

基于自监督与自适应感知关系网络的小样本图像分类

戴心杰^{1,2}, 郑家杰^{1,2}, 袁远飞^{1,2}, 王李进^{1,2}, 吴清寿³

(1. 福建农林大学计算机与信息学院, 福建 福州 350002)

(2. 福建农林大学智慧农林福建省高校重点实验室, 福建 福州 350002)

(3. 武夷学院数学与计算机学院, 福建 武夷山 354300)

[摘要] 关系网络是通过度量分析样本之间相似性的小样本分类方法,其固有的局部连通性限制了对样本全局特征的利用,并且在数据量较少时,模型的泛化能力不足. 提出一种混合自监督学习和自适应感知关系网络的小样本分类方法. 首先,通过结合自监督的实例级和场景级辅助任务、有监督的小样本分类辅助任务和自适应双相关注意任务提升模型特征表示和泛化能力. 其次,引入动态权重平均策略,用于自适应优化辅助任务之间的权重. 实例级辅助任务用于学习旋转样本未知类别的转移知识,场景级辅助任务确保不同旋转数据集的分类器预测结果一致性,小样本分类辅助任务则对扩展数据集进行有监督的分类预测平均,优化分类效能. 自适应感知关系网络任务通过自适应层对图像特征变化进行自动调节,通过双关联注意力机制增强特征间相互作用,促进关键特征辨识. 在数据集 miniImageNet、tieredImageNet 和 CUB-200-2011 上进行了验证,提出的方法在不同的骨干网络上都能较好地提升关系网络的分类性能,表明该方法是可行有效的.

[关键词] 小样本分类, 自监督学习, 自适应感知关系网络, 度量学习, 双关联注意力机制, 动态权重平均

[中图分类号] O643; X703 **[文献标志码]** A **[文章编号]** 1672-1292(2024)04-0068-11

Few-Shot Image Classification Based on Self-Supervised and Adaptive-Aware Relation Network

Dai Xinjie^{1,2}, Zheng Jiajie^{1,2}, Yuan Yuanfei^{1,2}, Wang Lijin^{1,2}, Wu Qingshou³

(1. College of Computer and Information Sciences, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

(2. Key Laboratory of Smart Agriculture and Forestry in Fujian Province University, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

(3. School of Mathematics and Computer Science, Wuyi University, Wuyishan 354300, China)

Abstract: Relation networks, as a method for few-shot classification through metric analysis of sample similarities, are limited by their inherent local connectivity which restricts the utilization of global features of samples. Furthermore, these networks demonstrate insufficient generalization ability when data is scarce. This paper proposes a hybrid method of few-shot classification combining self-supervised learning with adaptive perception relation networks. Firstly, it enhances model feature representation and generalization ability by integrating self-supervised instance-level and scene-level auxiliary tasks, supervised few-shot classification auxiliary tasks, and adaptive dual-relation attention tasks. Additionally, a dynamic weight averaging strategy is introduced to adaptively optimize weights between auxiliary tasks. Instance-level auxiliary tasks focus on learning transfer knowledge of unknown categories in rotated samples, scene-level tasks ensure consistency in classifier predictions across different rotated datasets, while few-shot classification auxiliary tasks average supervised predictions on expanded datasets, optimizing classification efficacy. The adaptive perception relation network tasks automatically adjust image feature variations through an adaptive layer, and enhance inter-feature interactions via a dual-relation attention mechanism, thereby promoting key feature recognition. The proposed method has been validated on the miniImageNet, tieredImageNet and CUB-200-2011 datasets, demonstrating its capability to significantly enhance the classification performance of relation networks across various backbone networks, proving the feasibility and effectiveness of the proposed approach.

Key words: few-shot classification, self-supervised learning, adaptive-aware relation network, metric learning, dual correlated attention mechanism, dynamic weight averaging

深度学习的成功归功于大量的有标注数据,然而在医疗、农作物病虫害等常见的生活领域中,常常存在着样本量少的问题,导致获取有标注的数据往往困难且费用昂贵.从有限经验中快速学习,小样本学习(few-shot learning, FSL)^[1-2]用来解决样本量少的问题.自小样本学习提出并成功应用于人脸识别^[3]、医学图像识别^[4-5]和语言处理^[6]等领域,与小样本学习相关的学习方法也相继提出.常见的小样本学习方法大致可分为:基于微调的方法^[7]、基于元学习的方法^[8-11]和基于度量学习的方法^[12-25].其中,基于度量学习的方法通过创建特征表示空间并在此空间内运用余弦或欧氏距离等度量手段来计算样本间的距离或者相似度,从而有效地区分不同类别.

关系网络(relation network, RN)^[14]是一种基于度量学习的小样本分类方法,通过度量分析样本之间相似性来进行分类.为了提升分类性能,一些改进性能的关系网络也相继提出. Abdelaziz 等^[20]提出了多尺度克罗内克积关系网络(multi-scale kronecker-product relation networks, MsKPRN),其采用多个尺度的信息来捕捉样本之间的不同关系特征,使网络能够更全面地理解和表征样本之间的关系. Li 等^[21]开发了具有自适应原型学习的关系网络(relation network with adaptive prototypical, ReNAP),提出的网络通过自适应的方式来学习类别的原型表示,而不是简单的求和或者平均,从而使网络可以得到更加准确的原型表示.另外一些工作将注意力机制^[22-24]引入关系网络,使网络能够有效地识别和加强对于分类任务关键特征的关注,从而提高分类的准确性和模型的适应性.

Wu 等^[19]提出的位置感知关系网络(position-aware relation networks, PARN),通过增加双相关性注意力机制来利用查询特征和支持特征每个空间位置的全局信息,提高了网络对关键特征的关注,并增强了特定任务上的性能.该注意力机制在处理具体问题时的效果显著,特别是在需要精细化特征理解和对齐的场景中.然而,现有学习方法在利用无标记数据和适应未见过类别方面仍有所限制. Zhang 等^[25]提出了实例级别和场景级别辅助任务(instance-level and episode-level pretext tasks, IEPT)框架,将自监督学习有效地融入到小样本学习中.在 IEPT 框架中,自监督学习不仅在实例级别进行(例如,识别图像的旋转角度),还涵盖了场景级别的学习.通过最大化来自不同场景的小样本分类预测的一致性来实现.

受 IEPT 的启发,本文在 PARN 算法的基础上引入自监督学习,提出了自监督与自适应感知关系网络(self-supervised and adaptive-aware relation network, SAARN)的小样本分类方法,通过无标记数据和适应未见过类别信息提升模型分类性能.

1 提出的方法

自监督与自适应感知关系网络,网络模型如图 1 所示.模型包括自监督辅助任务,小样本分类辅助任务,自适应感知关系网络任务,以及动态权重平均在调节辅助任务权重中的作用.

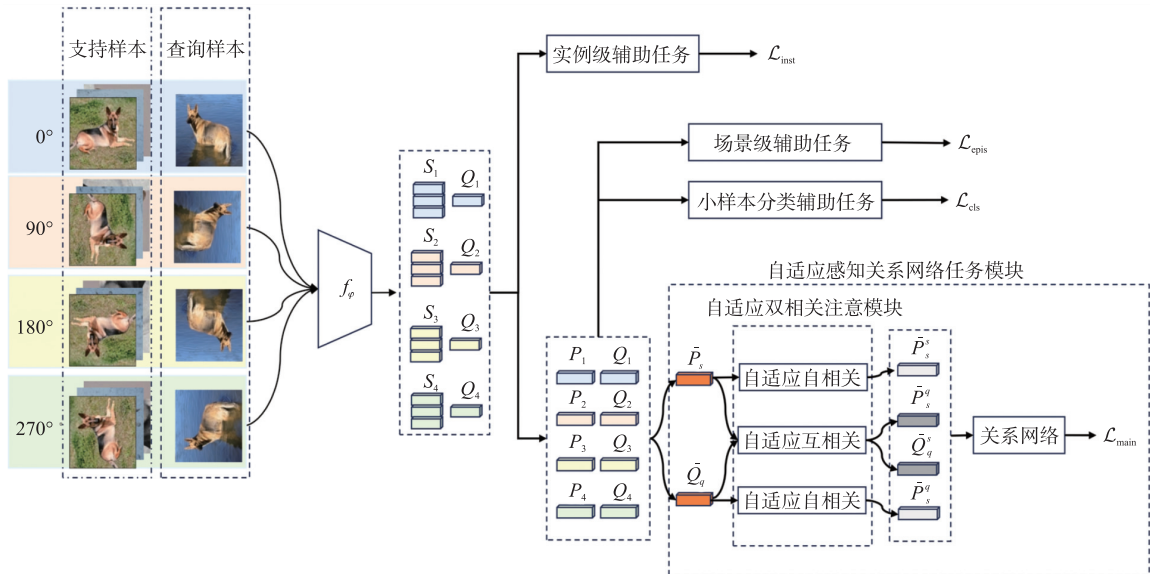


图 1 自监督与自适应感知关系网络模型

Fig. 1 Self-supervised and adaptive-aware relation network model

1.1 问题的定义

采用小样本图像分类问题常见的设置,问题被形容为 n -way k -shot 任务形式. 具体而言,在每个分类任务中,面临 n 个待分类的类别,每个类别中包含 k 个训练实例. 小样本学习的基本目的是在基本类别数据集 D_{base} 上进行模型训练,并通过验证类别数据集 D_{val} 来找到最优的参数类型,最后,在新类数据集 D_{novel} 上进行测试,以验证模型的有效性和适应性. 且 $D_{\text{base}} \cap D_{\text{val}} \cap D_{\text{novel}} = \emptyset$ 这一系列步骤确保了模型不仅能够学习现有的数据特征,还能够适应和处理之前未见过的新类别数据. 为提升模型在不同任务的泛化能力,小样本图像分类保持训练与测试过程一致. 因此,本文构建了一系列用于模型训练的小样本任务,每个任务集 $E = \{S, Q\}$ 包含一个支持集 S 和查询集 Q . 具体地,从基类数据集随机选取 n 个类别 C_e ,并从每个类别抽取 k 个样本作为支持集在剩下的样本中选 q 个样本作为查询集,

$$S = \{ (x_i^s, y_i^s) \mid y_i^s \in C_e, i = 1, \dots, M \} \quad (M = n \times k), \quad (1)$$

$$Q = \{ (x_i^q, y_i^q) \mid y_i^q \in C_e, i = 1, \dots, N \} \quad (N = n \times q). \quad (2)$$

在每次训练的迭代过程中,首先模型根据支持集 S 进行调整,以适应特定的任务或数据集. 然后,模型在查询集 Q 上进行性能测试,从而得出一个优化损失值. 这个损失值被用来更新模型的全局参数,以改善其在各种任务或数据集上的表现.

1.2 关系网络

采用关系网络^[14]作为基础的小样本学习模型. 该模型由特征嵌入模块 f_ϕ 和关系模块 g_ψ 组成. 特征嵌入模块负责从原始图像中提取复杂的特征,而关系模块则专注于学习一种非线性相似度度量,用以精确计算查询图像与类原型之间的相似度或关联得分. 该度量能够自动适应不同的数据,并捕获其独特的内在结构. 在每个场景的内循环中,关系网络利用这些模块来评估支持样本和查询样本之间的关联,从而动态调整度量标准,有效提高模型在小样本学习任务中的灵活性和效率.

首先,需要确定每个类别的原型. 可通过计算同一类别内所有支持集样本的特征平均值来实现,从而获得该类别的原型表示 P_c ,

$$P_c = \frac{1}{k} \sum_{x_i^s \in S} f_\phi(x_i^s) \cdot \mathbb{I}(y_i^s = c). \quad (3)$$

其中, \mathbb{I} 是指示函数,当 $y_i = c$ 时, $\mathbb{I}(y_i = c) = 1$, 否则为 0, 用于筛选属于类别为 c 的样本. 然后,需要对查询特征和类原型 P_c 进行特征拼接并且在关系模块 g_ψ 里面求得关系得分,

$$r_{q,c} = g_\psi(\text{concat}(f_\phi(x_i^q), P_c)). \quad (4)$$

其中 $\text{concat}()$ 表示两个特征进行拼接. 得到关系得分 $r_{q,c}$ 后,使用均方误差来计算关系得分和真实值之间的损失,

$$\mathcal{L}_{\text{rel}}(Q, P_c) = \frac{1}{|Q|} \sum_{i=1}^{|Q|} (r_{q,c} - y_i^q)^2. \quad (5)$$

其中, $|Q|$ 是查询集中的样本总数. 通过最小化损失函数,以便使模型学会更加准确地去预测新类别样本.

1.3 辅助任务模块

1.3.1 自监督辅助任务模块

为了增强模型对未标记数据的学习能力,引入自监督辅助任务模块^[25],该模块通过对图像执行一系列旋转变换来实现. 定义一个旋转操作集合 $R = \{r \mid r = 0, \dots, R-1\}$, 包含不同的旋转角度,例如 $0^\circ, 90^\circ, 180^\circ, 270^\circ$. 对于给定的任务集 $E = \{S, Q\}$, 其经过定义的所有旋转处理,形成扩展任务集以及旋转后的支持集和查询集,定义

$$E_{\mathcal{H}} = \{ \{S_{\mathcal{H}}^r, Q_{\mathcal{H}}^r\} \mid r = 0, \dots, R-1 \}, \quad (6)$$

$$S_{\mathcal{H}}^r = \{ (x_i^s, y_i^s, r) \mid y_i^s \in C_e, i = 1, \dots, M \}, \quad (7)$$

$$Q_{\mathcal{H}}^r = \{ (x_i^q, y_i^q, r) \mid y_i^q \in C_e, i = 1, \dots, N \}. \quad (8)$$

其中, $\{S_{\mathcal{H}}^0, Q_{\mathcal{H}}^0\}$ 代表了原始集,扩展任务集定义为

$$E_{\mathcal{H}}^r = \{ (x_i, y_i, r) \mid y_i \in C_e, i = 1, \dots, N, N+1, \dots, N+M \}, \quad (9)$$

且 $(i = 1, \dots, N) \in S_{\mathcal{H}}^r, (i = N+1, \dots, N+M) \in Q_{\mathcal{H}}^r$. 每个子任务 (x_i, y_i, r) 都存在着两个标签,分别是类标签 y_i

和旋转标签 r_i . 在生成扩展任务集 $E_{\mathcal{H}}$ 后,设计了两个自监督的辅助任务,旋转标签运用在实例级辅助任务,类标签运用在场景级辅助任务.

实例级辅助任务^[25]通过识别图像的旋转变化,有效地促进特征提取器学习到空间定位的关键知识,从而明显增强模型的泛化能力. 该任务采用的是自监督学习策略,模型通过识别图像所经历的旋转变化来判断物体是否被颠倒,从而获取对图像空间布局的基本认识. 当引入自监督旋转标签 r_i 时,定义了具体的映射 $f_{\varphi_{\text{rot}}}:x_i \rightarrow r_i$. 通过旋转分类器将输入图像 x_i 关联至其旋转标签. 基于对输入的子任务 (x_i, y_i, r) ,实例级的总旋转损失通过交叉熵损失计算.

$$\mathcal{L}_{\text{inst}} = \frac{1}{R(N+M)} \sum_{r=0}^{R-1} \sum_{(x_i, y_i, r) \in E_{\mathcal{H}}} - \ln \frac{\exp([f_{\varphi_{\text{rot}}}(f_{\varphi}(x_i))]r_i)}{\sum_{r'=0}^{R-1} \exp([f_{\varphi_{\text{rot}}}(f_{\varphi}(x_i))]r')}. \quad (10)$$

其中 $[f_{\varphi_{\text{rot}}}(f_{\varphi}(x_i))] \in \mathbf{R}^R$ 是旋转分量得分, $[]r$ 表示提取的第 r 个元素.

通过对小样本学习环境的探索,构建了场景级辅助任务^[25],其中主要目标是增强模型对图像旋转操作的不变性认知. 即使在多种场景下图像经历了旋转处理,但所属的类别标签仍然保持不变. 因此,关系网络应当对不同扩展任务每个实例生成一致的概率分布. 该一致性可通过 KL 散度来评估. 对于每个扩展任务集 $\{S_{\mathcal{H}}^r, Q_{\mathcal{H}}^r\} \in E_{\mathcal{H}}$,由式(3)、式(4)计算扩展后的类别原型 P_c^r 和关系得分 $r_{q,c}^r$,然后通过使用 Softmax 函数来计算查询集 $Q_{\mathcal{H}}^r$ 上的概率分布为 $\mathcal{P}_{\mathcal{H}}^r = [p_1^r, \dots, p_M^r] \in \mathbf{R}^{M \times n}$,其中 $p_i^r \in \mathbf{R}^n$ 表示 $Q_{\mathcal{H}}^r$ 中每个图像 x_i 的预测概率分布,它的第 c 个元素 $[p_i^r]_c (c=1, \dots, n)$ 为

$$[p_i^r]_c = \frac{\exp(\text{Softmax}((r_{q,c}^r)))}{\sum_{c'} \exp(\text{Softmax}((r_{q,c'}^r)))}. \quad (11)$$

计算出单个任务的概率分布后,可以得到 R 个扩展任务的平均概率分布,

$$\bar{p}_i = \frac{1}{R} \sum_{r=0}^{R-1} p_i^r. \quad (12)$$

最后,基于 KL 散度损失计算整体的场景级一致性正则化损失,

$$\mathcal{L}_{\text{epis}} = \frac{1}{RM} \sum_{r=0}^{R-1} \sum_{i=1}^M \text{mean}(p_i^r (\ln p_i^r - \ln \bar{p}_i)). \quad (13)$$

1.3.2 小样本分类辅助任务

在自监督辅助任务模块中,虽然两种辅助任务利用了无标签数据,但训练样本中的类别信息没有得到充分的应用. 本文采用有监督的分类任务,使用样本中的标注信息. 该任务基于关系网络的小样本图像分类流程,是计算扩展任务集上分类器的平均分类损失,

$$\mathcal{L}_{\text{cls}} = \frac{1}{R} \sum_{r=0}^{R-1} \mathcal{L}_{\text{isl}}(Q_{\mathcal{H}}^r, P_c^r). \quad (14)$$

1.3.3 动态权重平均

为了平衡各辅助任务的权重,引入了动态权重平均(dynamic weight average, DWA)^[26-28]. 该方法不依赖于网络的内部梯度,仅使用数值型的任务损失来动态调整任务权重,易实现. DWA 通过比较连续两个迭代的任务损失变化来更新权重,根据每个任务的学习进程调整其影响力,以此促进多任务学习中不同任务之间的平衡. 具体地,有 3 个辅助任务:实例级辅助任务(任务 1),场景级辅助任务(任务 2),小样本分类辅助任务(任务 3),对于任务 k ,权重为

$$\lambda_k(t) = K \frac{\exp\left(\frac{w_k(t-1)}{T}\right)}{\sum_{i=1}^3 \exp\left(\frac{w_k(t-1)}{T}\right)}, \quad (15)$$

$$w_k(t-1) = \frac{\mathcal{L}_k(t-1)}{\mathcal{L}_k(t-2)}.$$

其中, t 是一个迭代周期, $w_k(t-1)$ 是任务 k 在 $t-1$ 下的权重, $\mathcal{L}_k(t)$ 表示任务 k 在 t 下的平均损失值,参数 T

控制了权重调整的灵敏度, K 值为放缩系数. 通过这种方式, DWA 在多任务学习的环境中实现了多个任务权重的自适应调整. 最终, 辅助任务损失 \mathcal{L}_{aux} 是这三个任务损失的加权和,

$$\mathcal{L}_{\text{aux}} = \lambda_1(t) \cdot \mathcal{L}_{\text{inst}} + \lambda_2(t) \cdot \mathcal{L}_{\text{epis}} + \lambda_3(t) \cdot \mathcal{L}_{\text{cls}}. \quad (16)$$

1.4 自适应感知关系网络任务模块

1.4.1 自适应双相关注意模块

这个模块利用双相关性注意力和自适应权重机制, 使得网络能够全面利用特征在每个空间位置的全局信息, 并根据特征的实际重要性进行动态调整. 扩展后的类别原型 P_c^r 和查询集 $Q_{\mathcal{H}}^r$, 对其沿特征通道的方向求平均得到 \bar{P}_c^r 和 $\bar{Q}_{\mathcal{H}}^r$, 取两个特征分别为 \bar{P}_s 和 \bar{Q}_q .

自适应模块如图 2(a) 所示, 因为两个特征都需要输入自适应模块, 所以将其统一为 $Z \in \mathbf{R}^{C \times H \times W}$. 首先, 特征通过平均池化操作, 降低每个通道上的空间维度. 平均池化后, 特征张量被展平为一维向量后输入到接下来的全连接层. 展平后得到中间向量 Z_a , 通过全连接层进一步处理, 并经过 ReLU 激活函数来增加非线性. Z_a 代表经过初步维度变换和激活后的特征表示, 它的维度是 $H = C \times H_{\text{pool}} \times W_{\text{pool}}$, 全连接层输出维度为 N 的向量, 其中 $N = C/2$. 然后, 应用 Dropout 以减轻过拟合. 最后, 另一个全连接层和 Sigmoid 激活函数用于生成最终的输出自适应权重向量 O . 整个过程表示为,

$$Z_a = \text{ReLU}(FC(\text{AvgPool}(Z))), \quad (17)$$

$$O = \text{Sigmoid}(FC(\text{Dropout}(Z_a))). \quad (18)$$

自适应互相关性注意模块如图 2(b) 所示, 对于给定两个特征 $\bar{P}_s \in \mathbf{R}^{C \times H_1 \times W_1}$, $\bar{Q}_q \in \mathbf{R}^{C \times H_2 \times W_2}$. 首先, 通过运用两个共享的 1×1 卷积来调整通道数为 C' , 并生成两个特征 $\hat{P}_s \in \mathbf{R}^{C' \times H_1 \times W_1}$ 和 $\hat{Q}_q \in \mathbf{R}^{C' \times H_2 \times W_2}$. 然后, 通过重塑将特征调整为 $\hat{P}_s \in \mathbf{R}^{C' \times (H_1 W_1)}$ 和 $\hat{Q}_q \in \mathbf{R}^{C' \times (H_2 W_2)}$. 互相关操作 $g(\hat{P}_s, \hat{Q}_q)$ 通过计算两个特征 \hat{P}_s 和 \hat{Q}_q 中每一对的空间位置之间的关系, 生成了互相关注意力矩阵 A^c . 将两个特征用两个空间向量 $\{\hat{P}_{s,i}, \hat{Q}_{q,j}\} \in \mathbf{R}^{C'}$ 进行表示, 其中 $i \in \{1, \dots, H_1 W_1\}$, $j \in \{1, \dots, H_2 W_2\}$, 将原始的两个特征 \hat{P}_s 和 \hat{Q}_q 输入到自适应模块中, 得到自适应权重 O_s 和 O_q , 用于加权和调整互相关注意力矩阵, 使相关特征更加突出. 将重塑后的特征 \hat{P}_s 和 \hat{Q}_q 分别对进行自适应权重调整后的互相关注意力矩阵,

$$\bar{P}_s^q = c(O_s A^{cT}, \hat{P}_s), \quad (19)$$

$$\bar{Q}_q^s = c(O_q A^c, \hat{Q}_q). \quad (20)$$

其中, $c()$ 代表对两个特征图做矩阵乘法, 代表了输出特征 \bar{P}_s^q 综合了来自 \hat{P}_s 的全局信息并且分布到 \bar{P}_s^q 的每一个空间位置, \bar{Q}_q^s 对 \hat{Q}_q 也进行了相同的处理. 机制确保了在进行下一步卷积处理时, 网络能够对输入的每一个位置都有全面的理解, 包含了来自其他位置的信息, 从而提高特征的代表能力. 最后, 将两个特征重塑为 $\bar{P}_s^q \in \mathbf{R}^{C' \times H_2 \times W_2}$ 和 $\bar{Q}_q^s \in \mathbf{R}^{C' \times H_1 \times W_1}$ 并通过 1×1 卷积来调整通道数为 C .

自适应自相关注意模块如图 2(c) 所示, 该模块类似于自适应互相关性注意模块, 主要是对自身特征信息的关注, 对于给定的特征 \bar{P}_s , 输入一个 1×1 卷积, 该卷积与自适应互相关性注意模块的卷积是权重共享的. 通过自相关的操作生成自相关注意力矩阵 A^s , 并将特征 \bar{P}_s 输入到自适应模块中, 得到自适应权重向量 O_s , 由式(19)~式(20)计算得

$$A_{ii}^s = g(\hat{P}_{s,i}, \hat{P}_{s,i}) = \frac{\hat{P}_{s,i}}{\|\hat{P}_{s,i}\|} \frac{\hat{P}_{s,i}}{\|\hat{P}_{s,i}\|}, \quad (21)$$

$$\bar{P}_s^s = c(O_s A^{sT}, \hat{P}_s). \quad (22)$$

其中, \bar{P}_s^s 表示关注自身信息并分布到每个空间位置. 输入 \bar{Q}_q 可得 A^s . 通过重塑以及通过 1×1 卷积调整通道数为 C .

1.4.2 小样本分类主任务

通过上文中两个模块得到的四个特征 $\bar{P}_s^s, \bar{P}_s^q, \bar{Q}_q^s, \bar{Q}_q^q$, 进行级联操作得到特征 F_P^Q , 将得到的特征输入关系网络, 由式(4)、式(5)可得自适应感知关系网络模块构成的小样本分类主任务损失,

$$r_{Q,P} = g_\psi(F_P^Q), \quad (23)$$

$$\mathcal{L}_{\text{main}} = \frac{1}{|\bar{Q}_{\mathcal{H}}^r|} \sum_{i=1}^{|\bar{Q}_{\mathcal{H}}^r|} (r_{Q,P} - y_i^q)^2. \quad (24)$$

模型整体的训练损失为

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{aux}} + \mathcal{L}_{\text{main}}. \quad (25)$$

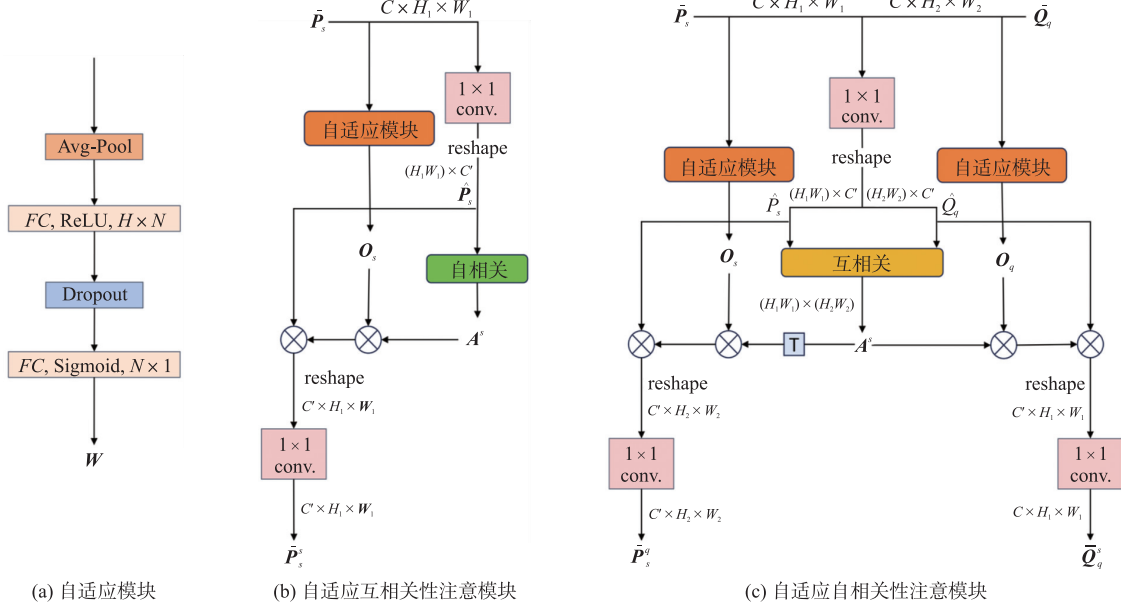


图2 自适应双相关注意模块

Fig. 2 Adaptive dual correlation attention module

1.5 算法流程

自监督与自适应感知关系网络训练流程如表1所示。

表1 自监督与自适应感知关系网络训练流程算法

Table 1 Training process algorithm for self-supervised and adaptive perception relation networks

算法1: 自监督与自适应感知关系网络训练流程

输入: 训练集 D_{base} , 初始化权重 O_1, O_2, O_3

输出: 可学习的参数集 M_θ

- 1: For epoch from 1 to Maxepoch do
- 2: 从 D_{base} 随机采样任务集 $E = \{S, Q\}$
- 3: 对 E 进行旋转生成扩展集 $E_{\mathcal{H}}$
- 4: For iter from 1 to IterationsPerEpoch do
- 5: 通过式(10)计算实例级辅助任务损失 $\mathcal{L}_{\text{inst}}$
- 6: 通过式(13)计算场景级辅助任务损失 $\mathcal{L}_{\text{epis}}$
- 7: 通过式(14)计算扩展集的平均分类损失 \mathcal{L}_{cls}
- 8: If iter% $t = 0$ do
- 9: For k from 1 to 3 do
- 10: 更新权重 O_k 为 DWA 计算得出的权重
- 11: end for
- 12: 通过式(23)求解主任务分类损失 $\mathcal{L}_{\text{main}}$
- 13: 更新总损失函数 $\mathcal{L}_{\text{total}}$ 为加权损失和
- 14: 使用 $\mathcal{L}_{\text{total}}$ 更新模型参数 M_θ
- 15: end for
- 16: end for
- 17: return M_θ

2 实验与分析

本文的实验验证是在3个被广泛采用的小样本图像分类数据集 miniImageNet、tieredImageNet 和 CUB-200-2011(简称 CUB)上进行的,提出的模型与现有的先进方法进行了比较分析. 通过一系列消融实验,验证了模型各个模块的有效性.

2.1 数据集介绍

本文为了评估模型性能,选取了两个源自广泛使用的大型图像数据库 ImageNet 的子集 miniImageNet 和 tieredImageNet,以及专门的鸟类图像数据集 CUB-200-2011. 数据集 miniImageNet 含有 100 个类别,每个类别 600 个样本,分为 64、16、20 个类别的训练集、验证集和测试集. TieredImageNet 则扩展了范围,提供了 779 165 张图像,平均每个类别约 1 281 张图像,训练集、验证集和测试集的类别数分别为 351、97 和 160. 另外,CUB-200-2011(简称 CUB)数据集,最初设计用于细粒度的鸟类图像识别,包含 11 788 张图像,涵盖 200 种不同的鸟类,其训练集、验证集和测试集分别由 100、50 和 50 个类别构成. 上述数据集的具体划分情况如表 2 所示.

2.2 实验设置和评价指标

本文沿袭了主流的研究方法,选取了两种广受认可的特征提取器 Conv4-64 和 ResNet-12 来进行 5way-1shot 和 5way-5shot 两种场景下的分类实验. 此外,引入了 Conv4-512 作为一个额外的特征提取器,以进一步增强实验. Conv4-64 以其简洁而有效的网络结构著称,包含 4 个卷积层,每层均配备了 64 个通道,以确保基础图像特征能被有效提取. 每个卷积层之后接有批量正则化和 ReLU 激活函数,增强了模型的泛化能力和非线性表达能力. 与此同时,Conv4-512 与 Conv4-64 在结构上相似,但每个卷积层包含 512 个通道,从而提供了更高维度的特征表示. 而 ResNet-12 则由 4 个残差块组成,每个残差块包括 3 个卷积层,分别配置有 64、128、256 和 512 个卷积核通道. 类似于 Conv4 系列,每个卷积层后有批量正则化和 ReLU 函数. 在此基础上,所有输入图像均被统一调整至 84×84 像素的尺寸,因而特征图的尺寸分别为 $64 \times 19 \times 19$ 、 $512 \times 19 \times 19$ 和 $512 \times 10 \times 10$.

本文在 miniImageNet、tieredImageNet 和 CUB 这 3 个数据集上进行了深入的实验分析. 实验采用 5-way 1-shot 和 5-way 5-shot 两种设置,每种设置下每个类别的查询样本数为 15 个. 在训练阶段,选用 Adam^[29] 作为优化器,初始学习率设为 0.001,权重衰减率为 0.5. 对数据集分别进行了 100、120 和 80 个训练周期,每个数据集在训练中使用了 5 000 个任务迭代优化模型. 在 DWA 中,辅助任务的初始权重 O_1, O_2, O_3 全部设置为 0.5,并将 DWA 中的温度系数 T 设置为 1, K 值设置为 1.5. 在测试阶段,从每个数据集中随机选取 600 个任务来评估模型性能,并报告 95%置信区间下的平均分类准确率.

2.3 实验环境

本文的编程实现采用 Pytorch 框架,所有实验环境为:16 vCPU Intel(R) Xeon(R) Platinum 8352V CPU @ 2.10GHz,NVIDIA RTX4090. 操作系统为 Linux.

2.4 实验结果分析

为了验证 SAARN 在小样本学习领域的有效性,将 SAARN 与一系列经典的小样本学习方法进行对比,包括度量学习的 MatchingNet^[12]、ProtoNet^[13]、RelationNet,元学习的 MAML^[9]、MetaOptNet^[10],以及融入自监督学习元素的 PN+rot^[24]、CC+rot^[24]、IEPT 等方法. 另外,为了公平实现,使用普通的特征提取器来替换 PARN 中的可变形卷积网络(即变成只在关系网络中加入双相关注意力模块). 在 miniImageNet 和 tieredImageNet 上进行常规图像分类,在 CUB 上进行细粒度图像分类. 使用 Conv4-64、Conv4-512 和 ResNet-12 作为骨干网络,对 SAARN 与其他主流小样本学习方法进行比较,包括基于自监督学习的 IEPT 和 CC+rot 等方法.

本文的 SAARN 模型融合了自监督学习和自适应感知机制,通过自适应双相关注意力模块,极大地提升了模型对图像全局信息及细节变化的处理能力. 这一机制在处理小样本学习场景中常见的不稳定性 and 类别间的细微差异方面表现出特别的适应性和效果. 其次,动态权重平均策略为模型提供了在多个辅助任务之间根据实际数据需要进行权重调整的能力,这不仅优化了学习过程,还显著提高了模型的泛化性能.

在数据集 miniImageNet 和 tieredImageNet 上,本文的模型展示了显著的性能优势^[30-35]. 具体而言,采用 Conv4-64 网络时,如表 3 所示,在 miniImageNet 的 5-way 1-shot 任务中,SAARN 模型比关系网络高出 5.68%,在 tieredImageNet 的 5-way 1-shot 和 5-way 5-shot 任务中性能提升分别为 6.46%和 5.24%. 当采用更大网络

表 2 数据集详情

Tabel 2 Dataset details

	miniImageNet	tieredImageNet	CUB
训练集数	64	351	100
验证集数	16	97	50
测试集数	20	160	50
总类别数	100	608	200
总样本数	60 000	779 165	11 788

Conv4-512 时, SAARN 模型在 miniImageNet 的 1-shot 和 5-shot 任务中分别提升了 4.3% 和 5.09%, 在 tieredImageNet 中则提升了 4.08% 和 4.99%. 此外, 使用 ResNet-12 作为骨干网络, 该模型在所有测试任务中维持了性能优势, 证明了本文的 SAARN 模型的适应性和鲁棒性, 即在较深的网络结构下仍能有效利用其特性.

表 3 数据集 iniImageNet 和 tirerdImageNet 上的准确率

Tabel 3 The accuracy on datasets miniImageNet and tieredImageNet

模型	骨干网络	miniImageNet		tirerdImageNet	
		1-shot	5-shot	1-shot	5-shot
MatchingNet	Conv4-64	43.56±0.84	55.31±0.73	—	—
ProtoNet	Conv4-64	52.61±0.52	71.33±0.41	53.33±0.50	72.10±0.41
MAML	Conv4-64	48.70±1.84	63.10±0.92	51.67±1.81	70.30±0.08
RelationNet *	Conv4-64	51.94±0.84	66.55±0.68	55.28±0.92	72.20±0.76
PN+rot	Conv4-64	53.63±0.43	71.70±0.36	—	—
CC+rot	Conv4-64	54.83±0.43	71.86±0.33	—	—
CPLAE	Conv4-64	56.83±0.44	74.31±0.34	58.23±0.49	75.12±0.40
IEPT	Conv4-64	56.26±0.45	73.91±0.34	58.25±0.48	75.63±0.46
CECNet	Conv4-64	54.45±0.47	70.57±0.38	56.59±0.50	72.86±0.42
PARN *	Conv4-64	54.78±0.91	71.02±0.67	58.12±0.96	74.45±0.75
SAARN	Conv4-64	57.62±0.88	73.01±0.70	60.52±0.94	77.46±0.72
RelationNet *	Conv4-512	54.88±0.87	68.96±0.68	59.34±0.97	76.26±0.82
ProtoNet	Conv4-512	53.52±0.43	73.34±0.36	57.88±0.50	76.82±0.40
MAML	Conv4-512	49.33±0.60	65.17±0.49	52.84±0.56	70.91±0.46
PN+rot	Conv4-512	56.02±0.46	74.00±0.35	—	—
CC+rot	Conv4-512	56.27±0.43	74.30±0.33	—	—
IEPT	Conv4-512	58.43±0.46	75.07±0.33	60.91±0.59	79.61±0.45
CPLAE	Conv4-512	57.46±0.43	75.69±0.33	61.56±0.50	80.03±0.38
SAARN	Conv4-512	59.18±0.93	74.05±0.75	63.42±0.97	81.25±0.86
RelationNet *	ResNet-12	60.59±0.85	74.89±0.66	64.52±0.96	78.75±0.76
MetaOptNet	ResNet-12	62.64±0.61	78.63±0.46	65.99±0.72	81.56±0.53
ProtoNet	ResNet-12	62.39±0.51	80.53±0.14	68.23±0.23	84.03±0.16
MTL	ResNet-12	61.20±1.80	75.50±0.80	65.62±1.80	80.61±0.90
Shot-Free	ResNet-12	59.04±0.43	77.64±0.39	66.87±0.43	82.64±0.43
DSN-MR	ResNet-12	64.60±0.72	79.51±0.50	67.39±0.82	82.85±0.56
PARN *	ResNet-12	62.01±0.91	77.01±0.69	65.98±1.03	80.74±0.82
SAARN	ResNet-12	65.23±0.92	81.02±0.71	68.91±1.08	84.65±0.84

注: SAARN 代表自监督与自适应感知关系网络, * 代表对代码复现的实验.

为了深入证明实验的有效性, 本文还在细粒度图像分类的标准数据集 CUB 上进行了测试^[33-36]. 如表 4 所示, 当使用 Conv-64 作为骨干网络时, SAARN 模型在 1-shot 和 5-shot 任务中分别比关系网络提高了 9.2% 和 8.68% 的准确率. 当转换到 ResNet-12 网络时, SAARN 模型在 1-shot 任务中的性能提升 9.7%, 在 5-shot 任务中提升了 5.89%. 这些显著的性能改进突显了 SAARN 模型在细粒度分类上的优势, 能够有效识别和处理细微的特征差异, 这对于如鸟类图像此类复杂数据集尤为关键. 通过以上实验, 证明了自监督与自适应感知关系网络在小样本学习中的先进性和有效性.

表 4 数据集 CUB 上的准确率

Tabel 4 The accuracy on dataset CUB

模型	骨干网络	CUB		模型	骨干网络	CUB	
		1-shot	5-shot			1-shot	5-shot
MatchingNet	Conv4-64	61.16±0.89	72.86±0.70	SAARN	Conv4-64	71.65±0.92	84.79±0.66
ProtoNet	Conv4-64	63.72±0.22	81.50±0.15	ProtoNet	ResNet-12	78.60±0.22	89.73±0.12
MAML	Conv4-64	55.92±0.95	72.09±0.76	RelationNet *	ResNet-12	72.53±0.98	84.67±0.63
RelationNet	Conv4-64	62.45±0.98	76.11±0.69	PARN *	ResNet-12	77.40±0.91	88.46±0.56
PARN *	Conv4-64	66.92±0.99	79.49±0.72	MIAN	ResNet-12	71.86±0.35	85.84±0.23
FEAT	Conv4-64	68.87±0.22	82.90±0.15	LPE-CLIP	ResNet-12	80.76±0.40	88.98±0.26
IEPT	Conv4-64	69.97±0.49	84.33±0.33	SAARN	ResNet-12	82.23±0.81	90.56±0.46

2.5 消融实验

表 5 展示了在数据集 miniImageNet、tieredImageNet 和 CUB 上,不同改进版本的关系网络模型的消融实验结果. 这些结果展现了两个主要观点:首先,引入自监督学习机制可以增强关系网络的性能. 其次,自适应机制的应用进一步提升了自监督和感知关系网络的性能. 从表 5 中可以看出,相比于标准的 RelationNet,加入自监督学习 (self-supervised learning, SSL) 的 RelationNet+SSL,其中自监督学习为上文中的自监督辅助任务模块,且在重新实现的 PARN 中,验证了双相关注意模块的有效性. 在各项任务上均实现了性能的提升. 这表明自监督学习能够有效地利用未标记数据来增强模型的特征表示能力,进而提高小样本分类的准确率. 此外,自监督与感知关系网络 (self-supervised and aware relation network, SARN) 模型,作为在 PARN 模型基础上进一步整合自监督学习的版本,在所有评估的任务上相比单纯的 RelationNet+SSL 显示出更显著的性能提升,进一步印证了自监督学习和自适应机制结合使用对模型准确率的正面影响.

表 5 数据集 miniImageNet、tieredImageNet 和 CUB 上不同模型的消融实验结果
Table 5 Ablation study results of different models on datasets miniImageNet, tieredImageNet and CUB

模型	miniImageNet		tireredImageNet		CUB	
	1-shot	5-shot	1-shot	5-shot	1-shot	5-shot
RelationNet	51.94±0.84	66.55±0.68	55.28±0.92	72.20±0.76	62.45±0.98	76.11±0.69
RelationNet+SSL	54.01±0.85	69.42±0.67	56.38±0.93	73.97±0.78	68.15±0.95	80.54±0.64
PARN *	54.78±0.91	71.02±0.67	58.12±0.96	74.45±0.75	66.92±0.99	79.49±0.72
SARN	56.91±0.86	72.31±0.68	59.56±0.94	76.85±0.72	71.23±0.93	84.03±0.66
SAARN (no D)	57.48±0.89	72.68±0.69	60.37±0.92	77.23±0.71	71.43±0.91	84.68±0.65
SAARN	57.62±0.88	73.01±0.70	60.52±0.94	77.46±0.72	71.65±0.92	84.79±0.66

注:表中 RelationNet+SSL 代表在关系网络的基础上加入自监督辅助模块,SARN 表示自监督与感知关系网络,SAARN (no D) 代表没有使用动态权重平均,SAARN 代表自监督与自适应感知关系网络,使用的骨干网络为 Conv4-64.

进一步地,SAARN 模型在 SARN 的基础上加入自适应感知关系机制,展示了更显著的性能提升,这表明自适应机制在增强模型解析特征间复杂关系以及提高小样本分类准确性中的重要作用. 此外,对比实验进一步考察了 DWA 在 SAARN 模型中的影响. 结果表明,引入 DWA 的 SAARN 模型在所有三个数据集上相比于未应用 DWA 的版本实现了性能上的进一步提升. 这些对比证明了 DWA 在平衡各辅助任务的学习重点以及在动态和多变的学习环境中调节任务权重方面的有效性. 通过这种精细化的权重管理,SAARN 模型能够更加精确地对复杂的数据分布进行建模,从而在小样本学习任务中取得更好的结果.

为了探究不同损失函数对模型性能的影响,本文在数据集 miniImageNet 和 tieredImageNet 上进行了一系列的消融实验. 为确保实验条件的公正性,所有参数初始值均设为 0.5,且未采用动态权重平均. 如表 6 所示,使用 Conv-64 作为骨干网络,模型结合 4 种不同的损失函数,这 4 种包括两种自监督损失,即实例级辅助任务损失 \mathcal{L}_{inst} 和场景级辅助任务损失 \mathcal{L}_{epis} , 以及两种监督损失,即小样本分类辅助损失 \mathcal{L}_{cls} 和主任务分类损失 \mathcal{L}_{main} . 实验首先只使用主任务分类损失 \mathcal{L}_{main} ,随后逐步加入其余 3 种损失. 从结果可以看出,随着更多损失函数的应用,模型性能呈现逐步提升的趋势,这表明每个损失函数都可以提升模型的性能.

表 6 数据集 miniImageNet 和 tieredImageNet 上损失函数的消融实验结果
Table 6 Ablation study results of loss functions on datasets miniImageNet and tieredImageNet

\mathcal{L}_{main} \mathcal{L}_{inst} \mathcal{L}_{epis} \mathcal{L}_{cls}	miniImageNet		tireredImageNet		CUB	
	1-shot	5-shot	1-shot	5-shot	1-shot	5-shot
✓	56.27±0.83	71.75±0.71	59.07±0.91	76.12±0.73	69.25±0.98	82.75±0.64
✓ ✓	56.67±0.81	71.96±0.73	59.50±0.93	76.54±0.75	69.63±0.99	83.14±0.66
✓ ✓ ✓	57.03±0.84	72.27±0.75	59.83±0.96	76.87±0.78	70.52±0.97	83.93±0.68
✓ ✓ ✓ ✓	57.48±0.89	72.68±0.69	60.37±0.92	77.23±0.71	71.43±0.91	84.68±0.65

注:消融结果均在 Con4-64 的骨干网络下运行.

2.6 参数分析

在本研究中,利用 Conv4-64 网络在数据集 miniImageNet 上进行了 5-way 1-shot 和 5-way 5-shot 分类实验. 图 3 展示了这两种实验设置下,模型性能随旋转变换次数 (标记为 R 值) 增加而变化的结果. 图中,不同颜色的柱状图分别代表 R 值从 1 至 4 的增长情况. 可见,在 5-way 1-shot 和 5-way 5-shot 任务中,随着 R

值的递增,模型的准确率明显提升. 这种趋势表明,通过增加输入图像的旋转变换,可以有效地增强模型对小样本类别的学习和泛化能力. 尤其在更复杂的 5-way 5-shot 任务中,较高的旋转变换数量更是显著地提升了分类准确率,证实了数据增强策略在复杂分类任务中的有效性.

3 结论

本文提出了一种改进的关系网络小样本分类方法,其通过混合自监督学习和自适应感知网络,并运用动态权重平均策略精确调节任务权重,显著增强了从有限数据中深入学习特征及其复杂关系的能力. 在 3 个数据集上进行的综合性实验表明,提出的关系网络在不同的主干网上都能够有效提升关系网络的分类性能. 其次,对比其他的一些小样本学习方法,提出的小样本学习呈现较有竞争性的性能. 消融实验进一步验证了不同的任务模块和动态权重平均策略在提升整体性能方面具有重要的作用,说明了本文提出的改进策略是可行有效的.

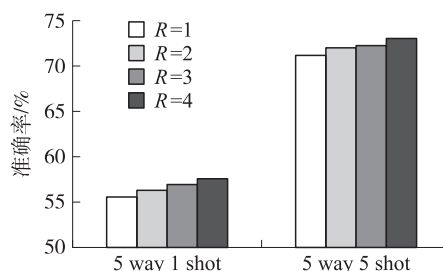


图3 数据集 miniImageNet 上 R 值分析

Fig.3 Analysis of R values on the dataset miniImageNet

[参考文献] (References)

- [1] LI F F,FERGUS R,PERONA P. One-shot learning of object categories[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2006,28(4):594-611.
- [2] LAKE B M,SALAKHUTDINOV R R,TENENBAUM J. One-shot learning by inverting a compositional causal process[J]. Advances in Neural Information Processing Systems,2013,26:2526-2534.
- [3] Yang J,LIU Y L. The latest advances in face recognition with single training sample[J]. Journal of Xihua University(Natural Science Edition),2014,33(4):1-5.
- [4] KOTIA J,KOTWAL A,BHARTI R,et al. Few shot learning for medical imaging[J]. Machine Learning Algorithms for Industrial Applications,2021,907:107-132.
- [5] CAI A H,HU W X,ZHENG J. Few-shot learning for medical image classification[C]//International Conference on Artificial Neural Networks. Bratislava,Slovakia:Springer,2020:441-452.
- [6] CHEN Z,EAVANI H,CHEN W,et al. Few-shot NLG with pre-trained language model[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg,PA:ACL,2020:183-190.
- [7] HOWARD J,RUDER S. Universal language model fine-tuning for text classification[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics. Stroudsburg,PA:ACL,2018:328-339.
- [8] RAVI S,LAROCHELLE H. Optimization as a model for few-shot learning[C]//International Conference on Learning Representations. Toulon,France,2017.
- [9] FINN C,ABBEEL P,LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks[C]//International Conference on Machine Learning. Sydney,Australia,2017.
- [10] LEE K,MAJI S,RAVICHANDRAN A,et al. Meta-learning with differentiable convex optimization[C]//CVF Conference on Computer Vision and Pattern Recognition. Long Beach,CA,USA,2019.
- [11] LIU Y Y,SCHIELE B,SUN Q R. An ensemble of epoch-wise empirical bayes for few-shot learning[C]//Proceedings of the European Conference on Computer Vision. Glasgow,Scotland,UK:Springer,2020:404-421.
- [12] VINYALS O,BLUNDELL C,LILICRAP T,et al. Matching networks for one shot learning[C]//30th Conference on Neural Information Processing Systems. Barcelona,Spain,2016.
- [13] SNEEL J,SWERSKY K,ZEMEL R. Prototypical networks for few-shot learning[J]. 31th Conference on Neural Information Processing Systems. Long Beach,CA,USA,2017.
- [14] SUNG F,YANG Y X,ZHANG L,et al. Learning to compare:Relation network for few-shot learning[C]//CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City,UT,USA,2018.
- [15] ZHANG C,CAI Y J,LIN G S,et al. Deepemd: Few-shot image classification with differentiable earth mover's distance and structured classifiers[C]//CVF Conference on Computer Vision and Pattern Recognition. Virtual,2020.

- [16] XIE J, LONG F, LV J, et al. Joint distribution matters: Deep brownian distance covariance for few-shot classification [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, LA, USA, 2022: 7972–7981.
- [17] HUI B, ZHU P, HU Q, et al. Self-attention relation network for few-shot learning [C]//Proceedings of the 2019 IEEE International Conference on Multimedia & Expo Workshops. Shanghai, China, 2019: 198–203.
- [18] 李晓旭, 刘忠源, 武继杰, 等. 小样本图像分类的注意力全关系网络 [J]. 计算机学报, 2023, 46(2): 371–384.
- [19] WU Z, LI Y, GUO L, et al. PARN: Position-aware relation networks for few-shot learning [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Long Beach, CA, USA, 2019: 6659–6667.
- [20] ABDELAZIZ M, ZHANG Z. Multi-scale kronecker-product relation networks for few-shot learning [J]. Multimedia Tools and Applications, 2022, 81(5): 6703–6722.
- [21] LI X, LI Y, ZHENG Y, et al. ReNAP: Relation network with adaptive prototypical learning for few-shot classification [J]. Neurocomputing, 2023, 520: 356–364.
- [22] HOU R, CHANG H, MA B, et al. Cross attention network for few-shot classification [J]. Advances in Neural Information Processing Systems, 2019, 32: 4005–4006.
- [23] LI Z, HU Z, LUO W, et al. SaberNet: Self-attention based effective relation network for few-shot learning [J]. Pattern Recognition, 2023, 133: 109024.
- [24] GIDARIS S, BURSUC A, KOMODAKIS N, et al. Boosting few-shot visual learning with self-supervision [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Long Beach, CA, USA, 2019: 8059–8068.
- [25] ZHANG M, ZHANG J, LU Z, et al. IEPT: Instance-level and episode-level pretext tasks for few-shot learning [C]//International Conference on Learning Representations. Vienna, Austria, 2021: 1–16.
- [26] GAO Y, FEI N, LIU G, et al. Contrastive prototype learning with augmented embeddings for few-shot learning [J]. Uncertainty in Artificial Intelligence, 2021, 21: 140–150.
- [27] YANG Z, WANG J, ZHU Y, et al. Few-shot classification with contrastive learning [C]//Proceedings of the European Conference on Computer Vision. Tel Aviv, Israel, 2022: 293–309.
- [28] LIU S, JOHNS E, DAVISON A J, et al. End-to-end multi-task learning with attention [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA, 2019: 1871–1880.
- [29] KINGMA D P, BA J. Adam: A method for stochastic optimization [C]//International Conference on Learning Representations. San Diego, CA, USA, 2015: 1–15.
- [30] LAI J, YANG S, ZHOU J, et al. Clustered-patch element connection for few-shot learning [C]//International Joint Conference on Artificial Intelligence. San Francisco, CA, USA, 2023: 991–998.
- [31] SUN Q, LIU Y, CHUA T S, et al. Meta-transfer learning for few-shot learning [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA, 2019: 403–412.
- [32] QIN Z, WANG H, MAWULI C B, et al. Multi-instance attention network for few-shot learning [J]. Information Sciences, 2022, 611: 464–475.
- [33] YANG F, WANG R, CHEN X, et al. Semantic guided latent parts embedding for few-shot learning [C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa, HI, USA, 2023: 5447–5457.
- [34] SIMON C, KONIUSZ P, NOCK R, et al. Adaptive subspaces for few-shot learning [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA, 2020: 4136–4145.
- [35] RAVICHANDRAN A, BHOTIKA R, SOATTO S, et al. Few-shot learning with embedded class models and shot-free meta training [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Long Beach, CA, USA, 2019: 331–339.
- [36] YE H J, HU H, ZHAN D C, et al. Few-shot learning via embedding adaptation with set-to-set functions [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, WA, USA, 2020: 8808–8817.

[责任编辑:陈 庆]