

矩阵的正则化器。受最近各向同性研究的启发，他们通过将特征空间正则化为各向同性来改进监督预训练。基于微调的方法操作简单，适用于源数据集与目标数据集分布类似的场景。然而此类方法容易在仅含有少量数据的目标数据集上产生过拟合现象，从而影响最终结果。

（2）基于元学习的方法：核心思想是利用从大量先验任务中学习到的元知识指导新任务的学习，即让模型具有学会学习的能力。根据使用技术的不同，基于元学习的方法可分为：基于模型的方法、基于优化的方法和基于度量的方法。

①基于模型的方法：此类方法通常利用某一模型来理解任务，并将学到的知识存储在模型中。当需要决策时，将学习好的模型进行快速泛化用于决策。Santoro 等人提出了基于记忆增强的神经网络 MANN^[4]。该模型以神经图灵机为基础，通过修改神经图灵机的训练模式和寻址机制来实现单样本学习。Munkhdalai 等人提出了元网络模型^[5]。该模型包含基学习器和元学习器，另有一个额外的记忆模块。基学习器在任务空间中学习，元学习器在与任务无关的元空间中学习。当新任务到来时，基学习器首先对新任务进行分析，然后给元学习器提供一种由高阶元信息构成的反馈。元学习器收到元信息后，再对自身和基学习器进行快速参数化。后续方法还有 SNAIL^[6]，CNP^[7]等。Cai 等人提出 MM-Net^[8]将键值记忆网络^[9]中的记忆模块集成到匹配网络中。通过记忆模块扩展了基于度量的元学习方法，以将整个支持集编码和泛化到记忆槽中，更好的编码查询集中的图像。基于模型的方法在灵活性方面有一些优势，但是在数据集规模较大时，此类方法通常表现较差。

②基于优化的方法：此类方法的主要目标是让模型在仅有少量样本的情况下完成优化，进而解决小样本学习问题。Ravi 等人使用了基于 LSTM 的元学习器来学习优化算法，利用优化算法的参数更新规则更新分类器的网络参数，使得分类器在小样本数据上取得了较好的分类效果^[10]。Finn 等人提出了未知模型的元学习方法 MAML^[11]。该方法可以找到一组可以快速适应新任务的初始化参数，使得模型在面对新任务时，能够利用梯度下降方法在很少的步数内完成收敛。MAML 是一种通用的优化算法，只需模型采用梯度下降进行优化即可。后续基于 MAML 的改进方法有很多，如 TAML^[12]、FTML^[13]、BMAML^[14]等。基于优化的方法在任务分布广泛时表现较好，然而由于此类方法需要针对每个任务优化学习器，因此计算复杂度非常高。