**北京工业大学学术学位研究生**

**开题报告**

**学位级别： 博士** √**硕士**

**学 号： S202074018**

**研究生姓名： 于高远**

**指导教师姓名： 贾熹滨**

**专业名称： 计算机科学与技术**

**所在学院： 信息学部**

**开题报告时间： 2021-12-24**

**北京工业大学研究生院制表**

**注意**：本表基本情况及报告正文由研究生本人填写，硕士不少于3000字，博士不少于5000字。格式要求：正文文字部分为5号宋体、单倍行间距排版，A4纸双面打印装订。

开题报告评价部分分别由指导教师及专家组书写。开题报告会结束后一周之内将报告原件交院（所）研究生教学秘书处。

**一、基本情况**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **研究生姓名** | | 于高远 | | **学 号** | S202074018 |
| **院、系** | | 信息学部 | | **指导教师姓名及职称** | 贾熹滨 教授 |
| **学科、专业** | | 计算机科学与技术 | | **入学年月** | 2020-09 |
| **填表日期** | | 2021-12-23 | |  |  |
| 1. **研究方向、论文选题范围：**   研究方向：计算机应用  论文选题范围：深度学习、小样本学习、医学影像分类   1. **拟定论文题目：**   面向多模态医疗影像小样本数据的分类算法研究   1. **论文科研课题属于哪一级科研项目，经费来源及金额（**课题来源选项分为国家计委、科委项目、国家经贸委项目、国家自然科学基金项目、国务院其他部门项目、主管部门（部委级）项目、省、市、自治区项目、国际合作项目、学校级项目、自选项目、其它）**：**   科研项目： 基于多参数磁共振深度表示学习及迁移学习的非酒精性脂肪肝病炎症、纤维化程度无创性评估研究（国家自然科学基金项目6187010150，23.5万）、基于多模态磁共振影像应用迁移及深度度量学习预判肝细胞癌微血管侵犯的研究（国家自然科学基金面上项目，15.4万）   1. **论文类型（基础研究、应用研究、开发研究、其它）**   应用基础研发 | | | | | |
| **摘**  **要** | **5.选题研究内容和意义简介（限400字）：**  因深度学习能自动分析大规模影像、挖掘人眼不易观察的潜在知识，故与深度学习结合的医学影像智能分析技术近年来逐渐兴起，为辅助医生诊断决策和提升病人无创诊断效率提供了有力支持与帮助。然而现有的有监督深度学习方法存在一个共识性问题，即网络参数需要通过大量有标注数据进行学习，当数据量过少时，神经网络往往不能表现出较优的性能。而因数据本身特点和标注工作繁琐等原因造成的小样本问题一直以来是医学影像智能诊断的瓶颈之一。因此如何在进行高效多模态融合的同时，减轻模型体量以贴合小样本数据，并利用迁移学习或元学习等小样本学习算法辅助目标数据学习，是目前多模态小样本医学影像诊断的关键。因此本研究将探究：1）轻量高效的特征提取网络，在不引入过多参数的前提下提升特征表示质量。2）渐进式的知识迁移策略，利用先验共性知识辅助目标小样本任务的学习，同时将多模态融合和特征原型校正增强引入知识迁移的过程中，以提升模型的诊断性能。 | | | | |
| **关键词（用分号隔开、最多5个）** | | | 小样本学习；图像分类；多模态医疗影像诊断 | | |

**报 告 正 文**

**（一）选题依据与研究内容**

**1、选题依据（专业学位类别或领域的研究意义、国内外研究现状等）**

**1.1 研究意义**

随着信息化时代的到来，人工智能、深度学习等前沿技术快速发展，在很多视觉识别任务中均取得了巨大的成功，如图像分类、图像分割、物体检测等。近年来人工智能也开始逐步向医疗、教育、金融等各个传统行业渗透，以“智能医疗”为代表的新兴交叉应用领域越来越受到人们的广泛关注，一些研究成果也展示了人工智能在医学影像智能辅助分析诊断方面的潜力[1-6]，相较于传统人工医疗诊断，结合深度学习的医学影像智能分析技术可以自动分析大规模影像并深度挖掘隐藏知识，从而提升诊断准确率，减轻病人就医压力，为辅助医生诊断决策和提升病人无创诊断效率提供了有力的支持与帮助。同时为了丰富数据多样性，且仿照医生根据多个影像模态综合进行诊断的临床经验知识，医学影像多模态融合技术不断发展，成为当下“智能医疗”的热点。

然而现有的有监督深度学习方法存在一个共识性问题，即大量有效的网络参数需要从与之体量相匹配的数据集中学习得到，当用于支持训练的有标注数据过少时，神经网络往往不能表现出较优的性能。而医学影像标准数据收集又常常受多种因素的制约，如因伦理道德、个人隐私问题，数据公开审核过程繁琐，科研单位难以公开标准医疗影像数据集；同时医学影像标准数据的标定工作往往也需要专业医生花费较高的时间成本，故因数据收集困难且标注缓慢等原因引发的小样本问题一直以来是医学深度学习的难点。

小样本学习，顾名思义就是在有标注样本数据量有限的前提下，通过特定的学习方法从有限的数据中发掘有效信息并得到较优的模型性能，近年来小样本学习越来越受到研究人员的关注。目前针对小样本学习主流解决方法包含有以下三种：基于数据增强的方法、迁移学习方法、元学习等方法；其中基于数据增强的方法主要通过增大训练样本数据容量来满足有监督深度学习的需求，从数据容量角度解决小样本问题；而基于迁移性学习和元学习的方法有一些相似之处，即都是先在大量丰富的基类数据集汲取表征知识，之后再将提取到的知识转移到目标任务数据集上，从而达到利用极少数有标注样本也可以快速泛化的目的，只不过两种方法的知识转移方式有所不同，迁移学习的知识转移是通过利用目标数据集对在基类数据集上得到的预训练模型的部分参数进行微调来实现的，而元学习的知识转移主要是通过构建相似的元任务并以任务的形式学习出好的泛化参数来实现的。

本研究主要探究小样本学习方法，并面向传统小样本图像分类任务和多模态医疗影像小样本数据的诊断任务，从网络模型、学习方法等多角度入手，在目标数据集标定数据不足的情况下，拟搭建轻量级高效的神经网络基础模型来契合小样本数据集，在减小模型体量同时保证模型性能，并利用改进的小样本学习方法使模型在目标数据集上实现较好的分类判断性能。

**1.2 国内外研究现状**

近年来，深度学习方法在人工智能的不同领域取得了令人瞩目的进展[7,8,9]。然而，大多数有监督的深度学习方法需要大量的有标签数据来支持模型的训练过程，而大规模标准数据集的收集工作需要花费大量的时间成本，如果采用深度学习方法和有限的有标注数据直接进

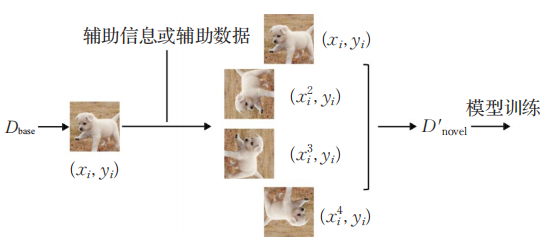
行模型训练，往往会导致模型过拟合于少数训练数据而无法推广到其他相似数据的情况。因此，是否能够从有限数量的有标注数据学习一个性能较优且鲁邦的模型成为了研究者们关注的又一重点。受人类快速学习能力的启发，更多的研究工作致力于如果仅通过的少量样本就可以对深度学习模型进行学习，即小样本学习[10,11]。如何让小样本学习模型掌握快速适应新环境和新任务的能力，是一个极具有挑战性的问题，因为只有少量的数据可用于调整模型，所以如何从少量的样本中学习新的概念知识，是小样本学习所面临的挑战[12]。对于深度神经网络下的小样本学习，按照方法可将小样本学习分为基于数据增强的小样本学习算法、基于迁移学习的小样本学习算法和基于元学习的小样本学习算法，目前国内外已经有了一些相关的研究工作。

同时，随着影像成像技术的发展，人体同一组织由不同的成像方式所得的多种图像各自有各自的优势，它们的信息通常是具有互补性的。找到合适的方案综合利用多种模态的医疗影像，综合多种成像结果的优点，发挥数据类型互补的优势在使用多个模态融合决策时，寻找互补的模态数据进行融合能有效提升诊断效果。基于多模态融合的医疗影像诊断工作近几年越来越受到人们的关注，研究人员通过数据融合、特征融合或决策融合等方式对多模态影像数据进行处理，并在病变定性诊断、病灶区域分割等方面取得了一定成果。

以下对各种小样本学习方法和医学影像多模态融合方法的研究现状进行介绍。

**1.2.1 基于数据增强的小样本学习算法**

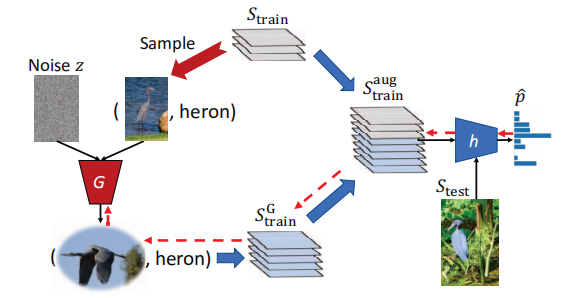
小样本学习所面临的最主要难题是有监督样本数量过少，较少的样本量容易直接导致样本多样性过低，而在这种前提下训练有监督的深度学习模型，则非常容易出现模型过度拟合于训练数据样本，即使得模型学习的信息过于具体而不够泛化，当使用其他同分布下的同类别数据进行测试时，模型也较难在这些数据上表现出较优的性能。因此扩充训练数据容量，丰富训练数据多样性就成为了一种小样本学习方法的解决思路，然而在许多应用场景中，因为特定数据的采集和标注价格十分昂贵，导致再扩充大量的有标记数据十分困难。故对已有的有标注数据进行数据增强(data augmentation)[13]等操作，可以有效便捷得提高样本多样性。数据增强也叫数据扩充，是指借助辅助数据、辅助信息和增强算法，在原有的小样本数据集上生成新的样本数据或特征向量，以此达到数据增广或特征增强的目的。

图1、基于几何变换的数据增强方法

在早期的工作中，很多训练方法的数据预处理步骤通过对现有的小样本数据集进行几何变换和随机剪裁等操作来生成新的样本，其具体操作方法是对数据集中的一个样本进行移位（translation）[14]、旋转（rotation）[15]一定角度、 缩放（scale）[16,17]、裁剪（crop）[16-18]、翻转（flip）[18,19]等操作，处理后生成的新图像与原始图像共享同一个标签。这种利用几何变换等操作生成数据的方式均是对单个原始样本处理生成多个样本的步骤，且增强过程不会改变数据的标签。近几年也逐步出现了同时针对多个样本进行操作的一些数据增强方法，即利用样本间信息的互补性生成新的样本，如 SMOTE[20]、SamplePairing[21]、Mixup[22]、CutMix[23]、FMix[24]等算法，这些方法在将两副图像按一定比例进行融合、融合图像的标签采用其中一张图像的标签或者两张图像标签融合的结果。以上方法均可以简单快速得获得大量新数据帮助神经网络进行训练，适用于许多神经网络训练任务。

然而利用数据自身或数据间信息直接做数据增强的方法，所生成的数据与原始数据在特征空间分布较近，即样本多样性的提高程度十分有限。因此有研究人员就提出利用神经网络来代替传统方法进行多样性数据的生成工作，得益于生成对抗网络的提出（Generative Adversarial Networks，GAN）[25]，一些工作将GAN加入到传统的小样本任务中，如Meta-GAN[26]，利用传统有监督分类器和生成图像判别器损失结合的方式训练整理网络。此外，一些工作移除了判断器的训练过程，只利用目标任务上性能的约束来完成生成器的训练，例如Wang等人[27]采用一个多层感知器作为生成器，输入原始图像和一个随机噪声，将输出的生成图像加入到原始训练集中，利用最终的小样本分类损失来更新小样本网络和生成网络的参数，如图2所示。还有一些工作[28]将待增强的小样本数据和一组特征映射关系输入到生成器中，希望通过生成器采用对应的映射关系对原始数据进行新样本的生成，公式如，其中和是类别a两个样本的特征向量，表示待增广的类别b中一个样本的特征向量。

无论是在数据层还是在特征层对样本进行增广，其目的都是为了丰富样本多样性并扩充数据集，从而减少模型过度拟合于个别特定样本的风险，基于数据增强的小样本学习方法作为一种直观简便的方式，已经取了一些显著性的成果。

图2、《Low-Shot Learning from Imaginary Data》网络结构图

**1.2.2 基于迁移学习的小样本学习算法**

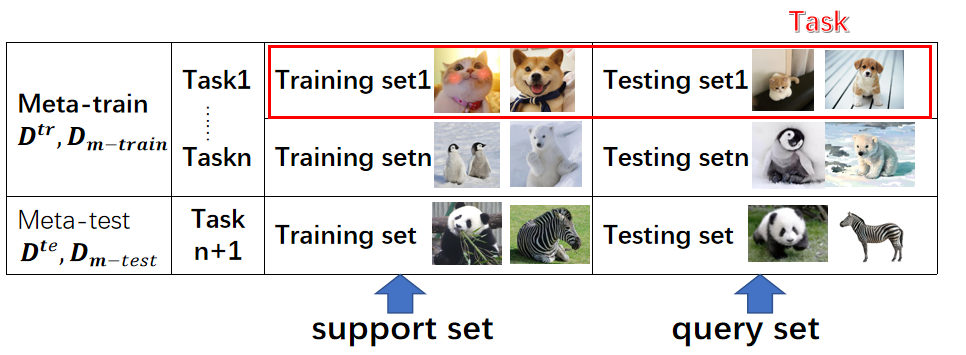
迁移学习的主要思想是利用以前学到的知识来辅助新知识的学习，主要目标是将已经学会的源领域知识快速地迁移到一个目标领域中，使得先验知识可以帮助模型在目标领域数据上进行快速泛化[30]。如何面对特定的源域-目标域搭建相应的迁移桥梁是迁移学习研究中的关键问题。一般来说，源领域和目标领域之间的关联性越强,那么迁移学习的效果就会越好[29]。目前主流的迁移学习方法可以分类两种：一种有效直观的方式是采用预训练模型再微调的迁移学习策略，即先在一个大规模数据集上对模型进行预训练，之后在少量目标数据集上进行微调[31]，一些研究[32]也表明预先训练的模型比随机初始化的模型具有更好的泛化性；另一种方式[33]是将预训练模型作为基础的骨架模型，并在其结构后增加特定的下游任务，如图像分类、目标检测、图像分割、自然语言处理等。

一些基于迁移学习的小样本学习任务已经取得了一定的成果，例如Hariharan等人[34]先在基础数据集上对整个网络进行预训练，之后冻结特征提取器的权重参数，并在目标数据集和基础数据集共同组成的数据集上对分类器进行微调，并在微调过程提出了一种平方梯度幅度损失来约束来建立目标数据微调参数和源域预训练参数的联系。Spyros Gidaris等人[35]利用一个注意力模块动态得为新类别分类器生成其权重参数。Qi 等人[18]提出了一种印记思想，即通过基类数据上训练的特征提取器和新类特征的样本均值生成新类别分类器的权重。Yu等人[36]在印记思想的基础上，利用性能较为稳定的半监督方法 MixMatch对分类器权重进行微调。

研究成果表明通过将大规模数据集上预训练模型迁移到样本量较少的目标数据集上，可以在极少有标注样本的前提下实现快速的泛化和较好的性能，且预训练的迁移学习思想也广泛得应用于许多深度学习任务的性能调优阶段。

**1.2.3 基于元学习的小样本学习算法**

元学习的小样本学习方法与迁移学习类似，只不过在学习过程中数据是以任务的形式引入的。元学习也被称为学会学习，使得机器也想人一样掌握快速学习的能力，其目标是学习一种仅通过的几个训练样本示例就可掌握一种新概念的学习范式。基于元学习的学习方法通常分为两个阶段，即元训练阶段和元测试阶段。通常元测试阶段就是仅含有几个有标注样本的小样本目标任务，根据目标任务中所含样本的类别个数和每种类别所含有标签样本的个数，可以将当前小样本问题成为N-way K-shot问题，其中N表示小样本问题中待识别的新类别总数，K表示每种新类别所含有标签可供网络学习的样本个数，一般的小样本问题N通常取5、10，K通常取小于等于20，当K=1时，小样本问题也叫one-shot learning。在训练之前，数据集会按类别划分为元训练集、元测试集，例如ImageNet数据集中800个类别用于元训练、200个类别用作元测试。在元训练阶段，每次迭代过程中会根据目标小样本任务的形式从大量元训练集的类别中采样N个类别，并且每个类别采样K个样本，由此构建出一个N-way K-shot的任务，而N个类别的剩余样本用来作为当前任务的测试集。这种每次迭代的训练资料都是以任务形式构建的学习策略也被称为Episode training，通常每个任务中有标签用以支持训练的数据集合称为支持集（Support set），也就是每个任务中的N\*K个样本，而用以预测评估的数据集合称为查询集（Query set）。也有一些文章，将元训练集任务中的支持集称为样本集（Sample set），将元测试集任务中的查询集称为测试集（Test set）。

图3、基于元学习的小样本学习方法数据集组织形式

通过在含有大量类别和数据量的基类数据上不断得模拟目标小样本任务形式采样相似任务，利用丰富的任务进行元学习，使得模型掌握学会学习的能力。按照模型参数的更新方式和模型架构，基于元学习的小样本学习方法又可以划分为三种：基于优化的小样本学习，基于外部记忆的小样本学习，基于度量的小样本学习。

**1.2.3.1 基于优化的小样本学习算法**

基于优化的小样本学习方法通常需要一个元学习器（Meta-learner）和一个基础学习器（Base-learner），其核心思想是学习一个好的元学习器，使得该元学习器在面对不同任务时，能针对特定任务快速得产生一个适用于当前任务的基础学习器（模型）。通俗来讲，对于许多不同的小样本学习任务，它们之间可能存在一些共性特征，而元学习的训练就是在捕获这些共性特征，当面对特定任务时，元学习器会根据每个任务的特性特征做出相应的调整，使得模型可以快速得适应新任务。基于优化的小样本学习算法的参数训练分为两个阶段（two-loop）：1、在内循环中（inner-loop），利用每个任务的支持集更新出每个任务的基础学习器；2、在外循环中（outer-loop），利用每个任务的查询集来评价之前更新出的基础学习器性能，利用所有任务基础学习器的性能指标更新元学习器参数。

MAML[37]作为一种代表性的基于优化的元学习方法，旨在利用元学习器学习一个适用于所有任务的初始化参数，当面对特定任务时，这个初始化参数可以根据极少的有标签样本和仅一到两次的迭代过程为目标任务更新出基础学习器的参数。在MAML的基础上，Li等人[38]提出了一种新型元学习优化器（MetaSGD）用于小样本学习，与MAML相比，Meta-SGD除了训练一个好的初始化参数外，还训练了参数的更新方向以及学习速率。Reptile[39]在MAML的基础上提出一种更为简单参数初始化的元学习算法，即对一次迭代中所有任务更新出一个基础学习器，并利用基础学习器参数和初始化参数的差异直接作为初始化参数的更新方向。LEO网络[40]则是在一个低维的参数隐空间中进行元学习，以此优化了MAML更新参数过多的问题。Meta-transfer [41]则是将基于优化的元学习方法和迁移预训练方法相结合，并指出MAML需更新维护参数较多导致更新计算代价昂贵的问题，Meta-transfer利用在保留预训练网络参数的基础上学习每层网络的缩放和平移参数，在缩小参数学习量的同时避免了灾难性遗忘的问题。LST[42]在Meta-transfer的基础上将半监督方法引入其中，利用置信度高的无标签样本帮助网络进行训练。

**1.2.3.2 基于记忆的****小样本学习算法**

基于记忆的小样本学习方法因其复杂的模型设计和需存储长期记忆的内存需要，在近两年的发展中较其他两种方法而言较为缓慢，其主要思想是通过LSTM或者RNN等模型，迭代性得将不同样本输入网络，利用其隐藏激活单元或外部内存积累以往知识并解决当前问题。

Santoro 等人[15]提出的基于记忆增强的神经网络 （Memory-Augmented Neural Network，MANN）借鉴了神经图灵机（Neural Turing Machine，NTM）的思想，采用显示的外部记忆模块保留样本特征信息实现短期记忆，并利用元学习算法优化 NTM 的读取和写入过程， 再结合神经网络自身具备的长期记忆能力共同实现小样本学习任务。Spyros等人[35]提出一种不会遗忘的动态小样本学习算法，核心思想是利用一个小样本权重向量生成器和以往基类的权重向量为新的类别生成其权重向量。Oreshkin等人[43]提出对每个任务的任务信息进行建模，故作者引入了任务嵌入网络（TEN），通过对原有的特征图进行放缩和平移的方式来引入任务信息。Gidaris等人[44]则是用基类类别上训练的特征提取器初始化新类类别的分类权重向量，并将基类类别权重向量和新类类别权重送入一个图神经网络中进行特征重构，得到最终分类器的权重参数。

**1.2.3.3 基于度量的小样本学习算法**

基于度量学习的小样本学习的目标是学习一个好的特征空间和度量函数，利用这个度量空间和度量函数使同类图像的相似度得分最大化，从而优化整个模型。这种基于度量的小样本学习方法通过迭代得将不同的N-way K-shot任务送入特征编码网络和度量网络中，利用端到端的学习方式得到一个适用于所有任务泛化的特征空间和度量函数，在元测试阶段可以直接应用于目标小样本任务而不需要对网络参数进行更新。

孪生网络[45]作为最早的度量学习网络，通过从图像对送入参数共享的特征提取网络，并选择与待预测图像相似度得分最高样本类别作为最终输出。匹配网络[46]使用注意力LSTM模型学习在整个支持集上下文的嵌入，并使用余弦距离计算相似度。原型网络[47]通过比较查询集与支持集中的类别原型进行分类，通过将支持集中同一类别的特征向量的均值作为相应类别的原型并采用欧氏距离度量相似度。以上三篇工作都是利用固定的距离度量函数来测量样本间的相似度，而简单的线性度量函数或最近邻分类器可能无法满足一个所有任务泛化度量的概念，因此RelationNet[48]提出利用一个神经网络来学习样本间的相似度，该工作使用一个三层的全连接层来实现网路。Satorras等人[49]则提出利用图神经网络（GCN），来传递样本特性向量之间的信息，再经过一到两层的图卷积后，度量待测节点与其他节点的距离并输出分类结果。Hou等人[50]在度量函数之前加入了一个Cross Attention模块，利用支持集和查询集之间的相似信息来增强待度量的特征向量。Lu等人[51]将Transformer的多头注意力机制引入分类器权重校正阶段，利用查询集的特性信息丰富支持集特征向量。Liu等人[71]认为基于原型度量的方法本质上忽略了每个类的分布，决策边界因类内方差的多样性而存在偏差，故提出了一种基于类的度量尺度机制，为每个类别学习一个度量标量，利用该度量标量优化度量距离，以帮助学习到更具鉴别性和可迁移性的特征表示。

基于度量的小样本学习方法，因其架构精炼效果显著，在小样本问题的发展过程中十分受到欢迎，目前很多小样本工作都是基于度量框架展开的。

**1.2.4 面向医疗影像诊断小样本学习算法**

因数据收集缓慢且收集成本昂贵而造成的小样本问题一直是因为医学人工智能的常见难题，随着众多小样本学习方法的出现，许多研究工作正试图将小样本学习算法融入到医学领域。但无论是基于迁移的小样本方法还是基于元学习的小样本方法，都依赖一个规模较大的先验数据集支持，理想情况下研究者均希望找到一个与目标任务相似且数据组织形式相同的先验数据集，但因为医学数据组织形式较多且较自然图像而言更为复杂，故构建一个标准的先验医学数据集是一个待解决的难点。因此医学小样本算法的进展并不像在自然图像上的小样本算法一样顺利。

即使如此，近几年的一些工作也展现出了医学小样本算法的发展潜力。Medela等人[52]利用基于迁移学习的小样本学习方法，将含有八种类型的原发肿瘤的数字化结直肠癌组织切片作为源域，将结肠、乳房和肺的健康或肿瘤样本作为目标域，利用源域上所学辅助目标域任务的学习，并证实了一个好的源域数据集对小样本迁移学习的重要性。Cai等人[53]将含有4种类别的OCT和胸部X光成像的数据集作为元训练集，将脑部肿瘤分类数据集作为元测试集，以解决脑膜瘤、神经胶质瘤、垂体瘤的三分类问题，但在元训练的过程中，研究人员也将元测试集的数据混入了元训练，虽然扩充了元任务的种类数且性能有所提升，但有些不符合元学习的标注做法。Li等人[54]在MAML框架的基础上进行了改进，根据每个任务的损失学习一个动态调节参数，使每个任务对最终的元损失贡献程度不同，该作者基于此改进架构将四类常见病的元知识迁移到了三种稀有病的诊断问题上。因为构建医学元训练集较为困难，研究人员就是否能将自然图像上学习到的知识迁移到医学数据上展开了一些讨论，Jiang等人[55]就利用自然图像小样本学习的Benchmark——miniImageNet作为元训练集，将含有4种类别的目标医学OCT图像作为元测试集进行训练，也取得了不错的效果。Matsoukas等人[57]的文章也指出，利用从大规模标注自然图像上迁移得到的目标医学模型可以取得与从少量纯医学数据上迁移得到模型相当的性能，也就是说明从自然图像上进行知识迁移是有一定价值的。受大规模标注自然图像数据集的启发，为了更好得帮助医学小样本学习的发展，一些研究人员开始着手构建医学大规模标注数据集，Chen等人[56]收集了众多不同形式的医学数据，利用这些组合数据预训练了一个网络框架MedicalNet，该研究指出此强大的预训练框架可以作为其他医学小样本问题的基础框架，可以直接衔接许多不同的下游任务，如分类、分割等。但目前的医学影像小样本任务多数是在单一模态组成的目标数据集上进行的，将多模态融合的方法引入小样本任务学习的工作鲜有存在，同时一些小样本知识迁移工作会在同一个数据集内部的不同类别间进行，跨多个先验数据集进行知识迁移的工作较少。

**1.2.5 医学影像多模态融合方法**

在多模态融合方法的应用方面，2017年，Xie等人[64]在使用深度学习方法对肺结节的良恶性进行判别的任务中，提出了一种TMME架构，即使用了3个不进行参数共享的ResNet-50分别用于提取三个模态的增强CT的特征，并在决策层进行融合，相比当年其他肺结节良恶性判别的模型中取得了很大的进步。2018年，Lee等人[65]提出了一种使用组卷积的方法用于提取不同模态的增强CT影像的特征用于肝脏病变区域的检测方法，将四种不同模态的增强CT图像沿着通道维进行拼接，再输入网络进行组卷积提取特征。2019年，Guo等人[66]将MR影像、CT影像和PET影像(正电子发射断层成像)进行融合并对软组织肉瘤病变进行分割，并证实了数据融合和特征融合效果要优于决策融合。Zhou等人[67]利用线性相关块来学习模态之间的相关性，同时利用双重注意的融合块沿着模态通道维度和空间维度对特征进行重加权，将融合后的特征表示输入解码器投影得到分割结果。Zhou等人[69]将SE Block和Dense Block结合，并对由DCE-MRI中动脉期、静脉期和延迟期三个模态组成的3D图像进行HCC分级工作，在计算3D数据中的每个通道的数据对最后分类任务贡献度的同时一定程度上缓解梯度消失的问题。He等人[68]针对不同模态设计了特定于模态的特征提取网络，即利用多尺度的注意力模块来提取纹理信息较多的眼底图像的局部和全局特征，利用区域引导的注意力模块提取背景区域较大的OCT图像特征，最终将两个模态的特征融合为多模态特征用于视网膜图像分类。Jia等人[70]将SE-Block和残差计算块结合并引入小样本度量学习损失，提出了一种新的网络架构MACT，该网络利用SE Block 计算T2WI、T1同反相位等多个模态对最终HCC分级任务的贡献程度，在关注重要模态通道的同时抑制非重要模态通道对最终诊断的影响。

然而现有的大多数多模态融合诊断方法，通过设计复杂的网络结构和注意力机制来计算每个模态的特征或多个模态的融合特征，而这种通过利用复杂结构化设计在提升特征表示质量的同时，也不可避免地增加了模型的复杂性。当面临小样本问题时，过多的模型参数易导致模型过拟合现象的出现，故现有的小样本网络框架的骨干网络大多采用参数量较少的浅层网络来实现，如Resnet12或WRN28。所以如何在不引入过多参数的前提下，有效得计算各个单模态特征表示和多模态融合特征表示，提升骨干网络提取到的特征表示质量，是现在影像多模态小样本网络待改进的一个研究点。

**2、选题的研究内容、研究目标以及拟解决的关键问题等**

**2.1 研究内容**

本课题针对医学影像小样本数据的分类任务，搭建轻量高效的神经网络骨架模型，利用较少的模型参数体量来贴合小样本数据集，将构建好的轻量级骨架模型与迁移学习和元学习方法相结合，利用相似任务数据集上汲取到的知识来帮助目标影像数据的学习，同时利用多模态融合和度量学习等思路，达到通过有限的有标注数据获得性能较好的模型的目的。因此本课题计划从两方面进行：

（1）基于轻量级多模态特征提取网络的医学小样本多分类方法研究

目前，传统的深度神经网络[8,9,58]会通过设计非常深层的神经网络结构来捕获表达能力更强的深度特征，或者通过设置大规模的特征通道数和较多的卷积操作来获取更好的模型性能。然而深层大规模模型的设计往往存在大量的冗余，较高的参数量严重限制了其在小样本数据上的发展与应用，且深度神经网络这对存储设备和计算资源的要求非常高。近年来轻量级神经网络模块和模型快速发展[59,60]，轻量级神经网络通过合理地减少卷积核的数量、减少目标特征的通道数、结合设计更高效的卷积操作等方式，搭建更加快速有效的神经网络模型，在保持神经网络性能的前提下，显著地减少网络的参数和计算量，并提升模型计算速度。

因此本研究拟探索适用于多模态小规模数据的轻量级神经网络架构，在充分利用影像学各个序列模态特征的同时减小网络参数，并引入度量学习、对比学习等必要算法，帮助网络模型在数据量较少的情况下学习有判别性的表征，从而搭建出轻量级多模态影像分类诊断框架。将构建的轻量级多模态框架与数据增强等小样本学习方法相结合，在已有的肝脏局灶性病变数据集上进行实验，对肝脏局灶性病变肿瘤区域的良恶进行区分，并进一步推广到局灶性病变的具体诊断任务上，达到仅利用少量有标注数据就可以实现好的分类诊断模型的目的。

（2）基于知识渐进与原型校正的元迁移小样本病变多分类方法研究

当有标注数据较少的，试图通过有监督学习学习到一个好的模型是非常贪心和奢侈的；虽然数据增强等方法可以有效得缓解过拟合，直观快速得提升模型性能，但有学者指出通过原始图像生成的扩充图像信息单一，不能完全起到丰富数据多样性的目的。目前大多数小样本学习方法大多是采用知识迁移的思路实现的，即迁移学习和元学习。在进行迁移学习或元学习时，在准备目标数据集的同时还需准备一个为网络提供可转移知识的源域数据集或元训练集。迁移学习的通用方法是先在源域数据上进行预训练，之后利用目标数据集对预训练模型的部分参数进行微调；而元学习是通过模拟目标任务从而对元训练集进行相似任务的构建，并利用Episode Training的思想进行迭代训练，例如一个5-way 1-shot的目标分类任务（即5个不同的类别，每个类别有标签的数据只有一例），在元训练阶段每次会从基础类别数据中随机采样出5个类别，每个类别仅采样一个有标签图像，剩余作为测试图像，由此来构造出一个5-way 1-shot的分类任务，通过学习不同的小样本分类任务（猫狗分类、鱼鸟分类等），使得模型掌握快速泛化的能力和通用的元知识（如初始化参数），这样在面对标注样本较少的目标数据集时，也可以实现快速适应泛化。小样本迁移学习旨在模拟人类利用先验知识不断学习的过程，而小样本元学习的设计初衷也是为了模拟人类仅通过少数有标注数据就可以进行快速学习和认知的能力，因此本研究拟采用迁移学习和元学习结合的方法对小样本学习任务进行建模。

然而迁移学习和元学习方法的实现，需要一个类别丰富且数据量较大的先验数据集作为支持，理想情况下先验数据集的数据分布形式应与目标任务尽可能得相似，当先验数据集的类别数据或是数据量不够充足时，模型提取到的可迁移知识往往不具有充足的泛化性。在自然图像视觉任务中，研究者常常使用ImageNet 等大型公开数据集中的几百个数据类别作为基类数据集，通过对基类数据的学习从而扩展到剩余未见类别样本数据的识别问题上。然而在医学影像小样本领域，目前并没有公开的标准基类数据集可直接供迁移学习和元学习使用，同时医学影像数据组织形式更为复杂多样，医学先验数据集和与目标相近的先验任务的构建始终是基于迁移学习和基于元学习的小样本方法的难点。

本研究拟在上一项研究的基础上，将轻量级神经网络作为特征编码模块，拟采用迁移学习和元学习结合的方式，根据公共的自然图像benchmark数据集和大规模相近医学图像数据集设置多个阶段的可转移泛化任务，实现从大规模自然图像->医学图像->易区分的良性肿瘤判别任务->难区分的恶性肿瘤判别任务的迁移过程。为了防止医学影像数据集类别较少而导致的构建任务单一的缺陷，本研究拟采用课程学习和簇对抗任务划分思想来扩充先验任务的多样性提升模型的泛化能力，同时在模型计算过程中引入多模态融合和原型特征校正等方法，提升特征的可判别性和准确性。最终在目标任务上实现良好的分类诊断性能。

**2.2 研究目标**

本课题旨在病变定性诊断目标任务数据量较少的前提下，搭建高效的轻量级特征提取框架，并结合多种小样本学习方法，达到仅利用少量有标注数据即可得到性能较优的目标模型的目的。

**2.3 拟解决的关键问题**

深度学习已在许多领域展现出了优秀的发展潜力，然而传统的深度学习方法均依赖于大规模数据的支持，当所用于深度网络训练的数据量不足且模型参数过多时，深度模型易因过拟合而出现性能不佳的现象。因数据收集缓慢和标注复杂等原因造成的小样本问题一直以来是医学影像智能分析技术的难点。因此，如果再可用目标数据较少的情况下，获得好的模型性能尤为重要。

因此本课题针对小样本医疗影像分类诊断任务，拟解决以下三点关键问题：

（1）如果搭建参数量较少的网络模型来匹配较少的数据体量

缓解过拟合的两个常用方法是减小模型参数量或增大数据容量。本研究需要在保证考虑各各序列模态信息和基本注意力权重计算的前提下，尽量利用更高效的计算方式缩小模型参数量，并保证模型的有效性。

（2）如何构建提供可转移知识的基类数据集并进行元迁移学习

小样本元学习和迁移学习方法需要丰富的先验数据集作为支持，然而与目标医学任务相近的基类数据集构建较为困难，如何找寻与目标任务尽可能相似的先验数据并设计算法实现基类数据集到目标数据集的知识迁移。

（3）如何保证利用元迁移学习后提取到的特征的无偏性

在目标数据集上泛化的过程中，如果保证由先验数据训练集学习得到的模型所提取到目标数据样本特征的无偏性，即在保留先验知识共性的同时需提取到目标数据的特性。

**3、拟采取的研究方案（研究思路、技术路线或研究方法）及可行性分析**

**3.1 研究方案**

本文依照上述研究内容，设计了如下研究方案，主要包含两部分：轻量级特征提取网络的构建、知识递进的元迁移算法研究。

（1）轻量级特征提取网络的构建

* 基于数据增强的小样本肝脏局灶性病变数据预处理方法：

本研究的目标数据来源于首都医科大学附属北京友谊医院，数据形式由T1同反相位、T2加权成像等多种模态核磁共振序列组成，数据容量共计310例病人，相关数据的采集、筛选和标注主要由经验丰富的放射科医生和研究生完成。该数据集包含6种常见的肝脏局灶性病变，其中包括良性病变：肝血管瘤（HEM）95例、局灶性增生结节（FNH）28例、肝脓肿（ABSCESS）40例、肝囊肿（CYST）78例，恶性病变：肝内胆管细胞癌（ICC）26例和肝转移性肿瘤（MET）43例。

由于每种模态序列从不同的角度和成像方式反映病变。例如，T1WI倾向于获取组织的解剖结构，T2WI对出血更为敏感，在确定病变范围方面起着重要作用，DCE-MRI则以动态、连续的增强对比特征反映病变的形态和灌注信息.每个模态都有其特定的诊断和判读价值。通过与专业医师的沟通，依据医生对肝脏局灶性病变影像的判读经验，本研究最终选取了8种序列模态进行实验，如下图所示。

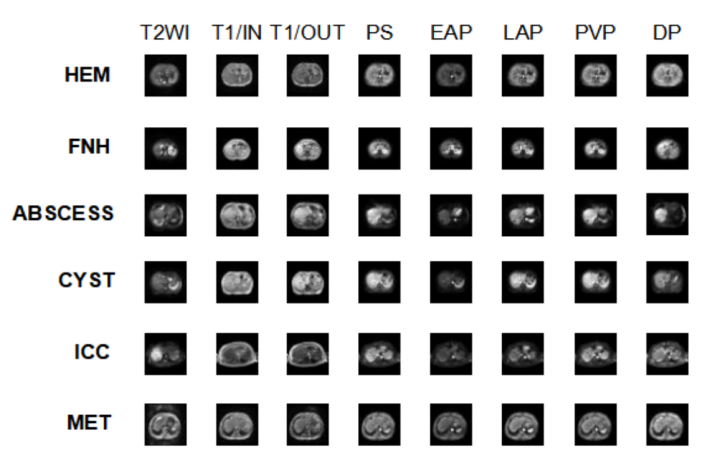


图4、肝脏局灶性病变种类和影像模态构成

T2WI表示T2加权成像、T1/IN表示T1同相位成像，T1/OUT表示T1反相位成像，PS、EAP、LAP、PVP、DP分别表示磁共振动态增强成像平扫期、动脉早期期、动脉晚期期、门静脉期和延迟期的影像，即造影剂注入肝脏组织后0s, 16s, 26s, 60s, 180s时得到的MR影像。通过对原有310位病人数据进行筛选，最终得到了同时含有8种全模态序列影像的病人191例。

本研究先根据医生的像素级标注对待判别的肝局灶性病变ROI区域进行提取，之后将同一病人各个模态的病灶ROI区域图像沿通道维度进行拼接。由于MRI具有12个灰度级，各像素点间数值差异大，且存在小部分数值较高的噪声点，因此需要进行灰度均衡化操作，如公式(1)所示：

 (1)

其中为归一化后的图像灰度值，是原始灰度值，为最小灰度值，为前1%的像素灰度值。

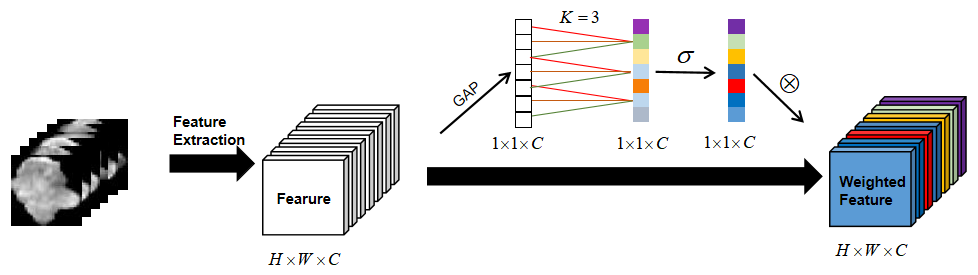
另外由于目标医疗影像数据集体量较小，为了以最直观得方式简洁直观的方式验证本研究提出的轻量级框架性能，防止因训练样本数据较少而产生的过拟合问题，同时在原始数据的基础上模拟临床病例的可能形式，本研究拟采用基于数据增强的小样本学习方法对原始数据集进行处理，包括顺时针旋转90°、180°、270°、随机上下翻转、随机左右翻转、对角转置、副对角线转置8种几何变化操作，以此达到扩充数据体量的目的。本研究整体的数据组成如表1所示。

**表1** 肝局灶性病变数据集的数据组成

|  |  |  |  |
| --- | --- | --- | --- |
| 病种 | 原始病例数 | 全模态病例数 | 扩充后病例数 |
| HEM | 95 | 45 | 360 |
| FNH | 28 | 21 | 168 |
| Abscess | 40 | 31 | 248 |
| CYST | 78 | 51 | 408 |
| ICC | 26 | 17 | 136 |
| MET | 43 | 26 | 208 |
| 总计 | 310 | 191 | 1528 |

* 轻量级多模态贡献度网络框架：

由于每种MRI模态包含与诊断相关的信息特征不同，各影像模态在诊断中的贡献程度是不同的，放射科医师在临床诊断过程中通常会根据病变类型对每种影像模态给予不同的关注。考虑到医生的这种临床判读经验，本研究采用了通道模态注意力机制对各模态影像学特征进行贡献度计算。2018年Hu 等人[61]提出SE-Block通道注意力计算模块，其主要思想是通过对全局平均池化后特征向量进行降维再恢复操作，从而得到各个特征通道的重要程度，使网络自适应学习重要通道上的特征，减低对不重要通道特征的关注，一些研究也在SE-Block的基础上开发更复杂的注意力模块或引入额外的空间注意力以获得更好的性能。然而2020年Wang等人[62]的研究工作指出，虽然基于SE-Block的通道注意力方法带了明显的性能改善，但是复杂的注意力计算方式和较多的参数量会带了严重的计算负担，且降维的计算方式会对通道注意力的预测带来负面影响，因此一种极轻量级的注意力模块ECA-Block被提出。ECA-Block的核心是通过自适应一维卷积来高效得获得跨通道交互信息，这样在维护较少参数量的前提下可以带来显著的性能提升和更高效的计算速度。

  
图5、高效模态通道注意力计算机制

高效的通道注意力计算方式可以总结为先通过一个输入的特征通道数，自适应得确认一维卷积大小，其公式如下：

 (2)

C表示输入特征的通道数，和是两个超参数，表示取最邻近奇数操作。即输入的通道数C与跨通道信息交互的范围k成正比。在确定一维卷积核大小之后，只要维护k个参数就可以对通道注意力进行计算：

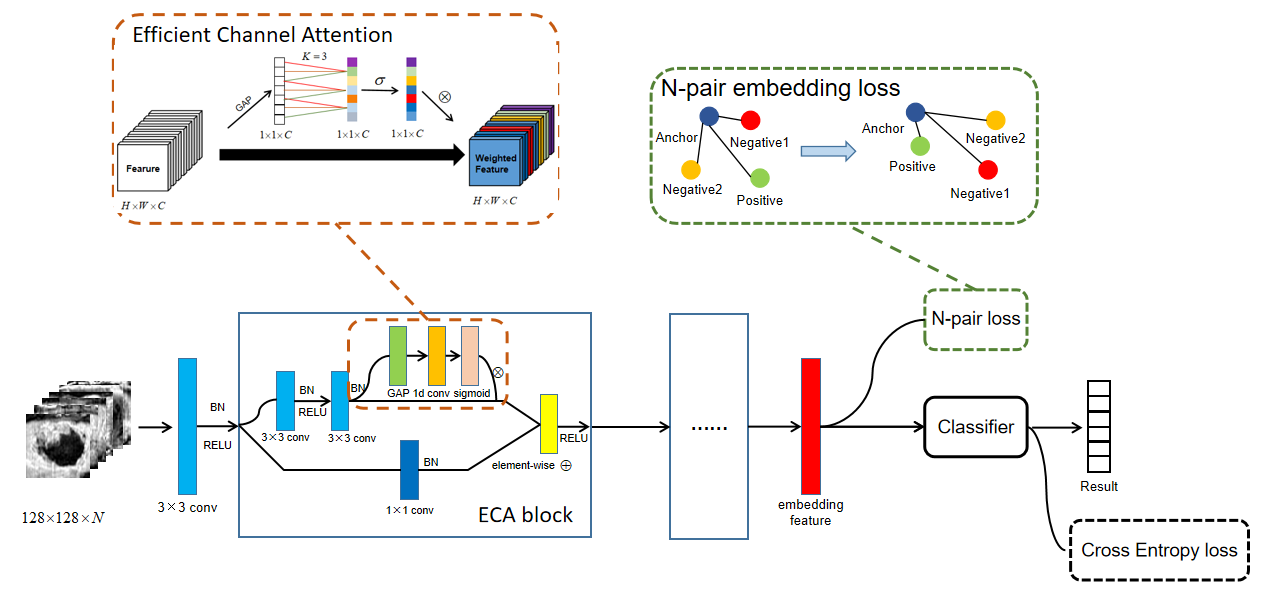
 (3)

C1D表示一维卷积操作，表示Softmax激活函数。这种轻量级的ECA通道注意力模块在降低模型复杂度的前提下适当得捕捉了局部跨通道交互信息，同时保证了模型的计算效率和有效性。

此外为了使学习到编码空间更具有可分辨性，本研究拟引入度量编码损失函数N-pair loss[63]，N-pair loss 可以使模型在面对多分类编码任务时，同类别样本的特征向量在编码空间中尽可能得近，不同类别样本的特征向量尽可能地远，其公式如下：

 (4)

其中，即对于一个batch中每个anchor样本均取一组N-pair样本对，N为当前batch中的类别数，该样本对中包含1个与anchor相同类别的正样本和N-1个与anchor不同类别的负样本，N-pair loss的目标就是在缩短anchor样本与正样本之间距离的同时增大其与其他负样本的距离。在面对多分类问题时，使用N-pair 编码损失相比于Triplet 编码损失更能考虑所有类别样本之间的距离，从而整体性得进行优化，而Triplet 编码损失在每次计算优化时只能考虑两个的类别，当任务中类别数较多时，易出现优化当前两个类别间距离的同时恶化了其他类别间距离的情况。故轻量级多模态贡献度诊断的网路框图如下图所示。

图6、轻量级多模态贡献度局灶性诊断框图

即提取ROI区域的多模态病灶影像，先通过一个3×3的卷积进行降维，再送入多个特征编码层中，本研究中设计了三层的特征编码层，每层含有三个高效通道注意力残差计算块，之后根据mini-batch中的类别标签随机采样N-pair样本对，利用N-pair样本对的编码特征向量间的距离计算N-pair loss，并将Anchor样本的特征向量送入后续的分类器中，利用分类器输出与Anchor样本标签计算交叉熵损失，利用交叉熵损失和N-pair loss共同约束网络参数更新。本研究任务最终的损失函数如公式(5)所示，其中a为加权系数，本研究设置a为0.2。

 (5)

本研究在目标数据集上进行了初步试验，试验结果如下：

**表2** 肝脏局灶性病变数据集上病变六分类定性诊断试验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 模型 | Paras | Acc(%) | Sens(%) | Prec(%) | F1 (%) |
| LeNet | 3.255M | 61.0256±8.3323 | 74.0000±8.0000 | 67.3333±11.6237 | 69.7846±11.0634 |
| Resnet20 | **0.275M** | 66.6667±4.2906 | 69.8333±9.0089 | 63.5000±9.3155 | 75.2024±7.2302 |
| Resnet110 | 1.733M | 62.5641±3.8376 | 67.1667±9.7094 | 65.2000±3.5961 | 71.5460±5.6682 |
| SeResnet20 | 0.280M | 68.2051±5.2798 | 74.5000±9.2736 | 74.3333±10.4137 | 81.3889±5.9964 |
| MCAT | 0.280M | 70.7692±3.4781 | 72.6667±9.0089 | 71.6667±8.8663 | 77.8657±5.8582 |
| ECAResnet20 | 0.278M | 71.2821±3.7684 | 78.6111±13.5104 | 78.2397±12.1413 | 81.7705±9.7618 |
| **EMCAN** | 0.278M | **75.3846±4.7557** | **77.0555±5.8889** | **71.2063±3.1055** | **82.1222±4.7738** |

消融实验：

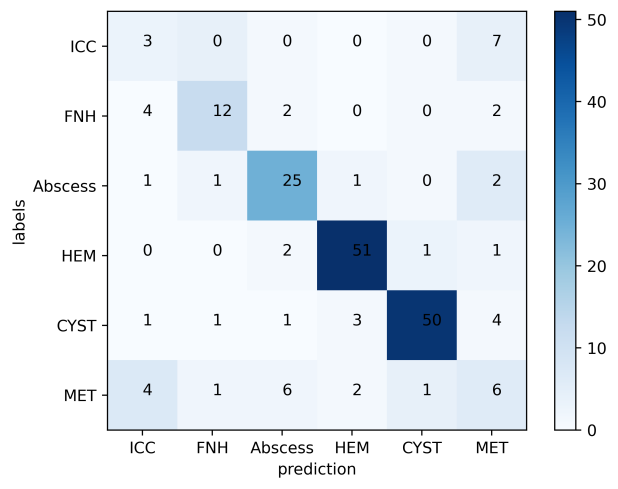
**表3** 肝脏局灶性病变数据集病变六分类任务对比消融实验

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型 | Acc(%) | Sens(%) | Prec(%) | F1 (%) |
| SE\_Triplet | 70.7692±3.4781 | 72.6667±9.0089 | 71.6667±8.8663 | 77.8657±5.8582 |
| SE\_N-pair | 70.7692±2.0513 | 77.2222±7.5359 | 76.0000±8.7939 | 80.2054±3.4481 |
| ECA\_Triplet | 71.2821±5.4754 | 81.7500±6.3047 | 74.5000±6.7823 | 82.9762±3.0023 |
| ECA\_N-pair | **75.3846±4.7557** | **77.0555±5.8889** | **71.2063±3.1055** | **82.1222±4.7738** |

不同分类任务：

**表4** EMCAN在肝脏局灶性病变数据上不同分类任务实验结果

|  |  |  |
| --- | --- | --- |
| 模型 | 分类目标 | Acc(%) |
| EMCAN | 良性肿瘤4分类 | 81.2121±3.5339 |
| 恶性肿瘤2分类 | 40.0000±8.1650 |
| **肿瘤良恶2分类** | **85.1282±3.4016** |
| 全类肿瘤6分类 | 75.3846±4.7557 |

图7、EMCAN肝脏局灶性病变六分类混淆矩阵

综上实验可见，高效的模态通道注意力机制和N-pair度量编码损失对局灶性病变的诊断性能有明显的提升，我们所提出的EMCAN模型在肝局灶性病变诊断任务的表现要优于其他对比方法，且相比于Resnet20而言，EMCAN可以在少幅增加参数量的情况下大幅提升特征提取质量和分类诊断性能。特别得是，EMCAN在肿瘤区域的良恶性判断任务上有着较高的诊断效能，且对良性肿瘤的定性诊断效果较为优异，这得益于肝囊肿和肝脓肿等良性肿瘤病灶表象特征较为明显等现场，然而该模型对于恶性病变的定性诊断效能并不理想，通过观察混淆矩阵可知，两类恶性局灶性病变容易相互混淆，且肝转移瘤易被模型识别为其他肝局灶良性肿瘤，除去恶性病变的行列，单纯观察良性病变的诊断混淆矩阵，其误诊数和漏诊数均在可接受范围内。这与医生诊断的临床场景相似，对于医生而言肝胆管细胞癌和转移瘤相对其他病变而言较难区分，病理学表征更加不明显。

因此针对以上初步试验情况，本工作考虑能否将良性病变和其他公共病变的诊断经验迁移到恶性病变的诊断过程中，从而进一步提升恶性肿瘤定性诊断的效能。

（2）知识递进的元迁移算法研究：  
 实现小样本元学习和迁移学习的前提是需要有一个丰富且与目标数据相近的基类数据集来帮助模型学习可转移知识。根据基类数据构建任务的多样性，直接影响着网络汲取到的可转移知识的质量，当基类数据集本身类别较少且构建出的任务种类不够充足时，网络面对目标任务时所体现的泛化性能会有所不足。本研究拟采用的解决方案为预训练迁移学习和元课程簇对抗任务的思想，利用已有的自然图像数据集和大规模医学图像数据集作为基类数据集，将上一研究的轻量级神经网络架构和改进的小样本学习方法相结合，从而提取出可转移的知识信息，最终泛化到目标任务肝脏局灶性特定病变的分类任务上。同时在目标数据泛化的阶段中，为了防止受由基类数据训练的网络参数影响而导致的特征有偏性，本研究拟采用原型特征校正等方法来提高特征表示的无偏性。本研究技术路线要点分为任务构建、模态无关的轻量级特征提取网络构建、多模态融合加权网络构建、特征原型校正模块四部分，以下是各个技术路线的拟实现细节。

* 任务构建：  
   课程学习的思路是指随着学习阶段的递进，渐进性得增大学习任务的难度，且研究表明先用简单的知识进行训练得到的模型比同时使用简单和困难知识得到的模型性能更加优异，并且简单的知识学得越好，模型最终的泛化性能也越有利。传统的N-Way K-Shot小样本分类问题，其任务形式固定，特别是在以元学习为基础的小样本学习中，其元训练阶段会根据N-way K-Shot的形式从基类数据集中构建多个子任务，但当基类数据集本身类别数不太充足时，构造出的子任务多样性也会相应不足。例如，当前基类数据集中有M个类别，传统的小样本问题会以N-Way K-Shot的形式构建任务，其构建的任务种类为种，当M较小时构建出任务的种类也十分有限，这可能会影响最终小样本模型的泛化性。  
   为了弥补这项缺陷，本研究不再局限于固定的任务形式，而是采用簇对抗思想和课程难度递进的思路构建子任务。即本研究不再局限于依据类别进行划分，而是根据任务目标的不同进行划分，可以同时将二分类目标、三分类目标、更多分类任务一起纳入模型优化过程中，即最终构建的任务种类数为。如Task1（病种1、病种2）、Task2（病种4、病种3）、Task3（病种1、病种2、病种6），Task中的病种类别可能有交叉，但Task之间两两独立没有重合，我的元学习器和迁移模型想要实现的目标是，即使面对不同、全新的病种分类任务，它依然可是很容易得泛化到当前目标任务上。这种任务划分方式摆脱了小样本分类任务的固定形式，丰富了任务多样性，且根据训练的不断进行，可以将任务按难度依次送入网络（如从二分类到多分类），达到递进式学习的目的。这种摆脱了固定形式的小样本学习任务最简单的实现方法可以通过特征编码模块和余弦距离分类器来完成。后续工作会探索更多其他可实现方法。
* 模态无关的轻量级特征提取网络构建

基于度量的小样本学习思路，是通过对不同由元训练集中的基础类别而构建的小样本任务进行学习，从而得到一个通用泛化的特征提取网络和度量网络，使得在面对未见过的新类数据时，在不需要调整模型参数的情况下，可以有效得提取特征并根据待查询样本特征与少量有标注的支持样本特征之间的距离对待查询样本的类别做出准确预测。这种基于度量的小样本学习方法的核心是学习一个好的通用特征编码网络。

因为本研究的目标数据集为核磁共振影像多模态数据，包含T1IN、T1OUT、T2WI、DCE等8个模态序列，目前开源的医学多模态数据集较少，想要找到与目标数据集相似的公共基类数据集十分困难，目前和目标数据集较为接近的数据集为BraTS 2021脑肿瘤多模态数据集，该数据集包含2000例病人的核磁共振影像，共计T1、T2、T1-FAR,T2-FAR4种不同的模态序列。因为公共影像数据集的模态种类和本研究的模态种类重合度较低，因此本研究希望利用RelationNet等基于度量学习元学习网络，对样本数相对较多的公共影像多模态数据集进行学习，得到一个模态无关的特征提取器，使得该特征提取网络可以提取各个模态的共性特征，为后续目标数据集上的特征提取提供基础。

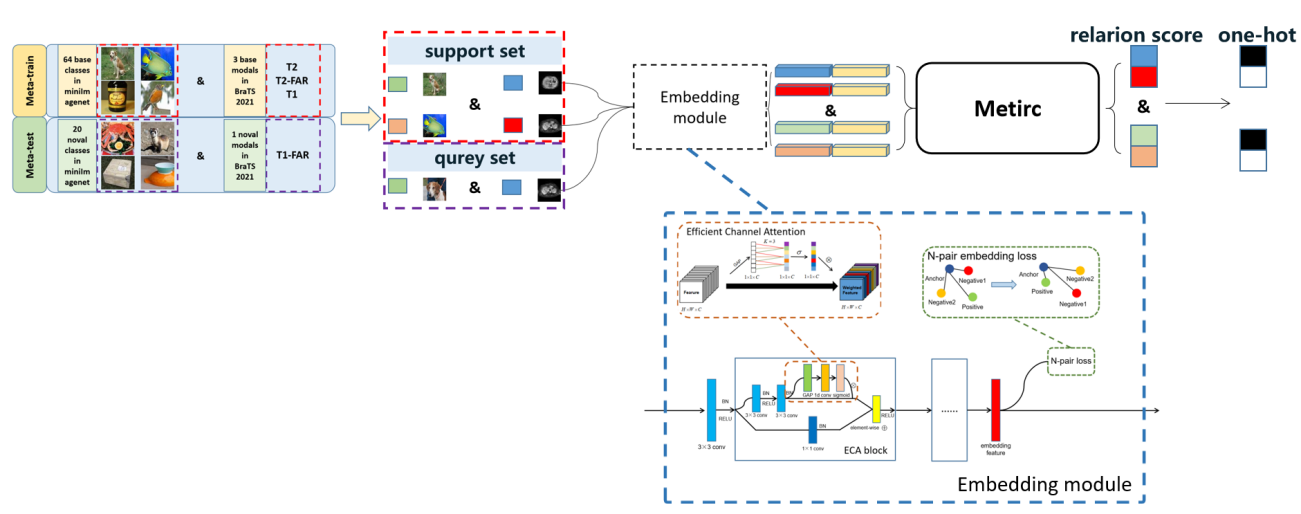


图8、轻量级特征提取网络框架

虽然自然图像和医学影像的组织形式差异较大，但有研究证实：从大规模自然图像上迁移到医学影像数据的模型可以取得和从医学图像迁移到医学图像相比有竞争力的性能；对此现象可以理解为，虽然自然图像和医学图像差别较大，但是网络在大规模自然图像上学习到了通用的特征表征能力和可转移知识，可以理解为人类幼年时期所学习到的认识世界的先验知识。该项研究表明，利用大规模自然图像数据集也可以为后续医学任务上的迁移学习提供一个初步的可迁移参数。因此为了增强通用特征提取器的泛化性，本研究拟采用大规模自然数据集和医学先验数据集相结合的方式，对特征提取器进行训练。  
 本工作拟采用的特征提取器训练方法主要分为3个步骤，具体过程如下：  
 第一步：因自然图像的通道数为三通道，而医学图像的通道为单通道，故需先将自然图 像数据集中的RGB图像转化为灰度图像，公式为I(x,y) = 0.3 \* I\_R(x,y) +0.59 \* I\_G(x,y)+ 0.11 \* I\_B(x,y)，其中I(x,y) 表示转换后的灰度图像像素强度， I\_R(x,y)、I\_G(x,y)、I\_B(x,y)分别表示彩色图像三个通道的像素强度值。  
 第二步：因为BraTS 2021数据中图像级的标签只有良性和恶性之分，故为了与目标任 务相似，元训练任务采用二分类任务；将ImageNet中1000个类别划分 meta-train/meta-val/meta-test为600/200/200的比例，通过在600个基础类别上 构建2-Way K-Shot 任务，在轻量级网络框架的基础上利用Episode 策略进行 学习，待模型收敛后保存模型参数。  
 第三步：利用BraTS 2021数据集中2000个脑部影像，对模型参数进行进一步微调。根 据数据集中4个不同的模态序列，划分其中3个序列为元训练集，1个序列为 元测试集，训练时每次从元训练集的其中一个序列中选出K张良性样本和K 张恶性样本，构建当前序列的2-Way K-shot任务。在进行一定轮数的再训练 后，保存最终的轻量级特征提取网络的参数。

* 多模态融合加权网络构建

在得到模态无关的通用特征提取网络之后，就可以进行第二步的知识转移工作，将自然图像和脑部图像上所获得的良好的模型参数迁移到目标肝局灶性病变诊断任务中。将肝脏局灶性病变数据集按照良性和恶性肿瘤分为元训练集和元测试集，即元训练集包含4个类别、元测试集包含2个类别。为了充分利用各个模态序列的信息，需要对各模态特征进行融合计算，以便考虑更加丰富的特征信息。同时在进行元训练的过程中，采用递进式课程学习的思路进行训练，具体网络结构如下图所示。

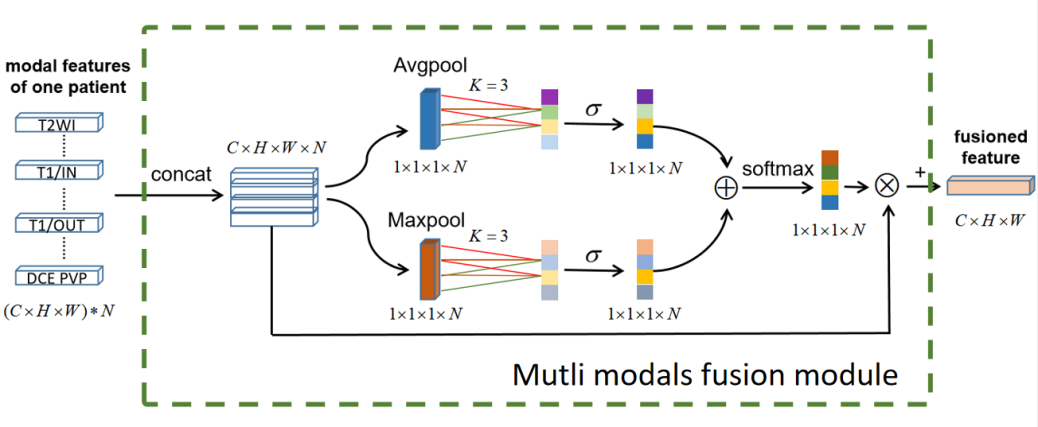


图9、多模态融合加权网络结构图

第一步：加载预训练好的轻量级特征提取网络参数  
 第二步：将输入病例的各个序列影像分别送入预训练好的特征提取网络，分别得到各模 态的特征

第三步：将各个模态的特征延扩展通道维度进行拼接，将拼接后的特征分别做全局平局 池化和全局最大池化操作如公式6、7所示，之后两个池化特征分别送入两个 ECA模块，将两个ECA模块得到的通道注意力权重加权求和并做softmax操 作，如公式8所示，最终将归一化后的模态权重与各模态特征加权求和得到多 模态加权融合特征，如公式9所示。

 (6)

 (7)

 (8)

 (9)

整个过程中冻结特征提取器的参数，只学习后续特征加权网络和原型校正模块的参数。特征原型校正工作会在下一部分介绍。

* 特征原型校正模块设计

由于特征提取器是通过自然图像和先验脑部数据图像元迁移训练得到的，所以虽然特征提取网络在大规模数据的学习下已经具有了一些通用共性特征的表征能力，但是本研究认为利用此特征提取器在不同器官的目标数据集上提取的特征存在一定的有偏性，即特征提取网络对于特定目标数据的特性表征不足。为此许多小样本学习研究也尝试了很多特征校正方法，因为查询样本和支持样本的数据分布是一致的，所以很多研究都是通过不同方法将查询集样本特征加入到支持集特征中去，以此达到提升特征表达和校正特征的目的。故本研究拟通过训练一个特征校正模块，利用查询集和支持集样本的互信息对彼此的多模态加权融合特征进行修正。在网络最终输出层面，采用简单的余弦距离分类器进行实验，通过度量查询集样本特征和校正后多模态加权融合特征之间的余弦距离，输出与查询样本特征相似度最高的支持集原型类别。

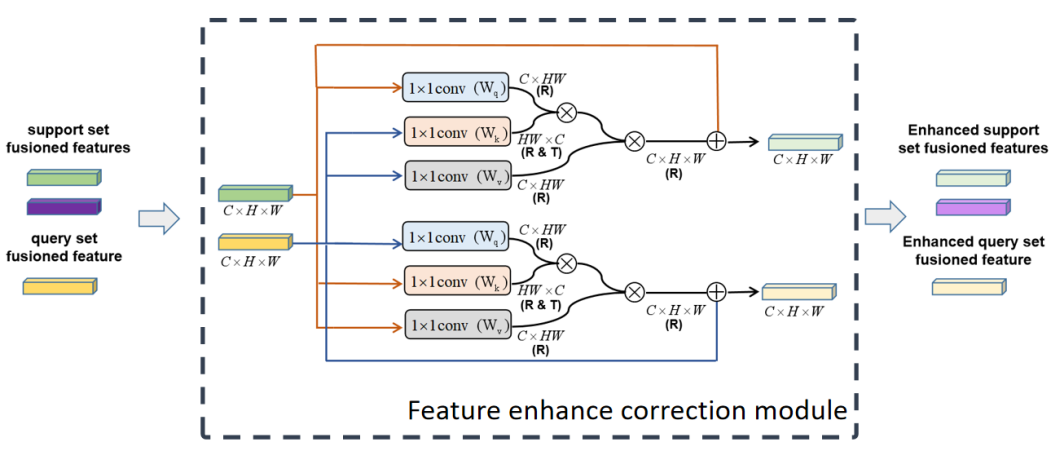


图10、特征原型校正模块结构图

计算方式如下：

 (10)

 (11)

其中分别表示查询集样本和支持集样本的多模态加权融合特征，，，为三个仿射变化的1×1卷积核参数，。表示增强校正之后的查询集样本特征和支持集样本特征。

即第二阶段知识迁移的整体框图如图所示。

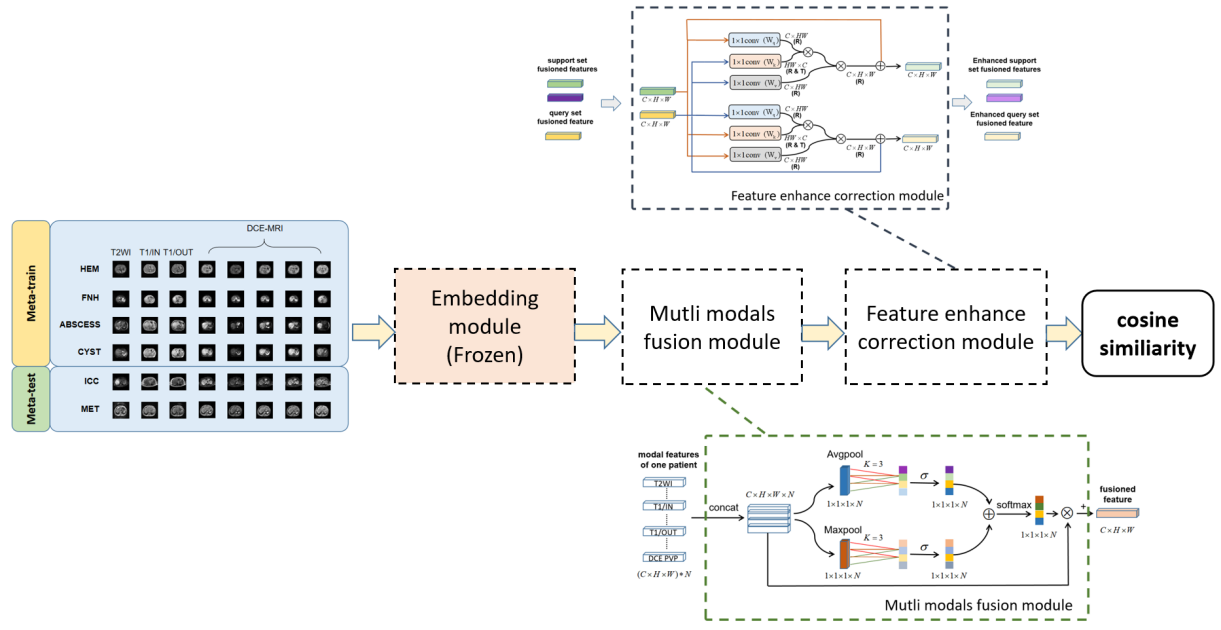


图11、医学公共数据集->局灶性目标数据集知识迁移框图

分类器采用余弦相似度分类器，其公式如下：

 (12)

其中表示分类器的权重向量，其由各各类别支持集样本特征的均值向量构成，表示查询集样本的特征向量。余弦相似度分类器将输出查询集样本加权增强特征和每个支持集类别原型加权增强特征的相似度分数，查询集样本会被预测为相似度最高的支持集样本所属的类别。

**3.2 可行性分析**

（1）理论基础

小样本学习是近年来发展的热点之一，大量顶会顶刊收录小样本学习方向的论文，本研究参考近年ICML、CVPR论文，具备一定的理论基础，同时已在局灶性病变诊断任务，HVPG回归拟合任务和HCC微血管侵犯任务中进行过编码实验，具备一定的代码基础。

（2）实验基础

目前有许多可供研究者直接使用的深度学习框架，包括Caffe、TensorFlow、Pytorch等，本实验前期工作已利用Pytorch框架为肝细胞癌微血管侵犯任务和肝脏HVPG回归预测任务编写实验代码，具有一定的实践基础。

**4、本课题的特色及创新之处（限博士论文开题，硕士可根据课题取舍）**

本课题的特色和创新处主要体现在以下三点：

1. 提出一种面向多模态小样本多分类问题的高效轻量级特征编码网络，提升特征编码质量。
2. 针对医学影像数据集规模较小的问题，引入小样本学习方法，并将多模态融合引入小样本迁移学习和元学习阶段。
3. 整理构建一个可用于医学小样本迁移学习和元学习的先验数据集，并建立标准的知识迁移策略，利用特征校正策略降低多次知识迁移带来的有偏性。
4. **论文研究进度计划（含研究进度计划安排、研究经费预算及解决办法）、预期研究结果**

课题研究进度计划安排如下：

|  |  |
| --- | --- |
| 2021.9～2021.12 | 收集阅读相关文献资料，制定标注方案，开展标定工作 |
| 2022.1～2022.2 | 与医院沟通逐步开展数据标定工作并整理数据集 |
| 2022.3～2022.8 | 搭建、调优、验证小样本学习算法 |
| 2022.9～2023.1 | 优化算法，将其应用于临床辅助诊断 |
| 2023.1～2023.6 | 整理相关资料，撰写毕业论文，完成论文答辩 |

预期研究结果：

构建面向小样本医学影像病变的诊断分类模型，设计高效的知识迁移算法。发表相关的软件著作、专利和论文。

**（二）论文研究工作基础及条件保障**

1. **工作基础（含入学以来取得研究成果、参与或承担的科研项目情况等）**
2. 入学以来已经完成了模式识别、机器学习、数理统计与概率论等课程的学习；
3. 在此基础上阅读了大量的相关文献，拥有进行此课题研究的理论基础；
4. 学校图书馆文献数据库提供了大量的文献资料，为实时把握科研动态提供了条件；
5. 作为项目参与者，参与首都卫生发展科研专项项目《基于多参数 MRI 创建肝静脉压力梯度无创性估测模型》的研究工作，帮助医生完成有关实验。
6. **工作条件（设备条件、实验场地条件、可能遇到的困难及应对措施）**

（1）设备条件：GTX1080型号GPU服务器一台；

（2）实验场地条件：北京工业大学信息楼321实验室；

（3）可能遇到的问题：

跨多个数据集进行元学习有一定的难度：泛化能力不足时，引入域泛化相关知识，查询相关论文，进一步分解图像风格信息；

通过多个数据集构建统一的数据读取形式较为复杂：花费较多时间成本；

单模态通用特征提取器性能不佳：引入在大型医学数据集上公开的预训练模型。

**（三）参考文献（参考文献目录用标准规范格式，在正文引用处须标注）.**

1. 黄子星, 宋彬. 人工智能在腹部影像的研究现状与展望. 放射学实践. 2018; 3(11):1216-1220.
2. Liu T, Liu J, Ma Y, He J, Han J, Ding X, and Chen C T, Spatial feature fusion convolutional network for liver and liver tumor segmentation from CT images. Med Phys, 2021; 48(1): 264-272.
3. Gao J, Jiang Q, Zhou B, Chen D. Convolutional neural networks for computer-aided detection or diagnosis in medical image analysis: An overview. Math Biosci Eng. 2019; 16(6):6536-6561.
4. Jia XB, Xiao YJ, Yang DW, Yang ZH【通讯作者】, Lu Chen. Multi-parametric MRIs based assessment of Hepatocellular Carcinoma Differentiation with Multi-scale ResNet. KSII Transactions on Internet and Information Systems. 2019; 13(10):5179-5196.
5. Yang DW, Jia XB, Xiao YJ, Wang XP, Wang ZC, Yang ZH【通讯作者之一】. Noninvasive Evaluation of the Pathologic Grade of Hepatocellular Carcinoma Using MCF-3DCNN: A Pilot Study. Biomed Res Int. 2019; 2019:9783106.
6. Jia X, Liu Y, Yang ZH【通讯作者】, Yang D. Multi-modality self-attention aware deep network for 3D biomedical segmentation. BMC Med Inform Decis Mak. 2020; 20(Suppl 3):119.
7. A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In NIPS, 2012. 1, 2
8. K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016. 1, 2
9. K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. ICLR, 2015. 1, 2, 7
10. L. Fei-Fei, R. Fergus, and P. Perona. One-shot learning of object categories. TPAMI, 2006. 1, 2
11. B.Lake, R. Salakhutdinov, J. Gross, and J. Tenenbaum. One shot learning of simple visual concepts. In CogSci, 2011. 1, 2, 4
12. Lake, B.M., Linzen, T., & Baroni, M. (2019). Human few-shot learning of compositional instructions. CogSci
13. Royle JA, Dorazio RM, Link WA. Analysis of multinomial models with unknown index using data augmentation. Journal of Computational and Graphical Statistics, 2007,16(1):67−85
14. BENAIM S,WOLF L.One-shot unsupervised cross domain translation[C]//Advances in Neural Information Processing Systems，2018：2104-2114.
15. SANTORO A,BARTUNOV S,BOTVINICK M,et al. Meta-learning with memory-augmented neural networks[C]// Proceedings of International Conference on Machine Learning, 2016：1842-1850
16. ZHANG Y, TANG H, JIA K. Fine-grained visual categorization using meta-learning optimization with sample selection of auxiliary data[C]// Proceedings of International European Conference on Computer Vision, 2018:233-248
17. LAKE B M, SALAKHUTDINOV R, TENENBAUM J B. Human- level concept learning through probabilistic program induction[J].Science, 2015, 350: 1332-1338.
18. QI H, BROWN M, LOWE D G.Low-shot learning with imprinted weights[C]// Proceedings of IEEE Conference on Computer Vision & Pattern Recognition, 2018: 5822-5830.
19. SHYAM P, GUPTA S, DUKKIPATI A.Attentive recurrent comparators[C]/ /Proceedings of International Conference on Machine Learning, 2017: 3173-3181.
20. CHAWLA N V, BOWYER K W.Smote: Synthetic minority over-sampling technique[J].Journal of Artificial Intelligence Research, 2002, 16(1): 321-357.
21. INOUE H.Data augmentation by pairing samples for images classification[J].arXiv: 1801.02929，2018
22. ZHANG H, CISSE M, DAUPHIN Y N, et al. Mixup: Beyond empirical risk minimization[C]// Proceedings of International Conference on Learning Representations, 2018
23. Yun, S., Han, D., Oh, S., Chun, S., Choe, J., & Yoo, Y.J. (2019). CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 6022-6031
24. Harris, E., Marcu, A., Painter, M., Niranjan, M., Prügel-Bennett, A., & Hare, J.S. (2020). FMix: Enhancing Mixed Sample Data Augmentation. arXiv: Learning
25. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A.C., & Bengio, Y. (2014). Generative Adversarial Nets. NIPS
26. Zhang, R., Che, T., Ghahramani, Z., Bengio, Y., & Song, Y. (2018). MetaGAN: An Adversarial Approach to Few-Shot Learning. NeurIPS.
27. Wang, Y., Girshick, R.B., Hebert, M., & Hariharan, B. (2018). Low-Shot Learning from Imaginary Data. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 7278-7286
28. Hariharan, B., & Girshick, R.B. (2017). Low-Shot Visual Recognition by Shrinking and Hallucinating Features. 2017 IEEE International Conference on Computer Vision (ICCV), 3037-3046.
29. Wang H. Research review on transfer learning. Computer Knowledge and Technology, 2017,13(32):203−205 (in Chinese with English abstract).
30. Liu XP, Luan XD, Xie YX, et al. Transfer learning research and algorithm review. Journal of Changsha University, 2018,32(5): 33−36,41 (in Chinese with English abstract).
31. Garcia-Garcia, A., Orts, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P., & Rodríguez, J.G. (2018). A survey on deep learning techniques for image and video semantic segmentation. Appl. Soft Comput., 70, 41-65.
32. D. Erhan, Y. Bengio, A. C. Courville, P. Manzagol, P. Vincent, and S. Bengio. Why does unsupervised pre-training help deep learning? Journal of Machine Learning Research,11:625–660, 2010
33. Tsimpoukelli, M., Menick, J., Cabi, S., Eslami, S.M., Vinyals, O., & Hill, F. (2021). Multimodal Few-Shot Learning with Frozen Language Models. ArXiv, abs/2106.13884
34. Hariharan, B., & Girshick, R.B. (2016). Low-shot visual object recognition. ArXiv, abs/1606.02819
35. Spyros Gidaris and Nikos Komodakis. Dynamic few-shot visual learning without forgetting. In CVPR, pages 4367– 4375, 2018
36. Yu, Z., Chen, L., Cheng, Z., & Luo, J. (2020). TransMatch: A Transfer-Learning Scheme for Semi-Supervised Few-Shot Learning. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 12853-12861
37. Finn, C., Abbeel, P., & Levine, S. (2017). Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. ICML
38. Li, Z., Zhou, F., Chen, F., & Li, H. (2017). Meta-SGD: Learning to Learn Quickly for Few Shot Learning. ArXiv, abs/1707.09835
39. Nichol, A., Achiam, J., & Schulman, J. (2018). On First-Order Meta-Learning Algorithms. ArXiv, abs/1803.02999
40. Rusu, A.A., Rao, D., Sygnowski, J., Vinyals, O., Pascanu, R., Osindero, S., & Hadsell, R. (2019). Meta-Learning with Latent Embedding Optimization. ArXiv, abs/1807.05960
41. Sun, Q., Liu, Y., Chua, T., & Schiele, B. (2019). Meta-Transfer Learning for Few-Shot Learning. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 403-412
42. Sun, Q., Li, X., Liu, Y., Zheng, S., Chua, T., & Schiele, B. (2019). Learning to Self-Train for Semi-Supervised Few-Shot Classification. ArXiv, abs/1906.00562
43. Oreshkin, B.N., López, P.R., & Lacoste, A. (2018). TADAM: Task dependent adaptive metric for improved few-shot learning. NeurIPS
44. Gidaris, S., & Komodakis, N. (2019). Generating Classification Weights With GNN Denoising Autoencoders for Few-Shot Learning. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 21-30
45. Koch, G.R. (2015). Siamese Neural Networks for One-Shot Image Recognition
46. Vinyals, O., Blundell, C., Lillicrap, T.P., Kavukcuoglu, K., & Wierstra, D. (2016). Matching Networks for One Shot Learning. NIPS
47. Snell, J., Swersky, K., & Zemel, R.S. (2017). Prototypical Networks for Few-shot Learning. NIPS
48. Sung, F., Yang, Y., Zhang, L., Xiang, T., Torr, P.H., & Hospedales, T.M. (2018). Learning to Compare: Relation Network for Few-Shot Learning. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 1199-1208
49. Satorras, V.G., & Bruna, J. (2018). Few-Shot Learning with Graph Neural Networks. ArXiv, abs/1711.04043
50. Hou, R., Chang, H., Ma, B., Shan, S., & Chen, X. (2019). Cross Attention Network for Few-shot Classification. ArXiv, abs/1910.07677
51. lu, Z., He, S., Zhu, X., Zhang, L., Song, Y., & Xiang, T. (2021). Simpler is Better: Few-shot Semantic Segmentation with Classifier Weight Transformer. ArXiv, abs/2108.03032
52. Medela, A., Picón, A., Saratxaga, C.L., Belar, O., Cabezón, V., Cicchi, R., Bilbao, R., & Glover, B. (2019). Few Shot Learning in Histopathological Images:Reducing the Need of Labeled Data on Biological Datasets. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 1860-1864
53. Cai, A., Hu, W., & Zheng, J. (2020). Few-Shot Learning for Medical Image Classification. ICANN
54. Li, X., Yu, L., Fu, C., & Heng, P. (2020). Difficulty-aware Meta-Learning for Rare Disease Diagnosis. MICCAI
55. Jiang, X., Ding, L., Havaei, M., Jesson, A., & Matwin, S. (2019). Task Adaptive Metric Space for Medium-Shot Medical Image Classification. MICCAI
56. Chen, S., Ma, K., & Zheng, Y. (2019). Med3D: Transfer Learning for 3D Medical Image Analysis. ArXiv, abs/1904.00625.
57. Matsoukas, C., Haslum, J.F., Soderberg, M.P., & Smith, K. (2021). Is it Time to Replace CNNs with Transformers for Medical Images? ArXiv, abs/2108.09038.
58. Huang G, Liu Z, Maaten LVD, Weinberger K. Densely connected convolutional networks. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2017. 4700−4708
59. Tan, M., & Le, Q.V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ArXiv, abs/1905.11946
60. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. ArXiv, abs/1704.04861
61. Hu, J., Shen, L., Albanie, S., Sun, G., & Wu, E. (2020). Squeeze-and-Excitation Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 42, 2011-2023
62. Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., & Hu, Q. (2020). ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 11531-11539
63. Sohn, K. (2016). Improved Deep Metric Learning with Multi-class N-pair Loss Objective. NIPS
64. Xie, Yutong, et al. "Transferable multi-model ensemble for benign-malignant lung nodule classification on chest CT." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2017.
65. Lee, Sang-gil, et al. "Liver Lesion Detection from Weakly-Labeled Multi-phase CT Volumes with a Grouped Single Shot MultiBox Detector." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2018.
66. Guo, Z., Li, X., Huang, H., Guo, N., & Li, Q. (2019). Deep Learning-Based Image Segmentation on Multimodal Medical Imaging. IEEE Transactions on Radiation and Plasma Medical Sciences, 3, 162-169.
67. Zhou, T., Canu, S., Vera, P., & Ruan, S. (2021). 3D Medical Multi-modal Segmentation Network Guided by Multi-source Correlation Constraint. 2020 25th International Conference on Pattern Recognition (ICPR), 10243-10250.
68. He, X., Deng, Y., Fang, L., & Peng, Q. (2021). Multi-Modal Retinal Image Classification With Modality-Specific Attention Network. IEEE Transactions on Medical Imaging, 40, 1591-1602.
69. Zhou, Q., Zhou, Z., Chen, C., Fan, G., Chen, G., Heng, H., Ji, J., & Dai, Y. (2019). Grading of hepatocellular carcinoma using 3D SE-DenseNet in dynamic enhanced MR images. Computers in biology and medicine, 107, 47-57 .
70. Jia, X., Sun, Z., Mi, Q., Yang, D., & Yang, Z. (2021). A Multimodality-Contribution-Aware TripNet for Histologic Grading of Hepatocellular carcinoma. IEEE/ACM transactions on computational biology and bioinformatics, PP.
71. Liu, G., Zhao, L., Li, W., Guo, D., & Fang, X. (2021). Class-wise Metric Scaling for Improved Few-Shot Classification. 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 586-595.

**三、开题报告评价（本项分别由指导教师及专家组填写）**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **指导教师对该生选题报告的简要评语（**本栏由指导教师在开题报告会之前填写完毕）**：**  医学小样本学习是计算机学科中一个前沿的研究领域，本开题报告针对小样本多模态医学影像的分类诊断问题，对主流方法进行充分的文献调研和分析，确定了探究轻量级多模态特征提取网络和知识渐进元迁移策略的研究内容，重点优化多模态特征计算方式、构建用于知识迁移的先验基类数据集并设计相应的知识迁移任务，拟对轻量级高质量特征计算机制和元迁移学习方法加以研究和实验验证。论文选题先进，具有良好的研究意义和应用价值，所提出了解决方案具有可行性。  鉴于以上情况，同意参加开题答辩。    **指导教师签名： 年 月 日** | | | | | | |
| **开**  **题**  **报**  **告**  **会** | **开题报告会时间：**2021/12/24 下 **午** 13 **时— 17 时 地点：**信息楼西224会议室 | | | | | |
| **评**  **审**  **专**  **家**  **组**  **成** | 姓名 | 职称 | 所在单位及学科专长 | 博导**/**硕导 | （出席者）签名 |
| 马伟 | 副教授 | 北京工业大学，计算机科学与技术 | 硕导 |  |
| 陈军成 | 讲师 | 北京工业大学，计算机科学与技术 | 硕导 |  |
| 郭建伟 | 副研究员 | 中国科学院自动化研究所，计算机科学与技术 | 硕导 |  |
| 米庆 | 讲师 | 北京工业大学，计算机科学与技术 |  |  |
| **评审意见：**（由评审专家组填写）  **√（正常）通过；**  **□不合格， 前重做开题报告。**  **其它评语：**  医学影像智能诊断技术是计算机学科中一个前沿的研究领域，本开题报告针对医学影像诊断中的小样本问题，对主流方法进行充分的文献调研和分析，确定了研究内容。论文选题先进，具有良好的研究意义和应用价值，所提出了解决方案具有可行性，开题答辩准备充分、重点突出、阐述清楚。建议在设计轻量化网络时，考虑网络参数的相关性，去除不重要参数，提取小样本数据更本质的特征。同时在进行多模态融合时，建议考虑模态之间的差异性问题，设计相应的融合策略，并尽快课题实施，确保课题进度。  鉴于以上情况，答辩委员会一致通过开题答辩。  **评审组长签名： 年 月 日** | | | | | |