CNN to Capsule Network Transformation

Takumi Sato
Department of Electrical and Electronic Engineering
Meijo University
Aichi,Japan
193427015@ccalumni.meijo-u.ac.jp

Kazuhiro Hotta
Department of Electrical and Electronic Engineering
Meijo University
Aichi,Japan
kazuhotta@meijo-u.ac.jp

***Abstract*—Capsule Network has been recently proposed which outperforms CNN in specific tasks. Due to the network architecture differences between Capsule Network and CNN, Capsule Network could not use transfer learning which is very frequently used in CNN. In this paper, we propose a transfer learning method which can easily transfer CNN to Capsule Network. We achieved by stacking pre-trained CNN and used the proposed capsule random transformer to interact individual CNN each other which will form a Capsule Network. We applied this method to U-net and achieved to create a capsule based method that has similar accuracy compared to U-net. We show the results on cell segmentation dataset. Our capsule network successfully archives higher accuracy compared to other Capsule Network based semantic segmentation methods.**

Keywords—Capsule Network, transfer learning, cell segmentation

I. Introduction

CNN has succeeded to accomplish high accuracy in numerous tasks. But CNN has some weak points such as disappearance of location information. Capsule Network [1] has been proposed to solve the problem. Capsule Network has been implemented to some tasks such as object segmentation [2], classification tasks [3] and so on. Capsule Network preserves location information of the feature map. In order to achieve this, Capsule network uses capsule structure, which is vector representation of feature map. In comparison, CNN uses scalar to represent feature maps. It also uses routing-by-agreement to perform pooling like operation and preserve location information at the same time. Recently, some works [4] show that routing-by-agreement algorithm behaves poor than uniform or random learning.

In this paper, we propose a transfer learning method that can easily transfer CNN pre-trained to capsule network. We aimed at the point where CNN uses scalars to represent feature maps and Capsule Network uses vector represent feature maps. Our method stacks the CNN pre-trained networks that uses scalar to represent feature maps. We also use transformer-based method called capsule random transformer to interact CNN pre-trained network each other that will form a vector and create a Capsule Network. Our proposed capsule random transformer creates the attention map between capsules. But calculation between all capsules will take a huge amount of memory which is very hard to calculate. Thus, we randomly took capsules and reduced the memory cost.

We used cell image segmentation dataset which consists of images of liver cell from mice. We show that our proposed method has higher mIoU than capsule network based method and has nearly as same mIoU compared to U-net [5].

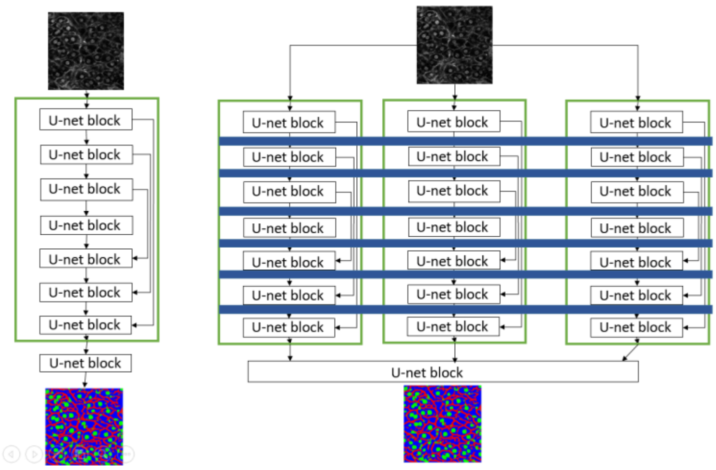

Figure 1 The left network represents the network architecture of U-net architecture used as pre-trained CNN. The right network represents proposed network that stacks pre-trained CNN and capsule random transformer represented in blue box**.**

This paper is organized as follows. In section 2,we introduce related works in capsule network. We introduce our proposed stacking pre-trained CNN and capsule random transformer in section 3. We show the experimental results of the propose method on cell image dataset which includes 3 classes in section 4. Conclusion and future works are described in section 5.

II. Related works

Capsule Network which has been proposed by S.Sabour et al. [1] to solve some of the weak points in CNN. SegCaps [2] has been proposed which uses only local capsules instead of all capsules. EnCapNet [3] has successfully achieved to accomplish high accuracy on ImageNet-1k using approximate routing. U-EnCapNet [6] uses approximate routing and U-net architecture and achieved higher mIoU than U-net. EnCapNet-3D [6] uses 3D convolution in approximate routing and achieved similar mIoU with very small parameters. SACN [7] uses self-attention as part of routing algorithm. Looking at recent researches of capsule network in segmentation, we can see that each one of them uses different routing algorithm and different network architecture. Thus, it is very difficult to apply transfer learning in capsule network. Our proposed method uses CNN network architecture as pre-trained network which has huge library of pre-trained network.

The most weak point of Capsule Network is routing-by-agreement. Some works [4,8] show that routing behaves poor than uniform or random learning. VB-routing [9] using Variational Bayes has been proposed, and it achieved the highest accuracy among all routing algorithms. We tackle to this problem by using self-attention used in SCAN and pre-trained weights which helps the routing-by-agreement be more close to supervised learning rather than unsupervised learning which is hard to train.

III. Proposed Method

Our proposed method uses U-net architecture as pre-trained CNN. We chose this network because the U-net architecture is very basic network in semantic segmentation. The proposed method is shown in Figure 1.First,we train a U-net and create the network illustrated on the left of Figure 1. Next, we stacked the networks together and connect it by capsule random transformer. Stacking CNN which uses scalar as feature map will form a vector and Creates a Capsule Network. The U-net block is constructed by 2D Convolution, Batch normalization, Dropout and ReLU function. This is repeated three times and constructs one U-net block.

Our another proposed method capsule random transformer is inspired by SACN [7]. Our biggest change to SACN is dividing with capsules and choosing capsule randomly. Dividing with capsule is used because location-wise self-attention cannot create an attention map between capsules. In order to accomplish this, we divided the input by capsules. Dividing by capsules creates the attention map between capsules but the memory usage is very huge. So we randomly took M capsule from C x H x W.

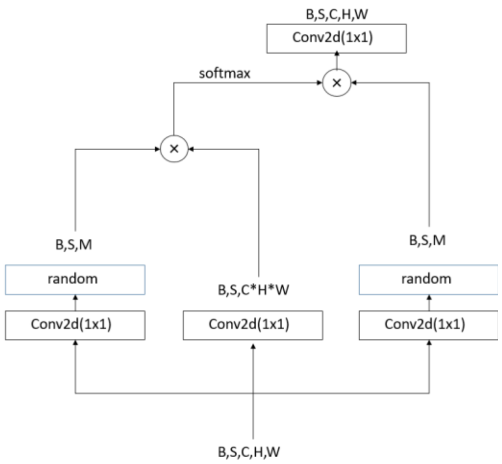

Figure 2 Network architecture of capsule random transformer. Each of the symbol corresponds to B: batch size, S: capsule number, C: capsule size, H: height, W: width, M: number of capsules chosen randomly

IV. Experiments

We used fluorescence image of the transgenic mice that expressed fluorescent markers on the cell membrane, cell nucleus and nucleus [8]. There are 50 images total and the size of the image is 256 x 256 pixels. We divided the dataset as 35 for train and 5 for validation and 10 for test. As an evaluation measure, we used intersection over union (IoU).We trained the networks with batch size 4 due to the memory consumption of the GPU. We also used class-balanced weight since the dataset is imbalanced.

Table 1 Comparison result on cell image dataset

	Membrane [%]	Nucleus [%]	Background [%]	mIoU [%]
SegCaps	11.67	25.13	55.88	30.89
U-EnCapNet	10.73	25.71	61.89	32.78
EnCapNet-3D	32.82	55.54	61.67	50.01
U-net	37.26	59.79	70.84	55.96
U-net-stack	36.28	59.63	73.85	56.58

Looking at Table 1, we can see that our proposed method has successfully achieved higher accuracy than the other capsule network based methods. Our method has similar mIoU compared to U-net.

V. Conclusion

In this paper, we proposed the transfer learning method that can easily transfer pre-trained CNN to Capsule Network. We also proposed capsule random transformer to interact CNN pre-trained network by each other. We applied this method to U-net and successfully achieved to outperform capsule network based segmentation methods and got similar performance compared to U-net.

Acknowledgment

This work is partially supported by JSPS KAKENHI Grant Number 18K111382 and 20H05427.

References

[1] S.Sabour, N.Frosst, G.E.Hinton Dynamic Routing Between Capsules, Advances in Neural Information Processing Systems, pp.3859-3869, 2017.

[2] Rodney LaLonde and Ulas Bagci. Capsules for object segmentation. In International Conference on Medical Imaging with Deep Learning, 2018.

[3] H.Li, X.Guo, B.Dai, W. Ouyang, X. Wang. Neural Network Encapsulation, European Conference on Computer Vision, pp.266-282, 2018.

[4] Inyoung Paik, Taeyeong Kwak, and Injung Kim. Capsule networks need an improved routing algorithm. ArXiv, abs/1907.13327, 2019.

[5] P.Fischer, and T.Brox. U-net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention, pp.234-241, 2015.

[6] T.Sato,K.Hotta, EncapNet-3D and U-EncapNet for Cell Segmentation Digital Image Computing: Techniques and Applications , 2019.

[7] A. Hoogi, B. Wilcox, Y. Gupta, and D. L. Rubin, ‘‘Self-attention capsule networks for object classification,’’ arXiv:1904.12483. [Online]. Available: http://arxiv.org/abs/1904.12483, 2019.

[8] A.Imanishi, T.Murata, M.Sato, K.Hotta, I.Imayoshi, M.Matsuda, K.Terai, A novel morphological marker for the analysis of molecular activities at the single-cell level, Cell Structure and function 43(2), pp.129-140, 2018.

[9] F. De Sousa Ribeiro, G. Leontidis, and S. Kollias, “Capsule routing via variational bayes,” in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp 3749-3756 ,2020..