

文章编号 1004-924X(2023)21-3145-11

类内-类间通道注意力少样本分类

杨利平^{1*}, 张天洋¹, 王宇阳¹, 辜小花²

(1. 重庆大学 光电技术及系统教育部重点实验室, 重庆 400044;

2. 重庆科技学院 电气工程学院, 重庆 401331)

摘要: 针对元学习少样本分类样本特征鉴别能力不足的问题, 提出了一种类内-类间通道注意力少样本分类方法 (Intra-inter Channel Attention Few-shot Classification, ICAFSC)。ICAFSC 在原型网络基础上设计了一个类内-类间通道注意力模块, 该模块通过类内-类间距离度量计算通道权重实现特征加权, 提高特征对类别的鉴别能力。为了克服直接在元训练阶段学习类内-类间通道注意力模块容易出现过拟合或欠拟合现象的问题, ICAFSC 在原型网络的元训练之前增加一个预训练阶段。该阶段设计具有大量标记样本的分类任务, 并利用这些任务充分训练类内-类间通道注意力模块, 促使该模块达到较优的状态。在原型网络的元训练和元测试阶段, ICAFSC 冻结类内-类间通道注意力模块的参数, 分别实现少样本分类经验的学习与迁移。在 MiniImagenet 数据集上分别开展了 1-shot 和 5-shot 的少样本分类实验。实验结果表明: 本文提出的类内-类间通道注意力少样本分类方法与原型网络相比, 在 1-shot 和 5-shot 条件下分类准确率分别提高了 1.93% 和 1.15%。

关键词: 深度学习; 少样本分类; 元学习; 原型网络; 通道注意力

中国分类号: TP391.4 **文献标识码:** A **doi:** 10.37188/OPE.20233121.3145

Intra-inter channel attention for few-shot classification

YANG Liping^{1*}, ZHANG Tianyang¹, WANG Yuyang¹, GU Xiaohua²

(1. Key Laboratory of Optoelectronic Technique & System of Ministry of Education,
Chongqing University, Chongqing 400044, China;

2. School of Electrical Engineering, Chongqing University of Science & Technology,
Chongqing 401331, China)

* Corresponding author, E-mail: yanglp@cqu.edu.cn

Abstract: As only one or a few training samples are used for few-shot classification tasks, the features extracted via a prototypical network cannot guarantee much discriminative power. Accordingly, this paper proposes an intra-inter channel attention few-shot classification (ICAFSC) method. This method uses an intra-inter channel attention module (ICAM) to calculate channel weights based on an intra-inter distance metric. The module is integrated into the prototypical network to make the embedded features more discriminative. To overcome the problems of overfitting or underfitting when directly learning the ICAM in the few-shot classification's meta-training stage, ICAFSC adds a pre-training stage before the meta-training of the prototypical network. We design adequate classification tasks with a large number of labeled

收稿日期: 2023-04-06; 修订日期: 2023-05-14.

基金项目: 国家重点研发计划课题 (No. 2022YFF0708001); 重庆市自然科学基金面上项目 (No. csts2021jcyj-msxmX0478)

samples to learn optimal parameters of the ICAM in the pre-training stage. Subsequently, in the meta-training and meta-testing stages of the prototypical network, ICAFSC first freezes the parameters of the ICAM to guarantee a stable channel attention relationship. It then achieves few-shot classification experience learning and transfer via meta-training and meta-testing. We conduct 1-shot and 5-shot few-shot classification experiments on the MiniImagenet dataset. The experimental results indicate that, compared to the prototypical network, the proposed ICAFSC method shows improvements of 1.93% and 1.15% for the 1-shot and 5-shot scenarios, respectively.

Key words: deep learning; few-shot classification; meta-learning; prototypical network; channel attention

1 引 言

在拥有足够标记数据的前提下,深度神经网络分类模型可在模式分类任务中取得优异的分类性能^[1-2]。然而,实际应用的类别往往是变化的,可用于训练新分类任务的标记数据非常少、甚至为零。由于欠拟合或过拟合现象的存在,简单的利用少量新类别样本更新模型参数并不能使预训练的深度神经网络分类模型很好地适应新分类任务,这即是少样本分类问题^[3-5]。

元学习是当前解决少样本分类问题的主流方法^[6-7]。该类方法将少样本分类作为元学习任务来处理,分别利用元学习框架的元训练和元测试阶段完成分类经验(也称为元知识)的学习与迁移。元训练阶段的目标是在已有基类数据集上构造大量少样本分类任务并从中学习分类经验;元测试阶段的目标是利用少量新类别样本更新分类模型,将元训练阶段获得的少样本分类经验迁移到新的分类任务中。

原型网络^[8](Prototypical Network, ProtoNet)是一种用于解决少样本分类问题的经典元学习方法。原型网络的基本思想是:在一个嵌入空间中,每个类支持集样本的特征向量取均值作为原型表示(Prototypes),该类别的样本在嵌入空间中聚集在该原型表示的周围,通过距离函数得到样本与各个原型表示之间的相似度。然而,由于原型网络采取简单平均的方法获得原型表示,噪声样本与正常样本对原型表示的影响力度是相同的,因此,极少数的噪声样本可能会对原型表示的计算带来较大的影响。为了缓解这种情况,现有方法基于原型网络在元训练阶段做了很多工作。文献[9]应用空间注意力机制关注

嵌入空间中较为突出的区域,以保留目标主体的空间关系。文献[10]应用通道注意力机制关注任务内的最大化差异,以增强模型对新任务的判别能力。文献[11]通过引入额外的自监督任务人为地增加基类数据集的多样性和复杂性。上述少样本分类方法主要通过引入注意力约束和自监督学习来改善特征提取器对关键信息的提取。尽管这些方法在少样本分类任务中是有效的,但是这些约束策略都没有考虑特征对类别的鉴别能力。由于少样本分类任务标记样本有限,所以约束策略容易出现欠拟合或过拟合现象,不易达到最优状态。

针对上述问题,本文提出了一种类内-类间通道注意力少样本分类方法(Intra-inter Channel Attention Few-shot Classification, ICAFSC)。该方法在原型网络基础上设计了一个类内-类间通道注意力模块(Intra-inter Channel Attention Module, ICAM),该模块通过类内-类间距离度量计算通道权重实现特征加权,促使同类别特征更加紧凑不同类别特征更加分散,提高特征对类别的鉴别能力。ICAM包含类内通道注意力和类间通道注意力两个部分。类内通道注意力计算同类别样本特征图的通道距离,寻找类内相似度高的通道,并增加这些通道的权重。类间通道注意力计算不同类别样本特征图的通道距离,寻找类间判别性大的通道,并增加这些通道的权重。为了使ICAM充分学习类内通道权重和类间通道权重,提高特征对类别的鉴别能力,ICAFSC在元训练之前增加一个预训练阶段。该阶段设计具有大量标记样本的分类任务,并利用这些任务充分训练ICAM,使其达到最优状态。在元训练阶段,ICAFSC冻结ICAM的参

数,利用少样本分类任务训练特征提取器并从中学习少样本分类经验。在元测试阶段,ICAFSC 利用少量的新类别标记样本更新少样本分类模型,将元训练阶段获得的少样本分类经验迁移到新的分类任务中。

2 原理

2.1 少样本分类问题设置

本文用 x 表示输入样本, y 表示类别标签; X 和 Y 分别表示输入样本和类别标签的分布空间。少样本分类任务 T_i 给定包含支持集 S_i 和查询集 Q_i 的数据集 $D_T = \{S_i, Q_i\}$, 其中支持集 $S_i = \{x_i^s, y_i^s\}_{i=1}^M$ 由少量带有类别标签的样本组成, 查询集 $Q_i = \{x_i^q\}_{i=1}^R$ 由无标签数据组成; $x_i^s, x_i^q \in X, y_i^s \in Y, M$ 和 R 分别表示支持集和查询集样本数量。通常, S_i 包含 N 个类别, 并且每个类别仅拥有 K 个样本 (例如 $K=1$ 或 $K=5$), 即 $M=N \times K$ 。包含上述 S_i 的少样本分类任务 T_i 被称为 N -way K -shot 任务。少样本分类利用支持集 S_i 中非常有限的标记样本学习映射关系 $F: X \rightarrow Y$, 并利用 F 对查询集 Q_i 中的样本进行分类。但是由于 K 往往很小, 少样本分类仅利用支持集 S_i 很难训练出一个高质量的映射关系 F , 因此, 许多学者引入元学习解决少样本分类。

元学习框架下的少样本分类给定一个拥有大量标记样本的数据集 $D_B = \{x_i^b, y_i^b\}_{i=1}^L$, 其中 $x_i^b \in X, y_i^b \in Y, L$ 表示 D_B 的样本数量。首先, 元训练阶段在 D_B 上随机采样出大量的少样本分类任务 T_i , 这些任务服从任务分布 $p(T)$, 并利用少样本分类任务 T_i 训练一个能够快速泛化的映射关系 F 。然后, 元测试阶段利用少样本分类任务 T_i 的支持集 S_i 更新映射关系 F 以适应新分类任务, 并在 Q_i 上测试映射关系 F 的分类效果。需要注意的是, 在少样本分类中 D_B 和 D_T 的类别是不相交的。

2.2 原型网络

在元训练阶段, 原型网络给定一个支持集为 $S_b = \{x_i^s, y_i^s\}_{i=1}^{N \times K}$ 和一个查询集 $Q_b = \{x_i^q, y_i^q\}_{i=1}^{N \times U}$, 其中 U 表示查询集每个类的样本数量; $S_b, Q_b \in D_B$ 。原型网络通过一个嵌入式网络将样本数据映射到嵌入空间中, 如图 1 所示。这里

采用 4 层卷积层作为嵌入式网络 $f(\theta)$, 其中 θ 为可学习参数。支持集与查询集的样本特征图 $F_{s,i} \in \mathbb{R}^{C \times W \times H}$ 和 $F_q \in \mathbb{R}^{C \times W \times H}$ 分别表示为:

$$F_{s,i}^m = f_\theta(x_i^s | y_i^s = m), \quad (1)$$

$$F_q = f_\theta(x_i^q), \quad (2)$$

其中: $F_{s,i}^m$ 表示支持集第 m 类的第 i 个样本特征图, F_q 表示查询集的样本特征图; C, W 和 H 分别表示特征图的通道数、宽度和高度; 类别 m 的原型表示 $P^m \in \mathbb{R}^{C \times W \times H}$ 是支持集所有样本在嵌入空间的均值, 即:

$$P^m = \frac{1}{K} \sum_{i=1}^K F_{s,i}^m. \quad (3)$$

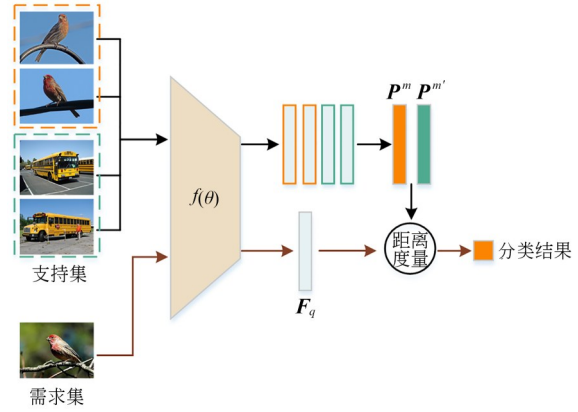


图 1 原型网络

Fig. 1 Prototypical Network

原型网络计算查询集样本 x_i^q 在嵌入空间与原型表示的距离, 并利用 softmax 函数得出查询集样本属于目标类别 m 的概率:

$$p_\theta(y = m | x_i^q) = \frac{\exp(-d(F_q, P^m))}{\sum_{m'} \exp(-d(F_q, P^{m'}))}, \quad (4)$$

其中 $d(\cdot)$ 为距离函数, 一般采用欧氏距离。最后, 原型网络利用梯度下降法最小化目标类别的负对数概率损失优化模型:

$$L_p = -\frac{1}{N \times U} \sum_{i=1}^{N \times U} \log(p_\theta(y = m | x_i^q)). \quad (5)$$

在元测试阶段, 原型网络则利用少样本分类任务 T_i 的支持集 S_i 构建适应新分类任务的原型表示, 将元训练阶段的分类经验迁移到新分类任务上; 然后, 利用式 (4) 给出查询集 (Q_i) 样本的预测分数, 最后将预测分数最高的类作为分类预测结果。

本文提出的 ICAFSC 在原型网络基础上设计了一个类内-类间通道注意力模块,该模块通过类内-类间距离度量计算通道权重实现特征加权,提高特征对类别的鉴别能力。为了使 ICAM 充分学习类内通道权重和类间通道权重,ICAFSC 在元训练之前增加一个预训练阶段。该阶段设计具有大量标记样本的分类任务,并利用这些任务充分训练类内-类间通道注意力模块,促使该模块达到最优状态。

2.3 类内-类间通道注意力少样本分类

2.3.1 少样本分类的类内-类间通道注意力原型网络

原型网络包含元训练和元测试两个阶段。其中,元训练利用许多任务训练模型对少样本分类任务的适应性;元测试则是利用少量的新类别样本将模型的先验知识迁移到特定的少样本分类任务。由于元训练阶段每个类别标记样本较少,因此 ICAFSC 在原型网络的基础上增加了预训练阶段。预训练阶段的目的是利用包含大量标记样本的分类任务 t_i 训练类内-类间通道注意力模块,从而能够学习更具有判别性的通道权重,提高特征对类别的鉴别能力。 t_i 是数据集 D_B 上随机采样出的分类任务,服从任务分布 $p(t)$,即 $t_i \sim p(t)$ 。ICAFSC 的主要训练与测试流程如下:首先,预训练阶段利用分类任务 t_i 训练 ICAM 以学习更具有判别性的通道权重,同时初步更新嵌入式网络的参数;然后,元训练阶段固定 ICAM 的参数,仅利用少样本分类任务进一步提高嵌入式网络对少样本分类的适应性;最后,元测试阶段利用训练后的模型对 D_T 数据集进行预测以测试模型的少样本分类性能。以上三个阶段采用的模型结构如图 2 所示。

ICAFSC 将类内-类间通道注意力模块嵌入原型网络中。在预训练阶段,分类任务 t_i 给定支持集 $S_b = \{x_i^s, y_i^s\}_{i=1}^{N \times G}$ 和查询集 $Q_b = \{x_i^q, y_i^q\}_{i=1}^{N \times U}$;其中 G 非常大(例如 $G=100$ 或 $G=200$); $S_b, Q_b \in D_B$ 。ICAM 利用支持集监督信息计算第 m 类的综合通道权重 $W^m \in \mathbb{R}^{C \times 1 \times 1}$,该权重由类内通道权重和类间通道权重线性组合得到。ICAFSC 通过综合通道权重 W^m 将支持集第 m 类的第 i 个特征图 $F_{w,i}^m$ 转换成任务自适应特征图(也

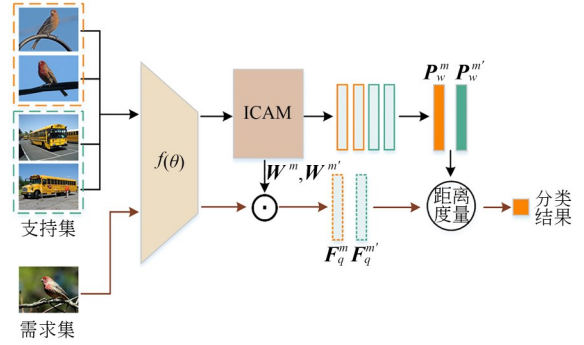


图 2 少样本分类的类内-类间通道注意力原型网络
Fig. 2 Intra-inter channel attention prototypical network for FSC

称为加权特征图) $F_{w,i}^m \in \mathbb{R}^{C \times W \times H}$:

$$F_{w,i}^m = W^m \odot F_{s,i}^m, \quad (6)$$

其中 $\odot$ 表示元素级乘法。计算中,综合通道权重值沿空间维度复制得到尺寸为 $C \times W \times H$ 的通道权重。然后,ICAFSC 利用转换后的任务自适应特征图计算任务自适应原型表示 $P_w^m \in \mathbb{R}^{C \times W \times H}$:

$$P_w^m = \frac{1}{G} \sum_{i=1}^G F_{w,i}^m. \quad (7)$$

虽然支持集可以根据监督信息计算更具有可靠性的原型表示,但是需求集缺少监督信息无法计算对应的通道权重。因此,在计算需求集与第 m 类原型表示的距离时,本文将综合通道权重 W^m 应用于需求集特征图 $F_q \in \mathbb{R}^{C \times W \times H}$:

$$F_q^m = W^m \odot F_q, \quad (8)$$

其中: $F_q^m \in \mathbb{R}^{C \times W \times H}$, 表示加权后的需求集特征图。需求集样本 $x_i^q \in Q_b$ 属于第 m 类的预测分数可由 softmax 函数获得:

$$p_\theta(y_i^q = m | x_i^q) = \frac{\exp(-d(F_q^m, P_w^m))}{\sum_{m'} \exp(-d(F_q^{m'}, P_w^{m'}))}. \quad (9)$$

类内-类间通道注意力模块主要由类内通道注意力和类间通道注意力组成,如图 3 所示。为了更加简明地阐述 ICAM,图 3 中只展示了二分类的示意图,更多类别的讨论方式根据公式类推。类内通道注意力关注类内的相似信息并计算类内通道权重,类间通道注意力关注类间判别信息并计算类间通道权重。然后,ICAM 通过线性组合(CE)将类内通道权重和类间通道权重组,形成一个综合通道权重对特征图实现加权。

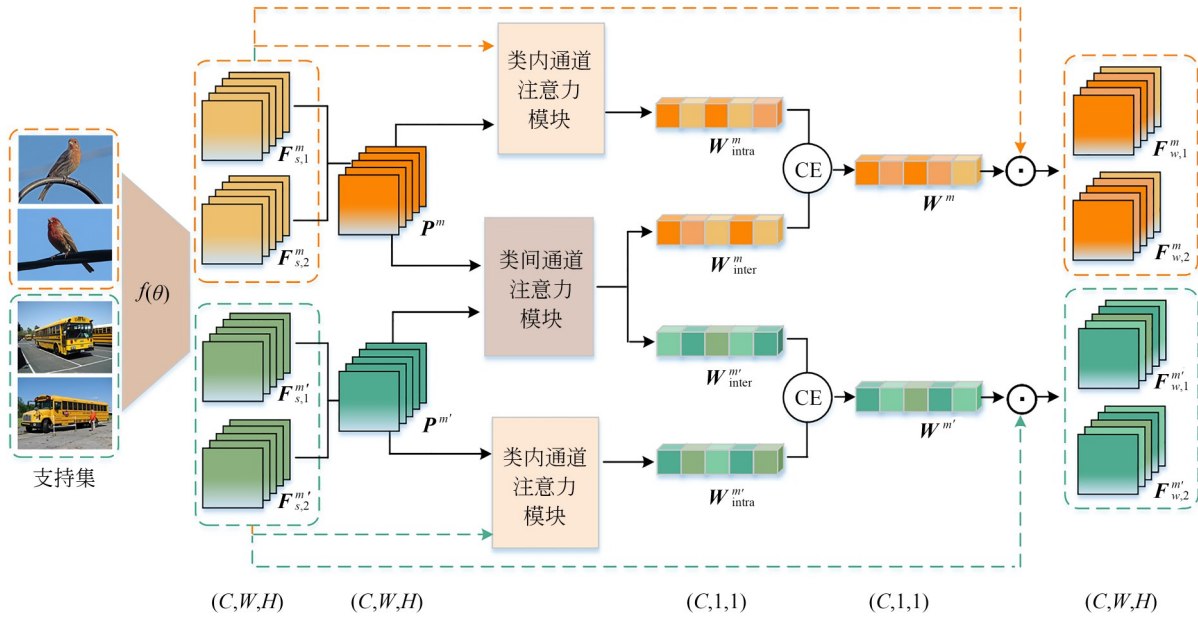


图 3 类内-类间通道注意力示意图

Fig. 3 Schematic diagram of intra-inter channel attention

2.3.2 类内通道注意力

通道注意力^[12-13]通常是对特征图做池化操作或挤压-激励操作以提取通道间的信息差异,并基于该信息差异对特征图施加通道级的权重。虽然这种注意力可以获取特征图自身每个通道的重要程度,但是对于具有一个或几个监督样本的小样本任务来说,样本的缺乏会极大地限制通道注意力对通道间信息差异的有效学习。为解决该问题,本文以同类别所有特征图为基础设计了类内通道注意力,如图 4 所示。类内通道注意力主要由通道距离计算和全连接映射组成,该注意力以第 m 类的特征图和原型表示作为输入计算第 m 类的类内通道权重。

当给定支持集中类别为 m 的特征图 $F_{s,i}^m$,可

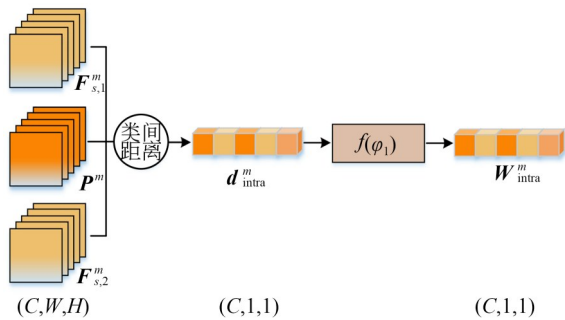


图 4 类内通道注意力示意图

Fig. 4 Schematic diagram of intra-class channel attention

以由式 (3) 计算得到第 m 类的原型表示 $P^m \in \mathbb{R}^{C \times W \times H}$ 。之后,对于第 m 类第 c 个通道的类内通道距离 $d_{intra}^{m,c}$ 的表达式为:

$$d_{intra}^{m,c} = \frac{1}{|W \times H|} \max_{1 \leq i \leq K} \|F_{s,i}^{m,c} - P^{m,c}\|^2. \quad (10)$$

该距离表示支持集样本高度激活的通道与原型表示的类显著性通道的相似程度。取通道 c 上 K 个样本的最大值作为第 m 类第 c 个通道的类内通道距离。距离越小,说明第 m 类所有样本与原型表示在通道 c 上的距离越小,通道相似程度越高,特征提取器提取的特征在通道 c 上更显著。

得到第 m 类的类内通道距离 $d_{intra}^m \in \mathbb{R}^{C \times 1 \times 1}$ 后,为了更好地整合通道间的相似信息,将类内通道距离作为输入送进两层的全连接网络中计算类内通道权重 $W_{intra}^m \in \mathbb{R}^{C \times 1 \times 1}$:

$$W_{intra}^m = f_{\varphi_1}(d_{intra}^m), \quad (11)$$

其中: f_{φ_1} 表示全连接网络, φ_1 表示可学习参数。假定预训练阶段每个分类任务的类别数 $N = N_{pre} \odot$ ICAM 将所有类别的类内通道距离拼接形成一个尺寸为 $N_{pre} \times C$ 的向量,并将该向量作为全连接网络的输入。全连接网络结构如表 1 所示。

从式 (10) 可以得出,类内通道距离越小,支持集样本与原型表示的通道匹配程度越高,也就

表 1 全连接网络的结构参数

Tab. 1 Structural parameters of fully connected network

层名称	输出尺寸
输入	$N_{\text{pre}} \times C$
全连接层	$N_{\text{pre}} \times C/2$
批归一化	$N_{\text{pre}} \times C/2$
ReLU	$N_{\text{pre}} \times C/2$
全连接层	$N_{\text{pre}} \times C$
Sigmoid	$N_{\text{pre}} \times C$
转换层	$N_{\text{pre}} \times C$

意味着这些通道包含类内相似特征。因此,为了增大具有类内相似特征通道的权重,在全连接网络最后一层增加了转换层。为了更简明地阐述转换层,假设 $N_{\text{pre}} = 1$, 即有(类别为 m):

$$W_{\text{intra}}^m = 1 - \text{Sig}_{\text{intra}}^m, \quad (12)$$

其中: $\text{Sig}_{\text{intra}}^m$ 表示 Sigmoid 的输出, 则 $\text{Sig}_{\text{intra}}^m$ 的元素均属于 0 到 1 之间。经过转换层之后, 类内通道距离越小, 类内通道权重越大。类内通道权重强调了类内的相似信息, 同时也抑制了类内的冗余信息。因此, 加权特征图所计算的原型表示更可靠, 在不断迭代更新中也能促进特征提取器对类别的鉴别能力。

2.3.3 类间通道注意力

虽然类内通道注意力可以增大具有类内相似特征通道的权重, 但对于少样本分类任务, 有效地学习类间判别性特征也尤为重要。图 5 为基于所有原型表示的类间通道注意力。类间通道注意力主要由通道距离计算和全连接映射组成, 该注意力以所有类别的原型表示作为输入计算类间通道权重。为了更加简明地阐述 ICAM, 图 3 只展示了二分类的示意图, 更多类别的讨论方式根据公式类推。类间通道注意力旨在提高类间判别性较大的通道的权重, 以此增加样本在嵌入空间与其他类别样本的判别性。

当给定支持集中类别为 m 的原型表示 P^m , 第 m 类在第 c 个通道与其他类别 m' 的类间通道距离 $d_{\text{inter}}^{m,c}$ 的表达式为:

$$d_{\text{inter}}^{m,c} = \frac{1}{|W \times H|} \min_{1 \leq m' \leq N, m' \neq m} \|P^{m,c} - P^{m',c}\|^2. \quad (13)$$

该距离表示类别 m 高度激活的通道与其他类别 m' 显著性通道的相似程度。取 c 通道上

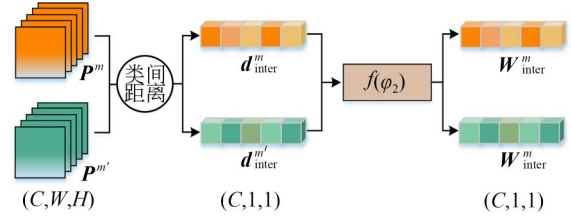


图 5 类间通道注意力示意图

Fig. 5 schematic diagram of inter-class channel attention

$N - 1$ 个类别的最小值作为第 m 类第 c 个通道的类间通道距离。该距离越大, 说明第 m 类原型表示在通道 c 上与其他类别原型表示的距离越大, 通道的差异程度越高。

得到第 m 类的类间通道距离 $d_{\text{inter}}^m \in \mathbb{R}^{C \times 1 \times 1}$ 后, 为了更好地整合通道间的判别信息, 将类间通道距离作为输入送进两层的全连接网络中得到类间通道权重 $W_{\text{inter}}^m \in \mathbb{R}^{C \times 1 \times 1}$:

$$W_{\text{inter}}^m = f_{\varphi_2}(d_{\text{inter}}^m), \quad (14)$$

其中: f_{φ_2} 表示全连接网络, φ_2 表示可学习参数。该全连接网络相比类内通道注意力的全连接网络仅缺少最后的转换层, 其他结构完全相同。

从式(13)可以得出, 类间通道距离越大, 类间通道权重越大。类间通道权重强调了类间的判别性信息, 因此嵌入空间中的加权特征图的可分性会提高, 在不断迭代更新的过程中, 特征提取器的鉴别能力会逐步提升。

2.3.4 类内-类间通道注意力线性组合

类内通道注意力与类间通道注意力分别关注类内相似通道和类间判别通道。为综合考虑样本在嵌入空间类别表征, 将两种通道权重线性组合形成一个综合通道权重, 即:

$$W^m = \alpha W_{\text{intra}}^m + (1 - \alpha) W_{\text{inter}}^m, \alpha \in [0, 1], \quad (15)$$

其中: $W^m \in \mathbb{R}^{C \times 1 \times 1}$, 表示第 m 类的综合通道权重; α 为平衡参数, α 取 0.5, 以保证综合通道权重对类内相似通道和类间判别通道的关注度相同。

ICAFSC 在原型网络的框架下增加了一个预训练阶段, 并利用 ICAM 提高特征对类别的鉴别能力。假定元训练阶段每个少样本分类任务的类别数 $N = N_{\text{train}}$ 。ICAFSC 冻结 ICAM 的参数(全连接层)来保留在预训练阶段中学习到的通道特性。由于 N_{pre} 与 N_{train} 的值可能不同, 并且全连接层的输入神经元不可更改, 在元训练阶段

和元测试阶段中对输入全连接层的通道距离施加拆分或复制的操作,如图6所示。

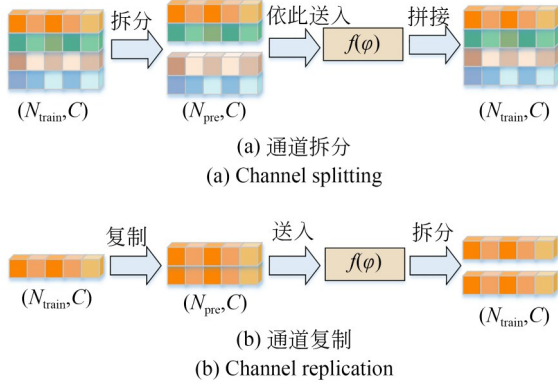


图6 通道距离拆分和复制示意图 ($f(\varphi)=f(\varphi_1)\cup f(\varphi_2)$)
Fig. 6 Channel distance splitting and replication ($f(\varphi)=f(\varphi_1)\cup f(\varphi_2)$)

以元训练阶段为例,当 $N_{\text{train}} > N_{\text{pre}}$ 时,ICAFSC 将通道距离拆分为多份符合全连接网络要求的输入,并将全连接的输出拼接在一起,如图6(a)所示;当 $N_{\text{train}} < N_{\text{pre}}$ 时,ICAFSC 将通道距离复制多份输入全连接网络,再将输出进行拆分,如图6(b)所示。元测试阶段与元训练阶段一致,当 $N_{\text{test}} > N_{\text{pre}}$ 时,ICAFSC 执行拆分操作;当 $N_{\text{test}} < N_{\text{pre}}$ 时,ICAFSC 执行复制操作。

3 实 验

3.1 数据集与预处理

本文采用 MiniImagenet 数据集^[14]进行实验,该数据集包含 100 个类别,每个类别有 600 个样本,共计 60 000 张图像。数据集主要包含鸟类、昆虫、汽车和飞机等类别。文献[15]将数据集按照 64:16:20 的比例划分为训练集、评价集和测试集。

根据文献[8],图像需预处理为 84×84 的分辨率。为防止图像变形,先对原始图像按短边等比例缩放到 84 的大小,然后再利用中心裁剪的方法将图像裁剪为 84×84 的分辨率。

3.2 实验过程与参数设置

骨干网络采用由 4 个卷积神经网络模块 (CNN Block) 堆叠的卷积神经网络^[16-17]。每个模

块包含卷积层、批归一化层、ReLU 激活层以及最大池化层。CNN Block 的参数设置如表 2 所示,其中 c 表示输出通道数, k 表示卷积核的尺寸, s 表示最大池化的步长,下标表示第几个模块。主干网络最后输出 $84 \times 5 \times 5$ 的特征图。

表 2 卷积神经网络的参数设置

Tab. 2 Parameter setting of convolutional neural network

模块名称	
卷积层	$c_1: 64, c_2: 64, c_3: 64, c_4: 64$
	$k_1: 3, k_2: 3, k_3: 3, k_4: 3$
批归一化	$c_1: 64, c_2: 64, c_3: 64, c_4: 64$
最大池化	$s_1: (2, 2), s_2: (2, 2),$
	$s_3: (2, 2), s_4: (2, 2)$

模型预训练与训练阶段均采用 Adam 优化器,初始学习率为 0.001。采用 ReduceLROnPlateau 的学习率调整策略对评价集损失进行监控,当 10 个迭代周期的评价集损失没有下降时,学习率降低到原来的 0.3 倍。在 MiniImagenet 数据集上的预训练周期为 50 次,训练周期为 400 次,当 100 个迭代周期的评价集损失没有下降时,结束模型训练。

模型测试阶段,从测试集中随机生成 600 组 N -way K -shot 任务,每类包括 15 个查询集样本用于评估模型少样本分类的性能。分类准确率取测试集中随机生成的 600 组任务的平均值,并给出 95 % 的置信区间。

针对不同 N -way K -shot,选择欧氏距离作为距离函数^[8]以及合适的训练阶段任务参数配置,具体设置如表 3 所示。

表 3 元训练阶段少样本分类任务的参数配置

Tab. 3 Parameter configuration of few-shot classification task in meta-train stage

N -way K -shot	元训练阶段		
	way	shot	query
5-way 5-shot	20	5	15
5-way 1-shot	30	1	15

3.3 结果与分析

首先在 MiniImagenet 数据集上进行消融实验,验证 ICAFSC 在不同训练阶段更新 ICAM 参

数对少样本分类结果的影响。然后,验证 ICAFSC 不同少样本任务配置对少样本分类结果的影响。最后,给出 ICAFSC 在 MiniImagenet 数据集与其他方法的对比。

3.3.1 消融实验与分析

为验证提出的 ICAM 在不同训练阶段更新参数对 ICAFSC 少样本分类结果的影响,实验采用 3 种 ICAM 参数更新方式,分别是仅在预训练阶段更新参数(ICAFSC-pre)、仅在元训练阶段更新参数(ICAFSC-tra)以及预训练阶段和元训练阶段均更新参数(ICAFSC-pre+tra),分类结果如表 4 所示。

表 4 ICAFSC 在 MiniImagenet 数据集上不同训练阶段更新 ICAM 参数的实验结果

Tab. 4 Experimental results for ICAFSC updating ICAM parameters in different training stages on MiniImagenet

模 型	准确率/%	
	5-way 1-shot	5-way 5-shot
ProtoNet	49.62±0.78	67.96±0.62
ICAFSC-pre	51.55±0.75	69.11±0.65
ICAFSC-tra	49.76±0.77	68.14±0.60
ICAFSC-pre+tra	50.70±0.75	68.87±0.63

表 4 分别展示了 5-way 1-shot 和 5-way 5-shot 两种少样本分类问题的实验结果。从 5-way 5-shot 和 5-way 1-shot 的实验结果可以看出,3 种 ICAM 参数更新方式的分类准确率均高于 ProtoNet。这说明 ICAM 有效地增强了模型对不同类别的鉴别能力。在 5-shot 和 1-shot 情况下,ICAFSC-pre 的分类准确率比 ICAFSC-tra 分别提高了 0.97% 和 1.79%。这说明预训练阶段大量的标记样本更有利于 ICAM 学习类内通道权重与类间通道权重,提高了特征的鉴别能力。相比于 ICAFSC-pre,ICAFSC-pre+tra 在 5-shot 和 1-shot 的情况下分别下降了 0.24% 和 0.85%,其原因在于元训练阶段少量标记样本对 ICAM 的更新会破坏 ICAM 在预训练阶段学习到的类内-类间鉴别能力。

3.3.2 不同少样本任务配置的实验与分析

为验证 ICAFSC 少样本分类方法的有效性并寻找最优的 ICAM 预训练任务参数配置,在

MiniImagenet 数据集上进行了不同少样本任务配置的实验,结果如表 5 所示。

表 5 ICAFSC 在 MiniImagenet 数据集上不同少样本任务配置的实验结果

Tab. 5 Experimental results for ICAFSC under different classification task configurations in pre-training phase on MiniImagenet

模 型	预训练任务			准确率/%	
	Way	Shot	Query	5-way 1-shot	5-way 5-shot
ProtoNet	—	—	—	49.62±0.78	67.96±0.62
ICAFSC	5	100	15	51.55±0.75	69.11±0.65
ICAFSC	10	100	15	50.30±0.77	68.79±0.60
ICAFSC	20	100	15	49.97±0.79	68.47±0.66

表 5 分别展示了 5-way 1-shot 和 5-way 5-shot 两种少样本分类问题的实验结果。为与少样本分类问题的类别数量匹配,表中 3 组预训练任务参数配置的类别数量均设置为 5 的整数倍。从 5-way 5-shot 的实验结果可以看出,不同参数配置的 ICAFSC 的分类准确率均高于 ProtoNet。当预训练类别数量设置为 5 时,ICAFSC 的准确率比 ProtoNet 高出 1.15%。5-way 1-shot 的实验结果也表现出类似的结果,即类内-类间通道注意力模块可以提高少样本分类的准确率。这说明类内-类间通道注意力模块能够改善特征提取器以获得鉴别性特征,提升少样本分类模型的分类准确率。从不同参数配置的 ICAFSC 的实验结果可以看出,随着预训练类别数量的增大,模型的分类准确率在逐渐下降。相比于 way=5, way=20 的 ICAFSC 的分类准确率在 1-shot 和 5-shot 条件下分别下降了 1.58% 和 0.64%。这说明预训练的类别数量与元测试的类别数量保持一致时,ICAFSC 的表现最优,其原因在于全连接网络的输入神经元不可更改,当预训练的类别数量与元测试的类别数量不一致时,通道距离的拆分与复制会破坏通道距离的完整性。

3.3.3 对比实验

为了更加客观地评价所提出方法的有效性,本文在 MiniImagenet 数据集上开展对比实验。表 6 对比了不同方法在 MiniImagenet 数据集上 5-way 5-shot 和 5-way 1-shot 的实验结果。文献[8]为 ICAFSC 的基线模型,是一种基于度量学习的

元学习方法,并采用欧氏距离作为距离函数。文献[16]是一种与模型无关的元学习方法,采用参数优化的方式寻找泛化性最好的模型参数。文献[17]在基于度量学习的元学习方法中引入注意力机制,来刻画图像之间的特征关系和位置关系。文献[18]是文献[16]的改进方法,在外部循环更新元初始化参数时冻结分类器,只在内部循环时更新分类器。文献[19]使用变分自编码器建立以跨类样本对为条件的多模态权重分布,并利用该权重为每个分类任务生成一个判别性指标。文献[20]提出使用自监督学习训练一个广义的嵌入网络,该网络通过从数据本身学习为新分类任务来提供稳定的原型表示。

从表6可以看出,ICAFSC在1-shot和5-shot条件下的分类准确率分别为51.55%和69.11%。与文献[8]的ProtoNet基线模型相比,ICAFSC在1-shot条件下提高了1.93%,在5-shot条件下提高了1.15%。这表明类内-类间通道注意力模块有助于提高模型的少样本分类能力。与文献[16]相比,ICAFSC在1-shot条件下提高了3.48%,在5-shot条件下提高了5.96%。与文献[17]相比,ICAFSC在1-shot条件下提高了2.01%,在5-shot条件下提高了0.77%。与文献[21]相比,ICAFSC在1-shot条件下提高了0.35%,在5-shot条件下提高了0.31%。文献[22]在基线网络上增加变分自编码器、权重生成器和注意力度量网络,利用高复杂度的网络产生复杂数据分布来提高样本多样性。与文献[19]相比,ICAFSC在5-shot条件下分类准确率仅低了0.20%,表明增加的ICAM对特征鉴别性的提升与复杂网络的作用相当,并且ICAM可以作为一个即插即用模块,嵌入特征提取网络和分类器之间。ICAFSC在1-shot条件下低了2.35%,这种现象是因为本文提出的ICAM在1-shot条件下无法计算类内通道信息造成的。与文献[20]相比,ICAFSC在1-shot条件下提高了0.70%,在5-shot条件下提高了1.03%。与文献[24]相比,ICAFSC在1-shot条件下的准确率降低了0.87%,在5-shot条件下提高了0.33%。以上实验结果表明,ICAM通过类内-类间距离度量计算通道权重实现特征加权,提高了特征对类别的鉴别能力以及模型的少样本分类能力。在5-shot

条件下,ICAFSC能够取得最高的分类准确率,在1-shot条件下的分类准确率优于当前大多数的方法。

表6 在MiniImagenet数据集上不同方法的分类得分(主干网络为Conv4-64)

Tab.6 Classification scores of different methods with Conv4-64 as backbone on MiniImagenet dataset

模 型	准确率/%	
	5-way 1-shot	5-way 5-shot
ProtoNet ^[8]	49.62±0.78	67.96±0.62
MatchNet ^[14]	43.40±0.78	51.09±0.71
MAML ^[16]	48.07±1.75	63.15±0.91
P-RelationNet ^[17]	49.54±0.82	68.34±0.70
BOIL ^[19]	49.61±0.16	66.45±0.37
D2D2 ^[20]	51.20±0.60	68.80±0.10
HyperProto ^[21]	50.21±0.31	66.48±0.71
DAM ^[22]	53.80±0.19	69.31±0.16
CSS ^[23]	50.85±0.84	68.08±0.73
HyperShot ^[23]	52.42±0.46	68.78±0.29
OVE ^[24]	48.00±0.24	67.14±0.23
IMP ^[25]	49.60±0.80	68.10±0.80
ICAFSC	51.55±0.75	69.11±0.65

4 结 论

本文设计了类内-类间通道注意力模块,并将它嵌入原型网络中构建了类内-类间通道注意力少样本分类方法。该方法利用类内-类间通道注意力模块改善了特征提取器以获得鉴别性特征。其中,类内通道权重在强调类内相似信息的同时抑制了类内冗余信息;类间通道权重强调类间判别性信息,使嵌入空间的特征分布更好地表征类别。ICAFSC在元训练之前增加一个预训练阶段,使类内-类间通道注意力模型能够充分学习类内通道权重和类间通道权重,提高了样本特征的鉴别能力。为了验证所提方法的有效性,在MiniImagenet数据集上分别进行了1-shot和5-shot的少样本分类实验。实验结果表明:ICAFSC少样本分类方法在1-shot和5-shot条件下分别取得了51.55%和69.11%的分类准确率。与原型网络相比,ICAFSC的分类准确率在1-shot条件下提高了1.93%,在5-shot条件下提高了1.15%。显著提升了少样本分类模型的性能。

参考文献:

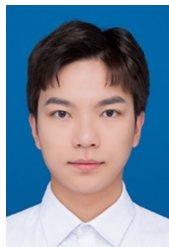
- [1] JIANG K, ZHU L, SUN Q D. Joint dual-structural constrained and non-negative analysis representation learning for pattern classification[J]. *Applied Artificial Intelligence*, 2023, 37(1): 2180821.
- [2] AGARWAL N, SONDHI A, CHOPRA K, *et al.* *Transfer Learning: Survey and Classification*[M]. Smart Innovations in Communication and Computational Sciences. Singapore: Springer Singapore, 2020: 145-155.
- [3] LI X, SUN Z, XUE J H, *et al.* A concise review of recent few-shot meta-learning methods[J]. *Neurocomputing*, 2021, 456: 463-468.
- [4] SONG Y, WANG T, MONDAL S K, *et al.* *A Comprehensive Survey of Few-shot Learning: Evolution, Applications, Challenges, and Opportunities*[EB/OL]. <https://arxiv.org/abs/2205.06743>, 2022-3-24.
- [5] LU J, GONG P, YE J, *et al.* *Learning from Very Few Samples: a Survey*[EB/OL]. <https://arxiv.org/abs/2007.15484>, 2022-7-30.
- [6] HOSPEDALES T, ANTONIOU A, MICAELLI P, *et al.* Meta-learning in neural networks: a survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(9): 5149-5169.
- [7] HUISMAN M, RIJN J N, PLAAT A. A survey of deep meta-learning[J]. *Artificial Intelligence Review*, 2021, 54(6): 4483-4541.
- [8] SNELL J, SWERSKY K, ZEMEL R. Prototypical networks for few-shot learning[C]. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. New York: ACM, 2017: 4080-4090.
- [9] YAN S P, ZHANG S Y, HE X M. A dual attention network with semantic embedding for few-shot learning[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, 33(1): 9079-9086.
- [10] LEE S, MOON W, HEO J P. Task discrepancy maximization for fine-grained few-shot classification[C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 5321-5330.
- [11] GIDARIS S, BURSUC A, KOMODAKIS N, *et al.* Boosting few-shot visual learning with self-supervision[C]. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. October 27 - November 2, 2019, Seoul, Korea (South). IEEE, 2020: 8058-8067.
- [12] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. June 18-23, 2018, Salt Lake City, UT, USA. IEEE, 2018: 7132-7141.
- [13] GUO M H, XU T X, LIU J J, *et al.* Attention mechanisms in computer vision: a survey[J]. *Computational Visual Media*, 2022, 8(3): 331-368.
- [14] VINYALS O, BLUNDELL C, LILLICRAP T, *et al.* Matching networks for one shot learning[C]. *Proceedings of the 30th International Conference on Neural Information Processing Systems*. New York: ACM, 2016: 3637-3645.
- [15] RAVI S, LAROCHELLE H. Optimization as a model for few-shot learning[C]. *Proceedings of the International Conference on Learning Representations*, 2017.
- [16] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks[C]. *Proceedings of the 34th International Conference on Machine Learning - Volume 70*. August 6 - 11, 2017, Sydney, NSW, Australia. New York: ACM, 2017: 1126-1135.
- [17] LIU X, ZHOU F, LIU J, *et al.* Meta-Learning based prototype-relation network for few-shot classification[J]. *Neurocomputing*, 2020, 383: 224-234.
- [18] OH J, YOO H, KIM C, *et al.* BOIL: Towards Representation Change for Few-shot Learning[EB/OL]. 2020: *arXiv*: 2008.08882. <https://arxiv.org/abs/2008.08882.pdf>.
- [19] ZHOU F, ZHANG L, WEI W. Meta-generating deep attentive metric for few-shot classification[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, 32(10): 6863-6873.
- [20] AN Y X, XUE H, ZHAO X Y, *et al.* Conditional self-supervised learning for few-shot classification[C]. *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*. August 19-27, 2021. Montreal, Canada. California: International Joint Conferences on Artificial Intelligence Organization, 2021: 2140-2146.
- [21] BERTINETTO L, HENRIQUES J F, TORR P

- H S, *et al.* Meta-learning with differentiable closed-form solvers [C]. *Proceedings of the International Conference on Learning Representations*, 2019.
- [22] DING N, CHEN Y, CUI G, *et al.* *Few-shot Classification with Hypersphere Modeling of Prototypes* [EB/OL]. [2022-11-10]. <https://arxiv.53yu.com/abs/2211.05319>.
- [23] SNELL J, RICHARD Z. *Bayesian few-shot Classification with One-vs-each Pólya-gamma Augmented Gaussian Processes* [EB/OL]. [2021-1-21]. <https://arxiv.org/abs/2007.10417>.
- [24] SENDERA M, PRZEWIEŻLIKOWSKI M, KARANOWSKI K, *et al.* HyperShot: few-shot learning by kernel HyperNetworks [C]. 2023 *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. January 2-7, 2023, Waikoloa, HI, USA. IEEE, 2023: 2468-2477.
- [25] ALLEN K, SHELHAMER E, SHIN H, *et al.* Infinite mixture prototypes for few-shot learning [C]. *Proceedings of the International Conference on Machine Learning*, 2019: 232-241.

作者简介:



杨利平(1981—),男,内蒙古鄂尔多斯人,博士,副教授,硕士生导师,主要研究方向为机器学习、模式识别和信号处理。E-mail:yanglp@cqu.edu.cn



张天洋(1998—),男,重庆綦江人,硕士研究生,2020年于重庆大学获得学士学位,主要从事少样本学习在图像分类和声音事件检测中应用的研究。E-mail:zhangty@cqu.edu.cn