

基于Swin Transformer和双聚焦相似度的图卷积标签传播网络小样本分类

李广生, 李 烨

上海理工大学光电信息与计算机工程学院, 上海

收稿日期: 2025年4月19日; 录用日期: 2025年5月12日; 发布日期: 2025年5月19日

摘 要

小样本学习旨在通过少量标注样本高效识别新类别, 然而, 现有方法在处理数据稀缺性、关系建模精准性以及图结构泛化能力等问题时仍存在局限性。为此, 文章提出了一种基于Swin Transformer和双聚焦相似度的图卷积标签传播网络(Swin Transformer-Based Graph Convolutional Label Propagation Network, ST-GCLPN)小样本分类方法。首先, 利用Swin Transformer提取输入样本的全局与局部特征, 缓解数据稀缺导致的特征不足问题。其次, 通过双聚焦相似度综合全局和局部特征关系, 构建精准的图结构, 并输入图卷积网络(GCN)以捕获局部邻域特征。随后, GCN的输出作为输入传递到标签传播算法(LPA)中, 通过全局图结构高效传播标签信息, 增强分类泛化能力。最后, 计算查询集的预测标签与真实标签之间的交叉熵损失, 优化模型参数。实验结果表明, 在miniImageNet数据集的5-way 1-shot和5-way 5-shot任务中, 所提方法的分类准确率均有提高, 显著优于当前主流方法。该方法有效解决了小样本学习中特征不足、关系建模不精准及泛化能力有限等问题, 为小样本分类任务提供了一种高效而鲁棒的解决方案。

关键词

Swin Transformer, 双聚焦相似度, 图卷积网络, 小样本学习, 标签传播

Few-Shot Classification with Graph Convolutional Label Propagation Network Based on Swin Transformer and Dual-Focus Similarity

Guangsheng Li, Ye Li

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 19th, 2025; accepted: May 12th, 2025; published: May 19th, 2025

文章引用: 李广生, 李烨. 基于 Swin Transformer 和双聚焦相似度的图卷积标签传播网络小样本分类[J]. 建模与仿真, 2025, 14(5): 488-502. DOI: 10.12677/mos.2025.145409

Abstract

Few-shot learning aims to efficiently recognize new categories with a limited number of annotated samples. However, existing methods still face limitations in addressing issues such as data scarcity, the precision of relational modeling, and the generalization capability of graph structures. To this end, this paper proposes a few-shot classification method based on a Graph Convolutional Label Propagation Network (GCLPN) that incorporates Swin Transformer and dual-focus similarity. Firstly, the Swin Transformer is utilized to extract both global and local features from input samples, alleviating the issue of insufficient features caused by data scarcity. Secondly, the dual-focus similarity integrates global and local feature relationships to construct an accurate graph structure, which is then fed into a Graph Convolutional Network (GCN) to capture local neighborhood features. Subsequently, the output of the GCN is passed as input to a Label Propagation Algorithm (LPA), which efficiently propagates label information through the global graph structure, enhancing the classification generalization capability. Finally, the model parameters are optimized by calculating the cross-entropy loss between the predicted labels and the true labels of the query set. Experimental results demonstrate that the proposed method significantly outperforms current mainstream methods, with improved classification accuracy on the miniImageNet dataset for both 5-way 1-shot and 5-way 5-shot tasks. This method effectively addresses the challenges of feature insufficiency, imprecise relational modeling, and limited generalization capability in few-shot learning, providing an efficient and robust solution for few-shot classification tasks.

Keywords

Swin Transformer, Dual-Focus Similarity, Graph Convolutional Network (GCN), Few-Shot Learning, Label Propagation

Copyright © 2025 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

卷积神经网络(Convolutional Neural Network, CNN)在计算机视觉领域取得了卓越成就,广泛应用于图像分类[1]、语义分割[2]、目标检测[3]和图像字幕生成[4]等任务。然而,这些方法的成功通常依赖于大量标注数据,例如每个类别需要数百甚至数千个样本。这种数据需求不仅耗时且成本高昂,对于某些罕见类别的数据标注更是难以实现,成为限制深度学习技术在实际场景中广泛应用的主要障碍。为了解决这一问题,小样本学习(Few-Shot Learning, FSL)逐渐成为研究热点。小样本学习的目标是通过少量标注样本(如每个类别仅 1~5 个样本)和从基础类别中迁移的知识,实现对新类别的高效识别。

小样本学习的关键在于样本关系度量和知识迁移的有效构建。例如,孪生网络通过可训练的加权距离度量样本间的相似性[5];匹配网络通过共享加权度量学习标注样本与未标注样本之间的关系[6];原型网络通过在嵌入空间中计算样本距离完成分类[7],在小样本分类任务中均展现了良好的效果。然而,这些方法在处理复杂样本关系时,难以充分利用样本之间的图结构信息,导致知识聚合效率较低,从而限制了它们在复杂场景中的适用性。

近年来,基于图神经网络(Graph Neural Network, GNN)的小样本学习方法因其卓越的知识聚合能力而受到关注。GNN 能够通过图结构中传递信息,有效聚合支持集和查询集的知识,天然适合小样本学习

任务。现有方法通常将 GNN 用作标签传播的核心模块, 通过图结构实现节点或边的分类预测[8], 并与特征嵌入网络协同优化以提升分类性能。然而, 标签传播网络(Transductive Propagation Network, TPN) [9] 在动态图结构和标签稀疏场景下表现有限: 其对固定图拓扑结构的依赖性导致在图结构变化时性能显著下降; 同时, 未标注节点的潜在信息利用不足, 也限制了其在标签稀疏场景中的分类能力。

为解决上述问题, 本研究提出了一种创新的小样本分类方法——基于 Swin Transformer 和双聚焦相似度的图卷积网络小样本分类方法。该方法首先通过提取输入样本的全局和局部特征, 从而生成高质量的特征表示; 随后, 利用双聚焦相似度计算, 构建能够准确捕获全局与局部特征关系的图结构。接着, 该图结构被输入到图卷积网络(Graph Convolutional Network, GCN)中, 通过聚合邻域节点特征捕获局部结构关系并生成高质量的节点嵌入表示。最后, 将 GCN 的输出作为输入, 传递到标签传播算法(Label Propagation Algorithm, LPA)中, 通过全局图结构实现高效的标签传播, 从而充分利用已标注和未标注样本之间的关系完成分类任务。这种设计将 GCN 的局部特征聚合能力与 LPA 的全局信息传播能力有机结合, 有效提升了小样本学习的分类性能和泛化能力。

实验结果表明, 在 miniImageNet 数据集的 5-way 1-shot 和 5-way 5-shot 任务中, 该方法的分类准确率均有提高, 显著优于当前主流方法, 验证了其优越性。

本文的主要贡献包括:

1) 该方法提出将 GCN 和 LPA 相结合, 充分利用 GCN 捕获的局部邻域结构信息和 LPA 传播的全局图结构信息, 实现了全局与局部特征的有机融合。GCN 负责生成高质量的节点嵌入表示, 而 LPA 则在全局范围内传播标签, 尤其在未标记节点与标记节点的交互中展现了出色的性能。

2) 双聚焦相似度计算方法综合了全局特征的相似性(全局聚焦)与局部特征的相似性(局部聚焦), 使得图结构能够同时捕捉输入样本的多层次关系。通过将这种相似度用作图的边权重, 构建了更加合理的图结构, 从而为后续的图卷积操作和标签传播提供了可靠的基础。

3) Swin Transformer 的引入增强了特征提取模块的能力, 通过从全局和局部视角生成具有判别力的特征表示, 有效缓解了小样本场景中特征不足的问题。其提取的特征为双聚焦相似度的计算和图结构的构建提供了更优质的支持。

2. 相关工作

2.1. 基于度量学习的小样本学习

在小样本学习的领域内, 度量学习方法特别强调在特征空间上学习相似性分类器的重要性[10], 其致力于掌握高质量且可迁移的特征, 并依赖于跨任务的神经网络骨干架构。匹配网络(Matching Network)采用了端到端可训练的最近邻算法, 利用少量标记样本(即支持集)的学习嵌入来预测未标记样本(即查询集)的类别。原型网络(Prototypical Network)进一步发展了这一概念, 通过构建每个类别的原型表示来进行分类。关系网络(Relation Network)通过一个简单的神经网络来学习非线性距离度量, 而不是依赖于传统的固定线性距离度量, 例如余弦或欧几里得距离。这些方法通常使用小批量的训练数据对端到端网络进行训练, 假设训练过程中获得的特征可以有效代表新的测试类别。使用神经网络学习非线性距离度量参数, 并将其与欧几里得距离结合, 以求得相似度并构建邻接矩阵, 可以更准确地度量数据之间的相似性, 尤其是在高维空间中。

2.2. 基于图的小样本学习

在小样本学习的研究领域, 基于图的方法常被视为度量学习的一个特殊分支。这是因为多数这类方法都依赖于使用基于径向基函数(RBF)的邻接矩阵来形成图, 以此来促进标签或特征的传播。例如,

Satorras 等人[11]通过在支持集和未标记数据之间构建亲和矩阵, 实现了标签的有效传播。wDAEGNN [12] 利用图神经网络(GNN)生成分类权重, 并结合去噪自动编码器(DAE)以正则化特征表示。此外, 嵌入传播的方法不仅传播标签, 还传播嵌入本身, 以减少类内距离, 从而提高分类准确性[13]。集合到集合的函数也被用于嵌入适配, 其中图卷积网络(GCN)被应用于实例化这种集合到集合的函数[14]。

2.3. 半监督小样本学习

在半监督的小样本学习中, 除了提供标记的支持集外, 还会包括未标记的数据, 假设这些未标记数据与目标类别具有相似的分布, 但也可能包含一些无关的噪声样本。在 LST [15]的研究中, 研究人员通过自标记和软注意机制对未标记样本进行间歇性处理, 并结合标记数据及自标记数据进行模型的微调。类似于 LST 的方法, Ren 等人[16]利用原型网络初始化的 K-means 迭代来更新类别原型, 并在处理未标记数据时降低那些可能不属于目标类别的干扰样本的影响。Simon 等人[17]则通过软标签传播技术利用未标记样本。此外, Saito 等人[18]探讨了半监督小样本领域自适应的问题。在其他研究中[19], 图神经网络被应用于半监督 FSL 设置, 以促进标记和未标记样本间的信息共享。在 TPN 中, 使用图构建网络来预测特定任务的图, 用于在半监督 FSL 任务中传播标签。Liu 等人[20]指出原型表示与理想表示之间的偏差, 并提出了一种基于类内和类间假设的简单策略来纠正这种偏差。

3. 方法

3.1. 问题定义

小样本分类旨在每个类别只有极少数训练样本的条件下学习分类器。因此, 每个小样本分类任务 T 包括一个支持集 S (一组已标记的输入 - 标签对), 以及一个查询集 Q (用于评估所学分类器的未标记集合)。当支持集 S 为每 N 个不同类别提供 K 个已标记样本时, 这种情况被称作 N -way K -shot 分类问题。

近年来, 元学习已成为解决小样本分类问题的标准策略。从理论上讲, 小样本学习可以仅依赖于任务的紧凑支持集, 训练分类器为查询集中的每个样本分配类别标签。然而, 有限的已标记支持样本数量不足以训练出能够充分反映类间和类内差异的模型, 常常导致分类性能不佳。通过从显式训练集中提取可迁移知识, 元学习解决了这一难题, 使得模型能在支持集上进行更高效的小样本学习, 并因此更成功地对查询集进行分类。

本文采用了情景训练作为元学习的高效方法, 情景训练的核心思想是在相对较大的已标记训练数据集的基础上抽样出训练任务(即情景), 模仿测试任务的少镜头学习环境。由于训练任务的分布被认为与测试任务相似, 通过让模型在训练任务上表现出色, 可以提升测试任务的性能。

具体地, 情景训练中的训练和测试任务都是按照 N -way K -shot 问题的形式构建的:

$$\Gamma = S \cup Q \quad (1)$$

其中: Γ 表示任务, $S = \{(x_i, y_i)\}_{i=1}^{N \times K}$ 和 $Q = \{(x_i, y_i)\}_{i=N \times K+1}^{N \times K+T}$ 分别是支持集和验证集, N 是类的数量, K 是从每个类中抽取数据样本的数量, T 是查询样本的数量。 $x_i, y_i \in \{C_1, \dots, C_N\} \subset C$, x_i 和 y_i 分别是第 i 个输入数据及其标签, C 是训练或测试数据集的所有类别集合。尽管训练和测试任务均来自相同的任务分布, 但它们的标签空间是互不相交的, 即 $C_{train} \cap C_{test} = \emptyset$ 。每个情景中的支持集 S 充当已标记的训练集, 模型在其上进行训练, 目的是最小化对验证集 Q 的预测损失。这个训练过程会重复进行, 直至模型收敛。若 $N \times K$ 个支持样本中存在未标记样本, 则该问题被称作半监督小样本分类。

3.2. 模型

本研究模型如图 1 所示, 首先, 通过 Swin Transformer 提高模型对输入数据的理解和表示能力, 生

成全局和局部特征的高质量嵌入表示[21]。其次,在图构造模块中,利用双聚焦相似度综合全局特征和局部特征间的关系,生成精准的节点特征和边权重,并构建图结构,以捕获输入样本的多层次关系。然后,将构建的图结构输入到图卷积网络(GCN)中,通过聚合邻域节点特征生成高质量的节点嵌入表示,并将其输出至标签传播算法(LPA),实现从支持集 S 到查询集 Q 的高效标签传播。最后,在损失生成阶段,通过计算传播标签与查询集 Q 上真实标签之间的交叉熵损失,直接优化模型参数。在联合训练框架中,交叉熵损失引导模型在支持集和查询集上的特征表示和标签传播效果不断提升。

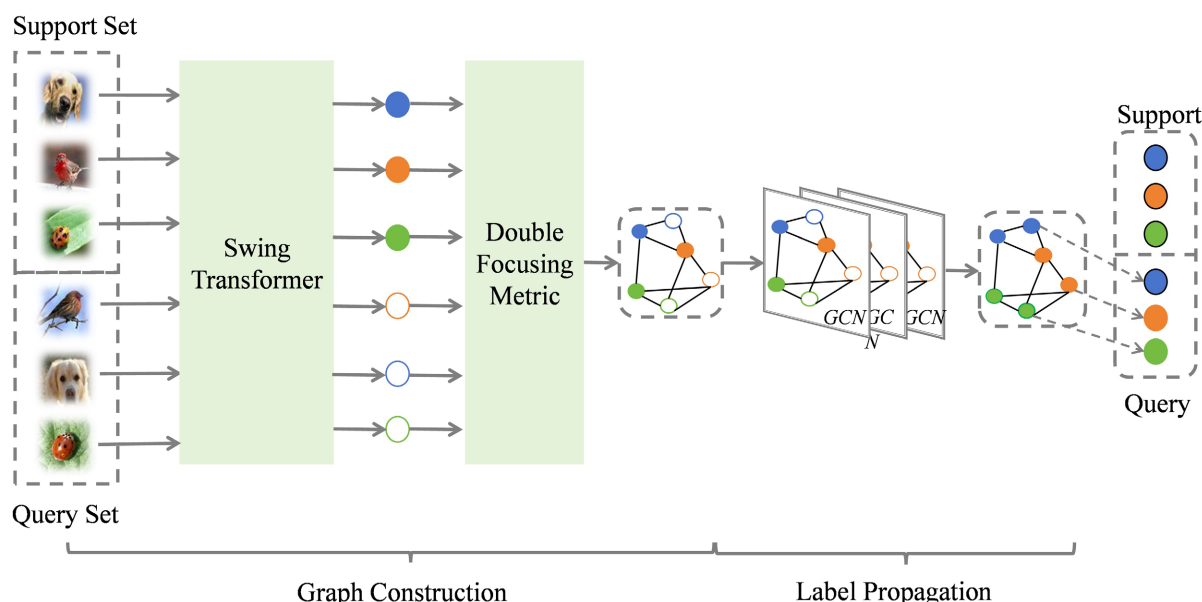


Figure 1. Model architecture diagram

图 1. 模型架构图

3.3. Swin Transformer 网络的特征提取模块

采用 Swin Transformer 作为图神经网络在小样本图像分类任务中的特征提取模块,将其输出的图像特征作为图神经网络的节点特征。Swin Transformer 的结构如图 2 所示,主要包括卷积层、线性层、Patch Merging 操作、Block、自适应全局池化层和全连接层。首先,输入图像通过卷积操作被划分成大小为 4×4 的不重叠图像块;随后,这些图像块经过线性嵌入层被转化为特征序列。在每个 Block 中,利用自注意力机制提取图像的特征,捕捉更丰富的局部与全局信息。通过 Patch Merging 操作,对特征图进行下采样以缩减宽度和高度,同时增加通道数量。在多次 Block 和 Patch Merging 操作的作用下,逐步提取图像的深层特征。最后,使用全连接层将这些特征映射到最终的特征空间。通过采用 Swin Transformer 作为图像特征提取器,模型能够更有效地捕捉图像的全局特征,从而显著增强图神经网络的节点特征表示能力。

Block 的结构如图 3 所示,包含四个关键模块:批归一化(Layer Normalization, LN)、窗口内多头自注意力(Window-Based Multi-Head Self-Attention, W-MSA)、窗口移动多头自注意力(Shifted Window Multi-head Self-Attention, SW-MSA)和多层感知机(Multi-Layer Perceptron, MLP)。其中, LN 模块对输入特征进行归一化处理,确保不同通道的特征分布一致性,从而为后续计算提供稳定的输入。W-MSA 模块通过计算窗口内的多头自注意力,利用注意力权重对特征进行加权聚合,促进不同位置间的特征交互与信息传递。SW-MSA 模块通过引入窗口移动机制,在局部区域内对特征进行平移和重组,从而进一步整合特征信息并捕捉更丰富的上下文关系。MLP 模块作为全连接前馈网络,由多层全连接层和非线性激活函数组

成, 用于对特征进行复杂的非线性变换, 进一步增强特征表示能力。

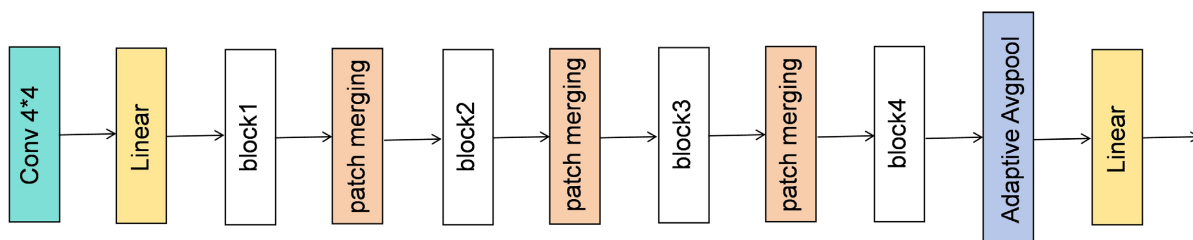


Figure 2. Swin Transformer structure diagram

图 2. Swin Transformer 结构图

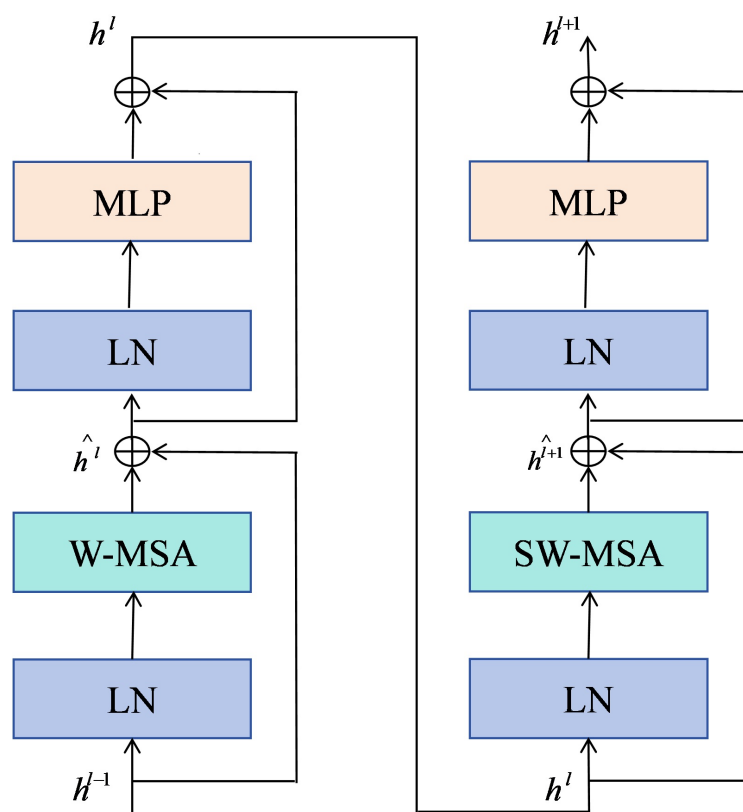


Figure 3. Block structure diagram

图 3. Block 结构图

通过 LN、W-MSA、SW-MSA 和 MLP 的协同作用, Block 不仅有效建模了窗口内的特征关系, 还保留了窗口之间的相对位置信息, 从而实现多尺度特征的提取与整合, 为后续任务提供了丰富的特征表达。

Block 的计算过程为:

$$\begin{aligned}
 \hat{h}^l &= W-MSA[LN(h^{l-1})] + h^{l-1} \\
 h^l &= MLP[LN(\hat{h}^l)] + \hat{h}^l \\
 \hat{h}^{l+1} &= SW-MSA[LN(h^l)] + h^l \\
 h^{l+1} &= MLP[LN(\hat{h}^{l+1})] + \hat{h}^{l+1}
 \end{aligned} \tag{2}$$

3.4. 双聚焦相似度模块

在 Swin Transformer 提取的全局和局部特征基础上, 本文设计了双聚焦相似度模块, 旨在综合全局特征和局部特征之间的多层次关系, 以构建更加精准的图结构。双聚焦相似度通过结合全局相似度和局部相似度, 全面刻画样本间的复杂关系, 为后续的图神经网络节点特征表示和标签传播提供了可靠支持。

3.4.1. 全局相似度计算

全局相似度衡量的是样本整体特征之间的关系, 反映了样本在高维特征空间中的整体语义相似性。在 Swin transformer 提取的全局特征为 $F_{global} \in \mathbb{R}^{N \times d}$, 其中 N 表示样本数量, d 为特征维度。全局特征矩阵 W_{global} 的计算过程为:

$$W_{global}(i, j) = \exp\left(-\frac{\|F_{global}^i - F_{global}^j\|^2}{2\sigma^2}\right) \quad (3)$$

其中, $\|F_{global}^i - F_{global}^j\|^2$ 表示样本 i 和样本 j 的欧氏距离, σ 是控制距离对相似度影响的平滑参数。

3.4.2. 局部相似度计算

局部相似度衡量的是样本在局部区域上的特征相似性, 能够捕捉细粒度的特征关系。Swin Transformer 提取的局部特征为 $F_{local} \in \mathbb{R}^{N \times K \times d}$, 其中 K 表示局部区域数量。局部特征矩阵 W_{local} 的计算过程为:

$$W_{local}(i, j) = \frac{1}{K} \sum_{k=1}^K \exp\left(-\frac{\|F_{local}^{(i,k)} - F_{local}^{(j,k)}\|^2}{2\sigma^2}\right) \quad (4)$$

其中: 该公式表示逐区域计算样本间的相似度, 并对所有局部区域的相似度进行平均。 $F_{local}^{(i,k)}$ 表示样本 i 在第 k 个局部区域的特征。

3.4.3. 双聚焦相似度融合

为综合全局相似度和局部相似度的优点, 本文提出了双聚焦相似度融合策略。通过加权组合的方式, 将全局和局部特征的相似性融合为一个统一的相似度矩阵 W , 其计算公式为:

$$W(i, j) = \alpha \cdot W_{global}(i, j) + \beta \cdot W_{local}(i, j) \quad (5)$$

其中, α 和 β 为可调的权重参数, 分别控制全局相似度和局部相似度对最终结果的贡献。

双聚焦相似度的融合策略确保了样本间关系的全面建模, 既能捕捉全局语义信息, 又能充分利用局部特征的细粒度差异, 从而构建更加精准的相似性图。

3.5. 图卷积网络和标签传播模块

在完成双聚焦相似度矩阵 W 的构建后, 本文基于图卷积网络的训练模块, 通过图卷积网络对节点特征进行学习与更新, 提取更加高效的图结构表示。随后, 利用标签传播算法对未标注样本进行标签传播, 从而生成概率分布, 完成小样本分类任务。

通过双聚焦相似度模块, 综合全局特征相似度和局部特征相似度, 构建相似度矩阵 W 。 W 表示节点之间的边权重, 其值反映了样本间的相似性。为保证图结构的稀疏性, 本文采用 k -近邻方法, 保留每个节点与其 k 个最相似节点的边, 并对矩阵进行对称化处理, 以构建稀疏的相似度图结构。

3.5.1. 图卷积网络的训练

在构建图结构后, 本文利用图卷积网络对图的节点特征进行学习和优化。图卷积网络通过消息传递

机制(Message Passing)更新节点的特征表示, 其基本计算公式为:

$$H^{(l+1)} = \sigma(\tilde{D}^{-1/2} \tilde{W} \tilde{D}^{-1/2} H^{(l)} W^{(l)}) \quad (6)$$

其中: $\tilde{W}$ 是 W 加上单位矩阵 I 得到的邻接矩阵; $\tilde{D}$ 是 $\tilde{W}$ 的度矩阵; $H^{(l)}$ 是第 l 层的节点特征表示; $W^{(l)}$ 是第 l 层的可学习参数; σ 是非线性激活函数。

通过多层图卷积操作, 模型逐步聚合邻域节点的特征, 同时保留全局和局部的图结构信息, 从而生成高质量的节点嵌入表示。

3.5.2. 标签传播算法

在获得经过图卷积网络更新的节点特征后, 采用标签传播算法对未标注样本进行标签预测。标签传播的目标是利用已标注样本的标签, 通过图结构上的信息传播, 预测未标注样本的标签概率分布。定义 F 为一个包含非负元素的 $(N \times K + T) \times N$ 矩阵集合。在标签矩阵 Y 中, Y 的函数形式为:

$$Y = \begin{cases} 1 & x_i \in S, y_i = j \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

标签传播基于图结构迭代更新中 $S \cup Q$ 实例的未知标签。原始的标签传播公式为:

$$F = (I - \alpha S)^{-1} Y \quad (8)$$

而在迭代的标签传播过程中, 公式为:

$$F(t+1) = \alpha S F(t) + (1 - \alpha) Y \quad (9)$$

其中 α 为平滑系数, Y 是已标注样本的真实标签。

经过多次迭代后, 标签分布 F 收敛, 未标注样本的标签分布即可用于分类任务。

3.6. 损失函数

为了进一步优化模型性能并将标签传播结果与真实标签对齐, 本文定义了监督损失函数, 以衡量标签传播输出的标签分布 F 与标注样本真实标签 Y 之间的差异。本文采用交叉熵损失作为主要的监督信号。交叉熵损失的表达式为:

$$L = - \sum_{i \in L} \sum_{c=1}^C Y_{ic} \log F_{ic} \quad (10)$$

其中: L 表示标注样本的集合; C 是类别数量; Y_{ic} 是样本 i 在第 c 类上的真实标签, 采用独热编码形式; F_{ic} 是样本 i 在第 c 类上的预测概率。

在计算损失后, 模型通过反向传播机制更新所有模块的参数, 包括特征提取器(Swin Transformer)、图卷积网络(GCN)的权重, 以及相似度矩阵 W 的构造方式。这种端到端的优化方法确保模型能够充分利用标注样本的信息, 并提高未标注样本的分类性能。

4. 实验

4.1. 数据集

实验中采用了三个广泛认可的 FSL 标准数据集: miniImageNet、tieredImageNet 和 CUB-200-2011。miniImageNet 包含 100 个类别, 每个类别有 600 张图像, 按照 Vinyals 等人(2016) [6]介绍的标准分为 64 个训练类、16 个验证类和 20 个测试类。tieredImageNet 是 ImageNet ILSVRC-12 的更大子集, 包括 608 个类别的 779,165 张图像, 分为 351 个训练类、97 个验证类和 160 个测试类。与其他两个数据集不同, CUB-

200-2011 是一个细粒度分类数据集, 包含 200 种鸟类的 11,778 张照片, 分为 100 个训练类、50 个验证类和 50 个测试类。所有数据集中的图像都标准化为 84×84 大小。

4.2. 设置

与许多其他基于 CNN 的视觉识别任务类似, 小样本学习模型中需要一个强大的特征嵌入网络, 而主干网络的选择对模型性能有着显著的影响。为了公平地与现有方法进行比较, 实验中采用了广泛使用的 ResNet-12 和基于 Swin Transformer 的特征提取器作为主干。ResNet-12 主干包含四个残差块, 其最终输出特征维数为 640, 是大多数先进 FSL 模型的标准配置。与之相比, Swin Transformer 提供了一种更灵活且表达力更强的特征提取机制, 通过多层次的局部与全局特征建模, 进一步增强了嵌入特征的判别能力。Swin Transformer 的输出特征不仅能捕捉局部区域的细粒度模式, 还能通过窗口移动机制整合全局语义信息, 为后续的图卷积和标签传播模块提供高质量的节点表示。所有实验均采用 Adam 优化器, k -最近邻图的超参数 k 设置为 20, 标签传播的平滑参数 α 设置为 0.99, 初始学习率为 10^{-3} 。针对不同数据集, 学习率调整策略如下: 对于 miniImageNet 和 CUB-200-2011 数据集, 每训练 10,000 次后将学习率减半; 对于 tieredImageNet 数据集, 由于其类别数量更多, 且每个类别包含的样本数量更大, 训练需要更多的迭代以达到学习效果, 因此每训练 25,000 次后将学习率减半。采用较长间隔的学习率调整策略, 可以更充分地利用 tieredImageNet 数据集中丰富的样本信息。

4.3. 实验结果

4.3.1. 主要结果

本文将提出的模型与多种先进模型的性能进行了比较, 包括基于图的方法, 如 TPN [11]、DPGN、EGNN [19]、HGNN [22], 以及基于非图的方法, 如 RelationNet [10]、MatchingNet [6]、ProtoNet [7]、MAML [23]、MetaGAN [24]、SNAIL [25]、Meta-Transfer [26]、TapNet [27]、CloserLook [28]、FEAT [29]、E3BM [30] 和 MetaOptNet [31]。

表 1 是不同算法在 miniImageNet 数据集上的实验结果。其展示了 miniImageNet 数据集上不同方法的 5-way 1-shot 和 5-way 5-shot 任务的准确率。可以看出, ST-GCLPN 在 ResNet-12 作为主干网络的情况下, 1-shot 和 5-shot 任务的准确率分别达到了 64.32% 和 78.55%, 超越了当前的方法, 如 TapNet (61.65%/76.36%) 和 Meta-Transfer (61.20%/75.53%)。当采用 Swin Transformer 作为主干时, ST-GCLPN 在 miniImageNet 上的 1-shot 和 5-shot 任务的准确率进一步提升至 67.45% 和 80.44%, 相较于 ResNet-12 提高了 3.13% (1-shot) 和 1.89% (5-shot), 说明 Swin Transformer 能够更有效地建模局部与全局信息, 从而提升特征表示能力。

表 2 是不同算法在 tieredImageNet 数据集上的实验结果。其展示了 ST-GCLPN 在 tieredImageNet 数据集上的实验结果。相比于 miniImageNet, tieredImageNet 具有更大的类别空间, 因此对模型的泛化能力提出了更高的要求。在 ResNet-12 作为主干的情况下, ST-GCLPN 取得了 71.84% (1-shot) 和 86.64% (5-shot) 的提升, 相较于 E3BM (70.00%/85.00%) 和 ProtoNet (69.63%/84.82%), 实现了更优的分类性能。进一步采用 Swin Transformer 作为主干后, ST-GCLPN 的 1-shot 和 5-shot 任务的准确率分别提高到 73.28% 和 87.35%, 相比 ResNet-12 分别提升了 1.44% (1-shot) 和 0.71% (5-shot), 表明 Swin Transformer 能够更好地适应大规模数据集, 提高小样本学习模型的泛化能力。

表 3 是不同算法在 CUB-200-2011 数据集上的实验结果。其展示了 ST-GCLPN 在 CUB-200-2011 细粒度分类任务上的实验结果。与 miniImageNet 和 tieredImageNet 相比, CUB-200-2011 主要关注细粒度类别分类, 因此更依赖于模型的局部特征提取能力。在 ResNet-12 作为主干时, ST-GCLPN 的 1-shot 和 5-shot 任务的准确率分别为 81.54% 和 88.06%, 相比 ProtoNet、DPGN [32]、CTX 等方法均有较大提升。当

采用 Swin Transformer 作为主干时, ST-GCLPN 在 1-shot 和 5-shot 任务的准确率分别达到 85.78%和 92.46%, 表明 Swin Transformer 的窗口注意力机制能够更有效地学习局部细节信息, 从而提升小样本学习的性能。

Table 1. Accuracy of the 5-way 1-shot and 5-way 5-shot tasks on the miniImageNet dataset

表 1. miniImageNet 数据集上 5-way 1-shot 和 5-way 5-shot 任务的准确度

Method	Backbone	5-way 1-shot	5-way 5-shot
MatchingNet [6]	Conv4	43.56 $\pm$ 0.84	55.31 $\pm$ 0.73
ProtoNet [7]	Conv4	49.42 $\pm$ 0.78	68.20 $\pm$ 0.66
MAML [25]	Conv4	48.70 $\pm$ 1.84	55.31 $\pm$ 0.73
DPGN [10]	Conv4	53.22 $\pm$ 0.31	65.34 $\pm$ 0.29
EGNN [20]	Conv4	51.65 $\pm$ 0.55	66.85 $\pm$ 0.49
HGNN [24]	Conv4	55.63 $\pm$ 0.20	72.48 $\pm$ 0.16
MetaGAN [27]	ResNet12	52.71 $\pm$ 0.64	68.63 $\pm$ 0.67
SNAIL [28]	ResNet12	55.71 $\pm$ 0.99	68.88 $\pm$ 0.92
Meta-Transfer [29]	ResNet12	61.20 $\pm$ 1.80	75.53 $\pm$ 0.80
TPN [11]	ResNet12	59.46 $\pm$ n/a	75.65 $\pm$ n/a
TapNet [30]	ResNet12	61.65 $\pm$ 0.15	76.36 $\pm$ 0.10
ST-GCLPN	ResNet12	64.32 $\pm$ 0.23	78.55 $\pm$ 0.14
ST-GCLPN	Swin transformer	67.45 $\pm$ 0.36	80.44 $\pm$ 0.39

Table 2. Accuracy of the 5-way 1-shot and 5-way 5-shot tasks on the tieredImageNet dataset

表 2. tieredImageNet 数据集上 5-way 1-shot 和 5-way 5-shot 任务的准确度

Method	Backbone	5-way 1-shot	5-way 5-shot
MatchingNet [6]	Conv4	54.02 $\pm$ 0.00	70.11 $\pm$ 0.00
ProtoNet [7]	Conv4	50.89 $\pm$ 0.21	69.26 $\pm$ 0.18
MAML [25]	Conv4	51.67 $\pm$ 1.81	70.30 $\pm$ 0.08
DPGN [10]	Conv4	53.99 $\pm$ 0.31	69.86 $\pm$ 0.28
EGNN [20]	Conv4	47.40 $\pm$ 0.43	62.66 $\pm$ 0.57
HGNN [24]	Conv4	56.05 $\pm$ 0.21	72.82 $\pm$ 0.18
TPN [11]	ResNet12	59.91 $\pm$ 0.94	73.30 $\pm$ 0.75
TapNet [30]	ResNet12	63.08 $\pm$ 0.15	80.26 $\pm$ 0.12
Meta-Transfer [29]	ResNet12	65.62 $\pm$ 1.80	80.61 $\pm$ 0.90
MetaOptNet [31]	ResNet12	65.81 $\pm$ 0.74	81.75 $\pm$ 0.53
ProtoNet [7]	ResNet12	69.63 $\pm$ 0.53	84.82 $\pm$ 0.36
E3BM [30]	ResNet12	70.00 $\pm$ n/a	85.00 $\pm$ n/a
ST-GCLPN	ResNet12	71.84 $\pm$ 0.16	86.64 $\pm$ 0.12
ST-GCLPN	Swin transformer	73.28 $\pm$ 0.14	87.35 $\pm$ 0.28

Table 3. Accuracy of the 5-way 1-shot and 5-way 5-shot tasks on the CUB-200-2011 dataset
表 3. CUB-200-2011 数据集上 5-way 1-shot 和 5-way 5-shot 任务的准确度

Method	Backbone	5-way 1-shot	5-way 5-shot
MatchingNet [6]	Conv4	61.16 $\pm$ 0.89	72.86 $\pm$ 0.70
MAML [25]	Conv4	55.92 $\pm$ 0.95	72.09 $\pm$ 0.76
RelationNet [12]	Conv4	62.45 $\pm$ 0.98	76.11 $\pm$ 0.69
CloserLook [31]	Conv4	60.53 $\pm$ 0.83	79.34 $\pm$ 0.61
FEAT [32]	ResNet12	68.87 $\pm$ 0.22	82.90 $\pm$ 0.15
ProtoNet	ResNet12	81.02 $\pm$ 0.20	91.93 $\pm$ 0.11
CTX	ResNet12	80.39 $\pm$ 0.20	91.01 $\pm$ 0.11
DPGN	ResNet12	75.71 $\pm$ 0.47	91.48 $\pm$ 0.33
ST-GCLPN	ResNet12	81.54 $\pm$ 0.46	88.06 $\pm$ 0.39
ST-GCLPN	Swin transformer	85.78 $\pm$ 0.35	92.46 $\pm$ 0.45

4.3.2. 消融实验

针对 miniImageNet 数据集, 进行了两组消融实验以验证 ST-GCLPN 的有效性。第一组实验, 验证了双聚焦相似度计算(结合全局和局部信息)相较于传统相似度计算方法(如欧几里得距离和余弦相似度)的优势; 第二组实验, 验证了 GCN 是否能够有效提升节点特征、增强分类能力, 以及 LPA 是否能够充分利用全局信息传播, 提高分类效果。

第一组实验结果如表 4 所示。实验结果表明, 仅使用欧几里得距离或余弦相似度计算时, 模型的分类准确率明显下降。在 5-way 1-shot 任务中, 去除双聚焦相似度, 仅使用欧几里得距离进行相似度计算时, 分类准确率下降 4.54%(从 67.45%降至 62.91%)。进一步分析发现, 仅使用全局相似度或局部相似度计算也会导致分类性能下降。在 5-way 1-shot 任务中, 仅使用全局相似度计算的准确率下降 1.23%(从 67.45%降至 66.22%), 而仅使用局部相似度计算的准确率下降 4.41%(从 67.45%降至 63.04%)。在 5-way 5-shot 任务中, 这两种方法的准确率分别下降 2.08%和 4.29%。这表明, 全局特征能够提供整体分布信息, 但在细粒度任务中缺乏局部细节的支持; 而局部特征能够捕捉细节信息, 但缺少全局上下文, 影响泛化能力。

相比之下, 双聚焦相似度计算同时结合了全局和局部特征信息, 有效提升了分类性能。由于全局相似度关注整体特征分布, 而局部相似度刻画细粒度结构信息, 两者的融合能够构造更精准的相似度矩阵, 进一步优化特征关系建模。最终, 双聚焦相似度在 5-way 1-shot 和 5-way 5-shot 任务中的准确率分别达到 67.45%和 80.44%, 相比于其他方法均取得最优结果。

第二组实验结果如表 5 所示。在去除 GCN 的情况下, 5-way 1-shot 任务的分类准确率下降 5.11%, 5-way 5-shot 任务的分类准确率下降 4.51%, 说明 GCN 在优化节点特征方面起到了关键作用, 可以更好地聚合邻域信息, 提升节点特征的可分性。进一步地, 在去除 LPA 后, 5-way 1-shot 任务的分类准确率下降 2.35%, 5-way 5-shot 任务的分类准确率下降 2.89%, 说明 LPA 能够充分利用全局信息传播, 提高未标注样本的分类精度, 增强模型的泛化能力。这表明, GCN 和 LPA 在 ST-GCLPN 中相互协同, 有效提升了小样本分类任务的准确性。

综上所述, 消融实验结果验证了双聚焦相似度计算、GCN 以及 LPA 在 ST-GCLPN 框架中的有效性, 证明其在 miniImageNet 数据集上的分类性能得到了显著提升。

Table 4. Ablation experiment 1: classification accuracy on the miniImageNet dataset
表 4. 消融实验 1：在 miniImageNet 数据集上的分类准确率

Similarity	miniImageNet	
	5-way 1-shot	5-way 5-shot
欧几里得距离	62.91 ± 0.34	75.30 ± 0.25
全局相似度	66.22 ± 0.26	78.36 ± 0.32
局部相似度	63.04 ± 0.16	76.15 ± 0.19
双聚焦相似度	67.45 ± 0.36	80.44 ± 0.39

Table 5. Ablation experiment 2: classification accuracy on the miniImageNet dataset
表 5. 消融实验 2：在 miniImageNet 数据集上的分类准确率

Method		miniImageNet	
GCN	LPA	5-way 1-shot	5-way 5-shot
✓		65.10 ± 0.25	77.55 ± 0.29
	✓	62.34 ± 0.30	75.93 ± 0.27
✓	✓	67.45 ± 0.36	80.44 ± 0.39

4.3.3. 相似度融合权重参数和邻接图构建参数敏感性分析

为评估双聚焦相似度融合策略中超参数 α 和 β 对最终模型性能的影响，本文在 miniImageNet 数据集上进行了参数敏感性分析实验。为便于实验分析，设定 $\alpha + \beta = 1$ ，仅调整 α 的取值，从而隐式确定 $\beta = 1 - \alpha$ 。实验在 5-way 1-shot 和 5-way 5-shot 两种典型小样本分类任务下进行， α 的取值范围为{0.0, 0.2, 0.4, 0.6, 0.8, 1.0}，对应 β 的值分别为{1.0, 0.8, 0.6, 0.4, 0.2, 0.0}。

实验结果如表 6 所示，当 $\alpha = 0.6$ 、 $\beta = 0.4$ 时，模型在 1-shot 和 5-shot 上均达到最优性能，说明融合适量全局信息与局部结构能够更准确刻画样本间的关系，从而提升分类性能。相比之下，仅使用单一相似度($\alpha = 0.0$ 或 $\alpha = 1.0$)时表现相对较差，表明两者互补性较强。

Table 6. Hyperparameters α and β accuracy on the miniImageNet dataset
表 6. 超参数 α 和 β 在 miniImageNet 数据集上的准确度

α	β	5-way 1-shot	5-way 5-shot
0.0	1.0	64.28	78.90
0.2	0.8	65.73	79.46
0.4	0.6	66.89	80.01
0.6	0.4	67.45	80.44
0.8	0.2	66.23	79.72
1.0	0.0	64.97	78.88

在基于图神经网络的小样本学习方法中， k -最近邻作为构建图结构的关键参数，其取值直接决定了每个样本节点所连接的邻居数量，从而影响图的稀疏性与连通性，进而影响标签传播效果。因此，本文进一步分析不同 k 值对 ST-GCLPN 模型性能的影响，以评估模型对该超参数的敏感性。

在 miniImageNet 数据集上，我们固定其他设置不变，将 k 值设为{1, 3, 5, 7, 9}，在 5-way 1-shot 与 5-

way 5-shot 两种任务下测试模型表现。实验结果如表 7 所示, 由表中结果可以看出, 当 $k=5$ 时模型在两个任务上均达到最优性能。当 k 较小时(如 $k=1$), 图结构较为稀疏, 限制了信息传播范围; 而 k 过大时(如 $k=9$), 会引入更多无关或噪声邻居, 反而可能导致关系建模不准确, 影响最终分类精度。实验表明, $k=5$ 是在本任务中兼顾信息传播能力与图结构纯度的最优选择, 同时也验证了 ST-GCLPN 对 k 值的设置具备一定的稳定性和鲁棒性。

Table 7. Accuracy of different k values in the miniImageNet dataset
表 7. 不同 k 值在 miniImageNet 数据集的准确度

k	5-way 1-shot	5-way 5-shot
1	60.32	75.18
3	64.74	78.02
5	67.45	80.44
7	66.90	79.89
9	65.81	79.47

4.3.4. 可视化分析

为了进一步验证所提出模型在较小的特征空间中是否能够学习到具有区分性的特征分布, 本研究采用类激活可视化方法(Grad-CAM)进行分析。具体而言, 我们从 miniImageNet、tieredImageNet 和 CUB-200-2011 三个小样本图像分类数据集中随机选取 6 张图像进行实验, 并通过 Grad-CAM 以热力图的形式可视化模型的判别性区域。

如图 4 展示了 Swin Transformer 与 ResNet-12 在小样本图像分类任务上的可视化结果。其中, 第 1 行为原始图像, 第 2 行为 ResNet-12 生成的热力图, 第 3 行为 ST-GCLPN 生成的热力图。从中可观察到, 相比于 ResNet-12 提取的图像特征热力图存在判别区域较为分散的问题, ST-GCLPN 能够更加精准地关注关键区域。

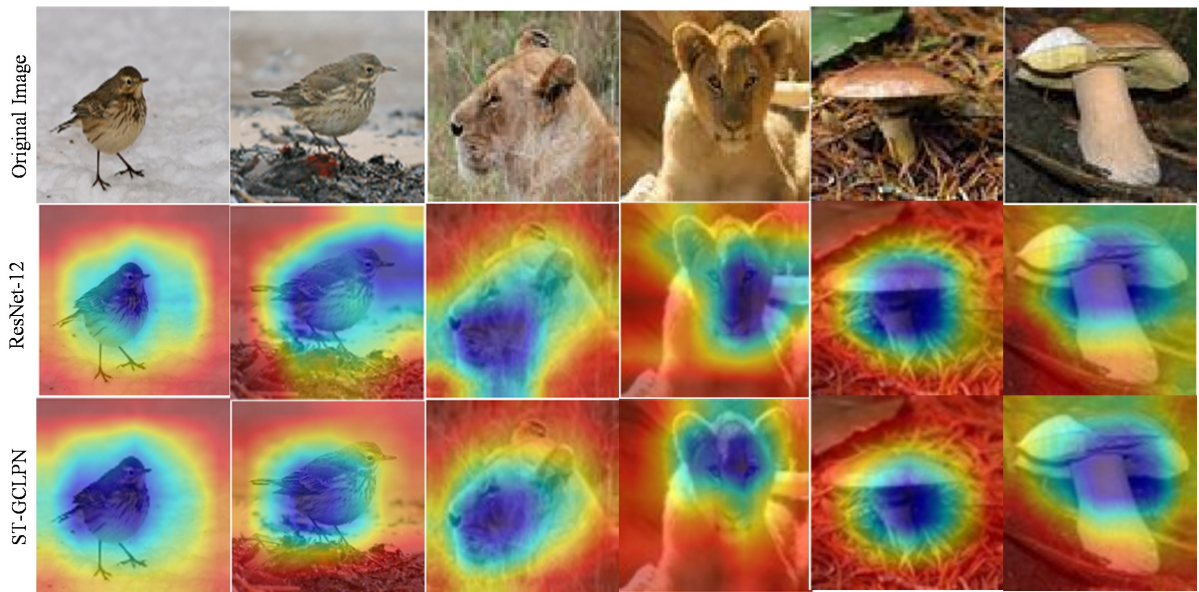


Figure 4. Visualize the results
图 4. 可视化结果

实验结果表明, 所提出的方法能够有效捕捉更全面、精细且富有区分性的特征信息, 增强模型的判别能力和特征表征能力。Grad-CAM 可视化的应用进一步证明了所提模型在特征学习和关键区域定位上的优势, 同时为其在小样本图像分类任务中的有效性和应用潜力提供了直观的支持。

5. 结论

所提出的方法将图神经网络(GCN)与 Swin Transformer 相结合, 应用于小样本图像分类任务。通过利用 Swin Transformer 强大的特征提取能力, 模型能够捕捉更丰富的全局特征和局部特征, 从而增强样本表征能力。为了进一步提升相似度计算的准确性, 我们设计了双聚焦相似度计算模块, 分别计算全局特征相似度和局部特征相似度, 并结合两者构建更精确的相似度矩阵, 以更准确地刻画样本间的关系。在相似度矩阵的基础上, 我们构建了图卷积网络(GCN), 以优化样本的特征表示, 并结合标签传播算法(LPA)进行全局信息传递, 充分利用已标注样本与未标注样本的关系, 提高分类决策的稳定性。所提出的方法在 miniImageNet、tieredImageNet 和 CUB-200-2011 三个小样本分类数据集上进行了实验验证, 实验结果表明, ST-GCLPN 显著优于其他方法, 在小样本学习任务中取得了更高的分类准确率, 进一步验证了本方法的有效性和泛化能力。

参考文献

- [1] 赵晓, 张懿丹, 章为川, 等. 基于通道先验感知的多尺度细化小样本细粒度图像分类[J]. 陕西科技大学学报, 2025: 1-11.
- [2] 张浩, 曹磊, 马利亚. 基于交叉协同注意力网络的小样本肠道息肉图像语义分割[J]. 中国数字医学, 2025, 20(1): 39-44.
- [3] 王新广, 李辉. 结合 Swin Transformer 与 MobileNetV3 的多源无人机影像目标检测方法[J]. 城市勘测, 2025(1): 27-32.
- [4] 周峻宇, 施水才, 王洪俊. 基于深度学习的图像字幕生成综述[J]. 软件导刊, 2025, 24(1): 211-220.
- [5] Chopra, S., Hadsell, R. and LeCun, Y. (2005) Learning a Similarity Metric Discriminatively, with Application to Face Verification. 2005 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, Vol. 1, 539-546. <https://doi.org/10.1109/cvpr.2005.202>
- [6] Vinyals, O., Blundell, C., Lillicrap, T., et al. (2016) Matching Networks for One Shot Learning. *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016*, Barcelona, 5-10 December 2016, 3637-3645.
- [7] Snell, J., Swersky, K. and Zemel, R. (2017) Prototypical Networks for Few-Shot Learning. *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017*, Long Beach, 4-9 December 2017, 4080-4090.
- [8] Luo, Y., Huang, Z., Zhang, Z., Wang, Z., Baktashmotlagh, M. and Yang, Y. (2020) Learning from the Past: Continual Meta-Learning with Bayesian Graph Neural Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34, 5021-5028. <https://doi.org/10.1609/aaai.v34i04.5942>
- [9] Liu, Y., Lee, J., Park, M., et al. (2018) Learning to Propagate Labels: Transductive Propagation Network for Few-Shot Learning.
- [10] Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P.H.S. and Hospedales, T.M. (2018) Learning to Compare: Relation Network for Few-Shot Learning. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-22 June 2018, 1199-1208. <https://doi.org/10.1109/cvpr.2018.00131>
- [11] Satorras, V.G. and Estrach, J.B. (2018) Few-Shot Learning with Graph Neural Networks. *International Conference on Learning Representations*, Vancouver, 30 April-3 May 2018.
- [12] Gidaris, S. and Komodakis, N. (2019) Generating Classification Weights with GNN Denoising Autoencoders for Few-Shot Learning. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 21-30. <https://doi.org/10.1109/cvpr.2019.00011>
- [13] Rodríguez, P., Laradji, I., Drouin, A. and Lacoste, A. (2020) Embedding Propagation: Smoother Manifold for Few-Shot Classification. *Computer Vision—ECCV 2020 16th European Conference*, Glasgow, 23-28 August 2020, 121-138. https://doi.org/10.1007/978-3-030-58574-7_8

- [14] Ye, H., Hu, H., Zhan, D. and Sha, F. (2020) Few-Shot Learning via Embedding Adaptation with Set-to-Set Functions. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 14-19 June 2020, 8808-8817. <https://doi.org/10.1109/cvpr42600.2020.00883>
- [15] Li, X., Sun, Q., Liu, Y., et al. (2019) Learning to Self-Train for Semi-Supervised Few-Shot Classification. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, Vancouver, 8-14 December 2019, 10276-10286.
- [16] Triantafillou, E., Larochelle, H., Snell, J., et al. (2018) Meta-Learning for Semi-Supervised Few-Shot Classification. *International Conference on Learning Representations*, Vancouver, 30 April-3 May 2018, 10276-10286.
- [17] Simon, C., Koniusz, P., Nock, R. and Harandi, M. (2020) Adaptive Subspaces for Few-Shot Learning. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 14-19 June 2020, 4136-4145. <https://doi.org/10.1109/cvpr42600.2020.00419>
- [18] Saito, K., Kim, D., Sclaroff, S., Darrell, T. and Saenko, K. (2019) Semi-Supervised Domain Adaptation via Minimax Entropy. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 8050-8058. <https://doi.org/10.1109/iccv.2019.00814>
- [19] Kim, J., Kim, T., Kim, S. and Yoo, C.D. (2019) Edge-Labeling Graph Neural Network for Few-Shot Learning. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 11-20. <https://doi.org/10.1109/cvpr.2019.00010>
- [20] Liu, J., Song, L. and Qin, Y. (2020) Prototype Rectification for Few-Shot Learning. *Computer Vision—ECCV 2020 16th European Conference*, Glasgow, 23-28 August 2020, 741-756. https://doi.org/10.1007/978-3-030-58452-8_43
- [21] Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021) Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows. 2021 *IEEE/CVF International Conference on Computer Vision (ICCV)*, 11-17 October 2021, 10012-10022. <https://doi.org/10.1109/iccv48922.2021.00986>
- [22] Yu, T., He, S., Song, Y. and Xiang, T. (2022) Hybrid Graph Neural Networks for Few-Shot Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, **36**, 3179-3187. <https://doi.org/10.1609/aaai.v36i3.20226>
- [23] Finn, C., Abbeel, P. and Levine, S. (2017) Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. *International Conference on Machine Learning*, Sydney, 6-11 August 2017, 1126-1135.
- [24] Zhang, R., Che, T., Ghahramani, Z., et al. (2018) Metagan: An Adversarial Approach to Few-Shot Learning. *Advances in Neural Information Processing Systems 31: Annual Conference on Neural Information Processing Systems 2018*, Montreal, 3-8 December 2018, 2365-2374.
- [25] Mishra, N., Rohaninejad, M., Chen, X., et al. (2017) A Simple Neural Attentive Meta-Learner.
- [26] Sun, Q., Liu, Y., Chua, T. and Schiele, B. (2019) Meta-Transfer Learning for Few-Shot Learning. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 403-412. <https://doi.org/10.1109/cvpr.2019.00049>
- [27] Yoon, S.W., Seo, J. and Moon, J. (2019) Tapnet: Neural Network Augmented with Task-Adaptive Projection for Few-Shot Learning. *International Conference on Machine Learning*, Long Beach, 9-15 June 2019, 7115-7123.
- [28] Chen, W.Y., Liu, Y.C., Kira, Z., et al. (2019) A Closer Look at Few-Shot Classification.
- [29] Ye, H.J., Hu, H., Zhan, D.C., et al. (2018) Learning Embedding Adaptation for Few-Shot Learning.
- [30] Liu, Y., Schiele, B. and Sun, Q. (2020) An Ensemble of Epoch-Wise Empirical Bayes for Few-Shot Learning. *Computer Vision—ECCV 2020 16th European Conference*, Glasgow, 23-28 August 2020, 404-421. https://doi.org/10.1007/978-3-030-58517-4_24
- [31] Lee, K., Maji, S., Ravichandran, A. and Soatto, S. (2019) Meta-Learning with Differentiable Convex Optimization. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Long Beach, 15-20 June 2019, 10657-10665.
- [32] Yang, L., Li, L., Zhang, Z., Zhou, X., Zhou, E. and Liu, Y. (2020) DPGN: Distribution Propagation Graph Network for Few-Shot Learning. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, 14-19 June 2020, 13390-13399. <https://doi.org/10.1109/cvpr42600.2020.01340>