Few-Shot Learning Using Class Augmentation in Metric Learning

Susumu Matsumi
Meijo University, Japan
193426018@ccmailg.meijo-u.ac.jp

Keiichi Yamada
Meijo University, Japan
yamadak@meijo-u.ac.jp

Abstract

We propose a few-shot learning approach based on metric learning in which the number of classes in the training data to perform metric learning is increased. The number of classes is augmented by synthesizing samples of imaginary classes at a feature level from original training data. The approach is evaluated using miniImageNet, and the effectiveness is demonstrated.

1. Introduction

Few-shot learning is a problem wherein new categories are learned from only a few samples. Modern few-shot learning uses prior knowledge from training data in addition to the target few-shot training samples. The main approaches are metric-learning, meta-learning, and data augmentation.

This paper proposes a few-shot learning method based on metric learning, focusing on increasing few-shot learning performance by increasing the number of classes in the training data to perform metric learning. The data augmentation approach used in previous few-shot learning works expands the number of few-shot training samples. In contrast, we expand the number of classes in the training data used for metric learning.

2. Method

Figure 1 outlines the proposed approach. The number of classes C_{trainset} in the original training set is expanded and the class-augmented data set is used in metric learning. By combining samples x_a in original training set class c_1 and x_b in class c_2 ($\neq c_1$), we generate a new sample x_{ab} of an imaginary class $l(c_1, c_2)$ that does not exist in the original training set.

A convolutional neural network (CNN) is used for metric learning and the trained CNN is used as the embedding function for few-shot learning. A new sample is synthesized by combining original samples in a feature map of the CNN during the CNN training. In the embedding space obtained

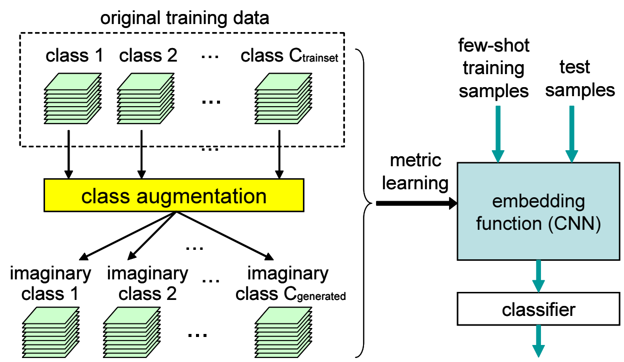

Figure 1. Overview of the proposed approach.

via metric learning, with the few-shot samples serving as the training data, the testing set class is identified with a classifier.

In the experiment, we use a CNN consisting of four convolution layers followed by three fully connected layers. Each convolution layer has a 3×3 convolution with 64 filters, followed by batch normalization, a ReLU activation function, and a 2×2 max-pooling. The first and second fully connected layers have 4096 and 512 units, respectively.

The samples are combined in the feature map of the fourth convolution layer by taking the element-wise maximum. The output from the first fully connected layer is used as an embedding vector. The nearest neighbor method using cosine similarity is used as the classifier.

The combining method is similar to the mixup [7], However, our method takes the maximum in the feature level and does not mix the class labels.

3. Experimental Results

The proposed method was evaluated using the miniImageNet [6] following the split introduced by [4]. The images were resized to 84×84 pixels.

First, the accuracy of few-shot learning when all classes in the original training set (C_{trainset} classes) were used in metric learning was compared with the accuracy for $m \times C_{\text{trainset}}$ classes ($m < 1$) selected at random from the original training set. We also compared it with the case

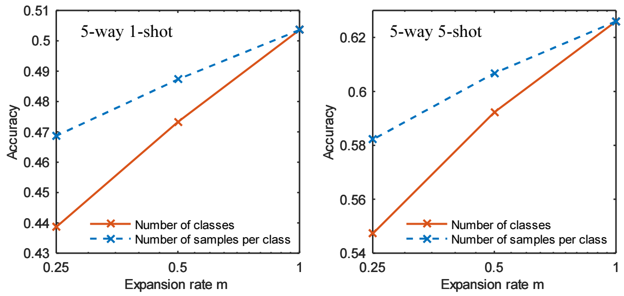

Figure 2. Accuracy of few-shot learning when the number of classes used in metric learning is decreased and when the number of samples per class is decreased by the same rate instead of reducing the number of classes. It can be seen that the number of classes has a greater impact on the accuracy of few-shot learning than the number of samples per class.

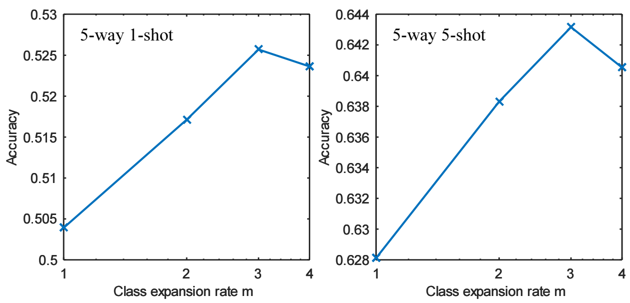

Figure 3. Accuracy of few-shot learning when the number of classes used in metric learning is increased by m times the original number of classes.

Table 1. Performance of the proposed method on miniImageNet.

method	5-way accuracy (%) 1-shot	5-shot
proposed ($m = 3$)	**52.57±0.57**	**64.32±0.47**
baseline ($m = 1$)	50.39±0.55	62.81±0.46

when the number of samples per class was decreased by the same rate instead of reducing the number of classes. Figure 2 shows the results. The solid lines represent the accuracy when the number of classes was reduced and the dashed lines represent the results when the number of samples per class was reduced. Results show that the number of classes has a greater impact on the accuracy of few-shot learning than the number of samples per class.

Next, we evaluated the performance of the proposed method on miniImageNet when the number of classes in the training set for metric learning was increased by m (> 1) times the original number of classes. Figure 3 shows the results. The accuracy of few-shot learning increases as the number of classes used in metric learning increases until $m = 3$. By increasing the number of classes by 3 times, the accuracy increases by 2.18 points from 50.39% to 52.57% for 5-way 1-shot learning. However, the accuracy decreases when the number of classes is increased by 4 times compared to when the number is increased by 3 times. This was probably because the difference between the classes was then smaller than the difference within each class. Table 1 summarizes the performance of the proposed method.

Table 2. Comparison of the performance on miniImageNet with some existing methods that used four convolutional layers with 64 filter channels.

Method	5-way accuracy (%) 1-shot	5-shot
DN4 [1]	51.24±0.74	71.02±0.64
TPN [2]	55.51±0.86	69.86±0.65
Dynamic Few-Shot [3]	**56.20±0.86**	**72.81±0.62**
Meta-Learn LSTM [4]	43.44±0.77	60.60±0.71
MAML [5]	48.70±1.84	63.11±0.92
Matching Nets [6]	43.56±0.84	55.31±0.73
Ours ($m = 3$)	52.57±0.57	64.32±0.47

Table 2 shows the comparison of the performance on miniImageNet with some existing methods that used four convolutional layers with 64 filter channels. Although our method used a simple classifier, it achieved an accuracy comparable to some of the existing methods.

4. Conclusion

This paper proposed a few-shot learning approach that uses an increased number of classes for the training data to perform metric learning. The approach was evaluated using miniImageNet. Although the proposed method is relatively simple, the method demonstrated good performance.

References

[1] W. Li, L. Wang, J. Xu, J. Huo, Y. Gao, and J. Luo, "Revisiting Local Descriptor Based Image-To-Class Measure for Few-Shot Learning," *CVPR*, 2019. 2

[2] Y. Liu, J. Lee, M. Park, S. Kim, E. Yang, S. J. Hwang, and Y. Yang, "Learning to Propagate Labels: Transductive Propagation Network for Few-shot Learning," *ICLR*, 2019. 2

[3] S. Gidaris and N. Komodakis, "Dynamic Few-Shot Visual Learning Without Forgetting," *CVPR*, 2018. 2

[4] S. Ravi and H. Larochelle, "Optimization as a Model for Few-Shot Learning," *ICLR*, 2017. 1, 2

[5] C. Finn, P. Abbeel, and S. Levine, "Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks," *ICML*, 2017. 2

[6] O. Vinyals, C. Blundell, T. Lillicrap, K. Kavukcuoglu, and D. Wierstra, "Matching Networks for One Shot Learning," *NIPS*, 2016. 1, 2

[7] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz, "mixup: Beyond Empirical Risk Minimization," *ICLR*, 2018. 1