元学习研究综述

李凡长 刘 洋 吴鹏翔 董 方 蔡 奇 王 哲

(苏州大学计算机科学与技术学院 江苏 苏州 215006)

摘 要 深度学习在大量领域取得优异成果,但仍然存在着鲁棒性和泛化性较差、难以学习和适应未观测任务、极其依赖大规模数据等问题.近两年元学习在深度学习上的发展,为解决上述问题提供了新的视野.元学习是一种模仿生物利用先前已有的知识,从而快速学习新的未见事物能力的一种学习定式.元学习的目标是利用已学习的信息,快速适应未学习的新任务.这与实现通用人工智能的目标相契合,对元学习问题的研究也是提高模型的鲁棒性和泛化性的关键.近年来随着深度学习的发展,元学习再度成为热点,目前元学习的研究百家争鸣、百花齐放.本文从元学习的起源出发,系统地介绍元学习的发展历史,包括元学习的由来和原始定义,然后给出当前元学习的通用定义,同时总结当前元学习一些不同方向的研究成果,包括基于度量的元学习方法、基于强泛化新的初始化参数的元学习方法、基于梯度优化器的元学习方法、基于外部记忆单元的元学方法、基于数据增强的元学方法等.总结其共有的思想和存在的问题,对元学习的研究思想进行分类,并叙述不同方法和其相应的算法.最后论述了元学习研究中常用数据集和评判标准,并从元学习的自适应性、进化性、可解释性、连续性、可扩展性展望其未来发展趋势.

关键词 元学习;深度学习;深度神经网络;泛化能力;自适应能力;扩展能力

中图法分类号 TP18 **DOI 号** 10.11897/SP.J.1016.2021.00422

A Survey on Recent Advances in Meta-Learning

LI Fan-Zhang LIU Yang WU Peng-Xiang DONG Fang CAI Qi WANG Zhe

(*School of Computer Science and Technology, Soochow University, Suzhou, Jiangsu* 215006)

Abstract With the substantial progress in parallel computing performance of computing devices and the continuous great breakthroughs of deep neural networks in various fields in recent years, several new fields of machine learning derived from deep neural network models have gradually matured. In fact, in the process of combining machine learning with other fields, not all fields have enough massive data for deep neural networks to fine-tune parameters to convergence, and quite a number of traditional deep learning models have the disadvantages of poor robustness and poor generalization performance, which make it difficult to learn and adapt to unobserved tasks quickly. The development of meta-learning on deep learning in the past two years has provided new ideas to solve the above problems. Meta-learning is a learning stereotype that mimics the ability of organisms to use prior knowledge and thus learn new unobserved things quickly. It is commonly believed that a meta-learning model must contain a learning subsystem and use previously learned meta-knowledge to extract the focus of a new task and dynamically select learning biases. The goal of meta-learning is to maximize the use of learned information to quickly adapt to new unlearned tasks, emphasizing the shortest possible adaptation time while pursuing the use of minimal new training data to achieve rapid learning on multiple fronts. This fits with

收稿日期:2019-09-12;在线发布日期:2020-12-04.本课题得到国家重点研发计划"变革性技术关键科学问题"重点专项(2018YFA0701700,2018YFA0701701)资助.**李凡长**,博士,教授,博士生导师,中国计算机学会(CCF)高级会员,主要研究领域为元学习、李群机器学习、动态模糊逻辑.E-mail:lfzh@suda.edu.cn.**刘 洋**,硕士研究生,主要研究方向为元学习.**吴鹏翔**,硕士研究生,主要研究方向为元学习.**董 方**,硕士研究生,主要研究方向为元学习.**蔡 奇**,硕士研究生,主要研究方向为元学习.**王 哲**,硕士研究生,主要研究方向为元学习.

the goal of achieving general artificial intelligence, and the study of meta-learning problems is also key to improving the robustness and generalization of models. The general meta-learning process is as follows: initial network parameters with strong generalization performance are obtained on training and validation datasets, and in the testing phase, the network model is subjected to a small number of learning processes on the test data as the learning process of that model for a task, followed by testing the model's effectiveness after learning. With the development of deep learning in recent years, meta-learning has once again become a new hot topic. At present, meta-learning research has been introduced into the fields of data prediction, image classification, reinforcement learning and autonomous robot control with a hundred schools of thought and flowers. In this paper, we start from the origin of meta-learning, introduce the development history of meta-learning systematically, including the origin and original definition of meta-learning, mainly including the initial definition of meta-learning in the field of education science and the later influence in physics education, and finally introduce the process of meta-learning into the computer field, then give the current general definition of meta-learning and training test paradigm, and also summarize some different directions of current meta-learning research results, including metric-based meta-learning methods, meta-learning methods based on strong generalization of new initialization parameters, gradient optimizer-based meta-learning methods, external memory unit-based meta-learning methods, data augmentation-based meta-learning methods, etc., and statistics on the effects and ideas of some representative meta-learning methods in the above directions, and a classification summary of the methods to better compare the different methods', in order to better compare the strengths and weaknesses of different methods and prepare the foundation for further selection of research directions. On the basis of summarizing their common ideas and problems, the research ideas of meta-learning are classified, and the different methods and their corresponding algorithmic contents are described. Finally, the common datasets and judging criteria in meta-learning research are discussed, and the future development trend of meta-learning is prospected in terms of its adaptiveness, evolution, interpretability, continuity, and scalability.

Keywords　meta-learning; deep learning; deep neural network; generalization ability; adaptive ability; expansion ability

1　引　言

随着计算设备并行计算性能的大幅度进步，以及近些年深度神经网络在各个领域不断取得重大突破，由深度神经网络模型衍生而来的多个机器学习新领域也逐渐成型，如强化学习、深度强化学习[1-2]、深度监督学习等.在大量训练数据的加持下，深度神经网络技术已经在机器翻译、机器人控制、大数据分析、智能推送、模式识别等方面取得巨大成果[3-5].

实际上在机器学习与其它行业结合的过程中，并不是所有领域都拥有足够可以让深度神经网络微调参数至收敛的海量数据，相当多领域要求快速反应、快速学习，如新兴领域之一的仿人机器人领域，其面临的现实环境往往极为复杂且难以预测，若按照传统机器学习方法进行训练则需要模拟所有可能遇到的环境，工作量极大同时训练成本极高，严重制约了机器学习在其领域的扩展，因此在深度学习取得大量成果后，具有自我学习能力与强泛化性能的元学习便成为通用人工智能的关键.

元学习(Meta-learning)的提出是针对传统神经网络模型泛化性能不足、对新种类任务适应性较差的特点.在元学习介绍中往往将元学习的训练和测试过程类比为人类在掌握一些基础技能后可以快速学习并适应新任务，如儿童阶段的人类也可以快速通过一张某动物照片学会认出该动物，即机器学习中的小样本学习(Few-shot Learning)[6-7]，甚至不需要图像，仅凭描述就可学会认识新种类，对应机器

学习领域中的(Zero-shot Learning)[8],而不需要大量该动物的不同照片.人类在幼儿阶段掌握的对世界的大量基础知识和对行为模式的认知基础便对应元学习中的"元"概念,即一个泛化性能强的初始网络加上对新任务的快速适应学习能力,元学习的远期目标为通过类似人类的学习能力实现强人工智能,当前阶段体现在对新数据集的快速适应带来较好的准确度,因此目前元学习主要表现为提高泛化性能、获取好的初始参数、通过少量计算和新训练数据即可在模型上实现和海量训练数据一样的识别准确度,近些年基于元学习,在小样本学习领域做出了大量研究[9-17],同时为模拟人类认知,在 Zero-shot Learning 方向也进行了大量探索[18-22].

在机器学习盛行之前,就已产生了元学习的相关概念.当时的元学习还停留在认知教育科学相关领域,用于探讨更加合理的教学方法.Glass 在 1976 年首次提出了"元分析"这一概念[23],对大量的分析结果进行统计分析,这是一种二次分析办法.Powell 使用"元分析"的方法对词汇记忆进行了研究[24],指出"强制"和"诱导"意象有助于词汇记忆.Maudsley 在 1979 年首次提出了"元学习"这一概念,将其描述为"学习者意识到并越来越多地控制他们已经内化的感知、探究、学习和成长习惯的过程",Maudsley 将元学习作为在假设、结构、变化、过程和发展这 5 个方面下的综合,并阐述了相关基本原则[25].Biggs 将元学习描述为"意识到并控制自己的学习的状态"[26],即学习者对学习环境的感知.Adey 等人将元学习的策略用在物理教学上[27],VanLehn 探讨了辅导教学中的元学习方法[28].从元分析到元学习,研究人员主要关注的是如何意识和控制自己学习的.一个具有高度元学习观念的学生,能够从自己采用的学习方法所产生的结果中获得反馈信息,进一步评价自己的学习方法,更好地达到学习目标[29].随后元学习这一概念慢慢渗透到机器学习领域.Chan 等人提出的元学习是一种整合多种学习过程的技术,利用元学习的策略组合多个不同算法设计的分类器,其整体的准确度优于任何个别的学习算法[30-32].Bensusan 等人提出了基于元学习的决策树框架[33].Vilalta 等人则认为元学习是通过积累元知识动态地通过经验来改善偏倚的一种学习算法[34].

Meta-Learning 目前还没有确切的定义,一般认为一个元学习系统需结合三个要求:系统必须包含一个学习子系统;利用以前学习中提取的元知识来获得经验,这些元知识来自单个数据集或不同领域;动态选择学习偏差.

元学习的目的就是设计一种机器学习模型,这种模型有类似上面提到的人的学习特性,即使用少量样本数据,也能快速学习新的概念或技能.经过不同任务的训练后,元学习模型能很好地适应和泛化到一个新任务,也就学会了"Learning to learn".

2 元学习的定义

2.1 元学习的概念

一般来说,元学习就是学会学习的学习.受当前计算资源与算法能力限制,元学习往往以小样本学习以及对新任务的快速适应作为切入点,因此当前研究也多以在小样本数据上识别准确率作为实验衡量标准.

元学习研究目前大致分为五个较为独立的方向:基于度量学习的方法、基于泛化性较强的初始化的方法、基于优化器,基于额外外部存储以及基于数据增强的方法,其中基于强泛化性的初始化方向上进展较多,目前最出色的实验准确率结果来自该方向上模型无关元学习算法(MAML)[14]的后续改进型[35],此方向也自然成为元学习领域的中坚力量,其中基于 MAML 发展出众多新颖的元学习模型[36-38].基于度量的方向倾向于最大程度上抽取任务样本内含的特征,使用特征比对的方式判定样本的种类归属,因此如何提取最能代表样本特点的特征便成为了该方向研究重点.

基于优化器的元学习模型则针对传统神经网络模型迭代速度慢、易过拟合的特性,提出将梯度下降过程替代为单独的神经网络模型,使用单独神经网络对梯度进行更准确、迅速的更新,以达到快速适应的目的.

基于数据增强的元学习模型通过为小样本任务增加额外的样本来解决元学习中数据缺乏的问题.该类模型具有较好的通用性,可以与其它元学习方法进行结合,提高这些方法的性能.下面简单介绍元学习的通用训练、测试过程.

元学习目前主要针对小样本学习问题,元学习训练和测试都以少样本任务为基本单元,每个任务拥有各自的训练数据集和测试数据集,也称为支持集和查询集.元学习在训练阶段使用大量少样本任

务进行训练，且在测试阶段可产生较好的效果. 为实现这一点，在元学习模型训练阶段和测试阶段所使用的数据均为只含有小样本数据的任务. 因为元学习的目标是通过在训练数据上的学习掌握快速学习新任务能力，因此元学习在训练时将整个任务集看作训练样例.

以元学习中的监督学习为例，考虑一个任务分布，目的是使这个模型可以适应这个任务分布 $\mathcal{P}(\mathcal{T})$. 在 k-shot 元学习问题（在训练和测试阶段，每一类样本仅使用 k 个样本参与训练和测试）中，经过训练后的模型在新任务上 $\mathcal{T}_i^{\text{test}}$ 上经过少量几次“学习”（通常为对模型参数的梯度下降）后验证学习后的模型在此测试数据上的效果，其中测试阶段使用的新任务 $\mathcal{T}_i^{\text{test}}$ 和训练阶段使用的任务数据 $\mathcal{T}_i^{\text{train}}$ 一样取自 $\mathcal{P}(\mathcal{T})$. $\mathcal{L}_{\mathcal{T}_i}$ 表示模型在任务 $\mathcal{T}_i$ 上的损失函数.

在训练阶段，从 $\mathcal{P}(\mathcal{T})$ 的训练数据集上采样一个样本 $\mathcal{T}_i^{\text{train}}$，模型在此训练任务上对网络参数进行少量几次优化，接着将在训练任务上优化后的模型结合验证数据集中的任务 $\mathcal{T}_i^{\text{val}}$ 得到损失函数，并通过例如 Adam、SGD、SVGD 等优化器对该损失函数进行最小化. 当训练结束后，从未参与训练的测试数据集 $\mathcal{D}_i^{\text{test}}$ 中生成一些测试任务 $\mathcal{T}_i^{\text{test}}$，使用这些测试任务对已经训练好的网络再进行少量几次训练，接着使用在测试任务上进行优化后的模型检验测试数据上的效果，如识别准确率、回归误差等.

模型结构上，与其它深度学习模型类似，通常从逻辑上分为特征提取部分（Feature extractor）以及分类部分（Classifier layers）. 特征提取部分使用主流深度网络框架，由卷积层组成，用于从数据中进行特征提取，将信息进行高度抽象，部分模型将已经学习的先验知识也作为特征的一部分进入模型参与训练. 分类部分则通常由带有非线性激活函数的全连接层组成，个别元学习模型的特征提取过程还包含从训练任务中随机采样，但总体上训练特征提取与主流深度学习相同，均是通过深层网络对输入信息进行降维，提取出更高层特征信息.

由此，一般元学习流程是：在训练数据集和验证数据集上得到泛化性较强的初始化网络参数，在测试时，将网络模型在测试数据上进行少量几次梯度下降操作，以达到“学习新任务”的目的，接着检验学习后的效果. 训练和测试一般流程如图 1 和图 2.

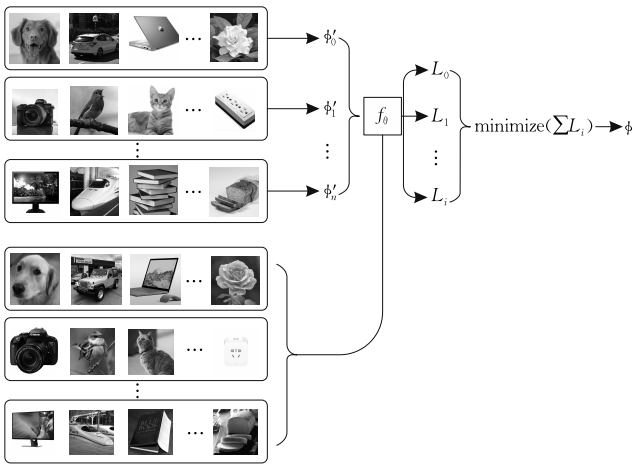

图 1　训练阶段

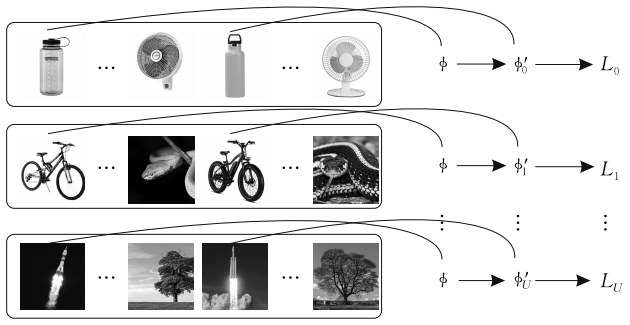

图 2　测试阶段

2.2　元学习的问题场景设置

Zero-Shot Learning（ZSL）、Few-Shot Learning（FSL）、One-Shot Learning（OSL）是目前元学习主要解决的三个问题，此小节我们给出 ZSL、FSL、OSL 问题的相关定义和说明，以便后面分析各模型适用的场景.

元学习以 task 为基本单元，一个 task 的数据集分为支持集和查询集，支持集为带标签的训练样例，将这些训练样例覆盖的类表示为 seen classes，未覆盖的类表示 unseen classes，标记 $\mathcal{S}=\{c_i^s \mid i=1,\cdots,N_s\}$ 为已见类集合，$\mathcal{U}=\{c_i^u \mid c=1,\cdots,N_u\}$ 为未见类集合，其中 $\mathcal{S}\cap\mathcal{U}=\varnothing$. 标记 $\mathcal{X}$ 为输入样例的特征空间，$D^{tr}=\{(x_i^{tr},y_i^{tr})\in\mathcal{X}\times\mathcal{S}\}_{i=1}^{N_{tr}}$ 为支持集，$X^{te}=\{x_i^{te}\in\mathcal{X}\}_{i=1}^{N_{te}}$ 为查询集样例.

在 ZSL 问题中，查询集样例 X^{te} 对应的真实标签 $Y^{te}=\{y_i^{te}\in\mathcal{U}\}_{i=1}^{N_{te}}$，属于未见类，零样本学习旨在学习分类器 $f(\cdot):\mathcal{X}\to\mathcal{U}$，即将属于未见类的查询样例进行类型预测. ZSL 识别依赖与每个未曾学习过的类（unseen classes）在语义上与所学习到的类（seen classes）存在的相关性知识，即 unseen classes

和 seen classes 都与一个高维语义空间相关,其中来自所学习到的类的知识可以转移到未学习到的类.考虑一个例子,如果一个孩子已经知道马长什么样子,并且知道斑马长得看起来像马,但有黑白条纹,尽管没有真正见过斑马,他也可以去判别一个斑马.

在 FSL 与 OSL 问题中,查询集样例 X^{te} 对应的真实标签 $Y^{te}=\{y_i^{te}\in\mathcal{S}\}_{i=1}^{N_{te}}$,属于已见类,少样本和单样本学习旨在学习分类器 $f(\cdot):\mathcal{X}\rightarrow\mathcal{S}$,即将属于已见类的查询样例进行类型预测.在 FSL 中,每一类的支持集样例 $D_{c_i}^{tr}$ 的数量通常设置为 10-shot、5-shot 或更少;而在 OSL 中,设置为 1-shot.

由于训练深层的神经网络通常需要大量的训练样例,基于零样本的 ZSL 和少样本的 FSL 和 OSL 很容易出现过拟合等问题,这也是元学习要解决的主要问题之一:在少样本情况下如何提高泛化能力.

3 元学习的研究方法

3.1 基于度量的元学习方法

3.1.1 孪生网络

Koch 等人提出了一种用于解决单样本学习图像分类问题的方法[13].该方法学习一个卷积孪生网络,通过该网络计算查询图像与所有单标注样本的相似度,其中相似度最高的支持样本所对应的类别即为预测类.孪生网络的输入是同一种类下的一对图片,通过相同的神经网络结构分别对两张图片提取特征,使得同一类别的图片在被映射到特征空间之后的差异不会太大.

模型首先用 L 层全连接层提取图片 x_1 和 x_2 的特征,之后通过一个全连接层和 sigmoid 激活函数,输出结果在[0,1]之间的相似度 $p(x_1,x_2)$.模型采用了交叉熵损失函数,并加入 $L2$ 正则化项减少过拟合程度,$y(x_1,x_2)$ 表示 x_1 和 x_2 是否为同类的标签值:

$$\begin{aligned}L(x_1,x_2)=&y(x_1,x_2)\log p(x_1,x_2)+\\&(1-y(x_1,x_2))\log(1-p(x_1,x_2))+\\&\lambda^{\mathrm{T}}|w|^2\end{aligned}\tag{1}$$

作者与现有的 Omniglot[39] 数据集上效果最佳的分类器进行比较,实验表明作为较早的元学习模型,孪生网络的效果远远优于一般的分类器.部分研究人员也基于此网络模型进行了深入研究.Hoffer 等人提出了类似的 Triplet Network,其网络输入是三个,一个正例和两个负例,或者一个负例和两个正例,效果略优于两个输入的孪生网络[40].Melekhov 等人将孪生网络用于计算机视觉,有效地提高了图像匹配的性能[41].Bertinetto 等人在目标跟踪领域,提出了一种基于孪生网络的端到端基本跟踪算法[42].

3.1.2 注意力机制匹配网络

Vinyals 等人在 2016 年提出了匹配网络模型[10],其主要创新体现在建模过程和训练过程.对于建模过程,文章提出了基于记忆和注意力机制的匹配网络,可以对参与训练的样本进行快速学习.对于训练过程的创新,论文基于传统机器学习的一个重要原则,即训练和测试应在同样条件下进行,作者提出在训练时每次仅使用每一类任务的少量样本参与对网络的训练,与测试的过程一致.

具体地,网络先通过若干层 CNN 提取支持集和查询集图像特征,接着利用两个 LSTM 网络结构(以下记为 $g(x)$ 和 $f(x)$)分别提取两集合图像的嵌入向量.计算测试样本嵌入向量和支持集中所有样本嵌入向量的余弦距离并使用 softmax 函数进行归一化.通过注意力机制,可以综合考虑到待测试样本和支持集中所有样本的相似性.表示为如下公式,其中 a 是注意力机制函数,$x\hat{}$ 是待测试样本,x_i 是支持集中的样本.$y\hat{}$ 是分类结果,是注意力值的线性组合:

$$y\hat{}=\sum_{i=1}^{k}a(x\hat{},x_i)y_i\tag{2}$$

$$a(x\hat{},x_i)=e^{\cos(f(x\hat{}),g(x_i))}\Big/\sum_{j=1}^{|S|}e^{\cos(f(x\hat{}),g(x_j))}\tag{3}$$

其中 $g(x)$ 是一个双向的 LSTM 网络,用于进一步提取支持集的嵌入向量.考虑到样本输入的先后顺序对网络输出结果应该是没有影响的,采用双向的 LSTM 网络结构提取每个样本的嵌入向量.h_i,c_i 表示第 i 个样本的输出值和状态,$g'(x_i)$ 是通过 CNN 层提取到的特征.

$$\vec{h}_i,\vec{c}_i=\mathrm{LSTM}(g'(x_i),h_{i-1},c_{i-1})\tag{4}$$

$$\overleftarrow{h}_i,\overleftarrow{c}_i=\mathrm{LSTM}(g'(x_i),h_{i-1},c_{i-1})\tag{5}$$

$$g(x_i)=\vec{h}_i+\overleftarrow{h}_i+g'(x_i)\tag{6}$$

$f(x)$ 是一个带有注意力机制的单向 LSTM 网络.$f'(x\hat{})$ 是通过 CNN 层提取到的特征,$\boldsymbol{g}(S)$ 是支持集的嵌入向量.$h_k\hat{}$,c_k 为第 k 步时的输出值和状态.

$$h_k\hat{},c_k=\mathrm{LSTM}(f'(x\hat{}),[h_{k-1},r_{k-1}],c_{k-1})\tag{7}$$

$$h_{k'}\hat{}=h_k+f'(x\hat{})\tag{8}$$

$$r_{k-1}=\sum_{i=1}^{|S|}a(h_{k-1},g(x_i))g(x_i)\tag{9}$$

$$a(h_{k-1},g(x_i))=\mathrm{softmax}(h_{k-1}^{\mathrm{T}}g(x_i))\tag{10}$$

嵌入函数 f 和 g 是对特征空间进行了优化，从而提升精度. 在计算机的其它领域，匹配网络也有着重要的作用. Wu 等人提出了一个顺序匹配网络来解决聊天机器人多轮对话中的上下文匹配问题[43]. Si 等人提出了一种端到端的双注意匹配网络用于典型人员再识别，相比于其它基于单个特征向量的方法，一定程度上解决了在实际场景中视觉模糊的问题[44].

3.1.3　原型网络

Snell 等人于 2017 年提出了原型网络[16]，该网络模型基于一个基本假设，即在数据集里，对于每种不同的类型都存在一个原型点. 数据集中距离该原型点越近的样本，其标签与该原型点对应的标签相同的概率就越大. 图 3 展示了共 3 类数据的对应原型点.

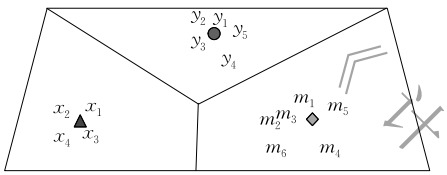

图 3　原型点示意图

原型网络通过一个深层神经网络（$f_{\varnothing}:R^D\to R^M$）将 D 维的样本数据映射到 M 维的嵌入空间，然后在该空间内计算每类样例的均值（原型点）C_k：

$$C_k=\frac{1}{|S_k|}\sum_{(x_i,y_i)\in S_k} f_{\phi}(x_i) \tag{11}$$

在测试时，从查询集中选取待分类向量 $\boldsymbol{x}$. 在嵌入空间中计算测试样例 x 和所有原型点 C_k 之间的距离，并用 softmax 函数对这些距离进行归一化：

$$P_{\phi}(y=k|x)=\frac{\exp(-\mathrm{d}(f_{\phi}(x),c_k))}{\sum_{k'}\exp(-\mathrm{d}(f_{\phi}(x),c_{k'}))} \tag{12}$$

训练过程中通过随机梯度下降算法优化目标损失函数 $J(\phi)=-\log(p_{\Phi}(y=k|x))$，其中 k 为样例 x 的真实标签. 文中还特别提出两个与匹配网络的不同之处：

（1）使用欧式距离计算样例与原型点之间的距离. 欧式距离是 Bregman 散度中的一种，而余弦距离不是. Bregman 有一个很好的性质，即给定集合 S，总体的 Bregman information（所有 x_i 到 $E(x)$ 的 Bregman 散度的均值）不变. 因此在聚类时只要使每一个簇（每一种类型的样本集合）的 Bregman information 值最小化，即可使得簇与簇之间的 Bregman information 值最大化. 而原型网络在判断未知样例的类型时，所采用的方式是 KNN（K-Nearest Neighbor），即选取最近原型点对应的标签作为预测结果. 若在计算距离时利用 Bregman 散度的性质，可以使得每个类型簇之间的 Bregman information 值，也就是每个类型之间的差异性最大，这样模型可以具有更好的分类效果.

（2）在以往的实验中发现，训练和测试时保持相同的 episode 设置往往会取得更好的效果. 例如在测试中采用 5-way、1-shot 训练方式，即 5 分类，每次使用一个样本进行训练的测试，那么训练的时候将 episode 设置为 $N_c=5$、$N_s=1$. 其中 N_c 为类别个数，N_s 为每个类中支持样例的个数. 但是在原型网络，使用比测试时更多的类别（$N_c>$way）对模型是有益的.

之后，Fort 提出了高斯原型网络，模型中编码器输出的一部分被解释为关于嵌入点的置信区域估计，并表示为高斯协方差矩阵，然后利用单个数据点的不确定性作为权重，在嵌入空间上构造一个和类相关距离度量[45].

3.1.4　关系网络用于小样本比较学习

Sung 等人[46]于 2018 年提出关系网络，关系网络由两个模块组成，嵌入模块 f_{φ} 和关系模块 $g_{\varnothing}$，其中 f_{φ} 提取图像特征，$g_{\varnothing}$ 为卷积网络，完成图像特征之间的相似性计算从而实现分类.

具体地，从查询集中取样 x_j，从支持集中取样 x_i，各自输入到 f_{φ} 得到对应的特征图 $f_{\varphi}(x_j)$ 和 $f_{\varphi}(x_i)$，使用拼接函数 C 拼接特征图 $f_{\varphi}(x_j)$ 和 $f_{\varphi}(x_i)$ 得到 $C(f_{\varphi}(x_j),f_{\varphi}(x_i))$. 拼接后的特征图输入关系网络 $g_{\varnothing}$，可得到 $f_{\varphi}(x_j)$ 和 $f_{\varphi}(x_i)$ 的相似度. 在 C-way 1-shot 任务中，得到待分类样本 j 关于其它种类样本 i 的 C 个相似度值 $r_{i,j}$.

$$r_{i,j}=g_{\varnothing}(C(f_{\varphi}(x_j),f_{\varphi}(x_i))),\ i=1,2,\cdots,C \tag{13}$$

对于 k-shot 任务，当 $k>1$ 时，得到支持集中每类的 k 个特征图，将 k 个特征图逐元素相加得到该类的特征图，其它过程与 1-shot 任务一致. 损失函数使用均方误差（MSE），算法最小化目标函数为

$$\varphi,\phi\leftarrow\arg\min_{\varphi,\phi}\sum_{i=1}^{m}\sum_{j=1}^{n}(r_{i,j}-1(y_i==y_j))^2 \tag{14}$$

本方法的有效原因一定程度上是其它基于度量的少样本学习模型使用固定的预先指定的距离度量方式，如欧式距离、余弦距离等，而本模型通过关系网络学习深度嵌入函数及深度非线性距离度量函数，以实现对样本间距离的更准确的表达，两者相互调整以便使关系网络在该少样本学习中表现更好.

3.1.5 任务依赖的自适应度量提高少样本分类

Oreshkin 等人[47]提出 Metric scaling 来提升少样本分类算法的性能，该方法学习一个距离缩放系数 α，让输出的度量在合适的范围内：

$$p_{\phi,\alpha}(y=k|x)=\mathrm{softmax}(-\alpha \mathrm{d}(z,c_k)) \quad (15)$$

通过度量缩放的方式，模型对不同的度量方法都可以有好的表现. 作者分析了交叉熵损失函数 $J_k(\varnothing,\alpha)$ 的梯度，进一步了解 α 的原理. 对于较小的 α，式(16)第一项最小化查询样例与对应类别原型点的距离，第二项最大化样例与非所属于类别的原型点之间的距离；对于较大的 α，式(17)第一项与式(16)相同，第二项最大化样例与最近的那个错误分类原型点的距离. 因此，α 的两种不同机制要么使样本分布的重叠最小化，要么按样本校正聚类分配. 对于给定的数据集、指标和任务组合，模型存在最优的比例参数 α：

$$\lim_{\alpha\to 0}\frac{1}{\alpha}\frac{\partial}{\partial_\phi}J_k(\phi,\alpha)=\sum_{x_i\in Q_k}\left[\begin{array}{c}\dfrac{K-1}{K}\dfrac{\partial}{\partial_\phi}\mathrm{d}(f_\phi(x_i),c_k)\\ -\dfrac{1}{K}\displaystyle\sum_{j\neq k}\dfrac{\partial}{\partial_\phi}\mathrm{d}(f_\phi(x_i),c_j)\end{array}\right] \quad (16)$$

$$\lim_{\alpha\to\infty}\frac{1}{\alpha}\frac{\partial}{\partial_\phi}J_k(\phi,\alpha)=\sum_{x_i\in Q_k}\left[\begin{array}{c}\dfrac{\partial}{\partial_\phi}\mathrm{d}(f_\phi(x_i),c_k)\\ -\dfrac{1}{K}\displaystyle\sum_{j\neq k}\dfrac{\partial}{\partial_\phi}\mathrm{d}(f_\phi(x_i),c_{j_i^*})\end{array}\right];$$

$$\text{where } j_i^*=\arg\min_j \mathrm{d}(f_\phi(x_i),c_j) \quad (17)$$

作者同时提出了一种任务依赖的度量空间学习方法，可以根据不同的任务进行特征提取. 具体的，文中定义了动态特征提取器 $f(x,\Gamma)$，其中 Γ 为任务表示获得的参数集合. 具体实现时 Γ 通过 FiLM 层生成的通道缩放和偏移系数影响基础特征提取器. 文中使用类原型的均值作为任务表示 Γ，从而根据每个任务的支持集 S 优化 $f(x,\Gamma)$ 的性能. 另外，该论文还提出了辅助任务协同训练，使得具有任务依赖性的特征提取更容易训练，并且具有更好的泛化能力.

实验部分，作者证明了原型网络中使用欧几里德距离带来的提升是由于尺度效应的假设. 通过度量缩放，余弦相似度与欧拉距离在少样本学习上的差距缩小了 14%. 另外，在没有辅助协同训练的情况下，使用动态特征提取器没有任何改善. 协同训练使初始收敛更容易，通过强制特征提取器在两个解耦的任务上表现良好，从而在少样本学习任务上提供正则化.

3.2 基于强泛化性的初始化参数元学习方法

3.2.1 模型无关层循环元学习框架算法

出自论文“Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks”[14]，主要思想为将对网络参数进行梯度下降优化过程和优化损失函数过程中所使用的任务数据进行分离，从而获得较好的泛化性能.

MAML 模型通常由两层循环构成，即内循环与外循环，其中每次外循环对应机器学习训练阶段的一轮训练过程. 模型所需数据分为两部分，D_{train} 与 D_{val}，将来自训练数据集与验证数据集的对应同一种类数据组成输入. 模型内置的网络生成副本并使用 D_{train} 进行多次梯度下降，在此过程中每一 task 产生对应的经过学习的网络参数副本主体 θ_i'，再将得到的网络副本用于与数据所属种类相同的 D_{val} 数据并输出预测结果，根据与给定真实标签产生对应损失函数 $\mathcal{L}_{T_i\sim p(T)}(f_{\theta_i'})$，进而可由优化器进行优化. 优化器作用在多个 task 产生的综合损失函数上，进而改变模型的真正网络参数数据，以完成一次训练过程. 训练过程可抽象为图 4.

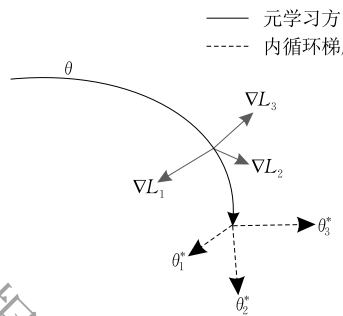

图 4 MAML 训练过程示意图

在测试阶段，测试数据来自与训练数据同属同一分布的数据集，测试操作类似训练阶段内循环过程，即使用对应位置类别相同的极少量数据对网络进行训练，再将未经训练且对应种类相同的数据通过该模型的网络得到预测或回归结果，与真实结果对比从而得到模型对新数据的适应能力.

MAML 是当前元学习的不同探索方向中比较有潜力的分支，且已有众多后续基于 MAML 的改进，如结合隐空间[35]、Relation Network[46]、Residual Network[4]、贝叶斯先验[48]、优化梯度下降过程[49]等方法，将模型在 mini-ImageNet 数据集上单样本学习识别准确率从 48.7% 提升至 62%. 同时由于 MAML 的开放性与灵活性，也可用于一系列基于梯度下降训练的模型，包括分类、回归、强化学习[37]等等，因此叫做模型无关元学习模型.

与传统图像识别中深度神经网络训练过程进行比较，传统深度神经网络的训练往往在同一样本上计算损失函数并以此修改网络参数，即使用样本 A

的预测输出结果直接与 A 的真实标签进行对比从而得到损失函数，这一过程中模型倾向于反复学习识别样本 A，而非样本 A 所属的类的内在分布. 虽然借助正则化项、滑动平均与 *dropout* 等手段可以提升一定泛化性能，但通过强迫网络不依赖特定单一路径得到的模型的泛化性能上限较低，且仅适用于可用于训练的数据量较大的情景. 考虑用于训练的数据集中每一类数据样本数较少的场景，难以通过多次利用新数据不断修正偏置从而使模型可适用于该类数据的绝大多数.

MAML 创造性地改造了训练过程的结构，其在训练阶段双层循环的设计使得网络参数不再过分倾向于训练所用数据，而是从训练数据上学习之后用于新数据的效果优劣. 因此经过大量训练后，优化方法从宏观上可以看作在学习图像之间最容易作为区分依据的关键特征，即尽可能学习如何快速学习新任务的能力. 因此其学习的不仅是模型内嵌的深度神经网络参数矩阵，更是可以通过网络本身表达的学习能力. 将损失函数与直接训练数据解耦使得其衡量标准是学习过程的效果本身. 从特征学习的角度看，这是建立适应多种任务的内部特征表达的过程，也就是力图找到一个可以尽量好的满足所有任务的网络初始化参数，使得其中一小部分参数可以被更快、更容易地微调，也可以将学习过程看作最大化损失函数对和新任务有关的参数的敏感度. 当训练结束后，面临未经训练的新数据种类，参数的局部小变动可以使损失函数的表现得到巨大提高.

与其它方向的元学习算法相比，MAML 在设计结构上便体现出更佳的通用性和潜力. 首先其模型内置的网络本身通常不需要额外参数，且自适应能力较强，抽象地看该类模型仅在传统端到端网络的输出之后加上由验证集生成的损失函数，在多数据集之间切换成本较低，由于其中的网络本身仅负责反向传播以及输出通过该网络的结果，因此在模型的结构方面具备组件化特点. 这些特点使得 MAML 类模型可以方便地将适用范围扩展至如强化学习、回归预测等领域. 目前主流定制化与优化思路大致分为两种：

(1) 针对模型内嵌的网络结构本身. 验证传统深度神经网络结构，如 Residual Network、Dense Network 是否需要针对小样本快速适应场景进行定制，修改思路包括减少内层循环过程中需要被微调的参数数量、减少网络后端全连接层数量、使用更高效的激活函数，该修改思路的基础假设为，传统的以图像识别为代表的网络结构为了应对大量标签分类任务，往往采取较深的卷积过滤器以及网络层数，这样复杂的网络在训练结束后往往参数被固定，而元学习需要在适应新场景时修改网络参数，过多的参数带来了过拟合的风险，也拖慢了适应速度，因此速度与计算空间开销优先的网络便成为优化方向之一. 未来该方向研究重点为更高效的激活函数、损失函数与标准化矫正算法，而缺点也较为明显，即目前对于使用者依然黑盒的深度神经网络系统，理论的更优解难以直接在实验中得到反馈与验证，且大量算法结构是否存在更优解也依然存疑.

(2) 针对训练过程中内循环本身. MAML 中使用基于微分的梯度下降算法对网络参数进行调整，受限于元学习的设定，通常包含少量次数的梯度下降过程，使用少量样本对高维参数进行修改本身容易导致梯度下降方向错误或过拟合至 D_{train} 数据集. 具体到计算步骤，该种思路不考虑使用基于微分的梯度下降作为更新网络参数的手段，考虑重视传统梯度下降过程中被浪费的信息细节，如参数空间的高维几何结构，如将神经网络矩阵结构看作李群的一个切面，使用基于流形测地线的梯度代替微分梯度，从而获得更准确的梯度下降方向和步长. 但高维空间上测地线的计算通常依赖最短路径搜索与群的左右作用，随着网络维度逐渐增高，计算路径与对应梯度时间复杂度增长极快，此时通常使用线性近似或降低高维空间结构假设复杂度，如使用通用 Stiefel 流形群组建正交李群，以使用群变换算法，并进一步实现线性近似；或直接使用一个经过训练的神经网络计算梯度，将微分过程替换为通过一个小型神经网络模型，代价是降低一定通用性，且该结构依然为黑盒，进一步减少了对模型的可控制性与可解释性.

作为元学习领域当前较为优秀的算法，MAML 作为众多元学习算法的基础框架，在后续与其它技术结合过程中衍生出一系列优秀元学习算法，如使用隐空间概念以避免高维参数梯度下降困难的问题[35]、现实机器人短时学习运动[50]以及将 MAML 异步思想用于 GAN 人脸替换技术[51]等.

3.2.2　迁移元学习 MAML

前文提到，MAML 作为优秀的框架，在提供独特训练模式的同时也存在相当多不足，如特征提取部分网络较浅，仅使用全连接卷积层面对图像数据时难以满足训练所需的精度需求，而直接替换为例如 Residual Network 或 Dense Network[52]等带有跳跃连接结构的残差网络则大幅提高了训练难度，由于 MAML 的训练阶段的内循环中需要将初始权

重网络复制为多个副本进行梯度下降操作,因此若直接更新整个网络,将产生严重过拟合现象,难以达到较为稳定的学习效果.大量 MAML 类论文均指出,在训练阶段、内层循环学习阶段、测试这三阶段中,需要被更新的变量数量应尽可能越来越少.

同时 MAML 的任务设计也在某种程度上限制着模型整体的学习能力,通常 MAML 中单个训练任务由多种数据构成,且在训练时对每一类数据不做额外区分,即默认网络对每一标签下的数据学习难度是相同的,这个假设仅在拥有大量训练数据以及足够次数迭代场景中正确.在小样本学习设定中,存在一部分类型的数据比其它种类更难以被区分,对所有标签数据一视同仁将限制学习能力的上限.

针对以上问题,MTL(Meta-Transfer Learning)[53] 对基础 MAML 模型在特征提取器、任务设计、网络结构方面进行改进.理论上更深的网络不会带来更差的学习能力,因此采用宽残差网络在预训练(Pre-train)阶段可进一步提升对当前数据集潜在统一分布的提取能力,这转移了元训练阶段的学习压力,但极为有限的训练数据更难以有效更新更深的网络,因此 MTL 在预训练结束后便将特征提取器部分参数固定,这基于假设:当前数据集中全部数据大致服从于某个确定的分布.在训练和测试阶段,需要被更新的参数则替换为单层全连接层,这极大的减少了每次梯度下降计算时需要被更新的参数规模.

在训练任务设计方面,前文提到将训练任务中每一类数据的被学习难度看作相同会降低学习能力的上限,因此 MTL 在每次内层循环后,即通过少量梯度下降更新网络后,统计模型在查询集(query set)上的预测结果,记录其中预测错误次数最多的标签并加入集合,记作 Hard Task(HT) meta batch.在元训练阶段结束后,将所有标记为难以学习的数据用作对模型的进一步训练,而相应的实验结果也证实,通过使模型学习数据集中一些最难以被区分的数据,可以使得最终测试准确率更加稳定,降低模型的不确定性.

对模型中网络的更新主要包含两部分,即梯度计算与参数更新.在参数更新方面,每次梯度计算中都更新一遍网络中全部参数明显代价过大,且极少数数据难以进行方向正确的梯度下降,与其匹配的学习率也难以确定.减少内层循环计算资源开销主要有两种方法:固定特征提取器中绝大部分参数,或仅更新模型最后的全连接层.前者需要手动指定被固定的网络层与其中的参数,这降低了模型的自动化程度,且仅可通过实验验证剩余参数选择效果,大大增加网络设计复杂度;后者建立在假设上:在预训练阶段固定下来的特征提取器不仅可有效对当前数据集进行特征提取,还可满足快速适应的实验场景,实验证明该思路由于将特征提取器完全固定,难以将用于完整数据集特征提取的特征提取器转换为快速响应型网络.MTL 通过为特征提取器中网络参数矩阵添加缩放系数(Scaling)与偏移系数(Shifting),记作 Φ_{S_1} 与 Φ_{S_2}.具体做法为,卷积层中过滤器卷积核中的每一层均添加一组缩放系数与偏移系数,以原始网络约 10% 的参数量控制整个特征提取器中的所有参数.由于卷积层的卷积核上每一层仅需添加一个标量 Φ_{S_1},因此在元训练阶段需要被更新的参数数量大幅减少.

将特征提取器中包含卷积层的网络部分记作 Θ,与全连接层部分记作 θ,则 MAML 类的参数更新方式可记作:

$$\theta' \leftarrow \theta - \beta \nabla_{\theta} \mathcal{L}_{T(tr)}([\Theta;\theta]) \tag{18}$$

而对卷积核添加额外控制系数后,元训练阶段梯度更新过程为

$$\begin{aligned} \phi_{S_1} &\leftarrow \phi_{S_1} - \gamma \nabla_{\phi_{S_1}} \mathcal{L}_{T(tr)}([\Theta;\theta'],\phi_{S_1}) \\ \phi_{S_2} &\leftarrow \phi_{S_2} - \gamma \nabla_{\phi_{S_2}} \mathcal{L}_{T(tr)}([\Theta;\theta'],\phi_{S_2}) \\ \theta &\leftarrow \theta - \gamma \nabla_{\theta} \mathcal{L}_{T(tr)}([\Theta;\theta'],\phi_{S_{\{1,2\}}}) \end{aligned} \tag{19}$$

可以看出,每次计算梯度时,求导对象由特征提取器的全体变为参数量极少的缩放与偏移系数 Φ_{S_1} 与 Φ_{S_2},这大大加快了训练速度并减少计算资源开销.考虑测试阶段与训练阶段数据量以及对模型参数修改次数的区别,在测试阶段 MTL 进一步减少内循环过程需要被更新的参数数量.由于模型在元训练阶段已经完成对当前数据集的适应,特征提取器可不经更新便应用至需要快速适应的测试场景.因此在测试过程中仅更新分类器参数 θ,这样便通过迁移学习的思路,使用系数完成对整个网络的更新,特征提取器中卷积核矩阵与其对应的偏置结构示意图如图 5.

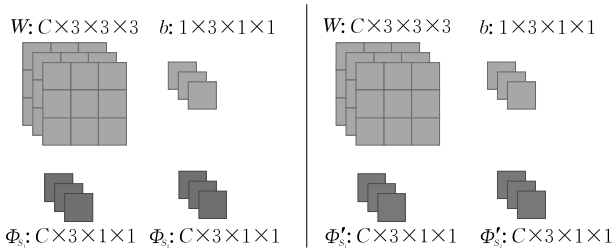

图 5 模型算法示意图

该方法着力于尽可能减少需要被改变的参数量，将卷积层中参数分组控制，为网络参数的更新方式提供了新思路，致力于提供规则统一且高效的训练资源分配逻辑. 对特征提取器进行与训练处理思路与当前机器学习领域对数据集的基本假设一致，即同一数据集内，所有数据大致符合某一分布，这也是神经网络可被训练的逻辑前提. 通常一个数据集包含训练(train)、测试(test)与验证(validation)三部分，且彼此无交集，使用训练集训练一个包含深层神经网络的分类模型，并将特征提取部分参数进行固定，可避免在元训练阶段模型同时学习提取特征与分类，即为模型提供有一定分类能力的起始点，通过预训练后固定大部分参数的设定，让模型在元训练和测试时更专注于对分类层的学习，加强对提取后数据间区分特征的敏感度. MTL 减少需训练参数量的方式高效但忽视了卷积核中每一层中不同单元的值与当前卷积层整体输出之间的协调性，即这种对网络的更新方式将卷积核维度重新划分，卷积核中每一层是否在经过整体缩放和偏移后是否依然可以与其它层协调还有待验证. 对于绝大部分深度神经网络，其中的卷积核包含了其中绝大部分参数，通常为四维矩阵，MTL 对卷积核添加的三维矩阵前三维度与对应卷积核相同，因而通过该方法减少的参数量并没有实现数量级的变化，进一步减少训练参数量，可以仅修改卷积核偏置或在预训练阶段便得到强泛化性能的特征提取器.

由实验结果可知，将迁移学习的线性变化思路应用于元学习领域可有效降低训练难度并提高模型稳定性，Hard task 思想的引入则可提高模型在当前数据集中的泛化能力.

3.2.3　不稳定环境条件下元学习的持续适应算法

近年来强化学习已经在众多领域取得此前监督学习无法实现的成果，如自动玩游戏[54]、人机对话系统[55]等，但相当多强化学习产生的优异成果存在前提是不变的环境参数，但现实世界中的环境往往是在持续变化且极为复杂的[56]，传统的处理连续变化环境变量的算法常基于上下文检测[57]，基于对已经发生的事件实时调整策略参数，然而现代大部分强化学习算法即使在一些项目上超越人类水平，但依然局限于一些狭窄领域.

论文“Continuous Adaptation via Meta-Learning in Nonstationary and Competitive Environments”[37]的思想是将持续变化的环境看作由一个个静止的瞬间组成，如同视频可视作由大量静态帧组成，期望提出一个可以不断对将要发生情况进行预测并提前调整参数的模型，而元学习的一个目标便是持续学习乃至终身学习[58-59].

模型基于 MAML[14]，将在多主角游戏中进行模拟[60]，对于 MAML，其核心思想便是修改网络架构使用内外层循环的设计进行训练，并将内外层循环任务样本设置为不同的，大大增加了模型的泛化性能. 本模型采用类似思想，即基于强化学习产生的轨迹进行训练，采用异步求损失函数的方式增加模型适应性和泛化性能. 其模型算法如图 6.

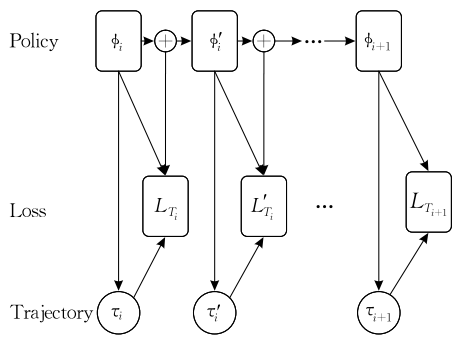

图 6　模型算法示意图

该算法将 t_1 时刻 w_1 在模型上结合产生的轨迹数据 τ_ϕ 进行梯度下降后的 w_1' 应用至 t_2 时刻的任务上，并求出在 w_1' 在 t_2 任务上的损失函数 $\nabla_\theta \mathcal{L}_{T_i,T_{i+1}}$ 以及 $\nabla_\alpha \mathcal{L}_{T_i,T_{i+1}}$，在随后训练中，最小化该损失函数，以达到尽可能让前一时刻基于环境进行梯度下降后的权重参数可以最好的适用于新任务和新环境，也就是让不断变化的网络参数可以在一系列任务、不同参数的环境中可以尽量平滑的过渡，相比起传统强化学习算法的更注重当前模型参数在已有数据上最小化损失函数，本模型更注重网络参数从一个任务到下一时刻任务的适应性，即元学习的泛化性特点. 因此其损失函数并不仅仅定义为单个任务的损失值和，而是临近的两个任务共同产生的过渡适应损失函数. 其优化目标为

$$\min_{\phi_0} \mathbb{E}_{P(T_0),P(T_{i+1}|T_i)}\left[\sum_{i=1}^{L}\mathcal{L}_{T_i,T_{i+1}(\phi_i)}\right] \tag{20}$$

其中目标损失函数为

$$\mathcal{L}_{T_i,T_{i+1}(\phi_i)}:=\mathbb{E}_{\tau_{i,\phi_i}^{1:K}\sim P_{T_i}(\tau|\phi_i)}\left[\mathbb{E}_{\tau_{i+1,\phi}\sim P_{T_{i+1}}(\tau|\phi_{i+1})}\left[L_{T_{i+1}}(\tau_{i+1,\phi_{i+1}})|\tau_{i,\theta}^{1:K},\phi_i\right]\right] \tag{21}$$

而在 ϕ_i^0 梯度下降到 ϕ_{i+1} 的过程中，每次都从 θ 开始，因此上式中 ϕ_i 为 θ，公式可写作：

$$\mathcal{L}_{T_i,T_{i+1}(\theta)}:=\mathbb{E}_{\tau_{i,\theta}^{1:K}\sim P_{T_i}(\tau|\theta)}\left[\mathbb{E}_{\tau_{i+1,\phi}\sim P_{T_{i+1}}(\tau|\phi)}\left[L_{T_i}(\tau_{i+1,\phi})|\tau_{i,\theta}^{1:K},\theta\right]\right] \tag{22}$$

在对损失函数的梯度下降部分，根据每个 T_i，对 θ 进行少量梯度下降得到 ϕ_{i+1}，接着使用该 ϕ_{i+1} 生成的策略 $\pi_{\phi_{i+1}}$ 作用于 T_{i+1} 得到 $\tau_{\phi_{i+1}}$，计算 $\tau_{\phi_{i+1}}$ 上的 loss，最终将所有 task 产生的 loss 期望最小化. 并得到 $\mathcal{L}_{T_i,T_{i+1}}(\theta,\alpha)$. 与 MAML 强调 θ 可代表全体 task 的一些共有属性不同，该算法假设不同 task 之间存在信息的迭代，即假设 task T_i 包含一些 T_{i+1} 的信息，且 T_{i-1} 的一些信息经过微调后可适用于 T_i. 该论文方法总目标是找到一个最佳初始参数，可以不断根据上一个 task 的参数 ϕ_i 调整用于下一时刻的参数 ϕ_{i+1}.

在实验部分，使用 3D 模拟环境 RoboSumo 中，此论文算法控制下的 agent 在学习如何进行与其它 agent 对抗过程中进步更快，且能快速适应本体 agent 的参数变化，如修改 agent 的体积、四肢数量等. 此论文作为将 MAML 与强化学习结合，且在连续动态环境中取得优秀结果的第一篇论文，为后续将该强化学习算法应用于真实世界机器人做好了理论准备，Nagabandi 等人已将该思想用于机器人控制[50]，并取得了优秀结果，实验表明采用该强化元学习的机器人适应新环境能力相比传统强化学习算法有巨大提升，且不易在某种地形环境上产生过拟合现象.

3.2.4 基于贝叶斯理论的元学习方法

3.2.4.1 基于 MCMC 的元学习

论文“Meta-Learning by Adjusting Priors Based on Extended PAC-Bayes Theory”[61] 提出一种基于泛化误差边界的元学习框架，通过该框架可以将 PAC-Bayes 边界扩展到元学习领域，同时给出贝叶斯泛化误差边界的证明. 作者提出的基于贝叶斯元学习方法构建基于已观察任务的假设的分布，再利用该分布来学习新任务. 先验知识通过为新任务设置一个依赖经验的先验来整合. 作者基于以上思想提出了一种基于梯度下降的算法，最小化了由边界导出的目标函数.

此文思想类似上述提到的 MAML 算法，均使用元学习器从样本中学习先验参数，再通过这个参数对模型进行初始化操作，此文基于贝叶斯理论，因此学习到的先验知识是一个分布. 在测试阶段对未参与训练的数据，即需要新学习的任务样本，通过先验知识学习到的参数初始化一个适用于新任务的分布，思想类似迁移学习中的 fine-tuning.

该论文另一重要内容为通过可计算理论(PAC)给出元学习的泛化误差边界，且作者给出了针对每个任务的泛化误差边界以及总体的任务边界，由于元学习基于多任务，作者还给出多任务学习中的联合优化算法. 通过可计算理论(PAC)给出 Mc Allester 提出的单任务误差边界：

$$er(Q,D)\leq er'(Q,S)+\sqrt{\frac{D(Q\|P)+\log\frac{m}{\delta}}{2(m-1)}} \tag{23}$$

推广到元学习，给出元学习(多任务)下的贝叶斯边界，并对其进行联合优化：

$$er(\mathcal{Q},\tau)\leq\frac{1}{n}\mathop{\mathbb{E}}_{P\sim\mathcal{Q}}er'(Q_i,S_i)+\frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{D(\mathcal{Q}\|P)+\mathop{\mathbb{E}}_{P\sim\mathcal{Q}}D(Q_i\|P)+\log\left(\frac{2nm_i}{\delta}\right)}{2(m_i-1)}}+\sqrt{\frac{D(\mathcal{Q}\|P)+\log\left(\frac{2n}{\delta}\right)}{2(n-1)}} \tag{24}$$

同时作者给出基于贝叶斯与梯度的算法 MLAP，该算法思想与 MAML 类似，针对每个任务进行训练，所有任务均进行一轮训练后更新模型网络的梯度与输出，将得到的先验分布用于后续新任务的初始化，并在新任务上经少量梯度下降微调以适应未经训练的新任务种类的分类任务.

3.2.4.2 基于变分贝叶斯的元学习方法

论文“Meta-Learning Probabilistic Inference for Prediction”[62] 给出了多任务的有向图模型，在此之上给出了元学习新的概率学解释，并基于此解释提出了一种新的算法示例 Versa.

本文将多任务学习解释为如下图 7 结构.

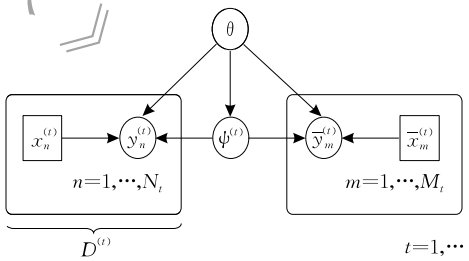

图 7 多任务学习有向图模型

其中 $x_n^{(t)}$、$y_n^{(t)}$ 为 t 任务的训练集的输入输出，$\tilde{x}_n^{(t)}$、$\tilde{y}_n^{(t)}$ 为 t 任务的测试集的输入输出，$\psi^{(t)}$ 为任务相关参数，θ 为元参数.

基于上述，本模型中的元学习的目标即如何快速地求解未观测任务的后验分布：

$$p(\tilde{y}_n^{(t)}|\tilde{x}_n^{(t)},\theta)=\int p(\tilde{y}_n^{(t)}|\tilde{x}_n^{(t)},\psi^{(t)},\theta)\,p(\psi^{(t)}|D^{(t)},\theta)\,\mathrm{d}\psi^{(t)} \tag{25}$$

本模型采用变分方法来近似后验分布，通过一个摊销分布 $q_\phi(\tilde{y}|D)$ 去近似预测后验分布，具体来说，通过用所有训练集和测试集训练一个参数是 ψ 的前馈推理网络来输出测试输出 $\tilde{y}^{(t)}$ 的近似后验分布

$q_\phi(\tilde{y}|D)$，通过这个分布来近似后验预测分布：

$$q_\phi(\tilde{y}|D)=\int p(\tilde{y}|\psi)q_\phi(\psi|D)\mathrm{d}\psi \tag{26}$$

在单任务模型中通常使用 KL 散度来比较近似分布和实际分布的距离. 由此元学习的目标及最小化所有任务的 KL 散度的期望：

$$\begin{aligned}&\phi^*=\arg\min_{\phi}\mathbb{E}_{P(D)}(T)=\arg\max_{\phi}\mathbb{E}_{P(\tilde{y},D)}(S)\\&T=\mathrm{KL}[p(\tilde{y}|D)\,\|\,q_\phi(\tilde{y}|D)]\\&S=\log\int p(\tilde{y}|\psi)q_\phi(\psi|D)\mathrm{d}\psi\end{aligned} \tag{27}$$

VERSA 与上述 MLAP 方法的不同之处：

MLAP 采用的是可计算理论通过扩展单任务的误差上界到元学习，最后通过 MCMC 采样来最小化误差边界来实现，是一种非确定的近似，VERSA 采用的是变分方法，通过引入一个简单分布 q，来优化 q 与真实分布之间的 KL 距离来近似后验分布，是一种确定性近似方法. VERSA 相比 MLAP 训练开销更小.

3.2.5　使用上下文参数快速适应元学习算法

本模型简称为 CAVIA，本方法以 MAML 作为基础框架，提出一种结合上下文参数的算法模型[63]，与 MAML 不同的是，MAML 在每个新任务上会更新所有参数，而 CAVIA 将模型的参数分为两部分，分别是上下文参数与共享参数. 其中上下文参数作为模型的额外输入，使其不仅可实现较好的泛化性能，也可以针对适用于单独的特殊任务，共享参数在任务间共享，并通过元学习训练过程优化. CAVIA 在每一个单独的新任务上仅更新模型的上下文参数. 这样可以在使用更大网络情况下避免在单一任务上过拟合，且减小了内存开销. 该模型结构如图 8 所示.

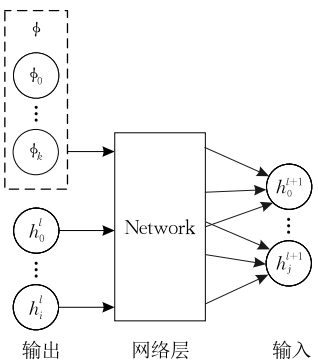

图 8　CAVIA 模型示意图

上述示意图中 $\phi_0\sim\phi_k$ 作为测试时的额外输入参数，来源是 MAML 模型中内层循环中产生的网络参数副本，也就是“上下文参数”. 其余 $h_0^l\sim h_i^l$ 为待测试样本本身输入.

为了使模型在未经训练的新任务上快速学习，即仅通过少数几步梯度下降便适应新任务，MAML 此时的内循环实际上等价于一个任务识别问题，而不是学习如何解决整个任务，因此如果模型中跨任务变化的部分是模型的额外输入，且独立于其它输入，便可在拥有较强泛化性的前提下实现对新任务的准确识别.

本方法与 MAML 相比的改进之处：

(1) 增加了模型的额外输入，即上下文参数 ϕ，可看作调整模型行为的嵌入任务或条件. 参数 ϕ 在元学习的内层循环中更新，其余参数 θ 在外层循环中更新，使得 CAVIA 优化过程更加准确，在任务间优化任务独立参数 θ，同时保证任务特异的参数 ϕ 可以快速适应新任务.

(2) 使任务求解与任务嵌入这两部分分离，任务求解与任务嵌入有如下优点：这两部分的大小可根据任务进行合适的调整，使得在内层循环中使用更深的网络进行训练时不会对某特定任务过拟合(MAML 在使用更深的网络时会过拟合). 模型设计和结构选择也从这种分离中受益，对于许多实际问题，往往任务之间哪些方面不同以及 ϕ 的容量是可知的. 同时由于仅对 ϕ 求高阶导数，避免神经网络对权重与偏差的操作. 由于不用像 MAML 一样在内层循环时复制参数产生网络参数的副本，因此减少了内存写入次数，加快了训练和学习的速度.

3.3　基于梯度优化器的元学习方法

一般认为，基于梯度的优化算法需要许多迭代步骤和许多样本才能很好的执行，论文“Optimization as a Model for Few-Shot Learning”[17] 中提出了一个基于 LSTM 的元学习者模型，学习精确的优化算法，用于在少样本问题下训练另一个学习者神经网络分类器.

通常的优化算法使用一些梯度下降的变体来训练深度神经网络，网络的更新形式为：

θ_{t-1} 为经过 $t-1$ 次更新后的参数，α_t 为 t 时刻的学习率，$\nabla_{\theta_{t-1}}\mathcal{L}_t$ 为损失函数关于 θ_{t-1} 的梯度，θ_t 为更新后的参数，在 t 时刻经过梯度下降的网络参数 θ 可表示为

$$\theta_t=\theta_{t-1}-\alpha_t\nabla_{\theta_{t-1}}\mathcal{L}_t \tag{28}$$

而 LSTM 隐藏状态的更新公式为

$$c_t=f_t\odot c_{t-1}+f_t\odot\tilde{c}_{t-1} \tag{29}$$

通过观察这两个公式，若将遗忘门 f_t 设置为 1，令隐藏单元值 $c_{t-1}=\theta_{t-1}$，输入门 $i_t=\alpha_t$，隐藏状态候选值 $\tilde{c}_t=\nabla_{\theta_{t-1}}\mathcal{L}_t$，则参数的更新过程和 LSTM 隐藏状态的更新过程相似，因此可以训练一个 LSTM 作

为“元学习者”来学习参数的更新规则,使用该 LSTM 来训练其它神经网络,即将 $t-1$ 时刻的网络参数 θ_{t-1} 作为该 LSTM 网络的输入,可输出比微分梯度下降更有效的更新后的网络参数 θ_t.

将 LSTM 隐藏单元设置为学习者的参数,令 $c_t=\theta_t$,候选值 $\tilde{c}_t=\nabla_{\theta_{t-1}}\mathcal{L}_t$,表示梯度中可用于更新的信息的数量.将遗忘门 f_t 和输入门 i_t 表示为参数的形式如下:i_t 与学习率相对应,是关于 $t-1$ 时刻的参数 θ_{t-1}、梯度 $\nabla_{\theta_{t-1}}\mathcal{L}_t$、损失值 $\mathcal{L}_t$、之前的学习率 i_{t-1} 的函数,根据这些信息,元学习者可以调节控制学习率,更快的训练待训练模型的参数值.

$$i_t=\sigma(W_I\cdot[\nabla_{\theta_{t-1}}\mathcal{L}_t,\mathcal{L}_t,\theta_{t-1},i_{t-1}]+b_I) \tag{30}$$

对于遗忘门 f_t,最优的选择不为常数 1,当学习者的参数状态置于局部最优,即损失值很大,但梯度接近于 0 的情况时,可以通过 f_t 来遗忘之前学到的 θ_{t-1} 的部分值,从而逃逸出该状态,f_t 的函数参数与 i_t 一致.

$$f_t=\sigma(W_F\cdot[\nabla_{\theta_{t-1}}\mathcal{L}_t,\mathcal{L}_t,\theta_{t-1},f_{t-1}]+b_f) \tag{31}$$

同时我们也可以学习 LSTM 隐藏单元的初始值 c_0,将 c_0 也作为元学习者的参数,通过训练这个初始化值让元学习者决定学习者的最优初始化参数.

通常学习者(神经网络)有上万的参数,若将 LSTM 的隐藏单元设置为上万个,LSTM 参数的数量将失控,难以进行训练,因此将 LSTM 的隐藏单元设置为一个,学习者的每个参数对应一个 LSTM,这些 LSTM 之间的参数共享,即它们的更新规则一样,但各自的历史隐藏信息不同.

由于每一个梯度和损失通常会有非常不同的量级,需要归一化这些值以便元学习者能正常使用,梯度和损失的预处理如下:预处理调节了梯度和损失的范围,同时将量级信息和方向信息分离.通过预实验可知 $P=10$ 时表现最好.

$$\begin{cases}\left(\dfrac{\log(|x|)}{p},sgn(x)\right), & \text{if } |x|\geq e^{-p}\\ (-1,e^p x), & \text{其他}\end{cases} \tag{32}$$

3.4 基于外部记忆单元的元学习方法

3.4.1 使用记忆增强的模型

论文“Meta-learning with Memory-Augmented Neural Networks”[11]提出,对于需要拥有较强泛化性能的模型,仅依赖深度多层神经网络容易对某些任务过拟合,因此尝试使用拥有外部存储的神经网络通过元学习的方法解决小样本学习的问题.论文参照神经图灵机(NTMs),引入外部存储模块,使用 LSTM 替换 RNN 作为控制器.论文中提出 LSTM 内部的存储不能满足接收新任务后能快速编码大量新信息的需求,因为 LSTM 本质上是时序比较长的短期记忆.但加上外部存储后,即 MANN 同时拥有长期和短期记忆能力,通过调用长短期记忆实现新信息的快速编码.

算法如图 9 所示,在每个时间步将待识别的样本 x_t 和上一个时间步的标签 y_{t-1} 输入模型,输出的结果是 x_t 的预测标签.在每个 episode 中,随机从数据集分布 $p(D)$ 中采样组成训练数据,因此模型的优化目标是:

$$\theta^*=\arg\min_\theta E_{D\sim p(D)}[L(D;\theta)] \tag{33}$$

模型为了防止对样本的 label 形成绑定记忆,将不同 episode 中样本标签打乱.

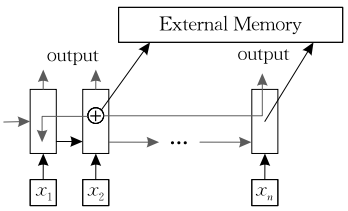

图 9 算法流程示意图

输入样本 x_t 通过控制器得到一个键值 k_t,键值 k_t 被用于访问外部存储内容或者写入外部存储.利用键值和存储单元之间的余弦相似度生成读权重:

$$w_t^r(i)\leftarrow\frac{\exp(K(k_t,M_t(i)))}{\sum_j\exp(K(k_t,M_t(j)))} \tag{34}$$

从外部存储读取的内容为

$$r_t\leftarrow\sum_i w_t^r(i)M_t(i) \tag{35}$$

将记忆写入外部存储使用 LRUA 方法(Least Recently Used Access),利用读权重,写权重和上一时间步的使用权重的线性插值得到当前时间步的使用权重:

$$w_t^u\leftarrow\gamma w_{t-1}^u+w_t^r+w_t^w \tag{36}$$

利用使用权重计算出最少使用权重:

$$w_t^{lu}(i)=\begin{cases}0, & \text{if } w_t^u(i)>m(w_t^u,n)\\ 1, & \text{if } w_t^u(i)\leq m(w_t^u,n)\end{cases} \tag{37}$$

最后计算出写权重:

$$w_t^w\leftarrow\sigma(\alpha)w_{t-1}^r+(1-\sigma(\alpha))w_{t-1}^{lu} \tag{38}$$

根据写权重将记忆写入外部存储:

$$M_t(i)\leftarrow M_{t-1}(i)+w_t^w(i)k_t,\forall i \tag{39}$$

3.4.2 surprise-based 的记忆方法

论文“Adaptive Posterior Learning: Few-shot Learning with a Surprise-based Memory Module”[64]介绍了近似后验学习(APL)算法:一种通过记住它

所遇到的“surprise”的观测来逼近概率分布的算法. 过去的观测来自外部存储器模块，并经过解码器网络处理，解码器网络可以组合来自不同记忆插槽的信息. 为了应对当前任务，APL 要学会对新的概率分布进行少量的近似，并且只存储尽可能少的上下文点. 此外，为了执行不同程度的复杂任务，它还要学习如何处理读取的经验.

模型由以下部分组成：为传入的查询数据生成表示的编码器；一个外部记忆存储器，包含以前见过的表示或者数据对；管理写入的记忆控制器；一个解码器，它结合查询表示和来自记忆存储器的数据来生成目标的概率分布. 编码器为输入的样本生成一个低维表示，通常为一个卷积网络. 外部存储是一个包含存储经验的数据库，每一列对应数据的一个属性. 当用于分类时，存储两列：样本的嵌入表示 e_m 和实际标签 y_m. 查询记忆时返回查询样本的 k 近邻. 记忆控制器最小化写入记忆的数据点的数量. APL 算法将“surprise”定义为与标签 y_t 预测相关的数量：$S=-\ln y_t$，这意味着模型分配给真实类别的概率越高，就越不惊讶. 如果数据点“surprise”，则应将其存储在存储器中；否则，可以被安全丢弃，因为模型已经能够正确地对其进行分类. 设置一个超参数作为“surprise”的基准值. 解码器的输入为查询表示和从外部存储返回的 k 近邻，论文设计了一个具有自注意的关系前馈模块，模块通过将每个邻居与查询表示进行比较来分别处理每个邻居，然后再用根据邻居距离计算的注意向量来减少激活之前使用自注意模块进行交叉元素比较.

APL 模型通过包含数据对的序列来训练，目的是学习一种序列更新算法，即它要最小化由一系列数据元素组成的序列的预期损失. 模型不需要反向传播，模型的参数在每个时间步独立地由最小化交叉熵损失来更新参数.

APL 使用一种简单的内存控制器设计，它使用基于“surprise”的信号将最具预测性的内容写入存储. 由于不需要学习写什么，避免了通过记忆进行反向传播的代价，这使得训练更容易设置并且更快. 这种设计也最小化了存储数据，使得方法更有内存效率.

3.4.3　记忆匹配网络模型

论文“Memory Matching Networks for One-shot Image Recognition”[65] 提出了 Memory Matching Networks(MM-Net)，MM-Net 将一组标记图像(支持集)的特征写入记忆，并在执行推理时从记忆中取出，从而全面地利用集合中的知识，同时将记忆槽输入到一个上下文学习器：循环神经网络(RNNs)中来预测未标记图像的 CNNs 参数. 深度 CNNs 用来学习图像的表示，通过写控制器来更新记忆，CNNs 的输出被看作是未标记图像的嵌入. 通过给定的未标记图像的嵌入与支持集中的每一幅图像之间的点积来计算它们之间的相似度，最相似的图像的标签被分配给这个未标记的图像.

模型采用键值记忆网络作为记忆模块，将类别上指定的编码上下文信息存储到键值结构的记忆中. 通过写控制器将整个支持集 S 编码入记忆后，最终的记忆 MN 被赋予了支持集中的上下文信息.

3.5　基于数据增强的元学习方法

3.5.1　MetaGAN 对抗模型

论文“MetaGAN：An Adversarial Approach to Few-Shot Learning”[66] 提出了一个简单通用的框架，称为 MetaGAN. 通过引入一个基于任务的对抗生成器(Adversarial Generator)，对传统的元学习方法进行了改进. 其主要思想是在区分真实数据与虚假数据的过程中，能够为模型找到一个更加紧致的决策边界，增强了模型的特征提取能力. 同时，该框架可以扩展有监督的元学习方法，自然地处理未标记的数据.

该模型将元学习方法和 GAN 生成模型进行结合. GAN 模型中的不完美生成器可以在不同真实数据类的流形之间提供假数据，从而为分类器提供额外的训练信号，并使决策边界更加清晰. 具体地，对于 N-way 的分类任务，模型将分类器输出的维度从 N 增加到 $N+1$(当标签为 $N+1$ 时表示虚假数据)，以建模输入虚假数据的概率. 在对抗的设置下同时训练分类器 D 和生成器 G.

分类器的损失分为有监督和无监督两个部分. 对于原始数据集中的带有标签的真实图片，分类器需要对应到真实的标签 y；对于生成的虚假图片，分类器需要将标签识别为 $N+1$；对于原始数据集中无标签的数据，由于不知道其真实的标签信息，只需使其标签 $y\leq N$ 即可.

$$\mathcal{L}_T^D=\mathcal{L}_{\text{supervised}}+\mathcal{L}_{\text{unsupervised}} \tag{40}$$

$$\mathcal{L}_{\text{supervised}}=E_{x,y\sim Q_T^s}\log p_D(y\mid x,y\leq N) \tag{41}$$

$$\mathcal{L}_{\text{unsupervised}}=E_{x,y\sim Q_T^u}\log p_D(y\leq N\mid x)+E_{x\sim p_G^T}\log p_D(N+1\mid x) \tag{42}$$

生成器要尽可能的“欺骗”分类器，让其将生成数据识别为真实数据. 因而其损失定义为

$$\mathcal{L}_G^T(D)=-E_{x\sim p_G^T}\log p_D(y\leq N\mid x) \tag{43}$$

论文将 MetaGAN 的思想与关系网络和 MAML 进行了结合，相比原始模型，均取得了一定的提升. 它与基于度量和基于初始化的元学习方法都能够很好的结合，并且为有些不具有半监督学习能力的模型提供了一个思路.

3.5.2 特征幻化模型

论文“Low-shot Visual Recognition by Shrinking and Hallucinating Features”[67] 提出了特征幻化的思想，通过生成结构将基础类映射为对应环境下的新类样本. 该模型分为表示学习和小样本学习两个部分. 在表示学习阶段，学习器在含有大量数据的基础类上提取精确的特征表示. 在小样本学习阶段，学习器在只含有少量数据的新类型和之前的基础类型的联合空间上学习分类器.

在小样本学习阶段，由于新类型数据缺乏，使得分类器难以捕捉到该类型内部的变化. 假设有一种特殊的鸟类，然而却只有其在飞行状态下的图片，这样分类器很可能会得出错误的结论：在树枝上、在地上的都不是此类型. 而通过生成结构可以得到不同环境下的样本，使分类器更加精确.

特征幻化方法主要模拟了人类所拥有的“类比”能力（当见过鸟类及其在飞机上的图片后，给定一种其它的飞行动物的图片，能够想象它在飞机上的场景）. 首先，在基础类中随机选取若干的图片对，每对图片 c_1^a, c_2^a 属于类 a，从其它类中依次选取 c_1^b, c_2^b，使得余弦距离 $D_{\cos}(c_1^a - c_2^a, c_1^b - c_1^b)$ 最小. 将四元组 $(c_1^a, c_2^a, c_1^b, c_2^b)$ 构成的集合作为生成数据集用来训练生成器 G. 在训练 G 时，每次以 (c_1^b, c_1^b, c_2^b) 为输入，由 G 生成 c_1^a.

论文还提供了一种基于特征的正则化方案 SGM，减少了在大数据集上训练的分类器和在小数据集上训练的分类器之间的差异，在小数据集上有更好的泛化能力.

3.5.3 CPAAN 模型

论文“Low-shot Learning via Covariance-preserving Adversarial Augmentation Networks”[68] 利用相似类的概念，由与新类型最为“相似”基础类来生成数据，同时限定了相似类之间的协方差矩阵，以保证其分布的一致性.

首先，通过原型网络的特征提取结构得到每张图片的特征表示，利用“原型点”的思想，计算每个新类型和所有基础类的欧式距离作为类型之间的相似度. 然后根据相似度选取最接近的 k 类作为相似类 $R(y_n)$.

$$\alpha(y_b, y_n) = \frac{\exp(-\| I_{y_b} - I_{y_n} \|_2^2)}{\sum_{y_b' \in Y_b} \exp(-\| I_{y_b'} - I_{y_n} \|_2^2)} \tag{44}$$

其中，Y_b 为基础类的标签集，y_n 为新类标签，I_y 表示类型 y 对应的原型点.

模型采用了 CycleGAN 作为生成结构来实现新类和对应相似类之间的相互转换，考虑新类所含样本数量较少，随机添加了噪声 z 使新类的分布更加多样化.

最后，模型添加了相似类和基础类的协方差矩阵距离作为整体损失函数的一部分.

$$\mathcal{L}_{\text{cov}}(G_n) = E_{y_n \sim Y_n}[E_{y_b \sim R(y_n)}[a(y_b, y_n) d_{\text{cov}}(y_b, y_n)]] \tag{45}$$

3.5.4 Imaginary data 模型

论文“Low-shot learning from imaginary data”[69] 提出生成虚假数据可以扩充样本的多样性，可以实现更好的元学习. 论文认为通过生成结构，能将拍照姿势、光照条件等环境因素迁移到新的样本上，从而生成具有不同变化的新样本，实现数据的扩充.

其数据生成方式比较简单：$x_{\text{new}} = G(x, z; wG)$，其中 x 为原始数据，z 为随机噪声. 模型将生成数据作为原始数据的补充，一齐放入分类器中进行训练.

作者还认为单纯通过生成模型生成的数据可能会造成模式崩溃，又或者不一定对能提升最终的分类效果. 因此，该模型将生成结构和分类结构视为一个整体，实现端到端的元学习方法，以保证生成的虚拟数据最终对任务本身是有利的.

4 元学习研究方法分析比较

4.1 元学习模型效果比较

4.1.1 符号解释

M：mini-ImageNet 数据集；

O：Omniglot 数据集；

5w-1s：5way，1shot，5 分类，每一类使用 1 个样本进行训练；

5w-5s：5way，5shot，5 分类，每一类使用 5 个样本进行训练；

孪生网络：Siamese Neural Networks for One-Shot Image Recognition；

匹配网络：Matching Networks for One Shot Learning；

原型网络：Prototypical Networks for Few-shot Learning；

关系网络：Learning to Compare：Relation Network for Few-Shot Learning；

TADAM：Task dependent adaptive metric for improved few-shot learning；

MAML：Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks；

MTL：Meta-Transfer Learning for Few-Shot Learning；

CMAML：Continuous Adaptation via Meta-Learning in Nonstationary and Competitive Environments；

MLAP：Meta-Learning by Adjusting Priors Based on Extended PAC-Bayes Theory；

CAVIA：Fast Context Adaptation via Meta-Learning；

LEO：Meta-learning with Latent Embedding Optimization；

MANN：Meta-learning with Memory-Augmented Neural Networks；

元学习器 LSTM：Optimization as a Model for Few-Shot Learning.

4.1.2　模型效果比较

表 1 为多种元学习模型在 mini-ImageNet 与 Omniglot 数据集上进行小样本学习测试结果.

表 1　元学习模型在 mini-ImageNet 数据集与 Omniglot 数据集上实验识别准确率比较

模型	方向	思路	M-5w-1s	M-5w-5s	O-5w-1s	O-5w-5s
孪生网络	度量	使用图片对的形式训练网络准确提取同种类图片共有特性的能力	None	None	92.00%	None
匹配网络	度量	使用多个网络同时对同一个样本进行特征提取后匹配，并提高样本特征提取稳定性	46.60%	60.00%	98.10%	98.90%
原型网络	度量	以每一类样本都存在一个最能代表该种类的样本作为原型点为基础，通过聚类思想进行分类	49.42%	68.20%	98.80%	99.70%
关系网络	度量	通过关系网络抽取样本特征并进行比对，以确定样本的分类	50.44%	65.32%	99.60%	99.80%
TADAM	度量	对度量尺度进行缩放；采用动态特征提取器	58.5%	76.7%	None	None
MAML	初始化	通过双层循环中损失函数计算方式中样本的区别，提高泛化性，得到优秀的初始化	48.70%	63.15%	98.70%	99.90%
MLT	初始化	在 MAML 基础上引入迁移学习思想，减少需梯度下降更新的变量数量	61.20%	75.50%	None	None
CMAML	初始化	将 MAML 用于环境持续变化的强化学习，改造连续两个时刻初始参数的损失函数以提高适应速度	None	None	None	None
MLAP	初始化	给出泛化误差边界并通过贝叶斯改造训练过程增强利用先验知识以及泛化能力	None	None	None	None
VERSA	初始化	将 MAML 学习系统以概率方式解释，通过摊销网络计算后验分布.	53.80%	67.37%	99.70%	99.75%
CAVIA	初始化	在 MAML 基础上增加额外输入，兼顾泛化性能以及新样本的特性	51.82%	65.85%	None	None
LEO	初始化	引入隐函数、关系网络等操作，通过梯度计算过程中实时降维减小高维网络中梯度计算困难问题	61.76%	77.59%	None	None
元学习器 LSTM	优化器	将梯度下降过程替换为额外网络，以达到梯度下降更快更准的目的	43.44%	60.60%	None	None
MANN	外部存储	通过添加额外的外部存储保存需要长期记忆的数据，以便随时调用	None	None	82.20%	98.10%
APL	Based-surprise	根据阈值来更新记忆损失			97.90%	99.90%
MM-Net	记忆匹配网络	模型采用键值记忆网络作为记忆模块，将指定上下文信息存储到记忆中.	53.37%	66.97%	99.28%	99.77%

4.2　基于度量的元学习方法分析总结

度量学习方法学习 $\mathcal{F}$，将 $x_i \in \mathcal{X} \subseteq \mathbb{R}^d$ 映射到维度更小的嵌入空间 $e_i \in \mathcal{E} \subseteq \mathbb{R}^p$，通过对比在嵌入空间中查询样本与各个支持集样本的相似度来分类.度量元学习方法通常包含三个组成部分：(1)映射 $f(\cdot)$：将支持集样本 $x^s \in D^s$ 映射到嵌入空间 $\mathcal{E}$；(2)映射 $g(\cdot)$：将查询集样本 $x^q \in D^q$ 映射到嵌入空间 $\mathcal{E}$；(3)度量方法 $s(\cdot,\cdot)$：在 $\mathcal{E}$ 嵌入空间中元素间的度量方式.该种方法假设提取的低维度嵌入能够捕获所有必要的区分性特征，再通过简单的最近邻算法避免在少量实例上过度拟合的情况.

本文所述的基于度量的元学习方法之间的比较如表 2 所示.

表 2 基于度量的元学习方法比较

方法	支持集嵌入	查询集嵌入	距离度量方法	特点	不足
孪生网络	全连接层	全连接层	带权 L1 距离	以 Pair 对的输入方式训练，在单样本问题中优于一般分类器	面对多 shot 情况，比较方式较低效；准确率与目前的度量方法相比差距较大.
匹配网络	CNN，带注意力的 LSTM	CNN，biLSTM	余弦距离	各支持样本的嵌入包含整个支持集的上下文信息	隐式加入支持集样本之间的顺序信息，导致输入的样本顺序影响分类结果；LSTM 的引入使模型较难训练
原型网络	CNN，四层卷积结构	CNN，四层卷积结构	L2 距离	同类嵌入向量的均值作为该类的原型点，只需要单次比较	仅使用支持集中每一类的均值信息，没有利用支持集上下文的更多信息
关系网络	CNN，四层卷积结构	CNN，四层卷积结构	可学习的 CNN 网络	使用可学习的度量函数，提升度量方法的精度和通用性	同时训练特征提取部分和度量函数部分，增加训练难度
TADAM	CNN，残差网络	CNN，残差网络	缩放的 L2 距离	使用缩放的距离尺度，减小各度量方法间的差距；使用支持集合的均值作为任务表示，动态调控特征提取器；辅助任务协同训练	任务的表示过于简单；特征提取器和任务表示两者耦合度太高，特征提取器决定任务表示的准确性，任务表示决定特征提取器的少样本分类性能

孪生网络结构为两个对称的神经网络，共享参数，即映射 $f(\cdot)=g(\cdot)$，结构为 CNN. 度量方法 $s(\cdot,\cdot)$采用带权重的 L1 距离. 孪生网络适用于 FSL 和 OSL，孪生网络在 Meta-train 阶段输入为同一类的一对图片，经过 Prior tasks 训练后，网络使得同一类图片经过映射 $f(\cdot)$后的低维嵌入差异不会太大. 在 Meta-test 阶段，通过将查询集样本与所有支持集样本一一配对输入网络，相似度最大的支持集样本的标签即为预测类. 但由于依靠成对匹配进行分类预测，相比于直接对样本分类显得低效.

匹配网络的映射 $f(\cdot)$由 biLSTM 构成，$g(\cdot)$由 CNN 和带注意力机制的 LSTM 构成，即 $f(\cdot)\neq g(\cdot)$. 度量方法 $s(\cdot,\cdot)$采用 Cosine 相似度. 匹配网络适用于 FSL 和 OSL，通过注意力机制，可以综合考虑到查询集样本和支持集中所有样本的相似性，并且 LSTM 作为特征提取可以引入支持集样本间的信息，但 $f(\cdot)$采用双向 LSTM 会隐式加入支持集样本之间的顺序信息，考虑到梯度消失等存在的问题，支持集样本间输入位置越接近，相互影响越大. 即样本顺序会很大程度上影响模型性能.

原型网络的映射 $f(\cdot)$和 $g(\cdot)$为相同的 CNN 结构，$f(\cdot)=g(\cdot)$，度量方法 $s(\cdot,\cdot)$采用欧式距离. 匹配网络适用于 FSL、OSL 和 ZSL，模型将相同类的支持集样本映射得到的低维嵌入向量的均值作为该类的类原型点，查询集样本只需与每个类原型点进行一次比较，而不像孪生网络存在多次比较的情况. 同时由于欧式距离是 Bregman 散度中的一种，相比采用余弦距离，使用欧式距离计算样例与原型点之间的距离可以一定程度上优化模型. 但是原型网络只利用了同类嵌入向量的均值信息而丢弃了方差信息.

关系网络的映射 $f(\cdot)$和 $g(\cdot)$为相同的 CNN 结构，$f(\cdot)=g(\cdot)$，而度量方法 $s(\cdot,\cdot)$采用 CNN 加全连接的结构，即由一个自学习的神经网络构成. 关系网络适用于 FSL、OSL 和 ZSL，将经过 $f(\cdot)$映射后的查询集样本与所有支持集样本一一拼接后输入 $s(\cdot,\cdot)$得到 k-way 分类的 k 个相似度值，值最大的对应支持集的标签即为预测类. 相比于预先指定的度量方式，自学习得到的度量函数会提升度量的精度和适应能力，但同时学习最优的嵌入函数和深度非线性距离度量函数会加大训练难度，反而降低模型性能.

TADAM 的 $f(\cdot)$和 $g(\cdot)$为相同的 CNN 结构，$f(\cdot)=g(\cdot)$，通过使用缩放系数 α 使得欧式距离和 Cosine 相似度两种不同度量方法间的差距缩小 14%. 相较之前的 4 层浅层结构，特征提取器选用 12 层的残差网络，提高特征提取能力. 同时将每个任务支持集嵌入的均值作为任务表示. 对于不同的任务，使用任务表示生成 Task-Specific 特征提取器，让每个任务都有自己不同的对比特征集合，进一步提高了泛化能力.

从上述各模型的发展和改进过程可以看出，度量方法通过先验的任务训练得到有判别力的视觉特征，再将这些知识用于目标任务. 模型的改进大体分为两部分：特征提取部分和度量部分.

首先对于特征提取部分，基本使用 CNN 卷积结构，匹配网络尝试使用 LSTM 来综合支持集间的信息，但 LSTM 引入的顺序信息对图像分类任务具有负面影响，LSTM 又较难训练，导致与结构简单的原型网络相比，匹配网络并没有取得较好的表现. 网络结构上，度量模型逐渐使用更深、更新的网络（如 Residual Network，Dense Network）来代替浅层网络，提高特征提取能力. 网络结构复杂的同时不免会增加训练难度，辅助预训练的提出很好地解决了

这一问题，不仅加快了收敛速度，实验证实还有利于提升少样本任务训练的最终性能. 预训练通过在训练集的所有基类上执行分类训练得到预训练模型，再将该模型移植到少样本任务上进行元学习. 除了主干结构和辅助训练，特征提取的方式也是提升少样本分类的关键. 关系网络，原型网络都假设有一个共同的嵌入空间，这意味着对于任何未见的新类，在基类上发现的区分性视觉特征都同样有效. 直观上，固定的比较特征集是存在局限的，从先验任务集上获得的比较特征可能并不适用于新任务，对某一任务适用的特征可能对其它任务而言是无关的，甚至是干扰性的. 为此，TADAM 使用支持集的均值动态调控特征提取器，选择适合的特征集，提高了模型对于未见任务的适应性.

综上所述，未来对于度量模型特征提取部分的改进大致可以分为三部分：主干结构，辅助训练，特征提取方式. 主干结构方面，可以设计更优的适合少样本任务的网络，或者直接选用最新的深度网络结构. 辅助训练方面，目前大多数论文采用相同的预训练方法，通过实验也证明了有效性. 尝试分析预训练为何有效对于提升少样本分类可能具有重要意义. 在特征提取方式上，如何利用每个任务的支持集上下文，并且用该上下文调整提取的特征是重要的研究点.

对于度量部分，实验和数学分析已证实无参度量方法可以通过缩放来提升小样本分类的性能，这意味着可能存在更多的传统度量方法上的待改进点；传统的人为设计的度量方式在某些复杂的情形下会失效，关系网络提出的可学习度量解决了这个问题. 对于可学习度量的提升，未来可以考虑针对度量的输入，网络的结构和如何训练做改进.

另一方面，预训练的使用往往能提高少样本分类的性能，在越来越多新的元学习模型被提出时，简单的基线可能被忽视了. Chen 等人[70] 由此提出了 Classifier-Baseline 和 Mate-Baseline 两个基准模型. Classifier-Baseline 在所有的基类（训练集的所有类）上使用交叉熵损失预训练分类器，然后移除最后一层全连接层，得到特征提取器 f_θ，之后使用余弦相似度直接执行少样本分类任务. Mate-Baseline 包含两个训练阶段，第一阶段是预训练阶段，与训练 Classifier-Baseline 方法相同. 第二阶段是元学习阶段，我们根据少样本任务的评估算法继续优化模型. 通过实验观察，采用 ResNet-12 主干网络的 Classifier-Baseline 的性能与当前许多最先进的元学习方法相当，甚至超过了它们，尽管模型没有针对元学习的小样本任务进行优化，也没有进行额外的微调. 在 mini-ImageNet 数据集上，Classifier-Baseline 在 5-way 1-shot 实验中达到 58.91%，5-way 5-shot 实验达到 77.76%；Meta-Baseline 在 5-way 1-shot 实验中达到 63.17%，5-way 5-shot 达到 79.26%. 除预训练外，数据集的规模，主干模型的大小都很大程度上影响着分类结果，以后的工作中应该注意这些因素以进行较公平的模型比较.

4.3　基于初始化的元学习方法分析总结

该类元学习算法以 MAML 为代表，其主要思想是将基于少步数梯度更新得到的学习效果作为损失函数进行优化，训练过程由内循环与外循环组成，其中内循环负责模拟小样本学习中模型学习新任务的场景，外循环则负责收集内循环学习后的效果，通常衡量标准为分类准确率. 通过双层循环的设计，使得模型不再拟合至任何训练种类数据，而是在所有参与训练的数据内部寻找其最具共通性的知识.

基于强泛化性初始化参数的模型大部分与传统深度学习模型类似，学习目标为网络参数的初始值，因此可按逻辑将网络参数分为特征提取器以及分类层，这使得以 MAML 为代表的该类模型可整合常用的主流网路模型，如随着特征提取技术的不断发展，MTL 模型便将 MAML 中用于特征提取的四层卷积层替换为残差网络，取得了较为优秀的提升. 对主流网络结构的兼容使得 MAML 类模型具有良好的通用性.

由于 MAML 类模型内层循环模型较为固定，因此在数据预处理、特征提取以及训练任务的安排上有比较多的发展空间，以近期实验表现较好的 LEO 模型与 MTL 为例，均在不修改模型整体结构的前提下，引入优秀的算法结构，如更优秀的特征提取器，测试阶段更少的训练变量，或对网络参数进行隐空间映射.

由实验结果对比表，在 MAML 的后续模型于 MAML 比较中，可知当元学习模型被应用至图像分类领域时，特征提取的效果对最终实验结果影响较大，而以 MTL 和 LEO 为代表的模型则证明了在测试阶段，即通过少步数学习新任务时，模型网络中需要被更新的参数越少，则对新任务学习能力越强，但如何在尽可能少更新参数前提下有效对网络整体进行修改更新依然是待解决的问题，目前已有的方法有对分类层部分参数进行降维处理，在维度更低的隐空间上进行更新，并通过非线性映射还原至可应

用至数据分类的网络参数本身;另一有效方法是MTL所采用的通过引入迁移学习思想,通过给卷积层添加缩放与偏移系数,实现更新少量数据以更新整个模型的需求.

目前在此方向,特征提取以及训练退化问题正逐渐被解决,训练主要瓶颈来自内层循环的梯度下降学习机制,内层循环的梯度下降严重依赖学习率的设置,目前依然需要大量人工干预,与其它元学习方向对比,此方向旨在避免网络拟合至任意数据种类,且Reptile[49]算法也证明MAML类模型对学习效果的提升并非来自其二次微分计算机制,而是通过双循环避免直接学习训练数据的特殊结构.

4.4 基于贝叶斯元学习方法分析总结

机器学习的训练过程是从已经观察的数据中学习一些模式和假设,并对未观察到的数据进行推断,但是由于输入数据的噪声、测量误差、参数设置等诸多因素,所以模型对预测有不确定性,因此从概率的角度来说,点估计作为权重来建立的任何分类都应该是不合理的,MAML等方法恰恰是采用点估计的方法,因此无法正确估计训练数据中的不确定性,而把贝叶斯理论引入元学习则可以把原来的点估计的变成估计一个后验分布,并通过对分布的采样来进行推断,从而解决了神经网络缺少预测中的不确定性的问题.对元学习来说重点的问题是先验信息如何运用,然而先验信息同样存在不确定性,因此将贝叶斯理论引入元学习浑然天成.

本文仅从参数初始化的角度介绍了这两个典型贝叶斯元学习方法MLAP和VERSA.因为这两个方法采用了贝叶斯的两大推理方法,较为典型.MLAP采用的是MCMC方法,用采样的方法来近似后验分布,通过引入PAC可计算理论来评估分布的误差.而VERSA采用一个简单的分布来近似实际分布.在mini-ImageNet数据集上,VERSA在5-way 1-shot任务中达到53.8%,在5-way 5-shot任务中达到67.37%.通过比较两者的KL散度来评价近似准确度.该方法是一种确定近似方法.目前在贝叶斯元学习方法中,因为确定性近似即变分贝叶斯方法相对计算开销较小准确度更加优秀,所以变分方法在元学习中应用更加广泛.本文因篇幅仅介绍了基于梯度的两种方法,其它例如变分自编码器(VAE)等贝叶斯方法也在元学习领域取得了不错的效果.

基于贝叶斯的元学习方法在继承了贝叶斯方法的同时也继承了其缺点,例如时空复杂度,准确率,参数的快速自适应,如何找到一个更好的分布函数的逼近方法都是贝叶斯元学习的问题,同时也是其接下来重要的研究方向.其次如何把贝叶斯元学习方法应用到实际问题和具体的环境中也是重要的研究课题之一.

4.5 基于外部记忆的模型分析总结

基于记忆的元学习模型将任务中有用的信息编码储存到外部记忆中,以便在下次遇到新任务时检索相关信息.基于记忆的元学习模型一般包括记忆控制器和外部存储记忆.记忆控制器用于将记忆写入外部存储和在外部存储中检索相关信息.

MANN模型使用LSTM作为记忆控制器,将输入样本x_t映射成键值k_t,利用键值k_t检索外部记忆中的相关信息和将样本信息写入外部记忆.在信息写入时,使用最近最少使用访问(LRUA)方法,将新信息写入最近最少使用的位置用于下次检索或者写入最新使用的位置用于更新记忆信息.

APL模型使用一个卷积网络作为编码器来为输入样本生成低维表示,记忆控制器使用一个基于"surprise"的机制来尽可能最小化写入外部记忆的数据点信息,定义一个代表"surprise"阈值的超参数来决定是否写入外部记忆.APL模型在检索记忆时与MANN不同,返回的是k个最近邻信息.最后将输入样本表示和k个最近邻信息一起输入到解码器得到概率向量.APL由于不需要学习写什么,避免了MANN通过记忆进行反向传播的代价,这使得训练更容易设置并且更快.这种设计也最小化了存储数据,使得方法更有内存效率.

MM-Net模型引入了支持集,即每一个批次每类取一个或少数几个有标签的图像组成支持集.为了更好地将网络推广到数据很少的新类别,模型构造了一个上下文学习器,它利用记忆插槽来预测未标记的图像的CNNs的参数.通过一个新的记忆模块的设计,集合中跨类别的整体知识在上下文上增强了支持集中图像的特征嵌入.

目前在基于记忆的元学习研究中,研究难点主要在:将哪些信息存储在外部存储中,如何快速在外部记忆中寻找相关信息,如何防止因外部记忆过大导致训练困难和减少硬编码策略,提高模型的自适应性.

4.6 基于数据增强的元学习方法分析总结

将先验知识合理利用到新任务上,是元学习的

核心目标. 而新任务往往缺乏大量的有效数据，使得传统的深度学习方法效果欠佳. 因此，通过生成数据来解决元学习中数据稀少的问题是一类显而易见的方法. 随着越来越多优秀的 GAN 和 VAE 模型产生，以及数据增强方法适用于大部分元学习模型的优良特性，使用先验知识来生成额外数据的生成模型也成为了元学习的主要研究方向之一. 根据生成数据的方式及其在训练阶段中的作用，我们将其分为基于任务的生成、基于类的生成、基于样本的生成三个部分.

基于任务的生成在任务级别上生成数据，对每一个元学习任务，会生成大量基于当前任务的额外数据. 这些生成数据不含标签信息，在训练过程作为虚假的数据. 元学习器通过鉴别真实数据与生成数据，能够提高最终的模型效果.

MetaGAN[66]将 GAN 和元学习模型结合，找到了一个更紧致的决策边界. 具体的，元学习器中的分类模块亦是 GAN 中的判别器，通过增加一个额外的输出维度来识别生成的虚假图片. 即在 N-way 的分类任务中，输出维度是 $N+1$. $y\leq N$ 表示真实图片的标签，$y=N+1$ 表示生成图片的标签. GAN 中的生成器，将经过特征提取后任务的特征向量与随机噪声 z 组合，生成当前任务下若干对应图片. 然而，该方法并未真正意义上生成“真实”数据，没有解决小样本任务下真实数据缺乏的问题.

基于类的生成是一种 class-to-class 的生成方式，其完成的是从基类到目标类的整体转换. 对新任务 $S=\{S_{\text{train}},S_{\text{test}}\}$ 上每一个 few-shot 类，通过对应类别上的数据，来生成相应带有标签的数据 S^{G}_{train}. 生成的 数据和原有数据一同作为真实数据 $S^{aug}_{\text{train}}=S_{\text{train}}\bigcup S^{G}_{\text{train}}$，放入分类器中训练.

Feature Hallucination[67]（FH）模拟了人类所拥有的“类比”能力. 模型认为在同一类型的内部，其不同样本之间的主要差异在于类型存在的客观环境不同，每对样本间存在的映射关系代表了环境的映射，并且这种环境的转换可以应用到其它相应类型中. 任意两例 z_1 和 z_2 属于同一类别，然后给出一个新类别示例 x_1，模型对 x_1 进行了 z_1 到 z_2 的转换，即完成了类比 $z_1\sim z_2=x_1:?$，这使得模型可以生成不同环境下的新样本. 其生成结构的输入是一个三元组，与原始数据集有较大差异，因此首先要构建用于生成的数据集，造成了训练之外的额外负担. 另外，如何挑选合适的样本对是影响模型效果的重要因素，那些环境转换更为合理的样本对无疑能提高模型的效果. 最后，新数据集中挑选的图片对是可数的，一定程度上限制了生成图片的多样性.

CPAAN[68]是相似类到目标类的生成方式. 该方法先用原型网络[16]结构计算每类图片的原型点，将原型点间的距离作为类与类间的相似度. 对每一个目标类，找到相似度最高的 k 个类，并将相似度作为权值，从这 k 类中挑选图片进行生成. 然而，这种转换是图片中对应的物体间的转换，并没有考虑到客观环境因素的影响. 此外，该方法受到相似类的挑选结果的影响，若相似类选择不恰当又或者数据集太小无法找到其它较为相似的类，会严重影响模型结果.

基于样本的生成不需要借助其它类的数据，如常用的旋转、翻转、裁剪等一样，只用当前目标类中的图片来生成对应标签数据. 生成图片亦是作为真实数据，和原有数据一同进行训练.

Imaginary data[69]为原始图片增加随机噪声 z，通过生成器 G 得到目标样本. 在生成的过程中，可以逐渐将光照、拍照姿势、周围环境等场景信息迁移到生成样本上，从而产生不同变化的图片. 然而，目标类的样本数量较少，也即这些样本中包含的场景信息不足，尽管模型中存在随机噪声，最后生成数据的模式仍是相对单一的.

基于任务的生成，其生成数据当作虚假数据处理，因此并不需要特别生成高质量的图片，对生成器的要求也相对较低. 对特定任务的表示是此类方法的关键，如何进行更为合理的表示以及考虑不同任务之间的相关性，是亟待解决的问题. 此外，由于生成模型并未增加 Few-shot 类中的样本数量，元学习缺少真实样本的核心问题没有得到很好的解决，也导致了此类方法取得的提高有限. 基于类的生成、基于样本的生成，将生成数据作为真实数据处理，对图片质量要求较高，因此生成器的优劣成为了主要的影响因素. 对生成器的设计，近年在生成领域取得了巨大成就的 GAN 及其改进模型是一个较好的选择. 基于类的生成中，如何从基础类中选取对应的类型用来转换为目标类及其重要. 而在同类生成中，由于仅使用了目标类中的少量源数据进行生成，比较容易产生模式崩溃现象.

基于数据增强的元学习模型通常包含两个部分：生成结构和分类结构. 生成结构用来生成额外数

据,分类结构利用原始数据和生成数据进行分类.由于模型真正的目的是正确识别小样本任务下的样本,而不是获得高分辨率的图片,所以分类结构的设计是影响模型效果的主要因素.数据增强只是一个辅助工具,需要与 MAML、度量学习、贝叶斯模型这些元学习方法结合才能发挥作用(当然也可以不用元学习方法直接分类,但这样做效果欠佳).因此,此类方法未来的提升空间较大程度地依赖于其它元学习方法的进步.当然,对于一个固定的分类结构,我们仍然希望通过改进生成结构来提升模型效果,这也是此类方法所需要研究的方向.其未来的发展方向,我们还是分为基于任务的生成、基于类的生成、基于样本的生成三个部分来讨论,如表 3 所示.

表 3 基于数据增强的元学习方法比较

方法	生成数据类型	生成数据方式	特点	不足
MetaGAN	当前任务下的虚假数据	基于任务的生成	分类器在识别生成图片时,能找到更紧致的决策边界.	需要先得到基于任务的表示,未解决小样本学习中样本数量少的问题.
Feature Hallucination	经过"类比"迁移环境后的真实数据	基于类的生成	将同一类型中,图片对存在的映射关系迁移到目标类中,并提供了 SGM 特征正则化方案.	训练生成器需要设计额外的数据集,图片对的数量有限,生成模式单一.
CPAAN	由相似类转换为目标类后的真实数据	基于类的生成	对每个目标类,找到对应的相似类,并限定了协方差矩阵的距离	对相似类的要求较高,数据集较小时,可能不存在较为相似的类.
Imaginary data	由目标样本迁移环境后的真实数据	基于样本的生成	对每个目标样本,添加随机噪声后直接由生成器生成数据,实现简单.	对生成器要求较高,目标类自身样本数量少,难以生成多样化图片.

对于基于任务的生成方法,可以用生成结构得到元学习任务更加精确的表示.例如在 TADAM 中,任务的表示是特征空间下样本的均值,相对比较简单,若是预先生成一些当前任务下的数据,然后与原始数据合并后一同计算均值,任务表示会更加合理.另外,可以对存在多个任务的元学习方法提供一个基于任务的权值.例如在 MAML 中,内循环包含多个小样本任务,可以通过生成结构得到每个任务的一个权值,使损失函数的计算更加合理.

对于基于类的生成方法,生成数据作为真实数据处理,因而生成是否精确至关重要,设计更合理的生成器无疑是一个不错的改进方向.例如,在转换过程中,不仅考虑从基础类到目标类的转换,还考虑由目标类的逆向转换,利用对偶的思想改善生成效果.另外,目前的方法要么是迁移了环境因素,要么是实现物体到物体间的变化,如何将二者结合起来,也是一个好的研究方向.

对于基于样本的生成方法,由于其实现简单,其发展方向比较有限.该类方法同样可以改进生成器的设计,另外可以考虑解决生成模式单一的问题.

5 总结与展望

5.1 总结

综上所述,本文对元学习的研究现状进行了总结,一给出了元学习的基本概念,二给出了元学习的研究方法,三给出了目前元学习方法的实验结果和比较分析.除了本文提到的理论研究外,元学习在实际应用上,也有很多惊艳成果,如结合 MAML 思想的单样本学习人脸替换算法[51],以及将元学习强化学习算法用于机器人控制,以达到快速适应机体以及新地形环境[50].

5.2 展望

元学习尽管取得了一些成绩,但还缺乏系统理论知识,还有许多问题有待进一步研究.如元学习的自适应性、进化性、可解释性、连续性、可扩展性等问题.

5.2.1 元学习自适应性

元学习算法的自适应性是未来研究的主要热点之一,自适应性要求元学习算法在面对不同的任务时都能有较好的表现.本文介绍的大多模型都旨在解决这个自适应性的问题,目前在这一问题上也取得了较大的成果.

未来在这一方面的研究难点在于如何用更简单、更快速的方式调整基础模型来适应各个不同的任务,甚至实现跨领域任务的适应.解决这些问题的关键在于引入任务的可度量属性,这其中又包括三个子问题:如何表示任务信息,如何获取任务信息,如何度量不同任务.

5.2.2 元学习进化性

进化算法(Evolutionary Algorithm,EA)是基于自然选择和自然遗传等生物进化机制的一种全局搜索和优化算法,不同于普通搜索算法,进化算法具有以下优点:(1)非线性,不需要函数梯度信息,也

不需要函数的连续性；(2) 全局寻优；(3) 并行性，从多个点开始寻优，容易获得最优解. 这些特点可以解决深度学习中遇到的问题.

同时进化算法和深度学习类似，也能够从可用的计算和大数据中得到提升. 然而，它解决了一个截然不同的需求：深度学习侧重于建模我们已知的知识，而进化算法则专注于创建新的知识. 从这个意义上讲，对特定的新任务，来创建对应的新知识正是元学习的目标. 因此，将进化算法与深度学习结合，是元学习未来的研究方向之一.

5.2.3　元学习可解释性

广义上的可解释性是指在学习或者解决一个问题时，可以获得所需要的足够的可以理解的信息. 如果我们无法获得足够的信息，那么这个学习或者问题是不可解释的. 元学习通过利用先验信息可以把部分不可解释的问题转化为可解释问题，研究元学习，是研究深度学习可解释性自然的方法.

2019 年图灵奖的获得者 Bengio 在 2019 年初新发表的文章“A Meta-Transfer Objective for Learning to Disentangle Causal Mechanisms”[71] 中提出了一种基于学习器适应稀疏分布变换速度的元学习因果结构. 这篇文章阐述了如何确定两个观察到的变量之间的因果关系. 而且还证明了因果结构可以通过连续变量和端到端的学习进行参数化. 这篇文章就是未来从可解释性方面研究的典型.

其次本文在贝叶斯学习中提到的 MLAP 模型的文章中作者结合 PAC-Bayes 即可计算理论对元学习也作出了相应的解释研究. 综上所述元学习的可解释性问题上取得了一定的成果但仍然只是初级阶段还有很多值得研究的方向.

5.2.4　元学习连续性

元学习的目的之一便是通过掌握一定具有共通性的基础知识，实现对新任务的快速学习，在对新任务的持续学习过程中，可将待学习数据看作已学习数据内容的延续或其子集，即新接触任务均由已有任务数据组合而成，如何有效利用已有数据本身而非通过已有数据学习的网络参数加强对新任务的学习能力是可提升元学习能力的自然思路.

5.2.5　元学习可扩展性

可拓展性实际上是和并行算法以及并行计算机体系结构放在一起讨论的. 元学习算法在某个机器上的可拓展性反映该算法是否能有效利用不断增加的 CPU. 研究元学习算法可扩展性的目的就是要使算法尽可能的利用最多的处理器，并且我们也可以预测当元学习算法移植到大规模处理机上后的运行效果(即问题规模扩大时对处理器的利用情况). 元学习未来在这一方面的研究重点在于如何有效利用更多的计算资源来解决更大规模的问题.

参　考　文　献

[1] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning. Nature, 2015, 518(7540): 529-533

[2] Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of go without human knowledge. Nature, 2017, 550(7676): 354-359

[3] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks. Communications of the ACM, 2017, 60(6): 84-90

[4] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA, 2016: 770-778

[5] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 1409.1556, 2014

[6] Lake B, Salakhutdinov R, Gross J, et al. One shot learning of simple visual concepts. Proceedings of the Annual Meeting of the Cognitive Science Society, 2011, 33(33): 2568-2573

[7] Li Fei-Fei, Fergus R, Perona P. One-shot learning of object categories. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4): 594-611

[8] Lampert C H, Nickisch H, Harmeling S. Attribute-based classification for zero-shot visual object categorization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 36(3): 453-465

[9] Edwards H, Storkey A. Towards a neural statistician. arXiv preprint arXiv:1606.02185, 2016

[10] Vinyals O, Blundell C, Lillicrap T, et al. Matching networks for one shot learning//Proceedings of the Advances in Neural Information Processing Systems. Barcelona, Spain, 2016: 3630-3638

[11] Santoro A, Bartunov S, Botvinick M, et al. Meta-learning with memory-augmented neural networks//Proceedings of the International Conference on Machine Learning. New York, USA, 2016: 1842-1850

[12] Kaiser Ł, Nachum O, Roy A, et al. Learning to remember rare events. arXiv preprint arXiv:1703.03129, 2017

[13] Koch G, Zemel R, Salakhutdinov R. Siamese neural networks for one-shot image recognition//ICML Deep Learning Workshop. Lille, France, 2015, 2

[14] Finn C, Abbeel P, Levine S. Model-agnostic meta-learning for fast adaptation of deep networks. arXiv preprint arXiv: 1703.03400, 2017

[15] Munkhdalai T, Yu H. Meta networks. Proceedings of Machine Learning Research, 2017, 70: 2554

[16] Snell J, Swersky K, Zemel R. Prototypical networks for few-shot learning//Proceedings of the Advances in Neural Information Processing Systems. Long Beach, USA, 2017: 4077-4087

[17] Ravi S, Larochelle H. Optimization as a model for few-shot learning//Proceedings of the International Conferenceon Learning Representations. Toulon, France, 2017

[18] Frome A, Corrado G S, Shlens J, et al. Devise: A deep visual-semantic embedding model//Proceedings of the Advances in Neural Information Processing Systems. Lake Tahoe, USA, 2013: 2121-2129

[19] Akata Z, Reed S, Walter D, et al. Evaluation of output embeddings for fine-grained image classification//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA, 2015: 2927-2936

[20] Zhang L, Xiang T, Gong S. Learning a deep embedding model for zero-shot learning//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 2021-2030

[21] Lei Ba J, Swersky K, Fidler S. Predicting deep zero-shot convolutional neural networks using textual descryiptions//Proceedings of the IEEE International Conference on Computer Vision. Santiago, Chile, 2015: 4247-4255

[22] Romera-Paredes B, Torr P. An embarrassingly simple approach to zero-shot learning//Proceedings of the International Conference on Machine Learning. Lille, France, 2015: 2152-2161

[23] Glass G V. Primary, secondary, and meta-analysis of research. Educational Researcher, 1976, 5(10): 3-8

[24] Powell G. A meta-analysis of the effects of "Imposed" and "Induced" imagery upon word recall//National Reading Conference (30th). San Diego, USA, 1980: 3-6

[25] Maudsley D B. A theory of meta-learning and principles of facilitation: An organismic perspective [M]//Mobile and Ubiquitous System. Computing, Networking, and Services. Thesis (Ed. D.). University of Toronto, 1979: 147-520

[26] Biggs J B. The role of metalearning in study processes. British Journal of Educational Psychology, 1985, 55(3): 185-212

[27] Adey P, Shayer M. Strategies for meta-learning in physics. Physics Education, 1988, 23(2): 97

[28] VanLehn K. Conceptual and meta learning during coached problem solving//Proceedings of the International Conference on Intelligent Tutoring Systems. Montreal, Canada, 1996: 29-47

[29] Zhang Qing-Lin, Wang Yong-Ming. Meta-learning ability and its cultivation. Chinese Education Journal, 1996, (3): 34-37(in Chinese)
(张庆林, 王永明. 元学习能力及其培养. 中国教育学刊, 1996, (3): 34-37)

[30] Chan P K, Stolfo S J. Experiments on multi-strategy learning by meta-learning//Proceedings of the 2nd International Conference on Information and Knowledge Management. Washington, USA, 1993: 314-323

[31] Chan P, Stolfo S. Meta-learning for multistrategy and parallel learning//Proceedings of the 2nd International Workshop on Multistrategy Learning. Washington, USA, 1993: 1

[32] Chan P, Stolfo S. Scaling learning by meta-learning over disjoint and partially replicated data//Proceedings of the 9th Florida AI Research Symposium. Florida, USA, 1996: 151-155

[33] Bensusan H, Giraud-Carrier C G, Kennedy C J. A higher-order approach to meta-learning//Inductive Logic Programming (ILP), International Conference, Work-in-Progress Reports. London, UK, 2000: 109-117

[34] Vilalta R, Drissi Y. A perspective view and survey of meta-learning. Artificial Intelligence Review, 2002, 18(2): 77-95

[35] Rusu A A, Rao D, Sygnowski J, et al. Meta-learning with latent embedding optimization. arXiv preprint arXiv: 1807.05960, 2018

[36] Lee Y, Choi S. Gradient-based meta-learning with learned layerwise metric and subspace. arXiv preprint arXiv: 1801.05558, 2018

[37] Al-Shedivat M, Bansal T, Burda Y, et al. Continuous adaptation via meta-learning in nonstationary and competitive environments. arXiv preprint arXiv: 1710.03641, 2017

[38] Grant E, Finn C, Levine S, et al. Recasting gradient-based meta-learning as hierarchical bayes. arXiv preprint arXiv: 1801.08930, 2018

[39] Ager S. Omniglot writing systems and languages of the world. Retrieved January, 2008, 27: 2008

[40] Hoffer E, Ailon N. Deep metric learning using triplet network //Proceedings of the International Workshop on Similarity-Based Pattern Recognition. Copenhagen, Denmark, Cham, 2015: 84-92

[41] Melekhov I, Kannala J, Rahtu E. Siamese network features for image matching//Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR). Cancun, Mexico, 2016: 378-383

[42] Bertinetto L, Valmadre J, Henriques J F, et al. Fully-convolutional siamese networks for object tracking//Proceedings of the European Conference on Computer Vision. Amsterdam, The Netherlands, Cham, 2016: 850-865

[43] Wu Y, Wu W, Xing C, et al. Sequential matching network: A new architecture for multi-turn response selection in retrieval-based chatbots. arXiv preprint arXiv: 1612.01627, 2016

[44] Si J, Zhang H, Li C G, et al. Dual attention matching network for context-aware feature sequence based person re-identification//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 5363-5372

[45] Fort S. Gaussian prototypical networks for few-shot learning on omniglot. arXiv preprint arXiv:1708.02735, 2017

[46] Sung F, Yang Y, Zhag L, et al. Learning to compare: Relation network for few-shot learning//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 1199-1208

[47] Oreshkin B, Rodríguez López P, Lacoste A. TADAM: Task dependent adaptive metric for improved few-shot learning. Advances in Neural Information Processing Systems, 2018, 31: 721-731

[48] Yoon J, Kim T, Dia O, et al. Bayesian model-agnostic meta-learning. Advances in Neural Information Processing Systems, 2018, 31: 7332-7342

[49] Nichol A, Achiam J, Schulman J. On first-order meta-learning algorithms. arXiv preprint arXiv:1803.02999, 2018

[50] Nagabandi A, Clavera I, Liu S, et al. Learning to adapt in dynamic, real-world environments through meta-reinforcement learning. arXiv preprint arXiv:1803.11347, 2018

[51] Zakharov E, Shysheya A, Burkov E, et al. Few-shot adversarial learning of realistic neural talking head models// Proceedings of the IEEE International Conference on Computer Vision. Seoul, Korea (South), 2019: 9459-9468

[52] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA, 2017: 4700-4708

[53] Sun Q, Liu Y, Chua T S, et al. Meta-transfer learning for few-shot learning//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA, 2019: 403-412

[54] Silver D, Huang A, Maddison C J, et al. Mastering the game of Go with deep neural networks and tree search. Nature, 2016, 529(7587): 484-489

[55] Li J, Monroe W, Ritter A, et al. Deep reinforcement learning for dialogue generation. arXiv preprint arXiv: 1606.01541, 2016

[56] Sutton R S, Koop A, Silver D. On the role of tracking in stationary environments//Proceedings of the 24th International Conference on Machine Learning. Corvallis, USA, 2007: 871-878

[57] Da Silva B C, Basso E W, Bazzan A L C, et al. Dealing with non-stationary environments using context detection// Proceedings of the 23rd International Conference on Machine Learning. Pittsburgh, USA, 2006: 217-224

[58] Silver D L, Yang Q, Li L. Lifelong machine learning systems: Beyond learning algorithms//Proceedings of the 2013 AAAI. California, USA, 2013: 49-55

[59] Mitchell T, Cohen W, Hruschka E, et al. Never-ending learning. Communications of the ACM, 2018, 61(5): 103-115

[60] Peng P, Yuan Q, Wen Y, et al. Multiagent bidirectionally-coordinated nets for learning to play starcraft combat games. arXiv preprint arXiv:1703.10069, 2017, 2: 2

[61] Amit R, Meir R. Meta-learning by adjusting priors based on extended PAC-Bayes theory//Proceedings of the International Conference on Machine Learning. Stockholmsmassan, Stockholm, Sweden, 2018: 205-214

[62] Gordon J, Bronskill J, Bauer M, et al. Meta-learning probabilistic inference for prediction. arXiv preprint arXiv:1805.09921, 2018

[63] Zintgraf L, Shiarli K, Kurin V, et al. Fast context adaptation via meta-learning//Proceedings of the International Conference on Machine Learning. Long Beach, USA, 2019: 7693-7702

[64] Ramalho T, Garnelo M. Adaptive posterior learning: Few-shot learning with a surprise-based memory module. arXiv preprint arXiv:1902.02527, 2019

[65] Cai Q, Pan Y, Yao T, et al. Memory matching networks for one-shot image recognition//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 4080-4088

[66] Zhang R, Che T, Ghahramani Z, et al. MetaGAN: An adversarial approach to few-shot learning. Advances in Neural Information Processing Systems, 2018, 31: 2365-2374

[67] Hariharan B, Girshick R. Low-shot visual recognition by shrinking and hallucinating features//Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy, 2017: 3018-3027

[68] Gao H, Shou Z, Zareian A, et al. Low-shot learning via covariance-preserving adversarial augmentation networks// Proceeding of the Advances in Neural Information Processing Systems. Montreal, Canada, 2018: 975-985

[69] Wang Y X, Girshick R, Hebert M, et al. Low-shot learning from imaginary data//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 7278-7286

[70] Chen Y, Wang X, Liu Z, et al. A new meta-baseline for few-shot learning. arXiv preprint arXiv:2003.04390, 2020

[71] Bengio Y, Deleu T, Rahaman N, et al. A meta-transfer objective for learning to disentangle causal mechanisms. arXiv preprint arXiv:1901.10912, 2019

LI Fan-Zhang, Ph. D., professor, Ph. D. supervisor. His main research interests include meta learning, Lie group machine learning, dynamic fuzzy logic.

LIU Yang, M. S. candidate. His main research interest is meta learning.

WU Peng-Xiang, M. S. candidate. His main research interest is meta learning.

DONG Fang, M. S. candidate. His main research interest is meta learning.

CAI Qi, M. S. candidate. His main research interest is meta learning.

WANG Zhe, M. S. candidate. His main research interest is meta learning.

Background

Meta-learning is currently a research hotspot that aims to solve the problems of traditional neural network model which requires a large amount of training data and iterative steps, and performs poorly on less sample training data tasks; cannot share the experience gained from training between different tasks; is poor to adapt to new types of tasks, and each time have to start training from scratch. The training and testing process of meta-learning can be analogized to the fact that humans can quickly learn and adapt to new tasks after mastering some basic skills. For example, children can quickly recognize the animal through one photo of an animal, corresponding to Few-shot learning in machine learning; Even without the need for image photographs, humans can learn to recognize new types just by description, corresponding to the Zero-shot Learning in machine learning. The basic knowledge of the world and the cognitive basis of behavior patterns that humans have mastered in early childhood correspond to the concept of "meta" in meta-learning, that is, an initial network with strong generalization performance and a rapid adaptive learning ability for new tasks. The purpose of meta-learning is to design a machine learning model that has a learning characteristic similar to that mentioned above, that is, using a small amount of sample data to quickly learn new concepts or skills, and learn "Learning to learn".

This paper is supported by the Key Special Project of the National Key Research and Development Program "Key Scientific Issues of Transformative Technology" (2018YFA0701700, 2018YFA0701701).

In this paper, first, we clarify the current meta-learning research advances in different directions, summarize the common thoughts and existing problems, then classify and describe the research ideas of meta-learning with corresponding algorithms. Finally, this paper discusses the commonly used datasets and evaluation criteria in meta-learning research, and forecasts the development trends of meta-learning.