

DOI:10.16356/j.1005-2615.2022.S.013

## 面向深度神经网络应用的小样本学习技术研究

徐惠灵, 尚政国, 董胜波, 苏琪雅

(北京遥感设备研究所, 北京 100854)

**摘要:** 深度神经网络相关研究中都面临着小样本数据集的挑战, 小样本学习技术研究逐渐受到广泛关注。目前, 国内关于小样本学习的相关研究综述缺乏全面性和系统性。本文针对面向深度神经网络应用的小样本学习技术进行综述。本文着眼于小样本学习算法框架的关键组成因素, 将其分为基于数据增强和基于网络模型两大类, 同时对代表性算法开展深入讨论与分析。最后, 通过对小样本学习研究现状的总结与分析, 对小样本学习的未来发展进行了展望, 期望为该领域后续的研究工作提供启示。

**关键词:** 小样本学习; 数据增强; 网络模型; 深度学习

**中图分类号:** TP391.41; TP181

**文献标志码:** A

**文章编号:** 1005-2615(2022)S-0080-07

### Review on Few-Shot Learning for DNN Applications

XU Huiling, SHANG Zhengguo, DONG Shengbo, SU Qiya

(Beijing Institute of Remote Sensing Equipment, Beijing 100854, China)

**Abstract:** In many application scenarios, the small number of samples imposes large challenges for deep neural networks (DNNs) and few-shot learning (FSL) has received widespread attention in recent years. However, there is a lack of comprehensive and systematic domestic review on this issue. This paper reviews data-augmentation-based and network-model-based approaches of DNNs. They are two key components of the algorithmic framework. Further, representative algorithms of each approach are elaborated. Finally, this paper summarizes current challenges for FSL and future development of FSL. It is expected to provide inspirations for the subsequent research works in this field.

**Key words:** few-shot learning (FSL); data augmentation; network model; deep learning (DL)

深度学习<sup>[1]</sup>(Deep learning, DL)算法研究是在神经科学、应用数学等多领域研究成果的基础上,为实现学习多层次组合任务开展的研究工作。随着大数据时代的到来,深度学习模型已经在图像分类、文本分类等任务中取得了先进成果。但深度学习模型的成功,很大程度上依赖于大量的训练数据。然而在真实应用场景中,存在许多只有少量数据或少量标注数据可供模型进行训练的情况。近年来,许多研究者针对小样本学习或其在相关领域的应用算法进行了综述,但由于小样本学习算法的多样性,出于不同的论述思路,研究者对其的分类

方式也各种各样。本文着眼于小样本学习算法框架的关键组成因素,将小样本学习技术分为基于数据增强和基于网络模型两大类,并针对代表性算法开展深入讨论与分析。最后,对小样本学习研究现状进行总结与分析,并对其未来发展进行了展望。

### 1 小样本学习技术分类

至今小样本学习方法的分类方式仍旧没有一个统一、完善的规范。刘颖等<sup>[2]</sup>根据小样本学习算法在图像分类领域不同的建模方式将其分为两类:卷积神经网络模型和图神经网络模型。也有国内

**收稿日期:** 2022-06-07; **修订日期:** 2022-07-11

**通信作者:** 尚政国, 男, 研究员, 硕士生导师, E-mail: ht\_shangzhengguo@163.com。

**引用格式:** 徐惠灵, 尚政国, 董胜波, 等. 面向深度神经网络应用的小样本学习技术研究[J]. 南京航空航天大学学报, 2022, 54(S): 80-86. XU Huiling, SHANG Zhengguo, DONG Shengbo, et al. Review on few-shot learning for DNN applications[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2022, 54(S): 80-86.

外的研究人员根据采用技术方法的不同对其进行分类<sup>[3-5]</sup>。结合前人的研究,本文认为小样本学习算法与一般深度学习算法在整体框架上是相同的,均是由数据及网络模型组成,其算法框架如图1所示。数据内部包含了任务实现所需的特征信息,多样的特征信息可为模型提供优异的泛化能力。网络模型可通过参数和架构设计,从有限信息中学习特征选择与提取能力,实现算法功能。小样本学习在面对数据多样性受限问题的同时,还在网络模型快速收敛能力、泛化能力等方面存在极大挑战。

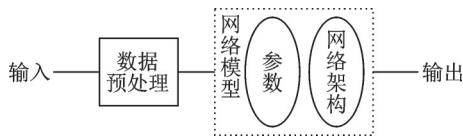


图1 小样本学习算法框架

Fig.1 Algorithmic framework of few-shot learning

为了实现对小样本学习更系统、更全面的论述,同时方便后续针对小样本学习的研究,本文将小样本学习算法分为基于数据增强和基于网络模型两大类,小样本学习技术具体分类结果如图2所示。数据增强方法旨在通过在样本空间或特征空间对数据进行增强操作,提升样本数据多样性,为模型提供充足的特征信息。网络模型方法从参数和网络架构设计两个层面入手,旨在有效提升包括快速收敛能力、泛化能力以及可解释性在内的各项模型性能。

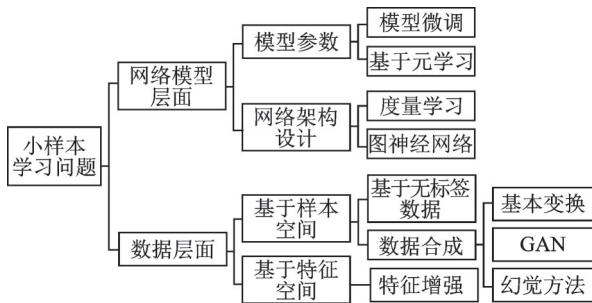


图2 小样本学习方法分类

Fig.2 Classification of few-shot learning approaches

## 2 基于网络模型设计的小样本学习

小样本学习算法研究的关键在于解决过拟合以及训练阶段的收敛问题,而模型的架构设计、参数初始化、超参数设置以及参数优化策略等都对其存在很大影响。故网络模型设计应作为小样本学习研究的重点,结合上述分析,本文将小样本学习中基于网络模型设计的方法按参数优化方法及网络架构设计方法两大类进行分别论述。其中基于模型参数的方法包含基于参数初始化的方法和基于参数优化的方法,基于网络框架设计的方法主要对小样本学习问题上的度量学习方法及图神经网络

方法进行论述。

### 2.1 基于模型参数的方法

由于小样本学习有标注数据数量少的特点,训练深度模型变成一个十分困难的问题。通用模型很大程度受到初始化参数选择以及参数优化策略的影响,导致其在小样本学习中不易收敛。目前小样本学习领域中最通用的参数优化方法主要基于迁移思想。迁移思想是指以从旧知识中学习到的内容为指导,对新知识进行学习的思想。主要目标是将已经学会的知识迁移到一个新的领域中,帮助模型拥有快速掌握新知识的能力,即在小样本目标数据集上快速收敛。基于迁移思想的模型参数优化方法主要包括针对参数优化策略设计的模型微调方法以及针对参数初始化的元学习方法。

#### 2.1.1 模型微调方法

模型微调是指在目标小样本数据集上通过相应的训练策略对经过大规模源数据集预训练的神经网络模型进行参数优化,其方法如图3所示。参数经过在源数据集上的预训练可以帮助模型在小样本目标数据集上快速收敛。若目标数据集和源数据集分布较类似,则可认为两数据集的顶层特征相似性高,故可采用模型微调方法,仅对顶层特征提取器及分类器进行微调即可实现算法收敛同时兼具泛化能力。模型微调需要重点关注架构约束、网络中待调参数的范围以及学习策略等问题。Howard等<sup>[6]</sup>针对文本分类提出了通用微调语言模型(Universal language model fine-tuning, ULM-Fit)。该算法从纵向和横向两个维度学习速率的变化对语言模型进行微调,让模型更快地在小样本数据集上收敛;同时,让模型学习到的知识更符合目标任务。Nakamura等<sup>[7]</sup>提出在小样本数据集上使用更低的学习率,在微调阶段使用自适应的梯度优化器的方法。同时作者还提出若源数据集和目标数据集之间存在较大差异性,可以通过调整整个网络来实现需求。

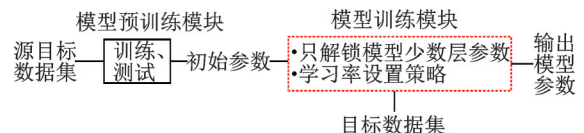


图3 模型微调方法示意图

Fig.3 Illustration of fine-tune based approach

#### 2.1.2 基于元学习的参数初始化方法

元学习的目的是让模型获得一种学习能力,这种学习能力可以让模型自动学习到一些元知识,即学会学习,其中元知识指在模型训练过程之外可以学习到的知识。元学习思想即利用基于任务的先

验知识学习到利于未来发展的初始参数,其方法示意图如图4所示。元学习中的数据集分为元训练集和元测试集,二者均包含了原始模型所需要的训练集和测试集,为了便于区分,二者一般又分别被称作支持集和查询集。其中支持集为小样本数据集。

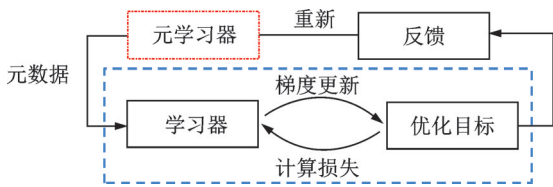


图4 基于元学习的参数初始化方法示意图

Fig.4 Illustration of meta-learning based approach to initialize parameters

基于元学习的参数优化策略是希望可以在源数据集上利用元学习思想对元测试阶段的初始化参数进行优化,得到稳定的初始参数,使模型在小样本目标数据集上快速收敛。不仅如此,元学习相较于一般迁移学习更在乎模型未来的发展潜力。Finn等<sup>[8]</sup>提出了未知模型的元学习方法(Model-agnostic meta-learning, MAML)。作者认为可以找到神经网络中对不同的N-ways-K-shot任务均较为敏感的参数,即这些参数具备提取共有特征的能力,而不考虑特定任务。将这些参数作为初始参数,通过微调这些参数,让模型在新任务小样本数据集上的损失函数快速收敛。Jamal等<sup>[9]</sup>对MAML进行了改进,在输出预测时加入了正则化项,避免了元学习器的偏移,提出了未知任务元学习方法(Task-agnostic meta-learning, TAML)。考虑到不同任务之间的差异,Yu等<sup>[10]</sup>提出了多任务聚类的元学习方法。该算法使相似的任务共享一套分类器参数,通过调整各分类器的权重得到当前任务的分类器。Xiang等<sup>[11]</sup>针对自然语言处理领域的应用需求,基于MAML提出了基于注意力机制的未知任务元学习方法(Attentive task-agnostic meta-learner, ATAML)。同时为解决MAML在深层网络中泛化性较差的问题,Sun等<sup>[12]</sup>和Liu等<sup>[13]</sup>在MAML的基础上融合了模型微调的方法对算法进行了改进。Wang等<sup>[14]</sup>融入了语义信息,提出了任务感知特征嵌入网络(Task-aware feature embedding network, TAFE-Net)。该算法将标签嵌入用于预测数据特征提取模型的权重。

## 2.2 基于网络架构设计的方法

以上算法均是在一定的网络架构上对模型参数进行优化,旨在帮助模型在小样本数据集上实现快速收敛并具备一定泛化能力。除此之外,本文还

可考虑结合小样本数据集自身数据量小的特点,将劣势转为优势,充分发掘小样本学习的特性,对网络架构进行重新设计。下文将对小样本学习问题中常用的基于度量方法和基于图神经网络方法的网络架构设计方案进行论述。

### 2.2.1 度量学习方法

一般的网络模型无法适应小样本学习问题很大的原因是因为其待优化参数数量过于庞大,仅有的数据不足以优化这些参数,造成过拟合现象。度量方法作为一种非参数化方法,通过对样本间的距离分布进行建模,使得同类样本靠近,异类样本分离。同时由于度量方法是对样本间的距离进行度量,若样本数量大,则训练和使用模型计算量将十分庞大,故而,小样本学习因为其本身的数据集小的特点,在此方法上反而具有一定优势。基于以上特点,度量方法近几年来被广泛用于小样本学习的研究。度量学习网络框架通用模型如图5所示。其中经典的网络模型包括孪生神经网络<sup>[15]</sup>、匹配网络<sup>[16]</sup>、原型网络<sup>[17]</sup>和关系网络<sup>[18]</sup>等。

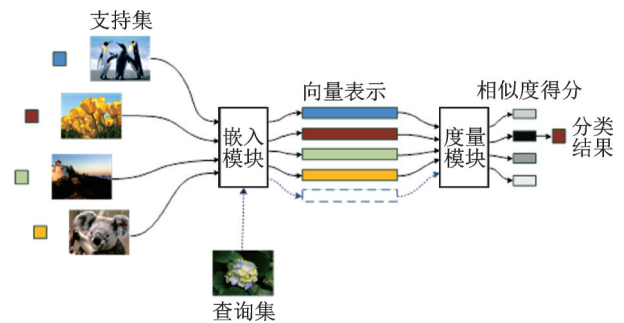


图5 基于度量学习的通用模型

Fig.5 Illustration of metric learning based approach

针对小样本学习问题,研究人员基于经典算法进行了大量改进研究。Wang等<sup>[19]</sup>基于匹配网络,使用GloVe编码方法将图像的标签嵌入向量空间,实现了样本的特征增强。Ren等<sup>[20]</sup>认为一个样本离原型越远,它就越容易被掩盖,而原型网络通过取平均的操作求得的类别原型并未考虑样本权重。针对此问题,Gao等<sup>[21]</sup>和Sun等<sup>[22]</sup>分别通过不同的注意力机制添加方式对原型网络进行了改进。Zhang等<sup>[23]</sup>在关系网络的基础上提出了深度比较网络(Deep comparison network, DCN)。该算法将嵌入学习分解为一系列模块,并将每个模块与一个关系模块配对。之前的度量方法都基于一阶统计量,考虑二阶统计信息,Li等<sup>[24]</sup>提出了一个协方差度量网络(Covariance metric networks, Cov-aMNet)。为了捕捉局部特征,Li等<sup>[25]</sup>对前者进行了改进,同时对样本的多个特征进行度量,提出了深度近邻神经网络(Deep nearest neighbor neural



network, DN4)。

在应用算法研究上,晏媛等<sup>[26]</sup>针对小样本 SAR 图像识别提出将原型网络与基于迁移思想的分类器微调方法有机结合的网络框架。作者认为在目标数据集与源数据集分布相近的情况下,原型网络相较迁移学习方法网络性能更佳,但同时两个数据集分布差异的增大对其造成的不良影响也更大。

### 2.2.2 图神经网络

图由节点和边构成,节点有节点的特征,边有边的特征。图神经网络<sup>[27]</sup>(Graph neural network, GNN)1个节点是1个样本,每条边分别代表着不同样本之间的关系。相较于一般神经网络,GNN在考虑样本自身信息的同时还考虑了样本之间的信息。图神经网络模型的学习过程如图6所示。同时,由于图神经网络的特殊结构,其还具有较好的性能和可解释性。类似于度量学习方法,考虑到计算量问题,图神经网络的特殊结构导致其不适用于大样本数据,小样本学习反而在这方面具有一定优势,可以利用图神经网络的特性为模型提供更好的可解释性性能。比较常用的GNN有图卷积神经网络(Graph convolutional network, GCN)、门控图神经网络和图注意力网络等。

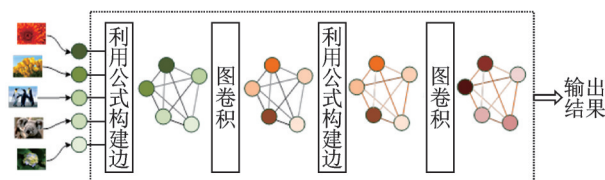


图6 图神经网络模型

Fig.6 Model of graph neural network

基于图卷积神经网络,Garcia等<sup>[28]</sup>将样本分别看作图的不同节点,通过对图中的节点进行分类实现了小样本图像分类任务。Kim等<sup>[29]</sup>则从另一角度入手,通过对图中的边进行分类操作,实现分类功能。Gidaris等<sup>[30]</sup>对传统的GNN进行了改进,在GNN基础上加入了降噪自编码器,通过对输入数据加入高斯噪声来提高模型泛化能力。

## 3 基于数据增强的小样本学习

基于数据增强的方法是利用辅助数据集或者辅助信息增强目标数据集在数据空间或特征空间的信息量,提高样本多样性,使模型能更好地提取特征。在小样本学习领域,基于数据增强的方法多用于在现有网络模型的基础上对模型进行优化,可视作一种辅助优化算法。本文将基于数据增强的小样本学习算法按数据增强操作所在空间的不同,

将其分为在数据空间的数据扩充方法和在特征空间的特征增强方法。

### 3.1 基于样本空间的方法

基于样本空间的数据扩充是指在样本空间上向原有数据集添加新的数据,可以是无标签数据或者合成的带标签数据,利用这些辅助数据来增强样本空间,故下文将数据扩充方法细分为基于无标签数据的数据扩充和基于数据生成的数据扩充方法进行论述。

#### 3.1.1 基于无标签数据的数据扩充方法

半监督学习是利用未标注数据来强化训练模型,使其符合聚类假设,即学习的决策边界必须位于低密度区域。许多研究者针对不同网络框架,巧妙利用无标注数据的信息,有效提升了模型性能。基于元学习中的MAML的模型,Boney等<sup>[31]</sup>在2018年提出可以利用无标签数据调整嵌入函数的参数,用带标签数据调整分类器的参数。基于度量学习的思路,Ren等<sup>[32]</sup>在2018年提出在原型网络的基础上利用无标注数据的算法,取得了更高的准确率。直推式学习是半监督学习的一个特例,当半监督学习中的无标注数据为待预测数据时,即为直推式学习。Liu等<sup>[33]</sup>基于图神经网络模型,结合直推式学习方法,提出了转导传播网络。该模型使用传导方法将标签传播给无标注数据,利用无标注数据的分布使模型实现了更好的预测。Hou等<sup>[34]</sup>提出了交叉注意力网络,使用了注意力机制和迭代的思想,充分利用了查询集中的样本信息实现了数据增强。

#### 3.1.2 基于数据生成的数据扩充方法

基于数据生成的方法是指为小样本类别合成新的带标签数据来扩充训练数据。除了对原始数据做最基本的拼接、旋转变换等操作,常用的方法还有生成对抗网络<sup>[35]</sup>(Generative adversarial net, GAN)、幻觉方法<sup>[36]</sup>。Mehrotra等<sup>[37]</sup>针对单样本学习问题,对GAN模型进行改进,提出了生成对抗残差成对网络,以生成更多新的有标注样本。Harisharan等<sup>[38]</sup>基于迁移学习的思想,将改进的幻觉方法用于目标数据集训练阶段,以提升模型在目标任务上的泛化能力。Wang等<sup>[39]</sup>将元学习与幻觉图像合成器相结合提出了针对小样本学习的网络模型。除了以上算法,Xian等<sup>[40]</sup>提出了f-VAEGAN-D2模型,该模型结合了变分编码器和GAN的优势,在完成小样本学习图像分类的同时,能够将生成样本的特征空间通过自然语言的形式表现出来,具有可解释性。

### 3.2 基于特征空间的方法

特征增强是在原样本的特征空间中添加便于

分类的特征,增加特征多样性。由于小样本学习的一个关键是如何得到一个泛化性好的特征提取器,显然恰当对数据进行特征增强将有助于针对小样本的学习,但特征空间增强的缺点是很难解释矢量数据。

特征增强方法一般是基于编码器-解码器的网络架构,将样本特征通过编码器映射到其他空间,利用其他空间的先验信息或相关关系对样本进行数据增强。Liu等<sup>[41]</sup>提出了特征迁移网络。该网络通过将样本特征通过编码器映射到具有外观和姿态参数的空间,实现了对目标外观及姿态变化引起的运动轨迹变化的描述。Schwartz等<sup>[42]</sup>提出了可以在特征空间合成新样本的Delta编码器。该模型通过提取同类训练样本之间的可转移类内变形来合成新类别样本。但此方法的特征增强过于简单,无法显著改善分类边界。Chen等<sup>[43]</sup>借助每个类别标签的语义信息,通过将样本特征映射到标签语义空间进行数据增强,再映射回样本特征空间的操作,提出了双向网络TriNet。但是上述方法的分类器大多存在只关注具有强判别性的特征,忽略了其他判别性较弱的区域,不利于网络的泛化。Shen等<sup>[44]</sup>利用不确定的注意力机制提取特征,借助对抗特征提升了模型的泛化能力。

## 4 结 论

由于现实条件下,很多领域均存在样本不易获得或不易标注的情况,小样本学习具有极高的现实需求和意义。目前小样本学习领域的研究工作已有很大进展,涌现了大量性能优良的算法模型,但其在准确率这一重要性能指标的表现上仍不理想,有待进行深入探索。

本文从数据和网络模型两个角度开展了小样本学习技术研究综述。数据层面可以考虑增强样本特征或在数据集的不同空间添加有助于模型学习的辅助数据,实现数据增强,提高数据多样性,为后续模型训练提供更多有效信息。网络模型的设计思想是为了在小样本条件限制下,有效提升包括快速收敛能力、泛化能力和可解释性在内的各项模型性能。同时还可通过对损失函数、训练策略、骨干网络等进行改进与设计,提升模型算法性能。

未来发展方向展望:

(1) 结合应用背景及任务需求,对小样本学习问题开展研究。深度学习算法是服务、植根于数据的,小样本学习亦然,其根本需求是要解决实际问题。在算法设计时,样本数据自身的特点会对算法产生本质影响。因此,针对不同应用背景下的样本

特点及任务需求进行算法深入研究,对于实际应用需求的满足和算法性能的有效提升都具有重要意义和指导价值。

(2) 各应用领域的小样本学习技术研究中,“小样本”的概念有待明确。小样本学习问题中的小样本数据集容量下限应与其面对的任务及数据自身特点相关。面向各领域的小样本技术理论研究和实际应用,对小样本数据集界定方法开展研究是十分必要的,且各领域内统一的定义方式有利于在相对统一的标准下进行算法性能评估,对小样本学习技术研究的未来发展具有重要意义。

(3) 充分挖掘小样本学习技术优势,对数据增强与算法模型设计融合技术开展研究。小样本学习在数据增强和模型参数优化、网络框架设计等方面的算法研究具有多样性,数据增强与模型设计的有机结合可为算法提供有效的特征提取和分类能力。因此,实现合理、有效的技术融合,对小样本学习算法的性能提升具有重要意义,值得深入探索。同时,考虑到各方法在结构设置、彼此特点等方面可能对算法性能造成消极影响,在技术融合时,除关注各自优势及劣势外,还需注意各方法之间的融合方式。

(4) 为充分提升算法性能,需深入开展模型特征提取能力提升方法研究。固定算法框架下,骨干网络设计、算法训练策略等对模型的特征提取能力及分类器性能有较大影响。因此,在一定算法框架下,为充分提升小样本学习算法性能,结合小样本问题数据量不足、对深度神经网络适应性差等特点,开展针对骨干网络设计、算法训练策略等环节的对比研究及改进设计是十分必要的。

## 参考文献:

- [1] IAN H, YOSHIOA B, ARION C. 深度学习[M]. 赵申剑,黎彧君,译.北京:人民邮电出版社,2017.
- [2] 刘颖,雷研博,范久伦,等.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. LIU Ying, LEI Yanbo, FAN Jiulun, et al. Survey on image classification technology based on small sample learning[J]. Acta Automatica Sinica, 2021, 47(2): 297-315.
- [3] LU J, GONG P, YE J, et al. Learning from very few samples: A survey[J]. arXiv Preprint arXiv: 2009.02653, 2020.
- [4] 李新叶,龙慎鹏,朱婧.基于深度神经网络的少样本学习综述[J].计算机应用研究,2020,37(8):2241-2247. LI Xinye, LONG Shenpeng, ZHU Jing. Survey of few-shot learning based on deep neural network[J].

- Application Research of Computers, 2020, 37(8): 2241-2247.
- [5] 赵凯琳,靳小龙,王元卓.小样本学习综述[J].软件学报,2021,32(2): 349-369.  
ZHAO Kailin, JIN Xiaolong, WANG Yuanzhuo. Survey on few-shot learning[J]. Journal of Software, 2021, 32(2): 349-369.
- [6] HOWARD J, RUDER S. Universal language model fine-tuning for text classification[J]. arXiv Preprint arXiv:1801.06146, 2018.
- [7] NAKAMURA A, HARADA T. Revisiting fine-tuning for few-shot learning[J]. arXiv Preprint arXiv: 1910.00216, 2019.
- [8] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks[J]. arXiv Preprint arXiv:1703.03400,2017.
- [9] JAMAL M A, QI G J. Task agnostic meta-learning for few-shot learning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern. [S.l.]: IEEE, 2019: 11719-11727.
- [10] YU M, GUO X, YI J, et al. Diverse few-shot text classification with multiple metrics[C]//Proceedings of the NAACL-HLT. [S.l.]:[s.n.],2018: 1206-1215.
- [11] XIANG J, HAVAEI M, CHARTRAND G, et al. On the importance of attention in meta-learning for few-shot text classification [J]. arXiv Preprint arXiv: 1806.00852,2018.
- [12] SUN Q, LIU Y, CHUA T S, et al. Meta-transfer learning for few-shot learning[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]:IEEE, 2019: 403-412.
- [13] LIU Y, SUN Q, LIU A A, et al. LCC: Learning to customize and combine neural networks for few-shot Learning[J]. arXiv Preprint arXiv: 1904.08479, 2019.
- [14] WANG X, YU F, WANG R, et al. TAFE-Net: Task-aware feature embeddings for low shot learning [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 1831-1840.
- [15] KOCH G, ZEMEL R, SALAKHUTDINOV R. Siamese neural networks for one-shot image recognition [C]//Proceedings of the ICML Deep Learning Workshop. [S.l.]:ICML, 2015.
- [16] VINYALS O, BLUNDELL C, LILLICRAP T, et al. Matching networks for one shot learning[J]. arXiv Preprint arXiv:1606.04080, 2016.
- [17] SNEEL J, SWERSKY K, ZEMEL R. Prototypical networks for few-shot learning[J]. Advances in Neural Information Processing Systems, 2017, 30: 4080-4090.
- [18] SUNG F, YANG Y, ZHANG L, et al. Learning to compare: Relation network for few-shot learning[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]:IEEE, 2018: 1199-1208.
- [19] WANG P, LIU L, SHEN C, et al. Multi-attention network for one shot learning[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]:IEEE, 2017: 2721-2729.
- [20] REN M Y, TRIANTAFILLOU E, RAVI S, et al. Meta-learning for semi-supervised few-shot classification[J]. arXiv Preprint arXiv: 1803. 00676, 2018.
- [21] GAO T Y, HAN X, LIU Z Y, et al. Hybrid attention-based prototypical networks for noisy few-shot relation classification [C] //Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.]: AAAI, 2019: 6407-6414.
- [22] SUN S L, SUN Q F, ZHOU K, et al. Hierarchical attention prototypical networks for few-shot text classification[C] //Proceedings of the Conference on Empirical Methods on Natural Language Proceeding and the 9th Int' Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Hong Kong, China: Association for Computational Linguistics, 2019: 476-485.
- [23] ZHANG X, SUNG F, QIANG Y, et al. Deep comparison: Relation columns for few-shot learning [J]. arXiv Preprint arXiv: 1811.07100, 2018.
- [24] LI W, XU J, HUO J, et al. Distribution consistency based covariance metric networks for few-shot learning [C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S.l.]:AAAI, 2019: 8642-8649.
- [25] LI W, WANG L, XU J, et al. Revisiting local descriptor based image-to-class measure for few-shot learning[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 7260-7268.
- [26] 晏媛,孙俊,孙晶明,等.雷达小样本目标识别方法及应用分析[J].系统工程与电子技术,2021, 43(3): 684-692.  
YAN Yuan, SUN Jun, SUN Jingming, et al. Radar few shot target recognition method and application analysis[J]. Systems Engineering and Electronics, 2021, 43(3): 684-692.
- [27] GORI M, MONFARDINI G, SCARSELLI F. A new model for learning in graph domains[C]// Proceedings of IEEE International Joint Conference on Neural Networks, [S.l.]:IEEE, 2005: 729-734.
- [28] GARCIA V, BRUNA J. Few-shot learning with graph neural networks[C]//Proceedings of the 6th International Conference on Learning Representation,

- (ICLR 2018).[S.l.]:[s.n.], 2018.
- [29] KIM J, KIM T, KIM S, et al. Edge-labeling graph neural network for few-shot learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]:IEEE, 2019: 11-12.
- [30] GIDARIS S, KOMODAKIS N. Generating classification weights with GNN denoising autoencoders for few-shot learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 21-30.
- [31] BONEY R, LLIN A. semi-supervised few-shot learning with maml[C]//Proceedings of International Conference on Learning Representations. [S. l.]: ICLR, 2018.
- [32] REN M, TRIANTAFILLOU E, RAVI S, et al. Meta-learning for semi-supervised few-shot classification [J]. arXiv Preprint arXiv: 1803.00676, 2018.
- [33] LIU Y, LEE J, PARK M, et al. Learning to propagate labels: Transductive propagation network for few-shot learning[J]. arXiv Preprint arXiv: 2018.10002, 2018.
- [34] HOU R, CHANG H, MA B, et al. Cross attention network for few-shot classification[J]. arXiv Preprint arXiv:1910.07677, 2019.
- [35] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge, MA: MIT Press, 2014: 2672-2680.
- [36] MILLE E G, MATSAKIS N E, VIOLA P A. Learning from one example through shared densities on transforms[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2000: 1464-1471.
- [37] MEHROTRA A, DUKKIPATI A. Generative adversarial residual pairwise networks for one shot learning[J]. arXiv Preprint arXiv:1703.08033, 2017.
- [38] HARIHARAN B, GIRSHICK R. Low-shot visual recognition by shrinking and hallucinating features [C]//Proceedings of the IEEE Conference on Computer Vision. [S.l.]:IEEE, 2017: 3018-3027.
- [39] WANG Y X, GIRSHICK R, HEBERT M, et al. Low-shot learning from imaginary data[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 7278-7286.
- [40] XIAN Y, SHARMA S, SCHIELE B, et al. f-VAEGAN-D2: A feature generating framework for any-shot learning[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2019: 10275-10284.
- [41] LIU B, WANG X, DIXIT M, et al. Feature space transfer for data augmentation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 9090-9098.
- [42] SCHWARTZ E, KARLINSKY L, SHTOK J, et al. Delta-encoder: An effective sample synthesis method for few-shot object recognition[J]. Advances in Neural Information Processing Systems, 2018, 31: 2845-2855.
- [43] CHEN Z, FU Y, ZHANG Y, et al. Semantic feature augmentation in few-shot learning[J]. arXiv Preprint arXiv: 1804.05298, 2018.
- [44] SHEN W, SHI Z, SUN J. Learning from adversarial features for few-shot classification[J]. arXiv Preprint arXiv: 1903.10225, 2019.

(编辑:刘彦东)