Attention-guided Feature Pyramid Network for few-shot learning

Kai He, Erwin Bakker, and Michael S. Lew

Leiden University, Leiden, the Netherlands

Keywords: Attention module · Feature Pyramid Network · Few-shot learning.

1 Introduction

Few-shot learning (FSL) [1, 5, 6] has been a challenging task in machine learning area due to limited samples are available during the training process. In order to boost the model's performance, recent FSL researches utilize many generic techniques or tricks, such as self-supervision [4] and knowledge distillation [7]. In this work, we focus on integrating spatial information of feature maps with different scales to train the model. For this purpose, we propose a novel attention-guided feature pyramid network (AFPN) for few-shot image classification problems, which improves the robustness of the model based on the multi-scale feature fusion.

2 Methodology

AFPN includes two attention modules, a simple parameter-free attention module (SimAM) [9] and a cross-layer spatial attention module (CSAM). The two attention modules help accomplish multi-scale feature interactions, reserving more high-semantic information for the model's training. The overall training process of our framework is illustrated in **Fig** 1 .

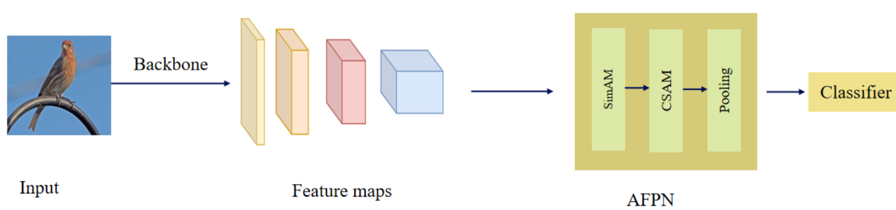

Fig. 1. Overall training process of AFPN.

Compared to other attention modules, such as Squeeze-and-Excitation Networks [2] and Convolutional Block Attention Module [8], SimAM has lower computational costs as it doesn't require extra parameters. Besides, SimAM can enhance the model's attention to crucial areas by adjusting weights of each neuron among feature maps. CSAM fuses feature maps output from different layers (with different resolutions) to form a pyramid structure, which can effectively perceive global information. Further, through feature fusion, we increase the resolution of low-level feature maps and more high-level semantic information can be kept at the same time, enabling the model to have a better performance in image classification tasks.

3 Results

We conduct experiments on a popular FSL benchmark dataset named FC100 [3], which is derived from CIFAR100 dataset, with 60 classes for training, 20 classes for validation and 20 classes for testing. We use ResNet12 as the backbone and use its results as the baseline. A N-way logistic regression classifier is trained for evaluation.The experimental results of baseline and our framework AFPN are presented in **Table** 1. The results show that our AFPN framework obtains 3% and 3.2% improvements on FC100 5-way 1-shot and 5-way 5-shot tasks, respectively. This demonstrates the effectiveness of AFPN for few-shot image classification problems.

Table 1. Experimental results on FC100 5-way tasks.

Model	1-shot	5-shot
Baseline	42.6	59.1
AFPN	45.6	62.3

4 Conclusion

In conclusion, in this paper we propose a novel AFPN framework for few-shot image classification problems, which aims at exploiting spatial information of feature maps to train the model and improving its performance. Besides, a meticulously designed attention module named CSAM is introduced to combine feature maps with different scales. By fusing these multi-scale feature maps, CSAM helps the model capture more detailed information. Overall, our framework achieves a good performance on the FC100 classification tasks and can be extended to other different domains. In the future, we will consider to use AFPN framework for object detection problems.

References

1. Gidaris, S., Komodakis, N.: Dynamic few-shot visual learning without forgetting. In: CVPR. pp. 4367–4375 (2018)
2. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: CVPR. pp. 7132–7141 (2018)
3. Oreshkin, B.N., López, P.R., Lacoste, A.: TADAM: task dependent adaptive metric for improved few-shot learning. In: NeurIPS. pp. 719–729 (2018)
4. Rajasegaran, J., Khan, S., Hayat, M., Khan, F.S., Shah, M.: Self-supervised knowledge distillation for few-shot learning. In: BMVC. p. 179 (2021)
5. Snell, J., Swersky, K., Zemel, R.S.: Prototypical networks for few-shot learning. In: NeurIPS. pp. 4077–4087 (2017)
6. Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P.H.S., Hospedales, T.M.: Learning to compare: Relation network for few-shot learning. In: CVPR. pp. 1199–1208 (2018)
7. Tian, Y., Wang, Y., Krishnan, D., Tenenbaum, J.B., Isola, P.: Rethinking few-shot image classification: A good embedding is all you need? In: ECCV. vol. 12359, pp. 266–282 (2020)
8. Woo, S., Park, J., Lee, J., Kweon, I.S.: CBAM: convolutional block attention module. In: ECCV. Lecture Notes in Computer Science, vol. 11211, pp. 3–19 (2018)
9. Yang, L., Zhang, R., Li, L., Xie, X.: Simam: A simple, parameter-free attention module for convolutional neural networks. In: ICML. Proceedings of Machine Learning Research, vol. 139, pp. 11863–11874 (2021)