**Meta-Learning in Neural Networks: A Survey**

**摘要**—近年来，元学习或学习学习领域的兴趣急剧上升。 与使用固定学习算法从头解决给定任务的传统AI方法相反，元学习旨在根据多次学习事件的经验来改善学习算法本身。 这种范例为解决深度学习的许多传统挑战提供了机会，包括数据和计算瓶颈以及泛化的基本问题。 在这项调查中，我们描述了当代的元学习环境。 我们首先讨论元学习的定义，并将其相对于相关领域（例如转移学习，多任务学习和超参数优化）进行定位。 然后，我们提出了一种新的分类法，该分类法为当今的元学习方法提供了更为全面的细分。 我们调查了元学习的有希望的应用程序和成功案例，包括小样本学习，强化学习和架构搜索。 最后，我们讨论了未来研究的突出挑战和有希望的领域。

**1简介**

现代机器学习模型通常使用手工设计的固定学习算法从头开始针对特定任务进行训练。 基于深度学习的方法在各个领域都取得了巨大的成功[1] – [3]。 但是，有明显的局限性[4]。 例如，成功之处主要在于可以收集或模拟大量数据以及可以使用大量计算资源的领域。 这排除了数据本来就很少或昂贵[5]或计算资源不可用[6]，[7]的许多应用程序。

元学习提供了一种替代范式，其中机器学习模型获得了多个学习阶段的经验（通常涵盖了相关任务的分布），并利用这种经验来改善其未来的学习表现。 这种“从学习中学习” [8]可以带来各种好处，例如数据和计算效率，并且可以更好地与人和动物学习[9]保持一致，在学习和学习中，学习策略在生命周期和进化时间尺度上均得到改善 [9] – [11]。 机器学习在历史上基于手工设计的特征来构建模型，而特征选择通常是最终模型性能的决定因素[12] – [14]。 深度学习实现了联合特征和模型学习的希望[15]，[16]，为许多任务[1]，[3]的性能提供了巨大的提高。神经网络中的元学习可以看作是旨在提供集成联合特征，模型和算法学习的下一步。神经网络元学习历史悠久[8]，[17]，[18]。 然而，其作为推动当代深度学习行业前沿的推动者的潜力已导致近期研究的爆炸式增长。 尤其是，元学习有可能减轻当代深度学习的许多主要批评[4]，例如，通过提供更好的数据效率，利用先验知识转移并实现无监督和自我指导的学习。成功的应用已经在小样本图像识别[19]，[20]，无监督学习[21]，数据有效[2​​2]，[23]和自我指导[24]强化学习（RL），超参数优化[25] 和神经体系结构搜索（NAS）[26] – [28]的领域得到证明。

关于元学习的许多不同观点可以在文献中找到。 特别是由于不同的社区对术语的使用略有不同，因此可能很难定义。与我们的观点[29]相关的观点将元学习视为一种处理“免费午餐”定理[30]的工具，并通过搜索最适合给定问题或问题家族的算法（归纳偏差）来提高泛化能力。但是，从广义上讲，该定义可以包括迁移，多任务，特征选择和模型集成学习，而今天通常不将其视为元学习。 元学习的另一种观点[31]广泛地涵盖了基于数据集特征的算法选择和配置技术，并且很难与自动机器学习（AutoML）区别开来[32]。 在本文中，我们专注于当代神经网络元学习。 我们将其表示为按照[29]的算法或归纳偏差搜索，但重点是通过端到端学习明确定义的目标函数（例如交叉熵损失，准确性或速度）来实现此目标。

因此，本文对神经网络元学习的快速增长领域进行了独特，及时和最新的调查。 相反，以前的调查在这个快速发展的领域中已经过时，并且/或者侧重于数据挖掘的算法选择[29]，[31]，[33]-[37]，AutoML [32]或元学习特定应用，如小样本学习[38]或神经体系结构搜索[39]。

我们同时介绍了元学习方法和应用程序。特别地，我们首先提供一个高级的问题形式化，可以用来理解和定位最近的工作。 然后，我们在元表示，元目标和元优化器方面提供了一种新的方法分类法。 我们调查了一些流行的和新兴的应用程序领域，包括安全，增强学习和体系结构搜索； 和相关主题的位置元学习，例如迁移学习，多任务学习和AutoML。 最后，我们讨论了未解决的挑战和未来研究的领域。

**2背景**

元学习已经被定义为各种各样的不一致的方法，即使在当代的神经网络文献中也很难定义。 在本节中，我们介绍我们的定义和关键术语，其目的是对理解大量文献有用。 然后，我们将元学习相对于相关主题定位，例如转移和多任务学习，层次模型，超参数优化，终身/持续学习和AutoML。

元学习最常被理解为学习学习，是指在多个学习情节中改进学习算法的过程。 相比之下，常规ML考虑在多个数据实例上改进模型预测的过程。 在基础学习期间，内部（或下层基础）学习算法可解决由数据集和目标定义的任务，例如图像分类[15]。在元学习期间，外部（或上层元）算法会更新内部学习算法，以使内部算法学习的模型可以改善外部目标。 例如，该目标可能是通用算法的性能或内部算法的学习速度。 基本任务的学习情节（即（基本算法，训练模型，性能）元组）可以看作是提供外部算法所需的实例，以学习基本学习算法。

如上所述，许多传统的机器学习实践，例如通过交叉验证进行的随机超参数搜索，都可能属于元学习的定义之内。当代神经网络元学习的显着特征是明确定义的元级目标，以及针对此目标的内部算法的端到端优化。通常，元学习是对从任务族中采样的学习情节进行的，从而导致基础学习算法经过调整，可以很好地执行从该族中采样的新任务。 当学习新任务时，这可能是一种特别强大的技术，可以提高数据效率。 但是，在有限的情况下，可以从单个任务中采样所有训练情节。 在以下部分中，我们将更正式地介绍这些概念。

**2.1形式化元学习**

常规机器学习在常规监督机器学习中，我们得到了训练数据集，例如（输入图像，输出标签）对。 我们可以通过求解下式来训练参数化的预测模型：



其中L是一个损失函数，用于测量真实标签与预测的标签之间的匹配。我们包括条件 以明确表明此解决方案对因素的依赖，例如的优化器选择或f的函数类选择，我们用表示。然后通过评估带有已知标记的多个测试点来测量泛化。

常规假设是针对每个问题D从头开始执行此优化； 而且那是预先指定的。 但是，“如何学习”的规格的使用会极大地影响概括性，数据效率，计算成本等。元学习通过学习学习算法本身来解决性能提升问题，而不是假设它是预先指定和固定的。 这通常（但并非总是）是通过重新审视上述第一个假设，并从分配的任务中学习而不是从头开始来实现的。

**元学习：任务分配角度** 元学习旨在通过学习“如何学习”来提高性能[8]。特别是，愿景通常是学习通用学习算法，该算法可以在各个任务之间进行概括，并且理想地使每个新任务的学习效果都比上一个更好。因此，指定了 "如何学习"，并且通常用任务分布上的性能来评估。在这里，我们松散地将任务定义为一个数据集和损失函数.因此，学习如何学习就变成了

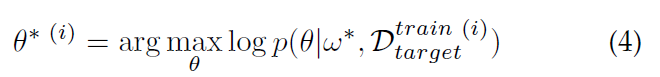


其中L（D ;）衡量使用训练的模型的性能 关于数据集D。知识“学习方法”通常被称为跨任务知识或元知识。

为了在实践中解决此问题，我们通常假定访问从p（T）采样的一组源任务，并从中学习ω。 形式上，我们将元训练阶段中使用的M个源任务表示为Dsource ，其中每个任务都具有训练和验证数据。 通常，源训练和验证数据集分别称为支持集和查询集。 将元知识表示为ω，则“学习方法”的元训练步骤为：



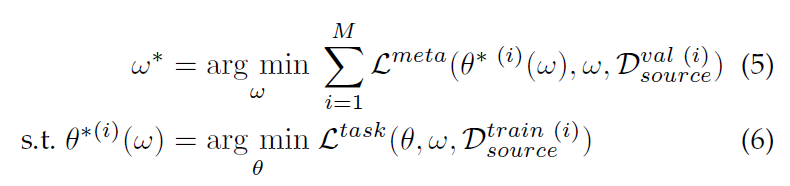
现在，我们将在元测试阶段使用的Q个目标任务表示为，其中每个任务都具有训练和测试数据。 在元测试阶段，我们使用学习到的元知识在每个先前看不见的目标任务i上训练基本模型：



与等式1中的常规学习相反，学习目标任务的训练集i现在受益于元知识ω 关于要使用的算法。 这可以采用估计初始参数的形式[19]，在这种情况下ω 和是指相同数量的相同大小的对象。但是ω 通常可以对其他对象进行编码，例如整个学习模型[40]或优化策略[41]。最后，我们可以通过在每个目标任务的测试拆分上的表现来评估我们的元学习器的准确性。

这种设置导致了传统的欠拟合和过度拟合的类比：元欠拟合和元过度拟合。 特别是，元过度拟合是一个问题，在源问题上学到的元知识无法泛化到目标任务。这是相对常见的，尤其是在只有少量源任务可用的情况下。就作为归纳偏向学习的元学习而言[29]，元过度拟合对应于学习归纳偏向ω 这就将的假设空间过于狭窄地限制在源任务的解决方案周围。

**元学习：双层优化角度** 前面的讨论概述了多任务场景中元学习的常见流程，但没有指定如何解决等式3.中的元训练步骤。这通常是通过将元训练步骤转换为双层优化问题来完成的。虽然可以说这观点仅适用于基于优化器的方法（请参阅第3.1节），但有助于更全面地可视化元学习的机制。双层优化[42]指的是分层优化问题，其中一个优化包含另一种优化作为约束[25]，[43]。 使用此表示法，可以将元训练形式化如下：



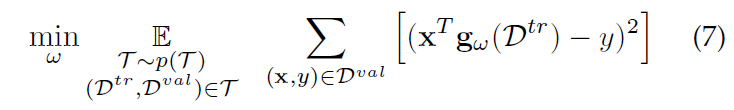
其中Lmeta和Ltask分别指的是外部目标和内部目标，例如在进行小样本分类时的交叉熵。双层范式的关键特征是外层和内层之间的leader-folloωer asymmetry：内层优化等式6以学习策略为条件由外部级别定义，但不能在训练中更改。

这里 ω 可以指示非凸优化[19]的初始条件，诸如正则化强度的超参数[25]甚至是损失函数的参数化以优化Ltask [44]。第4.1节详细地讨论ω的选择空间。 外层优化训练学习策略ω 这样就可以生成训练后在其验证集上表现良好的模型。4.2节详细地讨论了如何优化ω。注意，尽管Lmeta可以测量简单的验证性能，但我们将看到它也可以测量更细微的量，例如学习速度和模型鲁棒性，如第4.3节所述。

最后，我们注意到上述元训练的形式化使用了任务分布的概念，并使用了来自该分布的M个样本。 尽管它功能强大并且在元学习文献中得到了广泛使用，但这并不是元学习的必要条件。更正式地讲，如果我们给定一个训练和测试数据集，则可以拆分训练集以获取验证数据，以使进行元训练，对于元测试，我们可以使用。

我们仍然在几个情节中学习ω，尽管在元训练期间通常会使用不同的train-val分裂，但人们可能会认为M = Q = 1。

**元学习：前馈模型角度** 正如我们将看到的，有许多元学习方法以前馈方式合成模型，而不是像Eqs 5-6中那样通过显式的迭代优化来合成模型。尽管它们的复杂程度各不相同，但通过实例化等式2中定义了用于元训练线性回归的玩具示例[45]的抽象目标可能有助于理解这一系列方法。 。



在这里，我们可以看到我们通过优化任务分配来进行元训练。 对于每个任务，都绘制了训练和验证（又名查询和支持）集。训练集合Dtr被嵌入向量 它定义了线性回归权重，以预测从测试集中得出的示例x。优化上述目标，从而通过训练函数gω“学习如何学习”实例化将训练集映射到权重向量的学习算法。因此，如果从p（T）得出新的元测试任务T te，我们可能还会期望gω 提供良好的解决方案。该系列中的不同方法在所使用的预测模型的复杂性（它们实例化的参数g）以及支持集的嵌入方式（例如，通过简单的合并，CNN或RNN）方面有所不同。

**2.2元学习的历史背景**

元学习最早于1987年在文献中出现在J. Schmidhuber和G. Hinton [17] [46]的两篇独立的著作中。 Schmidhuber [17]为使用自我参照学习可以学习如何学习的新方法系列设定了理论框架。 自参考学习涉及训练神经网络，该神经网络可以接收自己的权重作为输入并预测所述权重的更新。Schmidhuber进一步提出，可以使用进化算法来学习模型本身。Hinton等[46]提出每个神经网络连接使用两个权重，而不是一个。 第一个权重是标准的慢速权重，它在优化程序更新后缓慢地获取知识（称为“慢知识”），而第二个权重或快速权重在推理期间快速地获取知识（称为“快速知识”）。快速权重的职责是能够消除或恢复过去因优化程序更新而被遗忘的慢速权重。 这两篇论文都介绍了一些基本概念，这些概念后来又分支出来，并引发了当代的元学习。

引入元学习后，人们可以看到该想法在多个不同领域的使用迅速增加。 Bengio等。[47]，[48]提出了尝试元学习生物学上合理的学习规则的系统。 Schmidhuber等人在后续工作中继续探索自我指称系统和元学习[49]，[50]。S. Thrun等人在[8]中用术语“学习来学习”作为元学习的替代方法，并继续探索和剖析元学习中的可用文献，以寻求通用的元学习定义。 2001年首次提出了使用梯度下降和反向传播训练元学习系统的建议[51]，[52]。 随后不久，元学习文献有了更多概述[29]。 Schweighofer等人首先将元学习用于工作中的强化学习。 [53]之后，Larochelle等人在零样本学习中首次使用了元学习[54]。 终于在2012年，Thrun等人[8]在深度神经网络的现代时代重新引入了元学习，这标志着本次调查中讨论的现代元学习的开始。

元学习还与用于分组数据的统计中的层次模型和多层模型的方法密切相关。 在这种分层模型中，使用组内模型对分组的数据元素进行建模，并使用组间模型对每个组之间的差异进行建模。 机器学习文献中的此类分层模型的示例包括主题模型，例如Latent Dirichlet Allocation [55]及其变体。 在主题模型中，从文档数据中学习新文档的模型； 该模型的学习以从整个语料库中已经学习到的一组主题为指导。 层次模型将在第2.3节中进一步讨论。

**2.3相关领域**

在这里，我们将元学习定位于相关领域，这在文献中经常引起混乱。

**转移学习（TL）** TL [34]利用源任务的过去经验来改善目标任务的学习（速度，数据效率，准确性）–通过从以前的任务解决方案中转移参数先验，初始条件或特征提取器[56]。 TL将此工作都提到问题区域。 在当代的神经网络环境中，它通常是指参数传递加上可选的微调的特定方法（尽管有许多其他方法可以解决此问题[34]）。

虽然TL可以指问题领域，但是元学习是指可以用来改善TL和其他问题的方法。 TL作为一种方法论与元学习不同，因为先验是通过对源任务进行原始学习而提取的，而无需使用元目标。 在元学习中，相应的先验将由外部优化定义，该外部优化评估先验在帮助学习新任务时的表现，例如，通过MAML [19]所示。 一般而言，元学习处理的元表示形式比单纯的模型参数要广泛得多（4.1节）。

**域自适应（DA）和域通用（DG）**域转移是指源任务和目标任务具有相同类别，但目标任务的输入分布相对于源任务发生了转移的情况[34]，[57]，从而导致转移时模型性能降低。 DA是转移学习的一种变体，它试图通过使用来自目标的稀疏或未标记的数据来适应源训练的模型来缓解此问题。 DG是指训练源模型对这种域偏移具有鲁棒性而无需进一步调整的方法。 已经研究了许多方法[34]，[57]，[58]来转移知识并提高目标领域的性能。 但是，就TL而言，原始DA和DG的区别在于，没有元目标可以优化跨域的“学习方法”。同时，元学习方法可用于执行DA和DG，这将在5.9节中介绍。

**持续学习（CL）** 持续学习和终生学习[59]，[60]是指学习从潜在的非平稳分布中提取的一系列任务的能力，尤其是在加速学习新任务的同时寻求学习的能力并不忘记旧任务。 它与处理任务分发有关，并且目标部分是为了加速对目标任务的学习。 但是，大多数持续学习方法不是元学习方法，因为没有明确解决此元目标。 尽管如此，元学习为继续学习提供了一个潜在的框架，最近的一些研究已经开始通过开发编码持续学习表现的元目标来做到这一点[61]-[63]。

**多任务学习（MTL）** 旨在共同学习几个相关任务，并受益于参数共享带来的效果正则化以及所产生的共享表示形式的多样性[64]-[66]。 像TL，DA和CL一样，常规MTL是没有元目标的单级优化。此外，MTL的目标是解决固定数量的已知任务，而元学习的重点通常是解决看不见的未来任务。 但是，可以引入金属收益以使MTL受益，例如 通过学习任务之间的相关性[67]，或如何在多个任务之间确定优先级[68]。

**超参数优化（HO）**在元学习的范围之内，因为“学习方法”的定义中可以包含诸如学习率或正则化强度之类的超参数。 在这里，我们专注于HO任务，该任务定义了使用神经网络进行端到端训练的元目标。 这包括HO中的一些工作，例如基于梯度的超参数学习[67]和神经体系结构搜索[26]。 但是我们排除了其他方法，例如随机搜索[69]和贝叶斯超参数优化[70]，这些方法很少被视为具有元学习。

分层贝叶斯模型（HBM）涉及先验下参数的贝叶斯学习。 先验是在其他变量上作为条件密度写的ω它有自己的先验p（ω）。 分层贝叶斯模型的主要特点是分组数据模型，每个组中i都有自己的

完整模型为。 层次结构可以进一步提高； 特别是 ω 本身可以被参数化，因此可以学习p（ω）。

学习通常是全流水线的，但是使用某种形式的贝叶斯边缘化来计算ω的后验：。 进行边缘化的难易程度取决于模型：在某些情况下（例如Latent Dirichlet分配[55]），由于选择了共轭指数模型，因此边缘化是精确的；在其他情况下（例如，在[71]中），随机变异方法是 用于计算近似后验，由此计算出边缘似然的下限。

贝叶斯层次模型为元学习提供了宝贵的观点，因为它们提供了建模而不是算法框架来理解元学习过程。 在实践中，贝叶斯层次模型中的先前工作通常集中于学习简单易处理的模型。 但是，大多数元学习工作都考虑复杂的内循环学习过程，其中涉及许多迭代。 但是，可以通过HBM [72]来理解一些元学习方法，例如MAML [19]。

**AutoML：**AutoML [31]，[32]是相当广泛的保护伞适用于旨在使机器学习过程中通常是手动的部分自动化的方法，例如数据准备和清理，特征选择，算法选择，超参数调整，架构搜索等。AutoML通常在此处定义的元学习范围之外利用很多启发式方法，并处理诸如数据清理之类的任务，这些任务对于元学习而言并不那么重要。 但是，AutoML有时会使用元学习，因为我们是根据元目标的端到端优化在此处定义的，因此元学习可以被视为AutoML的一种专业化。

**3分类法**

**3.1以前的分类法**

先前[73]，[74]的元学习方法分类倾向于在基于优化的方法，基于模型的方法（或黑盒方法）和基于度量的方法（或非参数方法）之间产生三向分类法 。

**优化** 基于优化的方法包括将内部任务（等式6）从字面上解决作为优化问题的方法，并且侧重于提取元知识ω 需要提高优化性能。其中最著名的也许是MAML [19]，这里的元知识ω 是内部优化中模型参数的初始化，即目标是学习，以使少量训练实例上的少量内部步骤产生对验证数据表现良好的分类器。 这也是通过梯度下降执行的，通过对基本模型的更新来区分。 更详尽的替代方法还可以学习步长[75]，[76]或训练递归网络以根据梯度[41]，[77]，[78]来预测步长。 通过梯度进行元优化会带来以下挑战：有效评估昂贵的二阶导数，并通过可能的数千个内部优化步骤的图表进行区分（请参阅第6节）。 由于这个原因，它通常应用于很少的内循环步骤就足够的小样本学习。

**黑盒/基于模型** 在基于模型（或黑盒）的方法中，内部学习步骤（方程6，方程4）被包裹在单个模型的前馈过程中，如方程7所示。模型将当前数据集D嵌入激活状态，并基于此状态对测试数据进行预测。 典型的体系结构包括递归网络[41]，[51]，卷积网络[40]或超网络[79]，[80]，它们嵌入训练实例和给定任务的标签以定义输入测试示例并预测其标签的预测变量。 在这种情况下，所有内部学习都包含在模型的激活状态中，并且完全是前馈。外层学习使用ω 包含CNN，RNN或超网络参数。 外部和内部优化紧密耦合为ω 直接指定。记忆增强神经网络[81]使用显式存储缓冲区，也可以用作基于模型的算法[82]，[83]。 已经观察到，与基于优化的方法相比，基于模型的方法通常不能推广到分布外任务上[84]。 此外，尽管它们通常非常擅长于数据有效的短时学习，但由于尚不清楚黑盒模型能否成功将大型训练集嵌入到丰富的基本模型中而被批评为渐近性较弱[84]。

**度量学习** 度量学习或非参数算法因此在很大程度上仅限于元学习的流行但特定的小样本应用（第5.1.1节）。这个想法是通过简单地将验证点与训练点进行比较并预测匹配的训练点的标签来在内部（任务）级别执行非参数“学习”。 按照时间顺序，这是通过诸如孪生网络[85]，匹配网络[86]，原型网络[20]，关系网络[87]和图神经网络[88]之类的方法实现的。 在这里，外层学习对应于度量学习（查找将数据编码为适合比较的表示的特征提取器ω）。 像之前一样 ω 在源任务上学习，并用于目标任务。

讨论上面讨论的常见故障并未涵盖所有感兴趣的方面，并且不足以理解当今可用的各种元学习框架之间的联系。 因此，在以下各小节中，我们将介绍元学习方法的新的跨领域细分。

**3.2拟议的分类法**

我们沿三个独立的轴引入了新的分类。 对于每个轴，我们提供一个分类法，以反映当前的元学习格局。

**元表示（“什么？”）**第一个轴是元知识ω表示的选择。 在程序归纳的情况下，这可以覆盖用于优化器初始化的模型参数[19]的估计，到可读代码[89]。 请注意，基本模型表示通常是特定于应用程序的，例如，在计算机视觉的情况下，是卷积神经网络（CNN）[1]。

**元优化器（“如何？”）**第二个轴是在元训练期间用于外部级别的优化器的选择（请参见等式5） 。外层优化器ω 可以采取多种形式，从梯度下降[19]到强化学习[89]和进化搜索[23]。

**元目标（“为什么？”）**第三个轴是元学习的目标，它由选择元目标Lmeta（公式5），任务分布p（T）和两个级别之间的数据流确定。 这些可以一起为不同目的自定义元学习，例如有效的小样本采样 [19]，[40]，快速多次优化[89]，[91]或对域移位的鲁棒性[44]，[92]，标签噪音[93]和对抗攻击[94]。

**4调查：方法学**

在本节中，我们将根据建议的新方法学分类法分解现有文献。

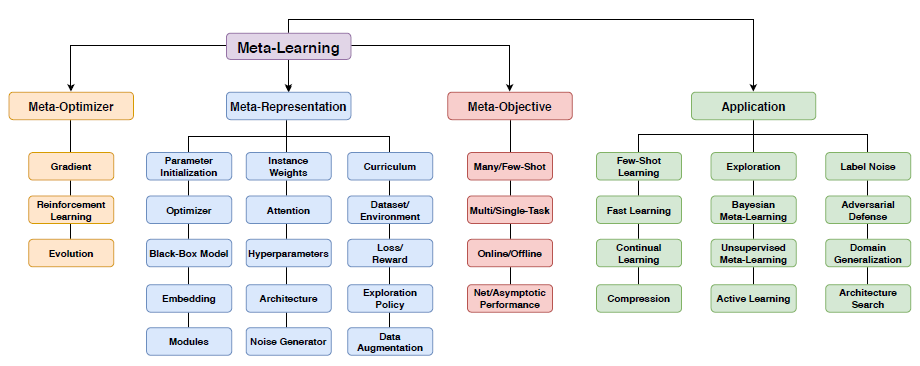


图1.元学习概况，包括算法设计（元优化器，元表示，元目标）和应用程序。

**4.1元表示**

元学习方法对ω应该是什么有不同的选择，即应该学习学习策略的哪些方面; （排除）哪些方面应视为固定的。

**参数初始化** 在这第一个方法系列中ω 对应于神经网络的初始参数。在MAML [19]，[95]，[96]中，这些被解释为内部优化的初始条件。 良好的初始化距离从p（T）提取的任何任务T的解仅几步之遥。 这些方法广泛用于小样本学习，在这种情况下，只要精心选择初始条件，就可以使用很少的示例来学习目标问题而不会过度拟合。 这种方法面临的主要挑战是，外部优化需要解决的参数与内部优化一样多（大型CNN中可能有数亿个参数）。 这导致了将参数子集隔离到元学习上的工作。例如，按子空间[74]，[97]，按层[79]，[97]，[98]或通过分离比例和移位[99]。 尽管内部循环初始化是元表示的一种流行而有效的选择，但这里的一个主要争论是单个初始条件是否足以为各种潜在任务提供快速学习，或者是否仅限于相当窄的分布p（T）。 这导致了可以在多个初始条件下对混合物建模的变量[97]，[100]，[101]。

**优化器** 上述以参数为中心的方法通常依赖现有的优化器（例如带有动量的SGD或Adam [102]）在给定一些新任务时优化初始化。 以优化器为中心的方法[41]，[77]，[78]，[91]而不是依靠手工设计的优化器，而是着重于通过训练将诸如和之类的输入优化状态作为函数的函数来学习内部优化器， 并产生优化步骤，以在每个基础学习迭代中执行。可训练的组件ω 可以将简单的超参数（例如固定步长[75]，[76]）扩展到更复杂的预处理矩阵[103]。最终ω 从定义输入梯度和其他元数据的复杂非线性转换的意义上讲，可以用来定义基于梯度的优化器[41]，[78]，[89]，[91]。如果优化器是跨权重坐标应用的，则此处要学习的参数可能会很少[78]。以初始化为中心的方法和以优化程序为中心的方法可以通过共同学习来合并，即让前者学习后者的初始条件[41]，[75]。优化器学习方法已应用于小样本学习[41]，并已加速和改进了多样本学习[78]，[89]，[91]。最后，人们还可以元学习黑盒零阶优化器[104]，该优化器仅需要评估Ltask而不是诸如梯度之类的优化器状态。 这些已显示[104]与常规贝叶斯优化[70]替代品竞争。

**黑匣子模型（递归，卷积，超网络）** 另一个模型家族训练学习者ω 直接从支持集提供前馈映射到分类测试实例所需的参数，即–而不是依赖于梯度[78]（或零阶[104]）迭代优化的。这些对应于常规分类法中基于黑盒模型的学习（第3.1节）。 嵌入支持集通常是通过递归网络[51]，[105]，[106]或卷积[40]实现的。

这些方法与Hypernetworks有很强的联系。超级网络[107]，[108]是产生以某些嵌入为条件的另一个神经网络的权重的网络，通常用于压缩或多任务学习。 通过对源（aka支持）数据集的嵌入进行条件化，超网络也可以用于合成预测模型[97]，[109]。在这种情况下 ω 是在前馈传递中给定支持集的情况下产生的权重合成超网络。最终，增强记忆的神经网络能够记住旧数据并快速吸收新数据，并且通常也属于黑盒模型类别。 在[82]中，作者通过更改其记忆检索机制使神经图灵机[81]适应元学习设置。然后，元网络[83]通过组合快速权重（由网络按任务预测）和较慢的权重（跨任务使用REINFORCE进行训练）以访问内存来改进此模型。我们注意到，有些方法在单个框架中同时实现了基于模型的条件和以初始条件[97]或​​以优化程序为中心的[98]元学习。

**嵌入特征（度量学习）** 这类方法的灵感来自常规机器学习中的度量学习方法，因此在常规分类法中也进行了归类（第3.1节）。 它们主要应用于小样本学习。 在这里，元优化过程将学习嵌入式网络ω 通过将查询和支持实例[20]，[79]，[86]，[110]之间的简单相似度比较（例如，具有余弦相似度或欧几里得距离）将原始输入转换为适合识别的表示形式。

但是，可以将度量学习方法视为上述前馈黑盒模型的一种特殊情况。对于基于支持和查询图像xs和xq的嵌入之间的内积生成logits的方法，这显然是这种情况[79]，[110]。在此，支持图片会生成权重来解释查询示例，这使其成为BBM的特殊情况，其中“超网络”会为查询集生成线性分类器。通过使嵌入任务条件式[98]，[111]或学习更详尽的比较指标[87]，[88]，可以进一步增强该家族中的原始方法。

损失和辅助任务 与优化器设计的元学习方法类似，这些方法旨在学习内部任务损失为基本模型。损失学习方法通​​常定义一个小型神经网络，该神经网络输入通常是损失输入的量（例如，预测，特征或模型参数），并输出要由内部（任务）优化器视为损失的标量。 这具有潜在的好处，例如，导致学习损失比常用的损失更易于优化（例如，局部最小值较小）[23]，[112]，[113]，从而导致学习速度更快，泛化程度提高[45]， [114]，[115]或其最小值对应于对域移位更健壮的模型[44]。此外，损失学习方法也已用于从未标记的实例中学习[98]，[116]。其他应用程序包括学习作为真实不可微任务损失的可微近似，例如精确召回曲线[117]，[118]下的面积。

自我学习[119]，[120]或辅助任务[121]的泛化也导致损失学习。在这些问题中，无监督的预测任务（例如，使视觉中的像素着色[119]或仅改变RL中的像素[121]）以主要任务的多任务方式进行定义和优化，但目的是改善表示形式 支持主要任务。在这种情况下，很难预先预测最佳的辅助任务（损失），因此可以使用元学习方法根据对辅助任务的改进对主要任务学习的影响来从中选择。即，ω 是每个辅助任务权重[68]。 更一般地，可以元学习辅助任务生成器，该辅助任务生成器使用辅助标签注释示例以预测主要的多任务模型[122]。

**结构** 架构发现一直是神经网络[39]，[123]中的重要领域，并且不适合简单的穷举搜索。 通过学习架构，元学习可用于使这一非常昂贵的过程自动化。早期尝试使用RL和LSTM来学习生成有关好的体系结构的描述[28]。进化算法[27]也用于尝试学习建模为图形的架构中的块，这些块可以通过编辑它们的图进行变异。 还以DARTS的形式访问了基于梯度的体系结构表示[26]，其中训练过程中的前向传递包括给定块中所有可能层的输出之间的softmax，并由要元学习的系数加权（即 ω）。在元测试期间，仅通过使层对应于最高系数来离散化体系结构。 通过交替一个内部步骤和一个外部步骤来更新体系结构系数和网络权重，从而贪婪地学习系数。 由于DARTS仍然相对较慢且精度有限，因此最近的工作集中在通过更好的可微近似[124]，易于学习的初始化[125]或体系结构先验[126]使体系结构学习更加有效。 有关神经体系结构搜索的更多详细信息，请参见第5.4节。

**注意模块** 已经显示出注意机制可以提高泛化性能和可解释性。这样的机制也已经形成了各种元学习模型的元表示的一部分。 例如，它们已被用作基于度量的转导金属学习者的支持和目标设置项目的比较器[127]，以及特征提取器，以防止几次连续学习中的灾难性遗忘[128]。最近，注意力也被用来总结输入文本分类任务的分布[129]。

**模块** 模块化元学习[130]，[131]假定任务不可知知识ω 定义了一组模块，这些模块以定义的任务特定方式重新组成，以解决遇到的每个任务。这些策略可以看作是知识共享的典型结构方法的元学习概括，这些方法在多任务和转移学习中得到了很好的研究[65]，[66]，[132]。

**超参数** 在这些方法中ω 包括基础学习算法的超参数，例如正则化强度[25]，每个参数正则化[92]，多任务学习中的任务相关性[67]或数据清理中的稀疏性[67]。 注意，诸如步长和方向[75]，[76]之类的超参数可以看作是优化器定义的一部分，因此会导致超参数和优化器学习类别之间的重叠。

**数据扩充** 在监督学习中，通常通过在现有数据上保留标签的变换来合成更多的训练数据，从而提高通用性。 数据扩充操作包含在内部问题方程6的优化步骤中,并且通常是手工设计的。 但是，当ω 定义了数据扩充策略，为了最大化验证性能[133]可以通过等式5中的外部优化来学习。由于增强操作通常是不可微的，因此需要增强学习[133]，离散梯度估计器[134]或进化[135]方法。一个尚待解决的问题是，基于强大的GAN的数据增强方法[136]是否可用于内部学习中并在外部学习中进行优化。

**最小批量选择，样本权重和课程学习** 当基本算法是基于最小批量的随机梯度下降时，学习策略的设计参数是批量选择过程。 存在各种手工设计的方法[137]来改进经典随机采样的迷你批次。 微型批次选择的元学习方法定义了ω 作为实例选择概率[138]或选择或排除实例[139]包含在下一个微型批处理中的小型神经网络，而在定义了微型批处理选择器的情况下，元损失可以定义为基本模型的学习进度 。

这样的选择方法还可以提供一种使课程学习自动化的方式。 在传统的机器学习中，课程是要手动设计的数据或概念序列，以产生比以随机顺序学习项目更好的性能[140]，例如，通过专注于正确的困难实例，而拒绝太难或太拒绝的情况 简单（已经学习）的实例。 通过将教学策略定义为元知识并对其进行培训以优化学生的学习进度，元学习有可能使这一过程自动化并选择正确难度的例子。[139]，[141]。

与小批量选择策略相关的是学习每样本损失权重的方法ω 对于训练集[142]，[143]。 可以通过打折嘈杂样本[142]，[143]，打折离群值[67]或校正类别不平衡[142]来在标签噪声下学习。

**数据集，标签和环境** 元表示的最奇怪的选择也许是支持数据集本身。 这不同于我们最初的元学习形式化，后者认为源数据集是固定的（第2.1节，方程2-3）。 但是，可以在等式5-6的双层视图中轻松理解它。如果上部优化中的验证集是真实且固定的，而下部优化中的训练集由ω进行参数化，则可以通过元学习来调整训练数据集以优化验证性能。

在数据集提纯[144]，[145]中，学习了支持图像本身，因此对它们的几个步骤可以对实际查询图像进行很好的概括。 这可用于将大型数据集汇总为少量图像，这对于无法存储流数据集的持续学习重放很有用。

除了学习固定标签y的输入图像x，还可以学习固定图像x的输入标签y。 这可用于半监督学习中，例如，直接学习未标记集的标签以优化验证集性能[146]或训练标签生成功能[147]。

在计算机视觉或强化学习中的sim2real学习[148]的情况下，人们使用环境模拟器来生成训练数据。 在这种情况下，还可以训练图形引擎[149]或模拟器[150]，以便在对由该环境模拟器生成的数据进行训练之后，优化下游模型的实际数据（验证）性能。

**讨论：归纳表示法和方法** 大多数表示法ω 上面讨论的是处理或生成数据的函数的参数向量。

但是，从ω的意义上讲，其中一些表示是可转导的。 字面上对应于数据点[144]，标签[146]或每个样本的权重[142]。 这意味着参数中ω 元学习规模随数据集的大小而变化。 虽然这些方法的成功证明了现代元学习的能力[145]，但该属性最终可能会限制其可扩展性。

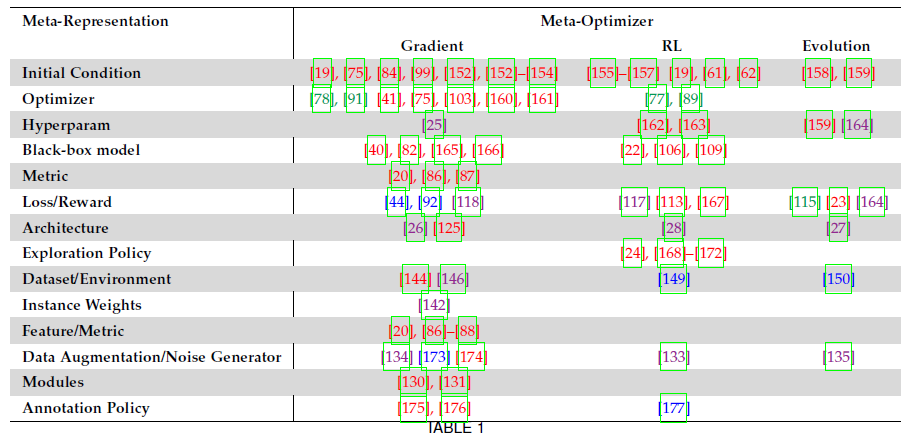
与转导表示形式不同的是，在设计为可对查询实例以及支持实例进行操作的意义上，它们是转导的方法[98]，[122]。

**讨论：可解释的符号表示形式** 可以在上面讨论的许多元表示形式之间进行跨领域的区分是在不可解释的（子符号）表示形式和人类可解释的（符号）表示形式之间。 次符号表示法，例如当ω 参数化神经网络[78]，是最常被研究的方法，它构成了上面引用的大多数研究。但是，也可以使用符号表示进行元学习，哪里ω 代表人类可以理解的符号功能，就像亚当[102]一样，是一段程序代码[89]。而不是神经损失函数[44]，可以训练象征性的损失ω 由与交叉熵相当的表达式定义[115]。人们还可以元学习优于ReLU等标准的新符号激活[151]。 由于这些元表示形式是不平滑的，因此元目标是不可微且难以优化的（请参见第4.2节）。所以上层优化ω 通常使用RL [89]或进化算法[115]。但是，符号表示在跨任务族的泛化能力方面可能具有优势[89]，[115]，[151]。即，在元训练中用带有一个ω来覆盖更宽的分布p（T）。或已学习的ω 在元测试中一般化为分发任务（请参阅第6节）。

**4.2 Meta-Optimizer**

考虑到要优化学习策略的哪个方面（如上所述），元学习器设计的下一个轴是用于调整的实际外部（元）优化策略ω。

**梯度** 许多方法对元参数都使用梯度下降ω [19]，[41]，[44]，[67]。这需要计算外物镜的导数 ，通常通过链式规则连接到模型参数，由于这些方法利用ω的解析梯度，因此可能是最有效的方法。但是，主要挑战包括：（i）通过内部优化使用许多步骤的长计算图有效区分，例如通过精心设计自动区分算法[25]，[178]和隐式区分[145]，[153]， [179]，并轻松处理所需的二阶梯度[180]。（ii）减少不可避免的梯度退化问题，该问题的严重性随着内环优化步骤的数量而增加。（iii）当基础学习器ω或Ltask包含离散或其他不可微分运算时，计算梯度。



表一 根据我们的分类法对研究论文进行分类。 我们使用颜色来表示显着的元目标或应用目标重点。 为了简单起见，我们专注于每篇论文的主要目标。 颜色代码如下：样本效率（红色），学习速度（绿色），渐近性能（紫色），跨域（蓝色）。

**强化学习** 当基础学习者包括不可微分的步骤[133]或元目标Lmeta本身不可微分[117]时，许多方法[22]会采用RL来优化外部目标方程5。这通常使用策略梯度定理来估计梯度。但是，以这种方式减轻对可区分性的要求通常非常昂贵。的高策略梯度估计意味着需要收敛许多外部级别的优化步骤，并且由于其中封装了任务模型优化，因此这些步骤中的每个步骤本身都是昂贵的。

**进化** 优化元目标的另一种方法是进化算法（EA）[17]，[123]，[181]。 许多进化算法与强化学习算法有很强的联系[182]。 但是，它们的性能并不取决于RL的内部优化的长度和奖励稀疏性。

EA具有吸引力的原因有几个[181]：（i）EA可以优化任何类型的基本模型和元目标，而无需差异性。 （ii）它们不依赖于反向传播，这既解决了梯度下降问题，又避免了上述常规基于梯度的方法所需的高阶梯度计算成本。 （iii）它们是高度可并行化的，从而使元训练更易于扩展。 （iv）通过保持解决方案的多样性，他们可以避免困扰梯度法的局部最小值[123]。 但是，它们具有许多缺点：（i）训练模型所需的总体规模随着可学习参数的数量而迅速增加。 （ii）它们可能对突变策略（例如，噪声的大小和方向）敏感，因此可能需要仔细进行超参数优化。（iii）它们的拟合能力通常不如基于梯度的方法，尤其是对于大型模型，例如CNN。

EA相对更普遍地应用于RL应用程序[23]，[158]（其中模型通常较小，内部优化又长又不可微）。 然而，它们也已被应用于在监督学习中学习学习规则[183]​​，优化器[184]，体系结构[27]，[123]和数据增强策略[135]。它们在学习人类可解释的符号元表示中也特别重要[115]。

**4.3元目标和情节设计**

最后一个组成部分是通过选择元目标Lmeta以及内部循环事件与外部优化之间的关联数据流来定义元学习方法的目标。在文献中，大多数方法都依赖于某种形式的性能指标，该指标在用ω更新任务模型并将此指标用作元目标之后，在验证集上计算得出。这与基于经典验证集的超参数调整和体系结构选择方法一致。 但是，在此框架内，有几种设计选项：

**多发和少发情节设计** 根据目标是提高几发还是多发性能，可以用许多[67]，[89]，[91]或很少[19]，[41]每个任务的示例。

**快速适应与渐近性能** 在内部学习阶段结束时计算验证损失时，元训练会鼓励基本任务的更好最终性能。 当将其作为每个内部优化步骤之后的验证损失总和进行计算时，元训练也鼓励在基本任务中更快地学习[76]，[89]，[91]。 大多数RL应用程序也使用后一种设置。

**多任务与单任务** 当目标是调整学习者以更好地解决给定家庭提出的任务时，内循环学习情节对应于p（T）的随机得出的任务[19]，[20]，[44]。 如果目标是让学习者更好地简单地解决一项特定任务，那么内循环学习情节都会从同一基础任务中提取数据[67]，[78]，[162]，[167]，[168]，[185]。

值得注意的是，这两个元目标往往具有不同的假设和价值主张。 多任务目标显然需要一个任务族p（T）来工作，而单任务则不需要。 同时，对于多任务，通过潜在地提高元测试期间多个目标任务的性能，可以分摊元训练的数据和计算成本； 但是单任务（没有用于摊销的新任务）需要改善当前任务的最终解决方案或渐进性能，或者需要足够快的元学习以使其在线。

**在线与离线** 尽管经典的元学习管道将元优化定义为内部基础学习者的外环[19] [78]，但一些研究已尝试在单个基础学习情节中在线进行元优化[ 44]，[167]，[185]，[186]。 在这种情况下，基础模型和学习者ω 在单个情节中共同发展。 由于现在没有可供分期偿还的学习操作，因此与基本模型学习相比，元学习需要快速进行，以提高样本或计算效率。

**其他情节设计因素** 其他操作符可以插入情节生成管道中，以自定义特定应用程序的元学习。 例如，可以在训练和验证之间模拟域转换，以在域转换下对最佳性能进行元优化[44]，[92]； 在训练和验证之间模拟网络压缩，例如量化[187]，以进行元优化以获得良好的网络可压缩性； 在元训练期间提供嘈杂的标签以优化标签噪声的鲁棒性[93]，或生成一个对抗验证集以进行元优化以进行对抗防御[94]。 在以下应用程序部分中将更详细地探讨这些机会。

**5应用程序**

在本节中，我们将讨论元学习的开发方式-在诸如计算机视觉和强化学习之类的应用领域以及诸如体系结构搜索，超参数优化，贝叶斯和无监督元学习等跨领域问题方面。

**5.1计算机视觉和图形**

计算机视觉是元学习技术的主要消费者领域之一。 这尤其受到元学习对小样本学习的影响的驱使，该学习有望应对由视觉识别的长尾概念带来的挑战。

**5.1.1小样本学习方法**

小样本学习（FSL）极具挑战性，尤其是对于大型神经网络模型[1]，[15]，其中数据量通常是性能的主要因素[188]，训练带有小数据集的大型模型会导致过拟合甚至不收敛。 基于元学习的小样本学习方法可以训练算法，使强大的深度网络能够在小型数据集上成功学习。 在小样本情况下，元学习可以帮助解决很多视觉问题，我们提供了以下非详尽的总结。

**分类** 到目前为止，元学习的最普遍应用是图像识别中的小样本分类，其中内部和外部损失函数通常分别是训练和验证数据上的交叉熵[19]，[20]，[41] ，[73]，[75]，[76]，[86]，[88]，[97]，[98]，[101]，[161]，[189]-[193]。 以优化器为中心的模型[19]，黑匣子[40]，[79]和度量学习[86]-[88]模型都已考虑在内。 相关基准在5.1.2节中介绍。

与早期方法[19]，[85]，[86]相比，该工作线导致了性能的稳定提高。但是，性能仍然远远落后于完全监督的方法，因此还有更多工作要做。 当前的研究问题包括具有更好跨域泛化性的小样本模型[173]，在元训练和元测试类定义的联合标签空间内进行识别[80]，以及新增加的小样本类[128]，[165]。

**对象检测** 基于小样本分类的快速进展，最近的工作也推广到了小样本对象检测[165]，[194]，经常使用基于前馈超网络的方法将支持集图像嵌入并在最终层分类权重中进行合成在基本模型中。

**地标预测** 地标估计的目的是找到图像内骨架关键点的位置，例如人体或机器人图像中的关节。 通常将其公式化为图像条件回归问题。 例如，一个基于MAML的模型可用于人体姿势估计[195]，模块化元学习已成功应用于机器人技术[130]，而一个基于超网络的模型则被应用于小样本适合新颖时尚服装的短裙 [165]。

**对象分割** 由于在此域中获取按像素标记的图像的成本，小样本对象分割很重要。 基于超网络的元学习方法已被证明可以在单样本状态下工作[196]，后来通过改编原型网络提高了性能[197]。 其他模型解决了分割密度低的情况[198]。

**图像生成** 在[199]中，分期付款的概率元学习器用于仅从单个图像生成对象的多个视图，并且通过学习对抗性模型的初始化以快速适应[200]，从很少的数据生成说话的面孔。

**视频合成** 在[201]中，作者提出了元学习权重生成器，该权重生成器将几帧作为输入并生成一个网络，该网络可以针对给定的任务在视频合成中获得强大的效果。

**密度估计** 由于自回归模型通常需要较大的深度才能捕获数据的分布，因此，仅凭少数情况就可能导致过度拟合，这尤其具有挑战性。 结合注意机制的元学习已显示使PixelCNN能够在这种情况下大放异彩[202]。

5.1.2小样本学习基准

人工智能和机器学习的进步通常由设计良好的基准来衡量和刺激[203]。 在机器学习中，基准由数据集和模型应执行的任务组成，同时从该数据集中的训练到测试实例进行概括。 在元学习中，基准设计更加复杂，因为我们经常与一个学习者打交道，该学习者应该在一组任务上进行（元）训练，然后应该推广到以前未见过的任务上进行学习。由于需要定义任务系列，可以从中得出元训练和元测试任务，因此基准设计更加复杂。 在本节中，我们将概述主要的基准测试。

**基准和设置** 大多数FSL研究都考虑了“set to set”的设置，在该设置中，模型必须学会在许多小的少量学习任务中表现出色。 每个这样的任务都由一个小型训练集（称为支持集）组成，该训练集由多个类别中的一些带有标记的示例组成；而一个较小的验证集（称为查询集）则由以前未见的实例组成 支持集中包含的相同类。 学习者应该能够从支持集中提取特定于任务的信息，然后生成可以在查询集中表现良好的模型。 可以通过学习可以很好地完成此任务的学习者来学习跨任务知识。 我们通常使用N路K射击任务的表示法来表示一个任务，每个任务有N个类别，每个类别有K个样本。

在此设置中使用了许多已建立的FSL数据集，例如miniImageNet [41]，tieredImageNet [204]，SlimageNet [205]，CUB-200 [110]和Omniglot [86]。 这些基准通过将它们分成许多较小的（较低的“方式”）识别问题来定义用于基准元训练和元测试的任务分布，从而将大量类别的先前数据集重新利用。

**数据集的多样性，偏差和一般化** 虽然上述方法很方便生成足够的训练和评估任务，但它缺乏多样性（狭窄的p（T）），这使得这些基准的绩效难以反映实际 世界上的小样本任务。 例如，在mini-ImageNet中的不同种类的动物或CUB中的鸟类之间进行切换是对传递性的相当微弱的测试。 理想情况下，我们希望跨越更多类别和类型的图像（卫星，医学，农业，水下等）； 甚至对元训练和元测试任务之间的域转换也很健壮。

这里仍然有很多工作要做，即使在多样本的情况下，将深层模型拟合到非常广泛的数据分布本身也不是一件容易的事[206]，就像推广到样本外数据[ 44]，[92]。 特别是，当在源任务分配和目标任务分配之间引入域转移时，元学习器的性能已显着下降[110]。 这激发了最近的元数据集[207]和CVPR跨域小样本挑战[208]。 元数据集聚合了许多个人识别基准，以提供更广泛的任务分布p（T），以评估适应广泛的任务分布并跨域转移进行概括的能力。

同时，[208]挑战了将图像网络的日常图像推广到医学，卫星和农业图像的方法。 最近的工作已经开始尝试通过元训练来解决域移位的鲁棒性和样本效率[173]。 在将模型应用于代表性不足国家的数据时，也会出现泛化问题[209]。 [210]另一个最近的数据集可以简化学习者的泛化研究，该数据集可以提供从模拟到高清模拟和真实环境的各种样本。

**现实世界中的小样本识别** 最常见的小样本问题设置是支持集[19]，[20]中各类之间的N向识别。 但是，这可能无法代表实际应用需求，因为在测试时，源和目标之间的识别都是很重要的。 越来越多的研究[128]，[165]和[211]都考虑到了这种通用的小样本设置。 在通用的小样本设置中，其他目标包括在不忘记基类或重新访问源数据的情况下有效地增量注册新颖的小样本类，[128]，[165]。 现实世界中的其他挑战包括在广泛研究的N = 1：：：20ωay识别范围之外扩大小样本学习的范围，这时流行且有效的度量学习者方法系列[20]，[87]开始陷入困境。

**小样本物体检测** 迄今为止，很少有研究[165]进行小样本物体检测，其重用了标准检测数据集，例如COCO和Pascal VOC。 但是，与分类基准相比，这些仅提供了一些用于元训练/测试的类，因此需要更多的基准。

**回归基准** 不幸的是，为小样本回归建立通用基准的工作比分类少。 在[19]，[212]中已经提出了玩具问题，例如一维正弦回归。 已经考虑了通过从像素坐标回归到RGB值来完成图像完成[166]，一些工作以人体姿势和时尚回归到兴趣点[165]，而[213]考虑了面部姿势回归的任务，另外进行了遮挡以引入歧义。总体而言，这些任务都是分散的，元学习社区尚未就回归基准达成共识。

**非元学习的小样本学习方法** 最近，许多非元学习的方法在少数测试基准上获得了竞争优势，这质疑在这种情况下学习的必要性。 在[110]中表明，立即对所有基本任务进行培训并对目标任务进行微调是比最初报告的更强的基准，主要是因为不公平地省略了扩充。此外，使用更深的主干可以缩小常规元学习方法之间的性能差距，并且基线可以在源任务分配和目标任务分配之间进行较大的域转换时胜过这些方法[207] –尽管最近的元学习方法在此设置取得好效果[173]。 在类似的主题上，[214]表明，简单的特征转换（如L2归一化）可以使最近邻分类器具有竞争力，而无需进行元学习。

因此，这里的辩论仍在进行中，但是精心设计的总体基线和更多样化的数据集非常重要，并且对于所有方法都应保持公平，一致的最佳实践。

**5.2元强化学习和机器人**

强化学习通常与学习控制策略有关，该策略使代理能够在环境中完成顺序动作任务时获得较高的回报，而监督学习则专注于给定数据集的准确性。 由于稀疏的奖励，对探索的需求以及高变异性[215]优化算法，RL通常遭受极端的样本低效。 但是，应用程序自然也很自然地需要元学习可以利用的任务系列，例如定位或到达不同的位置[172]，在不同的地图/环境中导航[40]或穿越不同的地形[63]，驾驶不同的地形 汽车[171]，与不同的竞争者竞争[61]，并应对不同的障碍，例如单个机器人肢体的故障[63]。因此，RL提供了一个肥沃的应用领域，其中与任务分配相关的元学习已成功地提高了标准RL算法的样本效率。可以直观地了解这些方法的功效。 例如，类人机器人的“如何站起来”的元知识是家庭中所有需要运动的任务的可转移技能，而迷宫布局的元知识可以在所有需要在迷宫中导航的任务中转移。

**5.2.1方法**

我们已经在RL中探索了我们已经看到的几种元表示，包括学习初始条件[19]，[159]，超参数[159]，[164]，步长方向[75]和步长[163]，这使基于梯度的学习能够以较少的环境交互作用来训练神经策略； 并训练快速卷积[40]或循环[22]，[106]黑盒模型，以迄今嵌入给定环境的经验，并使用它来综合前馈策略。 最近的工作为这些任务开发了改进的元优化算法[155]，[156]，[158]，并为Meta RL [216]提供了理论保证。

**探索**RL是一种非常独特的元表示形式。 RL的复杂性在于数据分发不是固定的，而是根据代理的操作而变化的。 此外，稀疏的奖励可能意味着代理必须采取许多行动才能获得可用于指导学习的奖励。因此，如何探索和获取用于学习的数据是任何RL算法的关键因素。 传统上，探索是基于对随机动作[90]或手工探索启发法[217]进行采样。 相反，一些元RL研究将探索策略或好奇心功能明确地视为元知识ω 并将其获取建模为元学习问题[24]，[170]，[171] –通过“学习探索”显着提高了样本效率。

**优化** 值得注意的是，与SL不同的是，优化通常会导致在训练组上以完美的精度获得良好的局部最小值。 RL通常是一个非常困难的优化问题，即使在“训练集”情节上，学习到的策略也远非最优。 这意味着，与meta-SL相比，meta-RL方法更普遍地用于提高渐近训练性能[23]，[164]，[167]以及样本效率，并且可以带来总体上更好的解决方案 。实际上，大多数元RL框架的元目标是代理人在整个训练过程中的净回报，因此，对样本有效学习和渐进式学习都进行了奖励。 优化难度还意味着，在学习损失（或奖励）上也有相对较多的工作[RL]，[167]，[218]，RL代理应代替而不是常规的稀疏奖励目标来进行优化。与真实目标[23]，[218]相比，这种元学习的损失可能更易于优化（更细，更平滑）。 这也可以作为奖励学习与探索联系起来，可以考虑实例化元学习方法来学习内在动机[168]。

**在线MetaRL** 我们注意到，很大一部分meta-RL研究是针对在线单任务设置的，其中元知识包括损失[113]，[167]，奖励[164]，[168]，超参数[162] ，[163]或探索策略[169]与基本策略一起在线学习，同时学习单个任务。 因此，这些方法不需要任务族，并且可以直接改善他们各自的基础学习者。

**On- vs Off-Policy Meta-RL** 传统RL方法学的一个主要二分法是在策略学习与非策略学习之间，例如PPO [90]与SAC [219]。 对于常规RL，非政策方法通常会大大提高样本效率。 但是，非策略方法很难扩展到meta-RL，导致大多数Meta-RL方法基于基于策略的算法构建，因此限制了Meta-RL的绝对性能。 最近的一些工作已经开始设计非政策方法的meta-RL概括，从而获得了强有力的结果[109]，[113]，[157]，[218]。 值得注意的是，脱离政策的学习还提高了元训练阶段的效率[109]，这在Meta-RL中可能非常昂贵。 它也提供了新的机会，可以通过从元训练阶段重播缓冲区样本来加速元测试[157]。

**其他趋势和挑战** 在本节中，我们将提到meta-RL的其他近期趋势。 [63]在演示真实的物理机器人上成功的meta-RL时值得注意。 机器人技术中的知识转移通常对于进行组成学习是有意义的[220]。 例如，步行，导航和物体拾取/放置可能是清理机器人房间的子例程。 但是，开发支持很好地传递知识的元学习器是一个悬而未决的问题，模块化元学习[131]是一个选择。 无监督的meta-RL变体旨在执行元训练而无需手动指定奖励[221]，或适应于对变化的环境进行元测试但没有新的奖励[222]。 持续适应利用元学习为代理提供了适应一个元测试情节[61]-[63]中与连续学习相关的一系列任务的能力。 最后，元学习也已应用于模仿[105]和逆强化学习[223]。

**5.2.2基准**

RL的元学习基准应该为代理商解决一系列问题，以便他们学习如何学习13，然后评估学习者。 这些可以是要实现的任务（奖励功能），也可以是域（不同的环境或MDP）。 RL基准可以根据它们是测试连续控制还是离散控制，以及根据状态或图像（例如图像）的致动来进行划分。

**离散控制RL** 街机学习环境（ALE）[224]是用于视觉驱动控制的早期meta-RL基准，它定义了一组经典的Atari游戏，可以将其分为元训练和元测试。 这里的典型协议是在元测试环境中经过固定数量的时间步长后评估收益。 Atari游戏的一个问题是它们的确定性，这意味着开环策略可能足以解决它们，从而导致人们努力增加随机性[224]。 另一个挑战是游戏之间存在很大的差异性（宽p（T）），这使得成功的元训练变得困难，并导致知识转移的收益有限[224]。 另一个基准[225]基于将Sonic-hedgehog水平划分为metatrain / meta-test。 这里的任务分布较窄，有益的元学习相对容易实现。 最近，Cobbe等。 [226]提出了两个专门设计的视频游戏，用于对Meta-RL进行基准测试。 CoinRun游戏[226]提供了个由程序生成的，具有不同难度和视觉外观的级别。 他们表明大约需要有10 000级的元训练经验才能可靠地推广到新的水平。 CoinRun主要设计用于测试直接概括而不是快速适应，并且可以看作是在MDP环境中提供了一个分发来测试概括，而不是在任务上提供了分发以测试适应。 为了更好地测试更广泛的任务分配中的快速学习，ProcGen [226]提供了一套包括CoinRun在内的16种程序生成的游戏。

**连续控制RL** 尽管使用gym [227]这样的通用基准极大地有益于RL研究，但对于metaRL的基准尚未达成共识，使得现有工作难以进行比较。 连续控制meta-RL的大多数研究都提出了自制的基准，这些基准是特定任务的低维参数变体，例如导航到不同的位置或速度[19]，[109]或穿越不同的地形[63]。最近已经提出了几种多MDP基准[228]，[229]，但是这些基准主要测试了在不同环境扰动下的泛化，而不是Meta-RL中感兴趣的新任务适应。 随着Meta-Ωorld基准测试[230]的发布，这种情况将得到改善，该基准测试提供了基于状态的致动的50个连续控制任务套件，不同于简单的参数变体，例如拉杆和开门。 该基准应该可以进行更可比的评估，并可以调查各种宽度的任务分布之内和之间的泛化。 元世界评估[230]表明，现有的Meta-RL方法难以在广泛的任务分布和元训练/元测试班次上进行概括，因此需要做更多的工作。 另一个适用于Meta-RL的基准是PHYRE [231]，它提供了一组50个基于视觉的物理任务模板，这些模板可以通过简单的动作来解决，但可能需要基于模型的推理才能有效地解决。 这些分为2个难度等级，并提供内部和跨模板概括测试。

**讨论** 视觉驱动的meta-RL的一个复杂之处是无法通过广泛地快速学习控制策略来实现视觉泛化和适应（与更广泛的计算机视觉相同）。 例如，CoinRun [226]评估显示了标准视觉技术（例如批处理规范）的巨大好处，这表明感知是主要瓶颈。

Meta-RL的一个热门问题是，在进行新任务的元测试性能之前，很难使用多任务或金属学习模型来适应广泛的meta-train任务分布。 这可能是由于我们的RL模型太弱和/或基准在任务数量上太小。 甚至Meta-world，ProcGen和PHYRE都有数十个而不是数百个任务，例如tieredImageNet等视觉基准。 尽管这些最新基准不断改进，但仍可以从具有可控泛化差距的更大基准中受益。 具有更高难度的基准（例如要求记忆和抽象推理），为跨任务跨元学习和利用更多抽象策略提供机会也是有益的。

**5.3环境学习和Sim2Real**

在Sim2Real中，我们感兴趣的是在模拟中训练能够推广到真实世界的模型，这是具有挑战性的，因为模拟与真实世界并不完全匹配。经典的域随机化方法模拟了域/ MDP上的广泛分布，其目的是训练一个足够强大的模型以在现实世界中取得成功，并且在视觉[232]和RL [148]方面均取得了成功。 然而，如何调整仿真分布是一个挑战。这自然导致了元学习的设置，其中，内部优化在模拟中学习模型，外部优化Lmeta在实际环境中评估模型的性能，和元表示ω 对应于仿真环境的参数。此范例已在RL [150]和计算机视觉[149]，[233]中使用。 在这种情况下，用于元训练任务的源任务不是预先提供的数据分发，而是由omega参数化的，但是，通过内部任务的昂贵且冗长的学习步骤来进行反向传播仍然存在挑战。 以及在针对RL的Sim2Real 元学习的情况下，最大程度地减少了实际的评估次数。

**5.4神经架构搜索（NAS）**

架构搜索[26] – [28]，[39]，[123]可以看作对应于一种超参数优化，其中ω 指定神经网络的架构。内部优化训练具有指定体系结构的网络，而外部优化则搜索具有良好验证性能的体系结构。 NAS方法通常根据“搜索空间”，“搜索策略”和“性能估算策略”进行分析[39]。这些对应于ω，元优化策略和元目标的假设空间。NAS尤其具有挑战性，因为：（i）全面评估内循环通常非常昂贵，因为它需要训练很多次神经网络才能完成。 这样就得出了近似，例如对训练组进行二次采样，内环的提前终止，并最终达到近似值，例如两者的交错下降 [26]就像在线元学习中一样。（ii。）搜索空间很难定义，并且对其进行优化是昂贵的。这是因为大多数搜索空间都很宽泛，并且代表了不可区分的架构。 这导致执行单元级搜索[26]，[28]来限制搜索空间的方法。 然后依赖RL [28]，即离散梯度估计量，它们为搜索空间[26]，[124]和演化[27]，[123]提供了可微的近似。

**示例** 一些著名的示例包括：（i）NASNet [28]，[234]，其中搜索空间仅限于单元级学习，并且定义为RNN生成的字符串，指示在单元树的哪些部分应进行哪些操作，使用RL优化。 （ii）规则化进化[27]，其中作者使用NASNet的搜索空间，但使用规则化进化进行优化，即基于标准锦标赛的进化，每次迭代后都会删除最老的个体。 （iii。）DARTS [26]，作者在许多预先选择的操作中精心将单元架构的空间转换为一系列softmax选择，从而使搜索空间具有可区分性。 学习架构然后对应于与网络参数一起学习softmax权重。 这使体系结构学习的计算开销和挂钟时间都加快了2-3个级别。 （iv）T-NAS [125]作者利用DARTS搜索空间，但使用数据流对其进行训练，该数据流强制使用很少的数据点和很少的更新来学习体系结构，同时保持较高的泛化性能 。 通过学习此类softmax权重，他们实现了小样本架构搜索。 一旦经过训练，这些砝码可以在几秒钟而不是几天内适应新任务。

NAS的一个有趣的特殊情况是激活功能搜索[151]。 尽管人工设计的激活功能（例如ReLU）在NN文献中占主导地位，但NAS元学习的成功示例是在符号激活功能空间中发现了带有RL的Sωish激活功能[151]。 Sωish继续为几种有影响力的最新和通用CNN架构做出了贡献[235]，[236]。

**多目标NAS** 除了验证准确性[7]之外，要部署在移动设备上的体系结构还具有其他约束条件，并且还可以部署NAS来生成紧凑而有效的模型[237]。 这可以通过定义一个多目标元目标来实现，该目标包含与验证性能以及模型产品的等待时间或大小相关给定，从而导致良好的性能成本权衡。

**话题问题** 虽然NAS本身可以看作是超参数或假设类元学习的实例，但它也可以与其他形式的元学习进行交互。 由于NAS成本高昂，一个热门问题是发现的体系结构是特定于数据集的，还是具有泛化为新问题的能力的通用[234]。 最近的结果表明，跨多个数据集进行元训练可以提高体系结构的跨任务概括性[126]。

虽然通常在手工架构[19]，[20]，[87]的背景下从参数学习的角度解决小样本元学习，但也可以定义NAS元目标来训练适合小样本的架构学习[238]，[239]。 此外，类似于MAML [19]等快速适应的初始条件元学习方法，人们可以训练易于适应特定任务的良好初始架构[125]或架构先验[126]。

**基准NAS** 通常在CIFAR-10数据集上进行评估。 但是，即使在这个很小的数据集上，架构搜索的执行成本也很高，这使得许多研究人员无法访问它。 而且由于其他混淆因素，例如超参数的调整，结果很难重现[240]。 为了支持可重复和可访问的研究，最近发布的NASbenches [241]，[242]为大量网络体系结构提供了预先计算的性能指标。

**5.5贝叶斯元学习**

贝叶斯元学习方法通​​过贝叶斯层次建模将元学习形式化，并使用贝叶斯推理进行学习，而不是直接优化参数。在元学习环境中，贝叶斯学习通常很棘手，因此可以使用不同的近似方法。 变异方法，尤其是随机变异方法，是最常见的方法，但也可以考虑采用抽样方法。

贝叶斯元学习的一个副产品是它为参数提供不确定性度量，从而为预测不确定性提供度量。 知道学习者预测的不确定性对于安全性至关重要的领域（例如，小样本医疗任务）至关重要，并且可以用于强化学习中的探索以及某些主动学习方法，其中模型可以搜索有关高度不确定性数据点的信息。

最近，许多作者探索了贝叶斯方法来获得具有竞争性结果的元学习复杂模型。 其中许多已经利用深度神经网络作为框架内的组件，例如扩展了变分自动编码器以明确地对任务变量建模[71]。 神经过程[166]的目的是将高斯过程的不确定性量化与神经网络的多功能性相结合，但并未证明它们可以在现代的一次性基准上工作。 深度内核学习也是一个活跃的研究领域，已经适应了元学习设置[243]，并且经常与高斯过程[213]结合在一起。 在[72]中，基于梯度的元学习被重铸为一个分层的经验贝叶斯推理问题（即先验学习），该问题对特定于任务的参数的不确定性进行建模。贝叶斯MAML [212]通过使用允许非高斯后验的贝叶斯合奏方法对该模型进行了改进，后来的工作消除了对昂贵的合奏的需要[199]，[244]。在概率MAML [95]中，这是元知识中的不确定性ω 被建模，而MAP估计用于。这些贝叶斯方法越来越能解决模棱两可的任务，主动学习和RL问题。

与上述方法不同，还提出了元学习来辅助贝叶斯推理过程本身。 举例来说，在[245]中，作者使用元学习框架来适应贝叶斯采样器，以提供有效的自适应采样方法。

**基准** 在贝叶斯元学习中，重点通常是对我们的元学习器的预测中的不确定性进行建模，因此标准的一次性快照分类基准的性能并不一定能反映我们的关注。 因此，文献中已经开发了不同的任务。 贝叶斯MAML [212]扩展了MAML [19]的正弦回归任务，使其更具挑战性。概率MAML [95]提供了一套一维玩具示例，能够显示模型不确定性以及如何在主动学习场景中使用该不确定性。 它还从celebA [246]创建了一个二元分类任务，其中通过两个面部属性的存在来确定阳性类别，但是训练图像显示了三个属性，因此引入了应在其中分类两个属性的歧义。据观察，采样ω 可以正确反映这种含糊。 [212]中还显示了主动学习玩具实验以及强化学习应用程序，[199]中也使用了模糊的单发图像生成任务。 最后，一些研究人员建议对准确性进行比较元学习者的信心（即他们的校准）[244]。

**5.6无监督的元学习和无监督的元学习**

在元学习文献中，元学习存在两种主要形式，即无监督学习。 在第一个中，外部循环的元目标是无监督的，因此学习者本身是在没有任何可用标签的情况下学习的。 我们将这种情况称为无监督元学习。 在第二个变体中，元学习被用作学习无监督内循环任务的一种手段。 在这种情况下，外部目标可以是基于监督，不受监督或基于增强的任何内容。 我们将此称为“元学习无监督学习”。

无监督元学习[247] – [249]的目的是放宽对元训练的源任务的注释集的常规假设，同时仍然为有监督的小样本学习提供良好的下游性能。 通常，合成源任务是在没有监督的情况下通过聚类或保留类的数据扩充来构造的。

元学习无监督学习旨在利用元学习来训练无监督学习算法，该算法很好地适用于下游有监督学习任务。 可以训练无监督的聚类算法[21]，[250]，[251]或损失[98]，[116]，以便优化下游的有监督学习性能。 通过将无监督学习问题转化为具有清晰（监督）目标的问题，这有助于解决无监督学习问题的不确定性。

**5.7主动学习**

元学习范式也可以用于培训主动学习，而不是到目前为止所讨论的监督或强化学习者。 主动学习（AL）方法包括监督学习，并定义了选择性数据注释的策略-通常在可以顺序获取注释的环境中。 AL的目标是找到要注释的最佳数据子集，从而以最少的注释来最大化下游监督学习的性能。 AL是一个经过广泛研究的问题，具有众多手工设计的算法[252]。元学习可以通过以下方式将主动学习算法设计转化为学习任务：到目前为止，将内部级优化视为带注释数据集上的常规监督学习任务，请考虑！ 成为一种查询策略，该方法选择最佳的未标记数据点进行注释，或者通过给定查询和注释的数据点，让外层优化训练查询策略以优化与下游学习性能相对应的元目标[175]-[177] 。然而，对于聚类，如果使用标签来训练AL算法，则它们需要在各个任务之间进行概括以分摊其训练成本[177]。

**5.8连续，在线和自适应学习**

**持续学习** 是指按顺序呈现的学习任务的类人能力。 理想情况下，这是在利用正向传递的同时完成的，因此，根据过去的经验，可以更好地学习新任务，而不会忘记以前学习的任务，也不需要存储所有过去的数据以进行排练以防忘记[60]。 深度神经网络在努力满足这些标准方面遇到了困难，特别是因为它们倾向于忘记在早期任务中看到的信息–这种现象称为灾难性遗忘。 元学习已被应用于改善深网的持续学习。 可以将持续学习的要求整合到一个元目标中，例如，通过定义一系列学习情节，其中支持集包含一个新任务，而查询集包含从迄今为止看到的所有任务中提取的示例[160]， [161]。 通过这种元目标设计，可以训练各种元表示形式，从而提高持续学习的性能。 例如：权重先验[128]，梯度下降预处理矩阵[161]或RNN学习的优化器[160]或特征表示[253]。

尽管没有直接应用于持续学习，但另一个有趣的想法是元训练表示法以支持本地编辑[254]，在该方法中，作者学习了一种可以在单个样本上快速改进自身的模型，而不会忘记已经学习的任何信息。

**在线和自适应学习** 还考虑任务到达流中，但是与记住旧任务相比，更关注有效地适应流中当前任务的能力。 为此，提出了MAML的在线扩展[96]，以在任务序列期间在线执行MAML风格的元训练。 同时，其他人[61]-[63]则考虑了对源任务预先进行元训练的设置，然后才对目标任务序列进行元测试适应能力。

**基准** 持续学习存在许多基准，它们可以与标准深度学习方法很好地配合使用。 但是，大多数这些基准测试无法轻松地与元学习方法一起使用。 他们中的大多数将需要调整其样本生成例程，以包括大量的显式学习集和显式评估集。 在[96]，[160]，[253]中为定义元学习就绪的持续基准做出了一些早期步骤，这些基准主要由Omniglot和受干扰的MNIST版本组成。但是，其中大多数只是为演示一种方法而构建的任务。在[205]中可以找到更明确的基准工作，其中连续小样本学习被定义为要处理的新型任务，并且该基准针对元和非元学习方法而建立。在这种情况下，一项任务由许多小型训练集组成，每个训练集可能由不同的类别组成，然后，学习的模型应该很好地概括从之前从其学习到的所有任务中看不见的样本中得出的结论。 该基准测试建议使用Omniglot和SlimageNet作为要使用的数据集。

**5.9域适应和域泛化**

当部署中遇到的数据统计信息与训练中使用的数据不同时，领域转移通常会在实践中妨碍机器学习模型。 为了在有监督，无监督和半监督的环境中解决此问题，已经研究了多种域自适应和泛化算法[57]。

**域泛化** 域泛化方法旨在通过设计[255]来训练具有更高鲁棒性的模型来训练测试域移位，这通常是通过利用训练域上的分布来进行的。 通过将外环验证集定义为相对于内环训练集具有域移位，元学习可以是支持该目标的有效工具。[58] 通过这种方式，可以学习（元）不同种类的元知识，例如正则化器[92]，损耗[44]和噪声增强[173]，从而最大程度地提高学习模型的典型鲁棒性以训练测试领域- 转移。

**领域适应** 尽管关于领域适应的大量先前工作是常规学习[57]，但最近的工作[256]也开始考虑元学习方法来促进领域适应。

**基准** DA和DG的流行基准围绕识别不同图像类型（例如照片/素描/卡通）而定。 通常使用具有多个域的数据集以提供用于元学习的域分布。 PACS [257]通过Visual Decathlon [44]，[206]，DomainNet [258]和Meta-Dataset [207]提供了一个很好的入门基准，提供了更大的替代方案。

**5.10超参数优化**

元学习可以通过考虑ω来解决超参数优化问题。 指定超参数，例如正则化强度或学习率。有两个主要设置：我们可以学习提高对任务分布的训练的超参数，或者可以提高对单个任务的学习的超参数。 前一种情况通常与小样本应用有关，特别是在基于优化的方法中。 例如，可以通过学习每层每一步的学习率来提高MAML [76]。 我们希望为单个任务学习超参数的情况通常与多次应用程序[145]更为相关，如第2.1节所述，可以从训练数据集中提取一些验证数据。

长期的内在元收益伴随着内存和计算扩展问题，这是第6节中讨论的一个活跃的研究领域。但是，值得注意的是，基于端到端的基于梯度的元学习已经证明了对数百万用户的可扩展性 参数（例如，由MAML [19]，[145]和数据集蒸馏[144]，[145]演示）与经典方法（例如通过网格或随机[69]搜索进行交叉验证或贝叶斯优化）形成对比 [70]）通常只有数十个超参数才能成功。

**5.11新颖且生物学上合理的学习者**

大多数对基础模型使用显式（非前馈/黑匣子）优化的元学习工作都是基于反向传播的梯度下降（与大多数传统的深度学习工作一样）。 元学习的一种有趣的可能性是定义学习规则的功能类别ω从而发现了可能不受监督的新颖有效的学习规则[21]，生物学上合理的[47]，[259]，[260]，它们利用了当代深度学习中较少使用的思想，例如赫比更新[259] ]和神经调节[260]。

**5.12语言和语音语言建模。**

小样本语言建模是展示元学习者多功能性的一种流行方法，早期的方法（例如匹配网络）在单发任务（例如填写遗漏单词）上表现出令人印象深刻的性能[86]。 此后，还解决了更多任务，包括神经程序归纳[261]和综合[262]，英语到SQL程序综合[263]，基于文本的关系图提取器[264]，机器翻译[265]以及快速适应 对话任务中的新角色[266]。

**语音识别** 现在，深度学习已成为最先进的自动语音识别（ASR）的主要范例。 元学习已开始应用于解决ASR内出现的许多短暂适应问题，包括学习如何训练低资源语言[267]，交叉口音适应[268]以及为单个说话者优化模型[269]。

**5.13社会公益的元学习**

元学习将自己置于AI应用于社会公益的各种挑战性任务中，例如医学图像分类和药物发现，而这些领域通常缺乏数据。 鉴于全球病理学家的短缺，在医学领域的进展尤为重要[270]。 在[5]中，LSTM与图神经网络相结合来预测分子在单次数据机制中的行为（例如，其毒性）。 在[271]中，MAML适用于弱监督的乳腺癌检测任务，并且任务的顺序是根据课程表而不是随机选择的。 MAML还与去噪自动编码器相结合以进行医学视觉问题解答[272]，而在学习权衡[204]中所做的支持样本时，将其调整为像素级加权，以解决具有嘈杂标签的皮肤病变分割任务[273] 。

**5.14抽象和成分推理**

抽象推理深度学习研究的最新目标是开发超越简单感知任务的模型，以解决更抽象的推理问题，例如以Raven的渐进矩阵（RPM）形式进行的IQ测试[274]。 解决RPM可以看作是要求从上下文面板到答案面板进行少量概括。 最近的元学习方法通​​过RPM进行抽象推理，通过元学习定义了面板数据生成分布的教师而获得了显着改善[275]。 教师与学生一起接受培训，并因学生的进步而获得奖励，从而自动定义最佳课程。

**组成式学习** 使人类擅长解决问题的特质之一是学习如何撰写概念的能力。 例如，能够采用新近学习的动词并将其与所有潜在副词一起使用。最近的元学习方法已经显示出通过在元训练期间要求查询和支持集之间的组成概括来提高这种概括能力[276]。 这样的元学习机制也可以使基本挑战受益，例如使序列模型能够推广到比训练期间观察到的序列更长的序列[276]。

**5.15系统**

**网络压缩** 当代的CNN需要大量内存，这在嵌入式设备上可能会被禁止。 因此，各种形式的网络压缩（例如量化和修剪）是热门的研究领域[277]，[278]。 元学习也开始应用于此目标，例如训练梯度生成器元网络以允许对量化网络进行训练[187]，以及权重生成器元网络允许以量化方式对量化网络进行训练[279]。

**通信** 深度学习最近在通信系统中掀起了波澜。 例如，通过学习超出针对实际渠道的最佳手工设计代码的编码系统[280]。 只要通过学习针对特定信道的特性而优化的编码方案来实现最佳性能，就可以使用小样本触发元学习来提供快速的在线编码适应变化的信道特性[281]。

当通过网络抓取或众包收集大型数据集时，使用**标签噪声学习**是当代深度学习中的一个挑战。 同样，虽然有针对这种情况手动设计的几种算法，但最近的元学习方法已经通过对小重量的有噪声样本进行转导学习来逐步降低样本重量，或者通过学习对噪声标签训练具有鲁棒性的初始条件来解决标签噪声 [93]。

**对抗性攻击和防御** 通过向数据添加精心制作的人为可见的扰动，可以很容易地欺骗深度神经网络，对应该易于识别的数据点进行错误分类[282]。 近年来，已经发布了许多方法，它们引入了更强大的攻击和防御方法。 典型的防御措施是精心设计的架构或培训策略。 类似于领域转移的情况，元学习的潜在研究不足之处在于通过在对抗性攻击下根据性能定义元损失来端到端训练学习算法的鲁棒性[94]，[283]。

最近提出了对抗性防御的新基准[284]，其中应将防御推广到无法预料的攻击。 有趣的是，未来的元学习方法是否可以在该基准上取得进展。

**6挑战和未解决的问题**

元通用化元学习面临跨任务的通用化挑战，类似于传统机器学习中跨实例的通用化挑战。 存在三个子挑战：（i）第一个挑战是使元学习器适合任务p（T）的广泛分布，正如我们所看到的，这对现有方法[206]，[207]，[230]具有挑战性 ]，并且可能部分是由于任务之间的梯度冲突[285]。 （ii）第二个挑战是从元训练推广到从p（T）得出的新型元测试任务。由于可用于元训练的任务数量通常很少（比传统的有监督学习中可用的实例数量少），这使情况更加恶化，这使其难以适应复杂的任务分配。 因此，到目前为止，元学习者的最大成功是在非常相似的任务系列中。（iii）第三个挑战是推广到与培训任务分布不同的元测试任务。 在元学习的许多潜在实际应用中，这是不可避免的，例如，将从ImageNet的日常训练图像到医学图像等专业领域的几次快照式视觉学习泛化[208]。 从学习者的角度来看，这是在监督学习中观察到的域转移问题的元级概括。通过正则化，转移学习，领域自适应和领域概括的元概括来解决这些问题是新兴的方向[173]。 此外，我们尚未了解在某些类型的域移位下，哪种元表示形式倾向于更好地概括。

另一个有趣的方向可能是研究引入另一层次的学习抽象如何影响泛化性能，即元元学习。 通过学习如何进行元学习，也许我们可以找到可以在各种类型和强度的域甚至模式转换方面非常强大地概括的元优化器。 当然，计算成本将成倍增加。

任务分配的多模式许多元学习框架[19]隐含地假设任务p（T）的分配是单模式的，并且是一种学习策略ω为他们所有人提供了一个很好的解决方案。 但是，实际上，任务分配显然可以是多模式的。例如，在计算机视觉中考虑医学，卫星与日常图像。 或者，可能要求机器人执行各种任务，从将钉子插入孔中到打开门[230]。 分发中的不同任务可能需要不同的学习策略，这可能会降低现有的元学习器性能。 在普通的多任务学习中，这种现象可以用将任务分组为集群[286]或子空间[287]的方法进行比较研究。然而，在元学习中这个领域才刚刚开始探索[288]。

任务族许多现有的元学习框架，特别是针对少量学习的元学习框架，都需要任务族进行元训练。 尽管这确实反映了人类的终生学习，但在某些应用程序中，此类任务族的数据可能不可用。 如何放松这个假设是一个持续的挑战。 无监督的元学习[247] – [249]和在线元学习方法[44]，[162]，[167]，[168]，[185]可以帮助减轻这种情况； 以及上面讨论的元通用化方面的改进。

**计算成本** 如2.1节所示，天真地实施双层优化会导致学习步骤数量二次增加，因为每个外部步骤都需要多个内部步骤。 而且，在多次实验中，内部步骤很多，这些步骤需要存储在内存中。 由于这个原因，大多数元学习框架在时间和内存上都非常昂贵，并且在少数情况下通常仅限于小型架构[19]。 但是，人们越来越关注解决这个问题的方法。 例如，可以替换内部和外部更新[44]，或训练替代模型[108]。 另一系列的最新方法通过内部循环中的闭合形式求解器来加速元训练[152]，[154]。 然而，代价仍然是很大的，并且不清楚前一套启发式算法对于收敛的意义。最近的一种使用隐式梯度来计算外环梯度的方法提供了一种更便宜的选择[153]，但只专注于学习MAML网络的初始化。 然后显示了隐式梯度可用于更一般的元学习任务，例如学习增强网络[145]，他们只能直接学习损失函数中涉及的参数，并做出一些假设（例如处的零训练梯度），常常导致不正确梯度变化。

**跨模式转移和异构任务** 到目前为止，研究的大多数元学习方法都考虑了所有任务均来自相同的模态，例如视觉，文本，本体感受状态或音频。 人类似乎能够跨模式转移知识（例如，通过视觉模仿学习）。 如何进行元学习以从一组可能跨越一个独特模态的任务中提取抽象知识是一个悬而未决的问题。 大多数研究都针对相同类型的任务之间的转移，例如对象识别，但是理想情况下，我们希望能够在异构任务之间进行转移，例如Taskonomy [289]中研究的任务。

**7结论**

最近，元学习领域的兴趣迅速增长。 关于它与相邻字段的关系，如何应用以及如何进行基准测试，这带来了一定程度的混乱。在本次调查中，我们试图通过从方法论的角度全面调查该领域来澄清这些问题-我们将其分解为元表示，元优化器和元目标的分类法； 并且从应用程序的角度来看。 我们希望这项调查将帮助新移民和从业者适应这个不断发展的领域，并突出未来研究的机会。