

基于元度量的小样本学习方法

朱唯一*, 尹伟石

长春理工大学, 数学与统计学院, 吉林 长春

收稿日期: 2023年2月11日; 录用日期: 2023年3月6日; 发布日期: 2023年3月14日

摘要

针对小样本图像分类任务中样本量不足导致的提取特征不精确、网络性能下降等问题, 基于元度量学习方法提出一个用于处理小样本图像分类的学习方法。该方法通过对同一类别的每个样本特征赋予权值, 利用加权求和的方式得到该类的原型。权值由样本间的曼哈顿距离、欧氏距离和切比雪夫距离综合决定。最后采用三元损失函数使网络拉近同种类之间的距离, 增加不同种类之间的距离。实验结果表明, 该方法在5-way 1-shot和5-way 5-shot的设置下, 在MiniImageNet数据集上实验准确率为68.33%和81.75%, 在CUB数据集上实验准确率为73.76%和87.35%。对比结果显示, 该方法能有效处理小样本问题。

关键词

小样本学习, 元学习, 深度学习, 卷积神经网络, 图像分类

A Metametric Based Few Shot Learning Method

Weiye Zhu*, Weishi Yin

School of Mathematics and Statistics, Changchun University of Science and Technology, Changchun Jilin

Received: Feb. 11th, 2023; accepted: Mar. 6th, 2023; published: Mar. 14th, 2023

Abstract

Aiming at the problems of inaccurate extraction features and reduced network performance caused by insufficient sample size in the few-shot image classification task, a learning method for processing few-shot image classification is proposed based on the metametric learning method. This method gives weights to each sample feature of the same category and obtains the prototype of the class by weighted summation. The weight is determined by the comprehensive distance between samples,

*通讯作者。

the distance between the Manhattan, the Euclidean distance and the Chebyshev distance. Finally, a ternary loss function is used to bring the network closer to the distance between the same type and increase the distance between different types. Experimental results show that under the setting of 5-way 1-shot and 5-way 5-shot, the experimental accuracy of this method is 68.33% and 81.75% on the MiniImageNet data set, and the experimental accuracy on the CUB data set is 73.76% and 87.35%. The comparison results show that this method can effectively deal with few-shot learning.

Keywords

Few-Shot Learning, Meta-Learning, Deep Learning, Convolutional Neural Network, Image Classification

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

实际应用中, 某些领域数据有采集困难、采集成本高、标注样本难度大、时间成本高等问题, 如医疗诊断、机器故障检测、汽车喷漆等领域。因此, 小样本学习就是研究在样本量小的情况下, 通过少量的样本训练泛化能力强的预测模型。小样本学习大多基于数据增强、基于迁移、元学习等思想来解决。

基于数据增强的方法主要研究通过辅助数据对原本的数据集进行扩充, 或者是进行特征增强。但该方法有可能会在数据集中引入噪声数据, 对特征提取进行干扰, 使得分类精度降低。基于迁移的方法研究寻找多个任务间的相似性, 将训练好的模型应用到新的任务上。Oquab 等[1]在 ImageNet 数据集上对模型进行预训练, 对分类层进行改进, 将新的网络用于数据集分类。但以上方法通常需要大量的计算, 且如果任务间差距很大或者没有相似性, 那么效果并不理想。

基于元学习的方法研究如何学习, 它旨在让模型能在不同的任务中获取学习的能力来指导完成新任务。常见的元学习方法是样本投影到度量空间中, 利用度量函数计算样本间的相似度, 之后通过最邻算法预测类别分类[2]。原型网络[3]正是基于这个思想构建的, 该网络提取每类样本的均值作为该类的原型, 在分类过程中, 计算待分类的样本与各类原型的距离, 将距离最小的作为类别预测结果。然而对样本特征不明显或分布散的情况, 用样本均值代替类别特征并不准确。

本文研究小样本图像分类任务, 基于元学习中原型网络的不足, 提出加权聚合原型的概念, 并构建 WPNet 模型。考虑样本特征不明显、分散、受噪声影响大等情况, 根据距离对每个类别的样本赋予权值, 利用加权求和的方式得到该类的原型。实验结果显示, 本文的方法提高了图像分类准确率。

2. WPNet

本节提出一个用于解决小样本图像分类的网络模型 WPNet (Weighted aggregation prototype Networks), 模型示意图如图 1。

在小样本学习中, 设我们当前的数据集为 D , 其内部的样本形式为:

$$\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$$

其中 x_i 代表样本, y_i 代表对应标签。在原型网络中提取类别特征是对类别样本求均值, 本文提出加权聚

合原型代替求均值的策略。

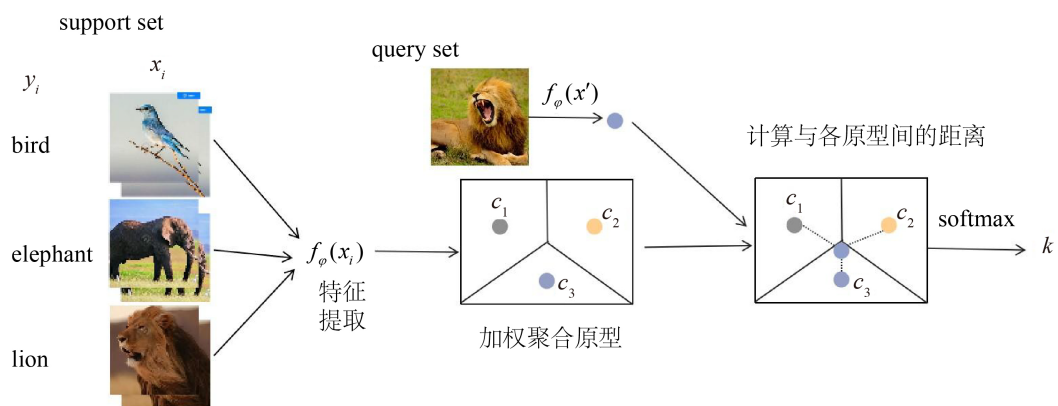


Figure 1. WpNet Model diagram
图 1. WpNet 模型示意图

2.1. 加权聚合原型

加权聚合原型是对同一类别的每一个样本特征赋予权值，最后再利用加权求和的方式得到该类的原型，见图 2。

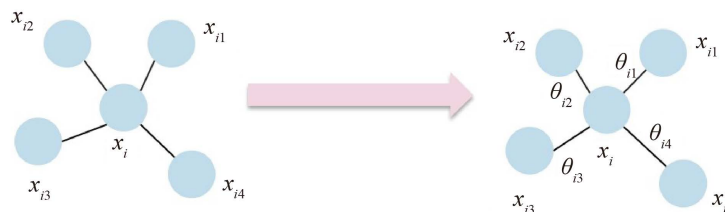


Figure 2. Weighted aggregation prototype diagram
图 2. 加权聚合原型示意图

权值的学习方法如下

$$\theta_i = h(l_{ij}, l'_{ij}, l''_{ij}), \quad (1)$$

其中

$$l_{ij} = |x_i - x_{ij}| + |y_i - y_{ij}| + |z_i - z_{ij}|, \quad (2)$$

$$l'_{ij} = \left(|x_i - x_{ij}|^2 + |y_i - y_{ij}|^2 + |z_i - z_{ij}|^2 \right)^{1/2}, \quad (3)$$

$$l''_{ij} = \max(|x_i - x_{ij}|, |y_i - y_{ij}|, |z_i - z_{ij}|). \quad (4)$$

h 为权值计算函数。通过特征提取网络得到样本的特征向量表示为 $f_\varphi(x_i)$ ，计算加权聚合原型：

$$c_n = \sum_{(x_i, y_i)} \theta_i \times f_\varphi(x_i) \quad (5)$$

其中 x_i 是输入样本， y_i 是对应的标签， $f_\varphi(x_i)$ 是样本特征向量， θ_i 是各样本的权重， n 表示第 n 类别的数据集。

2.2. 损失函数

本文的损失函数采用三元损失函数[4], 它可以使得网络拉近同种类之间的距离, 增加不同种类之间的距离。假设 x' 是数据集中要分类的一个样本, 则三元损失函数为:

$$L_{\text{large-margin}} = \frac{1}{N_l} \left[\sum_{x_s \in S_s} \|f_\phi(x') - f_\phi(x_s)\|_2^2 - \sum_{x_d \in S_d} \|f_\phi(x') - f_\phi(x_d)\|_2^2 \right] + \text{const}, \quad (6)$$

其中 $L_{\text{large-margin}}$ 表示三元损失函数, x_s 表示和 x' 是同类别的样本, x_d 表示和 x' 是不同类别的样本。则 (x', x_s, x_d) 构成了一个三元组, $L_{\text{large-margin}}$ 的目的是减少 x' 和 x_s 的距离, 增大 x' 和 x_d 的距离, const 表示一个常数, 最后的损失函数 L 为 softmax 与 $L_{\text{large-margin}}$ 的结合:

$$L = L_{\text{softmax}} + \lambda * L_{\text{large-margin}}. \quad (7)$$

2.3. WPNet 流程

下面描述下求解小样本图像分类问题的 WPNet 流程。对于 N -way, K -shot 任务, N 代表训练集中的类别数量, K 代表样本数量, $N_c < K$ 是每个任务批次中的类别数目, N_s 是每个任务中的支持集样本数量, N_q 是每个任务中的查询集样本数量。

首先从数据集 D 中随机选 N_s 个样本, 形成支持集 S 。同理, 随机选取 N_q 个样本, 形成查询集 Q 。

对于支持集 S 利用公式(1)~(5)求出每类的加权聚合原型。同理求出查询集 Q 中的加权聚合原型。

对于需要计算的是待分类样本属于哪一类, 使用 softmax 计算分类结果。

最终通过公式(6)和(7)计算损失函数 L 并更新参数。

以上就是整个算法的流程。

3. 实验结果与分析

本节为实验结果与分析, 通过在小样本数据集 MiniImageNet 和 CUB-200 上进行实验, 之后对比不同网络在数据集上的效果, 并对实验的结果做进一步的分析。

3.1. 数据集介绍

本节第一次实验采用的数据集为 MiniImageNet, 该数据集包含 60,000 张图片, 共 100 类, 每类下有 600 张图片, 训练集、验证集、测试集类别比例是 16:4:5。第二次实验采用的数据集是 CUB-200, 该数据集是一个鸟类数据集, 有 11,788 张图片, 共 200 类, 训练集、验证集、测试集类别比例是 2:1:1。实验前统一对数据集进行预处理, 保证样本的标准化。首先调整输入大小, 将图像 Resize 统一设置为 84×84 。接下来进行归一化处理, 图像像素值为(0, 255), 因此将图像中各位置的像素点值除 255, 得到了归一化后的数据集。

3.2. 实验设置

本文实验均在同一环境下使用同样的设置。操作系统为 CentOS Linux 7, 使用学习率为 0.001 的 Adam 优化器。在两个数据集上进行 5-way 1-shot 和 5-way 5-shot 的实验, 模型准确率为所有分类任务的平均准确率。由于本文是基于元学习提出的模型, 因此也是与基于元学习的模型进行对比实验。

3.3. MiniImageNet 实验结果与分析

本小节在 MiniImageNet 数据集上进行, 该数据集常用于检验小样本分类模型的性能。同时在相同的实验设置下, 对 8 个广泛应用的小样本模型进行实验对比, 结果如表 1 所示。

Table 1. Classification results on MiniImageNet
表 1. MiniImageNet 数据集上的分类结果

模型	5-way 1-shot (%)	5-way 5-shot (%)
MAML	52.80 ± 1.30	66.62 ± 0.91
Meta-learn LSTM	47.8 ± 0.70	60.8 ± 0.54
MTL	61.20 ± 1.80	75.50 ± 0.80
DC	62.53 ± 0.20	82.88 ± 0.42
Matching nets	63.08 ± 0.80	75.99 ± 0.60
TriNet	59.13 ± 0.95	76.72 ± 0.79
CAN	63.58 ± 0.48	79.44 ± 0.42
PN	63.39 ± 0.34	80.53 ± 0.14
Ours	68.33 ± 0.53	81.75 ± 0.96

对表 1 的实验结果分析可知,本文的方法在 5-way 1-shot 和 5-way 5-shot 上效果都很好,尤其是 5-way 5-shot 中,表现优秀,精确度最高为 81.75。对比其余网络, MAML [5]、Matching nets [6]、PN、Meta-learn LSTM [7]等都是基于元学习的网络模型。本文的 WPNet 比原型网络 PN 的精确度提高了约 5% 和 1.22%,这是由于原型网络中用均值代替原型的方法不适用类别样本分散或类别特征不明显的情况,WPNet 提出加权聚合原型,有效的改善了此情况,使得网络性能得到了提升。

3.4. CUB-200 实验结果与分析

本节实验设置与 3.2 中的一样,对 4 个广泛应用的小样本模型进行实验对比,结果如表 2 所示。

Table 2. Classification results on CUB-200
表 2. CUB-200 数据集上的分类结果

模型	5-way 1-shot (%)	5-way 5-shot (%)
MAML	55.36 ± 0.45	73.38 ± 0.56
Matching nets	51.87 ± 0.85	79.09 ± 0.57
TriNet	66.61 ± 0.45	82.10 ± 0.35
PN	66.09 ± 0.92	80.50 ± 0.58
Ours	73.76 ± 0.89	87.35 ± 0.76

分析表 2 可知,由于 CUB-200 都是鸟类图片,虽然不同种类的鸟不一样,但与在 MiniImageNet 数据集相比,分类任务相似,因此相同网络的模型准确率比表 1 中的高。本文的 WPNet 比 PN 在 5-way 1-shot 和 5-way 5-shot 任务上准确率分别提高了 7.67% 和 6.85%。在 5-way 1-shot 任务和 5-way 5-shot 任务中,网络表现最好的都是 TriNet [8],本文的 WPNet 比 TriNet 分别提高了 7.15% 和 5.25%。分析表明,本文提出的小样本学习方法能更好的提取类别特征,提高模型分类准确率,验证了 WPNet 能更好的处理小样本图像分类任务。

3.5. 与 PN 的浮点数对比实验

本文是基于原型网络提出的模型,是对原型提取策略进行改进,对网络结构没有修改。本节实验对

WPNet 和 PN 两个网络的浮点数进行估算对比, 对比浮点数相当于对比两个网络的计算量, 实验结果见表 3。

Table 3. Floating point number experiment

表 3. 浮点数对比实验

模型	WPNet	PN
浮点数(billion)	1.827	1.805

从表 3 中可以知道, PN 的浮点数为 1.804 billion, 本文提出的 WPNet 浮点数为 1.827 billion, 大约增加了 0.022 billion, 在计算机中增加的计算时间和难度约等于零。结合之前的两个实验结果, 说明本文提出的 WPNet 不仅效果好, 分类准确率高, 且几乎没有增加计算难度与时间成本, 有较高的实际应用能力。

4. 总结

本文考虑了小样本图像分类问题, 基于元学习中的原型网络, 提出加权聚合原型的概念, 优化了提取原型的方法。通过在 MiniImageNet 和 CUB-200 小样本数据集上进行实验, 与经典网络进行对比, 实验结果显示本文的模型表现优秀, 不论是 5-way 1-shot 任务或 5-way 5-shot 任务中准确率都最高。且本文通过与原型网络的浮点数对比, 表明本文模型复杂度并没有增加, 性能和实用性都很高。

基金项目

吉林省自然科学基金(20220101040JC)。

参考文献

- [1] Oquab, M., Bottou, L., Laptev, I., *et al.* (2014) Learning and Transferring Mid-Level Image Representations Using Convolutional Neural Networks. *Computer Vision and Pattern Recognition*, Columbus, OH, 23-28 June 2014, 321-334. <https://doi.org/10.1109/CVPR.2014.222>
- [2] 刘颖, 雷研博, 范九伦, 王富平, 公衍超, 田奇. 基于小样本学习的图像分类技术综述[J]. 自动化学报, 2021, 47(2): 297-315.
- [3] Snell, J., Swersky, K. and Zemel, R. (2017) Prototypical Networks for Few-Shot Learning. *Advances in Neural Information Processing Systems*, **30**, 1-10.
- [4] Yong, W., Wu, X.M. and Li, Q. (2018) Large Margin FewShot Learning. *CoRR*, **11**, 3-16.
- [5] Finn, C., Abbeel, P. and Levine, S. (2017) Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. *International Conference on Machine Learning, PMLR*, Sydney, 1126-1135.
- [6] Vinyals, O., Blundell, C., *et al.* (2016) Matching Networks for One Shot Learning. *Proceedings of the 30th International Conference on Neural Information Processing Systems*, Barcelona, 3630-3638.
- [7] Ravi, S. and Larochelle, H. (2017) Optimization as a Model for Few-Shot Learning. *International Conference on Learning Representations*, Toulon, 301-316.
- [8] Wang, F., Jiang, M., Qian, C., *et al.* (2017) Residual Attention Network for Image Classification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, HI, 21-26 July 2017, 3156-3164. <https://doi.org/10.1109/CVPR.2017.683>