38.27
	DenseUNet [20]	55.19	43.53
	UNet [2] +Ours	57.95 (†6.49)	46.13 (†6.26)
	UNet++ [4] +Ours	53.74 (†2.27)	41.89 (†2.01)
	MANet [18] +Ours	60.48 (†10.83)	48.35 (†10.08)
LiTs	DenseUNet [20] +Ours	62.07 (†6.88)	49.40 (†5.87)
	UNet [2]	61.2	–
	nnUNet [1]	68.4	–
	UNet [2] +Ours	63.5 (†2.3)	–
	nnUNet [1] +Ours	69.1 (†0.7)	–

Besides, we adapt the automatic data augmentation algorithm DADA [13] from the natural image domain to medical image segmentation. Then we choose the traditional data augmentation policies [1], [10] as well as DADA to compare with the proposed algorithm. In the implementation, MANet [18] is a relatively novel network for liver tumor segmentation, selected as the baseline on the liver tumor dataset, and \* represents the data augmentation policy in the paper in Table 5. Additionally, UNet is selected as the baseline on the LiTs dataset, and \* represents the combinations of traditional data augmentation transformations which include brightnesscontrast, randomgamma, elasticsearchtransform, griddistortion, opticaldistortion, and shiftsclerotate in Table 6. As shown in Tables 5 and 6, the performance of the proposed algorithm exceeds that of DADA and traditional data augmentation policies, which confirms the efficiency of the proposed algorithm. Moreover, we also plot the variation curve of the validation set loss in the search stage. As shown in Fig. 2, the proposed algorithm solves the problem that the DADA algorithm is not robust and also converges faster than DADA. These experiments demonstrate the superior robustness and convergence performance of the proposed algorithm.

**Conclusions:** The small scale and lack of diversity of medical image datasets, as well as the unique shapes and morphologies of different lesions, remain challenges in medical image segmentation.

Table 3. Comparison of Pneumothorax Segmentation Results on Pneumothorax Dataset

Method	Dice (%)	IoU (%)
UNet [2]	67.10	27.29
UNet++ [4]	68.85	25.39
FPN [19]	70.56	29.09
UNet [2] +Ours	67.91 ( $\uparrow 0.81$ )	26.03 ( $\downarrow 1.26$ )
UNet++ [4] +Ours	69.60 ( $\uparrow 0.75$ )	29.02 ( $\uparrow 3.63$ )
FPN [19] +Ours	71.94 ( $\uparrow 1.38$ )	30.70 ( $\uparrow 1.61$ )

Table 4. Comparison of GPU Hours

Method	Based on RL [11]*	ASNG [12]*	Ours
GPU hours (h)	768	100	5

Table 5. Comparison Results of Traditional and Automatic Data Augmentation Algorithms on Liver Tumor Dataset

Method	Dice (%)	IoU (%)
Baseline [18]	49.65	38.27
Baseline+ [1] *	57.83 ( $\uparrow 8.18$ )	46.20 ( $\uparrow 7.93$ )
Baseline+ [10] *	52.01 ( $\uparrow 2.36$ )	40.63 ( $\uparrow 2.36$ )
Baseline+ DADA [13]	58.95 ( $\uparrow 9.30$ )	46.60 ( $\uparrow 8.33$ )
Baseline+ Ours	60.48 ( $\uparrow 10.83$ )	48.35 ( $\uparrow 10.08$ )

Table 6. Comparison Results of Traditional and Automatic Data Augmentation Algorithms on LiTs Dataset

Method	Dice (%)	IoU (%)
Baseline [2]	61.2	59.8
Baseline+ *	60.3 ( $\downarrow 0.9$ )	63.4 ( $\uparrow 3.6$ )
Baseline+ DADA [13]	60.0 ( $\downarrow 1.2$ )	63.2 ( $\uparrow 3.4$ )
Baseline+ Ours	63.5 ( $\uparrow 2.3$ )	61.2 ( $\uparrow 1.4$ )

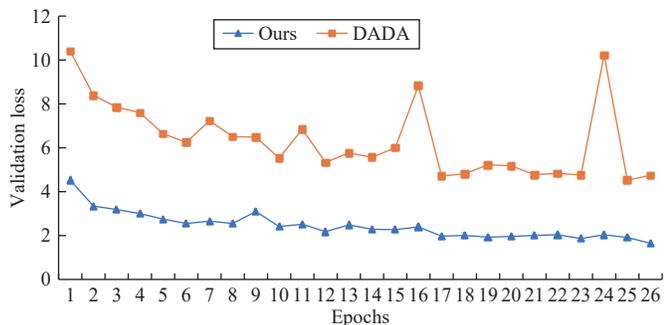


Fig. 2. Comparison results of validation set loss between ours (blue) and DADA (orange) in the search stage. Notably, the validation set loss of our method has been steadily decreasing, yet the loss curve of DADA oscillates during the search stage.

To solve the above problems, we firstly design a huge and novel search space that is suitable for most medical image segmentation tasks. Then, we propose a differentiable automatic data augmentation algorithm based on proximal update that searches for the optimal data augmentation policy. Finally, the comprehensive experiments demonstrate that the proposed algorithm outperforms state-of-the-art methods. And we could achieve the best segmentation accuracy by applying the searched data augmentation policy to basic network architecture. Additionally, the search speed of the proposed algorithm exceeds the best current automatic data augmentation algorithms in medical image segmentation.

**Acknowledgments:** This work was supported by the National Natural Science Foundation of China (62073126), the Hunan Provincial Natural Science Foundation of China (2020JJ2008), the Key Research and Development Program of Hunan Province (2022WK2011), the Science and Technology Program of Changsha (897202102345).

## References

- [1] F. Isensee, F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, 2021.
- [2] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Med. Image Comput. Comput. -Assisted Intervention*, 2015, pp. 234–241.
- [3] Y. Tan, M. Liu, W. Chen, X. Wang, H. Peng, and Y. Wang, "DeepBranch: Deep neural networks for branch point detection in biomedical images," *IEEE Trans. Med. Imag.*, vol. 39, no. 4, pp. 1195–1205, 2020.
- [4] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: Redesigning skip connections to exploit multiscale features in image segmentation," *IEEE Trans. Med. Imag.*, vol. 39, no. 6, pp. 1856–1867, 2020.
- [5] A. K. Bhandari, A. Ghosh, and I. Vinod Kumar, "A local contrast fusion based 3D Otsu algorithm for multilevel image segmentation," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 1, pp. 200–213, 2020.
- [6] X. Wang, M. Liu, D. S. Raychaudhuri, S. Paul, Y. Wang, and A. K. Roy-Chowdhury, "Learning person re-identification models from videos with weak supervision," *IEEE Trans. Image Process.*, vol. 30, pp. 3017–3028, 2021.
- [7] W. Jiang, M. Liu, Y. Peng, L. Wu, and Y. Wang, "HDCB-Net: A neural network with the hybrid dilated convolution for pixel-level crack detection on concrete bridges," *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5485–5494, 2021.
- [8] Y. Tang, B. Li, M. Liu, B. Chen, Y. Wang, and W. Ouyang, "AutoPedestrian: An automatic data augmentation and loss function search scheme for pedestrian detection," *IEEE Trans. Image Process.*, vol. 30, pp. 8483–8496, 2021.
- [9] T. Zhang, J. Wang, and M. Q.-H. Meng, "Generative adversarial network based heuristics for sampling-based path planning," *IEEE/CAA J. Autom. Sinica*, vol. 9, no. 1, pp. 64–74, 2022.
- [10] J. M. J. Valanarasu, V. A. Sindagi, I. Hacihaliloglu, and V. M. Patel, "KiU-Net: Towards accurate segmentation of biomedical images using over-complete representations," in *Proc. Med. Image Comput. Comput. -Assisted Intervention*, 2020, pp. 363–373.
- [11] D. Yang, H. Roth, Z. Xu, F. Milletari, L. Zhang, and D. Xu, "Searching learning policy with reinforcement learning for 3D medical image segmentation," in *Proc. Med. Image Comput. Comput. -Assisted Intervention*, 2019, pp. 3–11.
- [12] J. Xu, M. Li, and Z. Zhu, "Automatic data augmentation for 3D medical image segmentation," in *Proc. Med. Image Comput. Comput. -Assisted Intervention*, 2020, pp. 378–387.
- [13] Y. Li, G. Hu, and Y. Wang, "DADA: Differentiable automatic data augmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 580–595.
- [14] A. Buslaev, A. Parinov, E. Khvedchenya, V. I. Iglovikov, and A. A. Kalinin, "Albumentations: Fast and flexible image augmentations," arXiv preprint arXiv: 1809.06839, 2018.
- [15] E. Jang, S. X. Gu, and B. Poole, "Categorical reparameterization with gumbel-softmax," in *Proc. Int. Conf. Learn. Representations*, 2017.
- [16] Y. Bengio, N. Leonard, and A. Courville, "Estimating or propagating gradients through stochastic neurons for conditional computation," arXiv preprint arXiv: 1308.3432, 2013.
- [17] Q. Yao, J. Xu, W. Tu, and Z. Zhu, "Efficient neural architecture search via proximal iterations," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 6664–6671.
- [18] T. Fan, G. Wang, Y. Li, and H. Wang, "MA-Net: A multi-scale attention network for liver and tumor segmentation," *IEEE Access*, vol. 8, pp. 179656–179665, 2020.
- [19] T.-Y. Lin, P. Dollár, R. B. Girshick, K. He, B. Hariharan, and S. J. Belongie, "Feature pyramid networks for object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 936–944.
- [20] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 2261–2269.