HEURISTIC DROPOUT: AN EFFICIENT REGULARIZATION METHOD FOR MEDICAL IMAGE SEGMENTATION MODELS

22-27 MAY 2022
ICASSP 2022
SINGAPORE

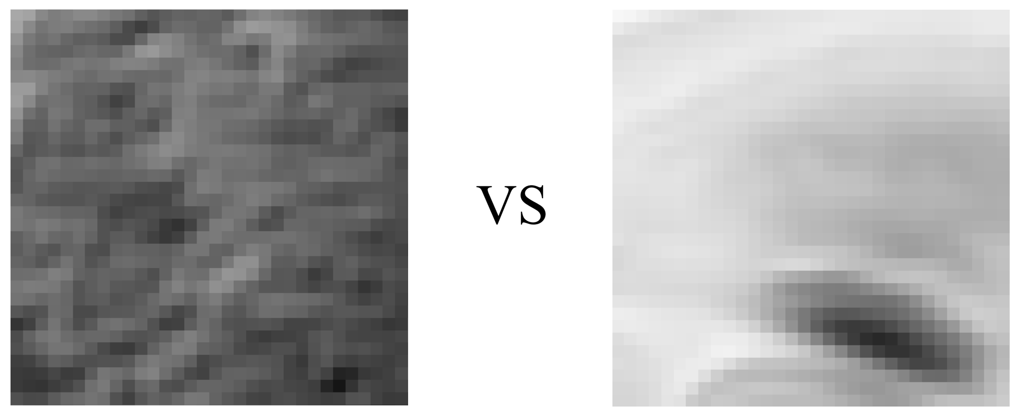
Dachuan Shi[1] Ruiyang Liu[2] Linmi Tao[2] Chun Yuan[1]
[1]Shenzhen International Graduate School, Tsinghua University, China
[2]Department of Computer Science and Technology, BNRist, Tsinghua University, China

PAPER ID
8799

INTRODUCTION

Deep learning models have been widely used in medical image segmentation. For some tasks, the annotation of the image is highly dependent on the expert knowledge, and thus many of these datasets are **small-scale**. The small-scale datasets tend to cause the **overfitting** problem.

This paper proposes an improved Dropout algorithm named Heuristic Dropout to tackle the overfitting problem for small-scale medical image segmentation datasets.



Traditional Dropout algorithm randomly drops units in the model with a certain probability, which mitigates the **co-adaptation** problem, and thus alleviate the overfitting. Here, we show two images in which the left one shows features with more severe co-adaptation problems compared to the right one.

In order to **efficiently drop** features suffering from more severe co-adaptation, we propose that **information entropy**

$$H(x) = -\sum_{x \in X} p(x) logp(x)$$

and **variance** can be used as heuristic rules to guide the drop operation. For more detailed explanation, please refer to our paper.

METHODOLOGY

Algorithm 1 Heuristic Dropout

Input: Output activations of a previous layer that contains c channels $A = [a_1, a_2, ..., a_c]$, drop rate p, k
Output: Activations after dropout $A^\star$

if *mode = Inference*, **then**
  Return $A^\star = A$
end
foreach a_i *in* A **do**
  Calculate information entropy of a_i by Algorithm 2 as e_i;
  Calculate variance of a_i as v_i;
  Calculate heuristic metric $m_i \leftarrow e_i + \frac{k}{v_i + \epsilon}$, where k is a hyperparameter, ϵ is a smoothing number;
end
Get rearranged A' by descending sort according to m_i;
for $j = 1; j \leq p \times c$ **do**
  $a_j^\star \leftarrow a_j' \otimes mask$
end
Return $A^\star$

METHODOLOGY Cont.

Combining the two heuristic rules of information entropy and variance, we can get the algorithm of Heuristic Dropout. We calculate the information entropy e_i and the variance v_i for each channel of input feature maps. We use

$$m_i \leftarrow e_i + \frac{k}{v_i + \epsilon}$$

as our heuristic metric, where k is a hyperparameter. We drop features according to the heuristic metric from largest to smallest.

Algorithm 2 Information Entropy for Quantized features

Input: A certain channel a_i of A, number of bins b
Output: Information Entropy e_i

Standardization by $tanh$, $a_i \leftarrow tanh(a_i)$;
Quantize by $round$, $a_i \leftarrow round(a_i \times b)$;
Calculate the corresponding histogram h_i according to the quantized a_i and bin number b;
Calculate information entropy e_i according to the definition, $e_i \leftarrow -\sum_{p(x) \in h_i} p(x) logp(x)$
Return e_i

Since the values of feature maps are continuous distributions, we have first to quantize the values and then calculate the information entropy based on the histogram as shown in the Algorithm 2.

EXPERIMENTS

To verify the effectiveness of the proposed algorithm, we conduct experiments on two different medical image segmentation datasets. These include:

1. A subset of Pancreas-CT Dataset.
2. A subset of BAGLS Dataset.

We adopt DICE value and IoU value as our evaluation metrics:

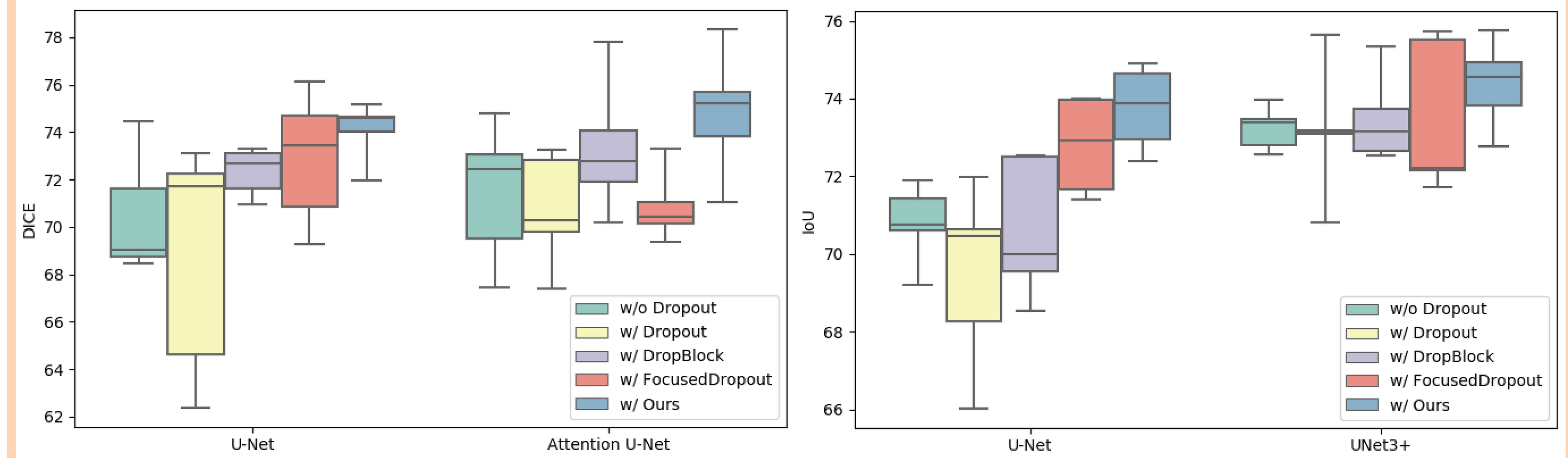
$$DICE = 2\frac{|X \cap Y|}{|X| + |Y|} \quad IoU = \frac{|X \cap Y|}{|X \cup Y|}$$

This table shows quantitative comparison results of our algorithm and other improved Dropout algorithms:

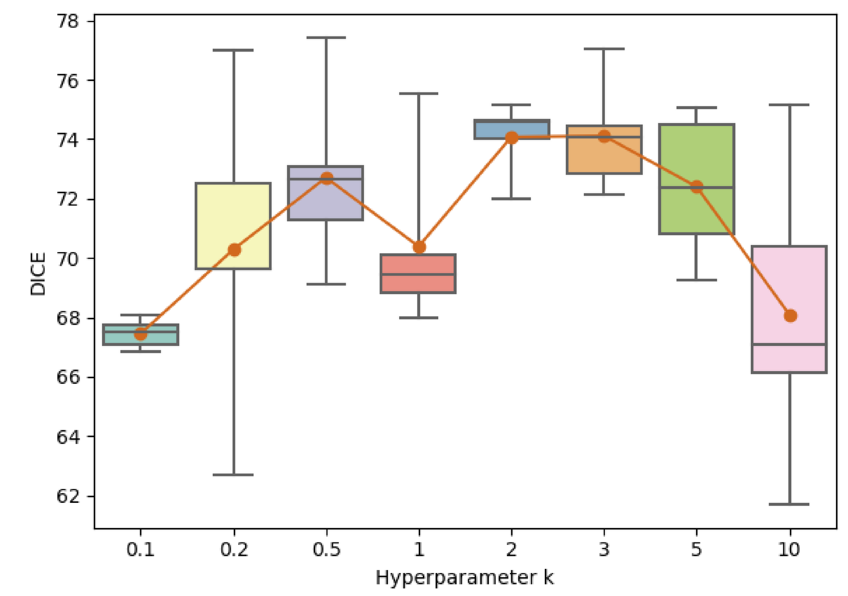
Method	Pancreas-CT Dataset				BAGLS Dataset			
	U-Net [1]		Attention U-Net [17]		U-Net [1]		UNet3+ [18]	
	Dice (↑)	Δ (↑)	Dice (↑)	Δ (↑)	IoU (↑)	Δ (↑)	IoU (↑)	Δ (↑)
Baseline [1, 17, 18]	70.40 ± 2.30	0	71.46 ± 2.62	0	70.78 ± 0.91	0	73.25 ± 0.50	0
+ Dropout [11]	68.82 ± 4.42	-1.58	70.7 ± 2.14	-0.76	69.48 ± 2.10	-1.30	73.17 ± 1.53	-0.08
+ DropBlock [13]	72.33 ± 0.91	+1.93	73.36 ± 2.56	+1.90	70.63 ± 1.10	-0.15	73.47 ± 1.78	+0.22
+ FocusedDropout [16]	72.88 ± 2.50	+2.48	70.85 ± 1.35	-0.61	72.79 ± 1.62	+ 2.01	73.49 ± 1.02	+0.24
+ Ours	$\mathbf{74.07 \pm 1.11}$	**+3.67**	$\mathbf{74.83 \pm 2.39}$	**+3.37**	$\mathbf{73.75 \pm 0.96}$	**+2.97**	$\mathbf{74.37 \pm 1.01}$	**+1.12**

The following figure illustrates the box plots of the experimental results corresponding to the above table:

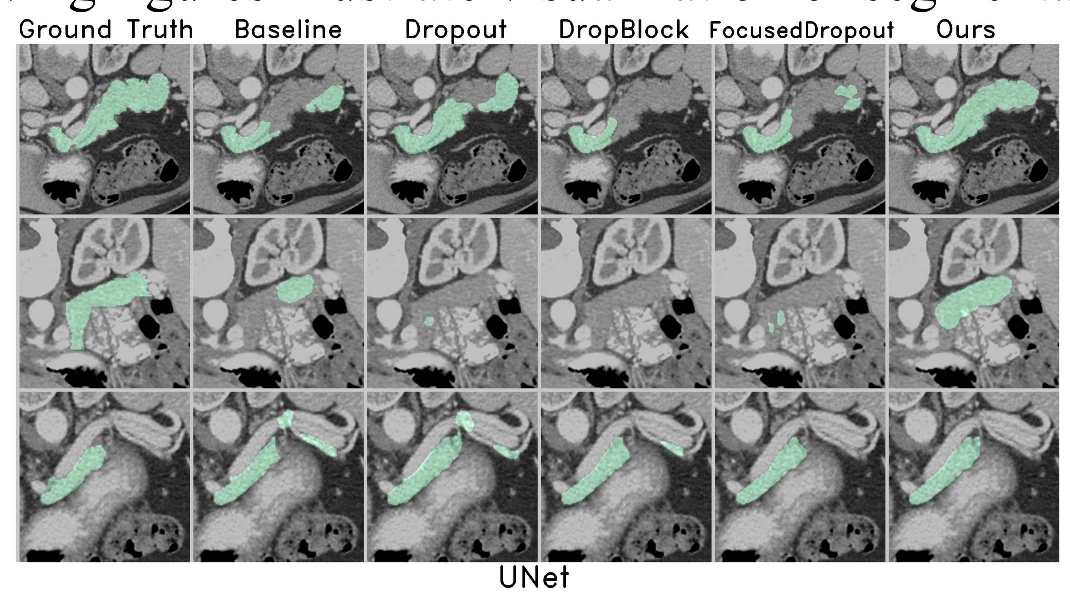
EXPERIMENTS Cont.

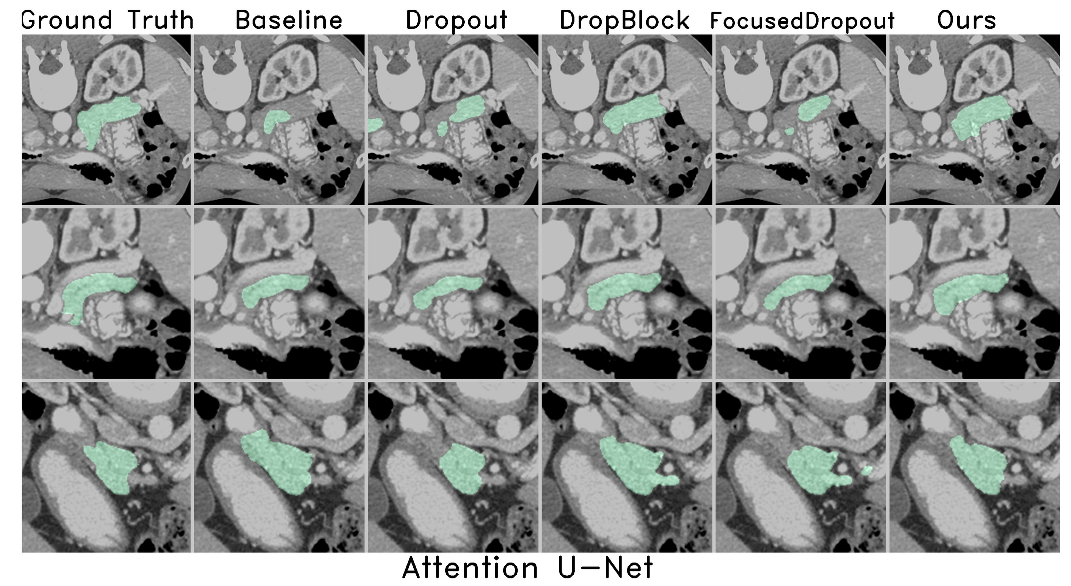


The effect of hyperparameter k in our algorithm is also investigated:



The following figures illustrate visualization of segmentation results:





CONCLUSION

- Our work suggests a novel Heuristic Dropout algorithm to address the **overfitting problem for small-scale medical image segmentation datasets**.
- Taking information entropy and variance as heuristic rules, our algorithm **alleviates the co-adaptation phenomenon more effectively** and thus **better mitigates the overfitting problem**.
- We will investigate the compatibility of our algorithm with natural images in future work.