**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： S202274129**

**研究生姓名： 张一鸣**

**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： 2023/12/19**

**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** | | 张一鸣 | | **学 号** | S202274129 |
| **院、系** | | 信息学部（原计算机学院） | | **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** | 2022年9月 |
| **填表日期** | | 2023年12月19日 | |  |  |
| 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、医学影像分类   1. **拟定论文题目：**   医疗影像域泛化分类算法研究   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”；“基于磁共振影像组学的大鼠肝纤维化无创性精准分期研究”；“基于磁共振影像及深度学习的肝细胞癌恶性程度无创性诊断研究”  经费来源： 国家自然科学基金项目，北京自然科学基金青年项目，中国肝炎基金会王宝恩基金   1. **论文类型（基础研究、应用研究、开发研究、其它）**   应用基础研究 | | | | | |
| **摘**  **要** | 选题研究内容和意义简介（限400字）：  在智能化医疗影像的业务流水线中，不同部位医疗影像有着对应方法，需要依据医疗影像切片中所包含的器官类别来确定其所在部位。因此，多标签器官分类成为了前处理过程中重要的一个环节。  肝细胞癌(hepatocellular carcinoma，HCC)是一种原发性肝脏疾病。而微血管侵犯（Microvascular Invasion, MVI）被认为是肝癌切除或移植患者早期复发和长期预后不良的重要因素。因此，在HCC患者术前评估是否存在MVI具有重要的临床价值。  近年来,深度学习技术在医疗影像分析方面取得了长足的进步。然而，由于成像原理、影像设备等因素不同，数据间存在着分布差异。如果直接将训练好的模型应用于存在分布差异的数据，会导致模型性能下降。因此，如何开发能够泛化到未见分布的模型成为了近年来的研究热点。  本课题针对CT-MR的多标签分类问题与跨序列的MVI分类问题，分别设计了基于多维度数据增强的方法和基于特征解耦的MVI分类模型。探讨了对医学影像分类问题中域间差异大的解决方案，以提升模型的泛化性。 | | | | |
| **关键词（用分号隔开、最多5个）** | | | 微血管侵犯诊断；多标签器官分类；数据增强；域泛化；计算机辅助诊断 | | |

**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**

在智能化医疗影像的业务流水线中，不同部位、区域的医疗影像（例如胸部、腹部、腿部）有着不同的处理规范与方法，在业务的前处理流程中需要将影像分发至对应的业务。一种比较直观的方法是依据医疗影像切片中所包含的器官类别来确定其所在部位。因此，多标签器官分类成为了前处理过程中重要的一个环节。

以肝细胞癌（HCC）为代表的肝脏局灶性病变是一种原发性肝脏疾病，临床认为手术切除和移植是目前治疗肝细胞癌的最佳选择。而微血管侵犯（Microvascular Invasion, MVI）被认为是肝癌切除或移植患者早期复发和长期预后不良的重要因素。因此，在HCC患者术前评估是否存在MVI具有非常重要的临床价值。

近年来,深度学习技术在医疗影像分析方面取得了长足的进步。传统的机器学习模型基于独立同分布的假设进行训练。然而，由于成像原理、影像设备、操作方法等因素的不同，影像数据间存在着分布差异。如果直接将训练好的模型应用于存在分布差异的数据，会导致深度学习模型的性能下降。但是收集所有可能域的数据来训练模型十分昂贵，不具有现实性。因此，如何开发能够泛化到未见分布的模型成为了近年来的研究热点。

本课题针对CT-MR的多标签分类问题与跨序列的MVI分类问题，分别设计了基于多维度数据增强的方法和基于特征解耦的MVI分类模型。探讨了对医学影像分类问题中域间差异大的解决方案，以提升模型的泛化性。

肝细胞癌（Hepatocellular carcinoma，HCC）目前最常见的肝脏恶性肿瘤疾病，也是全球第三大致死性癌症[1]。肝脏疾病种类很多，涉及范围广泛，在临床诊断中，肝细胞癌是大多数肝脏疾病的一个非常重要的诊断节点，肝细胞癌的治疗方案及预后在很大程度上取决于患者肿瘤的分化程度[2]。因此，在进行手术之前就对HCC的分化程度进行一个初步的预测，可以辅助医生为患者制定更加完善的治疗计划，以提升患者的生存率[3]。

近年来，人工智能技术在实际应用的各个方面如计算机视觉、文本分析，取得了非常大的进展[4]。随着医学成像技术的发展和各类疾病病例的积累，为了满足临床医学的需要并减轻医生的负担，目前许多医院和科研机构已经开始在一些疾病的诊断中部署基于人工智能的医学影像辅助诊断系统，很多基于电子计算机断层扫描(Computed Tomography, CT)的脑、乳腺、肺等病变区域的诊断和检测[5,6,7]应用已经成功落地。但是由于腹腔器官众多且结构复杂，扫描所得的影像受到各种因素的干扰较大。因此，有着多样的成像序列，对比度高，便于清楚地分辨各类软组织结构的磁共振成像(Magnetic Resonance Imaging, MRI)在肝脏疾病的判别中的应用前景很广[8,9]。但是与常见的自然图像不同， MRI影像有着数据差异性、数据分辨率不高、多模态的特点，并且样本相对较少。因此将人工智能技术应用于肝脏疾病的辅助诊断并不多见，但也是当前计算机辅助诊断的难点之一。

面对肝脏疾病MRI影像多模态，病灶区域多尺度，数据样本量较小的问题，采取相应的方法进行处理和解决对于辅助诊断模型的建立是非常重要的。首先，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MRI影像，综合多种成像结果的优点，发挥数据类型互补的优势，进一步提升临床治疗和评价的质量显得至关重要。其二，肿瘤的大小信息在临床的影像学诊断中有重要的作用，结合医生关于肿瘤尺度的先验知识，使用多尺度特征对模型进行训练。可以降低在数据归一化过程中丢失的肿瘤数据的原始尺度信息。最后，在医学图像辅助诊断方面，绝大部分临床扫描的数据都没有经过医生的标注。而医疗影像数据的标注工作只能由专业的放射科医生完成，这导致高质量的医学影像标注数据很难达到和普通数据相当的数据量，这给模型的训练带来很大挑战，因此小样本学习的策略在深度学习在医疗影像领域的应用是非常有前景的。

本课题旨在利用深度学习方法利用MRI影像对HCC分化程度进行判别，并针对数据多模态，病灶区域多尺度，数据样本量小的问题对模型加以改进，用于辅助医生对HCC的诊断，对于改善当前医疗服务状况，提高临床诊断效率和效果具有重要意义，并且对肝脏疾病的诊断也具有很大的临床应用价值。

**1.2 国内外研究现状**

在目前的实际应用中，CT和MRI在HCC分化程度的诊断中发挥关键作用，几乎所有临床检查和诊断等都把细胞外造影剂多期相CT和MR成像作为HCC分化程度判别和诊断的首选方式。其间，很多学者[11,12,13]的研究结果显示，相比较CT，MRI在HCC的检测和分级中具有更高的准确率和灵敏度。2016年，Mitsuru Nakanishi等人[10]在对HCC进行诊断时通过对MRI影像的各个模态的主要特征进行分析，找到了各个模态的征象与HCC分化程度之间的关系，并且得出了MRI在评价肝细胞癌组织学分化程度中具有重要价值的结论。但是与常见的自然图像不同， MRI影像有着数据差异性、数据分辨率不高、多模态的特点，并且与CT影像相比，数据较为稀缺。因此使用MRI作为深度学习模型的数据集时，数据的多模态、多尺度、小样本的特性也就成为了必须要解决的问题。

**1.2.1** **肝病的影像学诊断研究现状**

随着人工智能技术的发展，目前已经出现了一些基于深度学习的肝脏疾病诊断方法。2016年，Wang等人[14]提出了一种基于卷积神经网络架构的特征提取方法，使用动态对比度增强MRI（Dynamic contrast enhanced MRI，DCE-MRI）数据，获取用于HCC特征表征的深层特征，并设计了一种集成多通道分类器，将数据的冠状面，矢状面和水平面视图分别作为网络的输入用于提取深层特征，并在模型决策层进行特征融合用于HCC分级的决策。2017年，V. Alberotanza等人[15]提出了一种基于卷积神经网络的ConvNet用于对增强CT图像进行HCC分级。2018年，Xiao等人[16,17]对DCE-MRI中不同模态进行分析，提出了一种将根据对比造影剂注入时间不同的五种模态的DCE-MRI数据进行时间层面和空间层面上的融合的数据融合方式，并且根据这种数据融合方式设计了一种多通道级联的3D卷积神经网络，把五种模态的3D影像分别送入五个独立的3D卷积神经网络中并将五个网络的输出在决策层进行融合，这样可以同时利用DCE-MRI序列中所蕴含的时空信息，因此在HCC分化程度三分级的任务中取得了不错的成果。2019年，Wang等人[18]针对DCE-MRI中的动脉期、静脉期和延迟期三期图像，提出了一种结合了近几年在自然图像领域表现比较优异的SENets[19]和DenseNet[20]的3D-SEDenseNet，用于对三种模态的3D图像进行分级，SENets可以计算3D数据中的每个通道的数据对最后分类任务的贡献度，并根据贡献度调整每个通道的权值，以提升模型的效果。而DenseNet的使用即减少了网络的参数量，又可以在一定程度上缓解梯度消失的问题。因此在HCC分化程度二分类的任务中取得了不错的成果。现有深度学习方法对HCC分化程度的分级已经取得了一定的效果，**但是在多模态数据的使用方面和肿瘤多尺度方面的考虑还有所欠缺，并且针对样本数量少的问题，并没有提出有效的针对方法。**

**1.2.2医学影像多模态融合方法**

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MRI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。在多模态融合方法的应用方面，2017年，Xie等人[32]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，Xiao等人[16,17]通过分析DCE-MRI中不同模态中所反映的肿瘤在影像中强化程度随着造影剂注射时间发生变化的特点，提出了一种级联的3D卷积神经网络用于提取DCE-MRI中五个模态所蕴含的时空信息，并设计了一种特殊的数据融合方式，将处于不同时间段的volume对应的相同位置融合成新的volume，相当于每个volume说明了同一个层面的肿瘤随时间的变化，也就是说蕴含更多的时间信息**并反映了肿瘤的血流灌注信息，并且血流灌注信息比原本第三个维度上所反映的纹理信息（三维空间结构信息）更加重要**，该方法在针对肝细胞癌分化判别任务取得了不错的成果。同年，Lee等人[33]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，Jia等人[35]提出了一种将医生认为具有信息互补能力的多模态MR影像进行融合，在T1WI同反相位，T2WI和DCE-MRI序列中挑选三个最具“潜力”的模态组合成一个三通道数据用于模型的训练，作者通过**实验证明了在使用多个模态融合决策时，选择互补的模态数据进行融合能得到更好的效果**。

**1.2.3 医学影像多尺度融合方法**

自然界中的物体由于观测尺度不同， 在图像中会表现出不同的形态， 尺度空间中各尺度图像的模糊程度逐渐变大， 能够模拟人在距离目标由近到远时目标在视网膜上的成像过程。而由于在训练卷积神经网之间需要对一批数据的尺寸进行归一化，会损失数据的原始尺度信息，而对于肿瘤诊断来说尺度信息有一定的参考意义，模型提取多尺度特征有利于提升医学影像分类效果。在多尺度融合的应用方面，2018年,Wu等人[36]使用不同尺度的视网膜图像用于眼底血管的分割，取得了多个眼底血管分割数据集上的最佳结果，并证明了不同尺度的输入有助于改善分割结果。2019年，一些学者[35]使用了空间金字塔池化（SPP）来获得不同尺度肿瘤的特征，**并证明了使用多尺度特征的网络得到的特征和原始特征在空间上以原点为中心呈现发散的状态，相比于原始特征，多尺度特征不会集中在原点附近，更有利于分类，**但是这种方法只利用了网络的高层语义特征，存在一定的局限性，而特征金字塔网络 （Feature Pyramid Networks，FPN）可以提取高层低层融合下的多尺度特征，同时利用低层特征高分辨率和高层特征的高语义信息，通过融合这些不同层的特征达到更好的预测的效果。

**1.2.4 医学影像小样本学习方法**

医疗影像数据的标注工作只能依赖有专业的放射科医生，因此绝大部分从医院获取的数据都没有经过医生的标注，这就导致高质量的医学影像标注数据很难达到和普通数据相当的数据量。而使用少量的数据直接训练深度学习模型根据模型的大小可能引起过拟合和欠拟合的问题，这给模型的训练带来很大挑战。另一方面，由于不同患者和不同设备之间存在的差异性，以及设备存在不可避免的噪音等问题，使得不同个体即使是同一类型的疾病，在医学影像上也会表现出差异，这也增加了疾病辅助诊断的难度。**为了使用有限的训练数据得到有效的辅助诊断模型**，很多相关工作使用了小样本学习方法，例如，迁移学习、数据增强、元学习、度量学习等方法。

数据增强也叫数据扩增，意思是在不实际的增加数据的情况下，让有限的数据产生等价于更多数据的价值，充分利用数据中的信息。常用于医学影像方面的数据增强方法包括对图像进行几何变换，包括翻转，旋转，裁剪和转置等各类操作，对图像进行颜色变换，包括加入噪声和模糊，和使用GAN生成一些训练样本等等。**数据增强的本质是为了在不改变网络的容量且不涉及不增加计算复杂度和调参工程量的情况下去增强模型的泛化能力的手段**。在数据增强方面，2019年，Wang等人[18]使用了随机翻转，随机裁剪来充分的利用训练数据中的信息，并通过为训练数据添加椒盐噪声和高斯模糊来增加模型的泛化能力。类似的数据增强方法在[16,17]中均有涉及。同年，Eric等人[30]使用CycleGAN[31]用于心脏磁共振影像的生成，并证明使用生成的伪图像训练的分割模型可以有效地对真实的图像进行分割。

基于迁移学习的方法，是指有其他数据源时，利用这些辅助数据集去做迁移学习。以数据量比较多的辅助数据集作为源领域，而需要做分类工作的拥有少量数据的数据集作为目标领域，**通过解决源领域和目标领域的领域偏移来提升模型在目标领域小样本数据集的表现。**在迁移学习方面，2019年，Alfonso等人[21]使用了迁移学习中的领域自适应方法，并结合度量学习中的Triplet loss[22]，提出了一种可以在源领域中样本比较多的数据集中训练，并在目标领域样本数量比较少的数据集中进行微调的模型，将经过调整减少了参数VGG16[23]作为预训练和微调的特征提取器，再结合浅层分类器对组织癌进行分类，并取得了优于直接在目标领域上进行训练的模型的结果。同年，Zahil Shanis等人[24]针对不同模态的脑部MR影像的语义分割问题，提出了一种无监督领域自适应的方法，通过集成学习和对抗学习的方法有效的解决不同模态MR影响之间的领域偏移。在公开脑部MR数据集中取得了不错的效果。

基于元学习的方法，**主要希望利用以往的经验来指导新任务的学习，使得模型具有学会学习的能力。**因此，元学习的目标就是通过学习大量的任务，从而学习到内在的元知识，从而能够快速的处理新的同类任务，因此在医学影像应用方面，存在许多不同的病种或者同一病种的不同模态，每个类别的数据量特别少，但是类别的数目相对较多，因此可以组成不同的子任务，通过对大量子任务的训练来提升模型在目标任务上的表现。在元学习方面，2019年，Li等人[25]对MAML[26]加以改良，建立了一个DAML元学习模型并从常见疾病的数据中提取不同的子任务，使用常见疾病作为元训练数据来训练神经网络，这样当给定与少量数据样本相关的新任务(罕见病 元测试数据)时，模型可以通过几个梯度下降的步骤来快速调整网络模型来处理新任务。即通过对常见疾病的子任务的训练使模型适应对罕见疾病的分类任务。同年，Jiang等人[27] 在原型网络[28]和MAML的基础上，通过权衡两种方法的表现提出了一种任务自适应度量方法，利用基于梯度和基于度量的度量空间，用于少样本医学影像分类研究。

基于度量学习的方法**，这类方法是对样本之间的距离分布进行建模，使得属于同类样本靠近，不同类样本之间远离。**常见的无参估计方法，如最近邻算法（K-nearest neighbor，KNN）。KNN 算法不需要训练，但分类效果却依赖于距离度量方法的选取，目前也出现了完全使用神经网络来学习如何进行度量方法，例如关系网络[29]。在度量学习方面，Santi等人[30]提出了一种基于TripNet的针对有限的训练样本模型，用于脑部MR影像模态识别。通过对小样本数据集进行采样，生成用于训练TripNet的数据三元组（anchor，positive和negative），其中前两者为同一类，前两者与negative为不同类，在训练过程中使用Triplet loss在特征空间中拉近anchor和positive的距离并拉远anchor和negative的距离以完成embedding层的训练，之后使用训练好的embedding层输出用于分类器。作者通过实验证明在面对一个较小训练样本情况时，基于TripNet的模型远比一般的CNN好。目前，基于度量学习的医学影像分类方法在很多问题中均有涉及。

从以上介绍的方法可以看出，小样本学习方法已经被比较普遍的使用与医学影像分类任务中，在面对医学影像数据集图像匮乏的问题，它们通过从其他样本比较多的数据集中进行知识的迁移，从不同的子任务中获取足够内在的元知识或从数据层和特征层入手进行数据增强和度量学习，以增强模型在样本不足的情况下的泛化能力。因此，小样本学习方法在医学影像分类任务中非常具有前景。

综上所述，通过对国内外大量文献的分析，目前，基于深度学习的肝脏影像病变区域分类方法，研究者更多关注的是单个模态影像或者单个序列的影像，即仅采用MRI的一种模态或者一个增强序列的影像，对于多模态医疗影像的利用仍不够充分。而对于多尺度图像和多尺度特征的使用，多见与影像分割和检测，而对于肿瘤诊断来说尺度信息有一定的参考意义，模型提取多尺度特征有利于提升医学影像分类效果，因此在肝病影像分类中非常具有潜力。而面对医学影像数据集图像匮乏的问题，小样本学习方法在医学影像分类任务中应用也是非常值得期待的。根据以上分析，本课题研究结合多模态影像融合和多尺度特征融合并引入小样本学习方法用于解决多模态肝病影像的分类问题。

**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**

本课题拟通过基于卷积神经网络的分类网络，以此完成多模态肝脏影像HCC分化程度分类任务。本课题拟从3个方面进行：

1. 新的HCC多模态MRI影像数据集的建立

针对医院提供的不同模态的MRI影像，进行预处理，数据增强和数据归一化，建立一个用于HCC分化程度分类的数据集。

1. 对多模态肝脏影像和多尺度特征进行融合利用

在单模态肝脏影像HCC分级的基础上，通过充分利用多种模态间的互补信息关系，再结合包含肿瘤的大小信息的多尺度特征，提高模型分类性能。

1. 结合小样本学习方法的医学影像分类问题探究

针对医学影像训练样本不足的问题，为了使用有限的训练数据得到有效的辅助诊断模型**，**在多模态影像和多尺度特征融合的基础上，引入小样本学习方法，并探究小样本学习方法对小样本医学影像分类问题的帮助。

**2.2研究目标**

本课题旨在设计一种用于对肝细胞癌分化程度进行分类网络，通过不同模态间影像的融合，不同尺度特征的融合，以提高分类网络的准确率，并结合小样本学习方法，一定程度上解决训练样本不足的问题。

**2.3拟解决的关键问题**

为了构建一个完整高效的HCC分化程度分类网络，本课题拟解决的关键问题主要基于以下3个方面：

1. 如何解决多模态数据互补信息的利用问题

不同模态的肝脏数据提供不同的病变信息，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MRI影像，综合多种成像结果的优点，发挥数据类型互补的优势，是一个待解决的关键的问题。

1. 如何利用肿瘤的尺度信息的问题

肿瘤的大小信息在临床的影像学诊断中有重要的作用，结合医生关于肿瘤尺度的先验知识，使用多尺度特征对模型进行训练。可以降低在数据归一化过程中丢失的肿瘤数据的原始尺度信息。因此如何将肿瘤的尺度信息加入到网络是一个待解决的关键问题。

1. 如何部署小样本学习方法的问题

在医学图像辅助诊断方面，绝大部分临床扫描的数据都没有经过医生的标注，这导致高质量的医学影像标注数据很难达到和普通数据相当的数据量，这给模型的训练带来很大挑战，因此小样本学习方法的部署是非常必要的，本课题主要考虑如何将度量学习，元学习和迁移学习方法用于解决或减轻小样本问题对模型的影像，这些方法在模型中的部署也是一个待解决的关键问题。

**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研究思路**

1.首先针对从医院获取的医学影像数据，对ROI（HCC所在的位置）进行提取和整理，并且为了让有限的数据产生等价于更多数据的价值，充分利用数据中的信息，本课题使用几何变换的方式对数据集进行数据扩充，并使用重采样的方法减轻样本类别不均衡的问题。

2.对于多模态影像的使用，针对MRI所反映的血流灌注信息比图像的纹理信息（三维空间结构信息更重要），并且MRI图像纵向切面的二维纹理信息已经能够在一定程度上反映肿瘤的空间结构信息，本课题从每个病人的MRI序列中，提取肿瘤面积最大的一个层面，并将不同模态对应相同层面的图像在通道维进行拼接，并且改进SENets[19]，对不同模态影像对最后分类任务的贡献度进行评估，并将贡献度与对应层面的特征进行加权提升分类效果。

3. 肿瘤的大小信息在临床的影像学诊断中有重要的作用，而数据归一化过程中会丢失的肿瘤数据的原始尺度信息，因此为了适应不同尺度的HCC数据，提取多尺度特征用于网络训练是一个常用的方法。而在对对尺度特征的利用方面，拉普拉斯金子塔[37]和图像特征金字塔[38]是两种最早的方法，它们虽然对模型有很大的帮助，但是会带来巨大的计算量和内存需求，并且在深度网络中使用多个特征层做预测时，而且由于底层的语义特征比较弱，因此处理特征一般只出现在较低的特征层的小物体时效果表现得不够好。空间金字塔池化[39]是针对网络输入图像大小不一的情况，用于适应多种尺寸的特征输入的方法，但是这种方法只利用了网络的高层语义特征，存在一定的局限性。而特征金字塔网络（FPN）[40]同时利用低层特征高分辨率和高层特征的高语义信息，通过融合这些不同层的特征达到预测的效果, 它能较好的让各个不同尺度的特征都具有较强的语义信息。为了适应不同尺度的HCC数据，提取多尺度特征，本课题改进了FPN用于提取多尺度HCC影像的特征

4.针对数据样本量小的问题，本课题首先针对处理好的ROI区域组成的数据集进行数据增强，并且2，3步设计的模型加以改进，拟使用迁移学习，度量学习，元学习&度量学习的方式分别用于模型改良，用于提升模型在小样本医学影像数据集中的分类效果。

**3.2技术路线**

**(1) 构建用于HCC分化程度判别任务的多模态MRI医学影像数据集以及数据预处理**

数据分为两部分，一部分来源于首都医科大学附属友谊医院，共51例分化程度不同HCC患者的MRI影像，另一部分来源于首都医科大学附属佑安医院，共158例分化程度不同HCC患者的MRI影像。MRI影像数据包含：T1WI同反相位、T2WI、DCE-MRI类型的MRI影像数据，其中每种模态的数据都是包含3个维度（在人体解剖学中用冠状面、矢状面、水平面来区分）的3D数据。标注和验证工作分别由有10年工作经验的放射科医生和有30年工作经验的放射科主任完成。

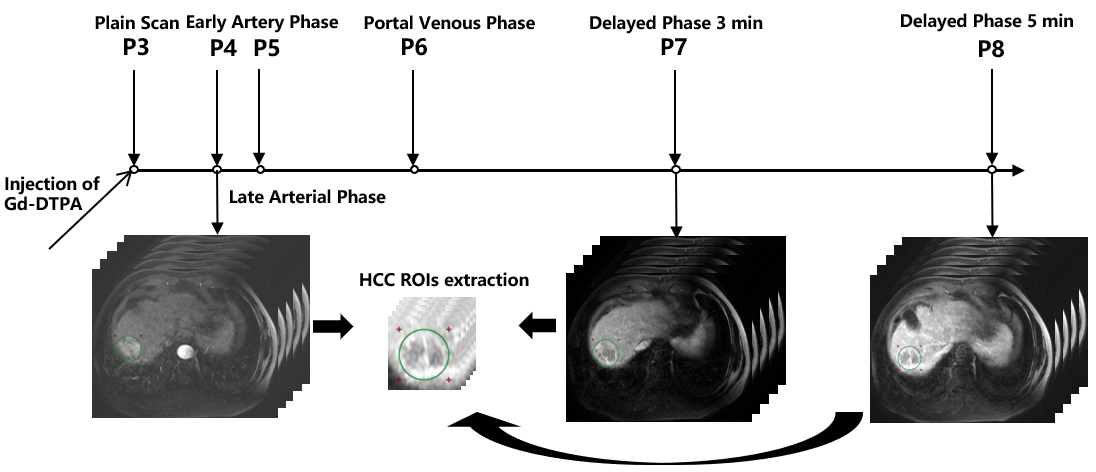


图1 DCE-MRI序列及ROI区域提取

根据世界卫生组织对HCC的分级，根据HCC分化的程度的高中低，将HCC分为Poorly，Moderately和Well三种分化级别，分别表示肝细胞癌的严重程度。下图展示了九个模态对应三种分类和背景的ROI区域。

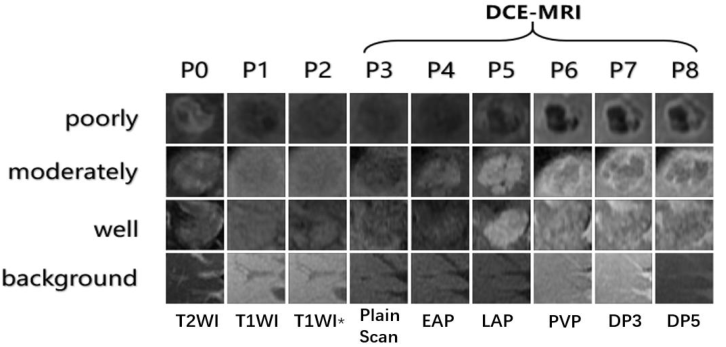


图2 9个模态的MRI影像中的ROI区域

其中P0到P9分别代表T2WI，T1WI同相位，T1WI反相位，平扫，动脉早期，动脉晚期，门静脉期和延迟期的MRI影像。

因为医学影像数据量受到病例数和标注困难等问题的限制，因此让有限的数据产生等价于更多数据的价值，充分利用数据中的信息非常重要。本课题使用了对图像进行几何变换，包括翻转，旋转，裁剪和转置等各类操作对数据进行扩充，总共对数据进行了8倍扩充，具体的扩充方法和扩充结果如下图所示。

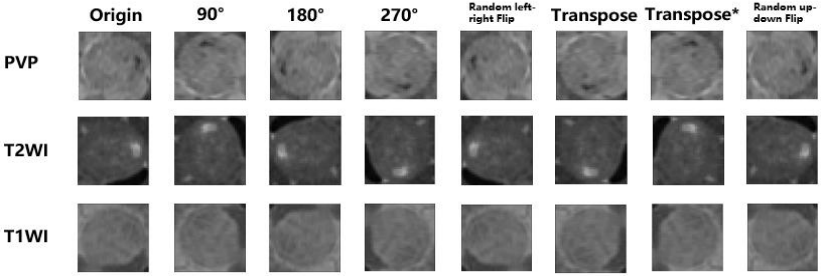


图3 数据扩充结果

**(2) 多模态影像的融合方法设计**

在多模态影像融合方面,本课题将提取每个模态3D序列中肿瘤面积最大的一个层面，并且将T1WI同反相位，T2WI和DCE-MRI共9个序列的单张最大图像在通道上拼接，，经过特征提取之后送入SEBlock中，SEBlock的结构如下图所示。

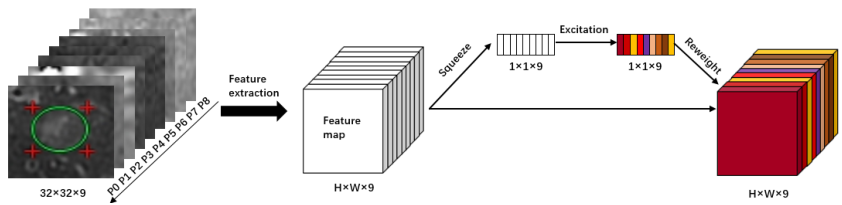


图4 SEBlock的结构

SEBlock总共分为Squeeze, Excitation和Reweight三个部分，其中Squeeze部分用全局平均池化将每个通道上的特征图压缩成一维，Excitation部分引入非线性计算每个维度的特征对于最后分类任务的贡献度，并去除用于减小参数量的瓶颈层。最后的Reweight操作使用Excitation操作所得的特征与Squeeze操作之前的特征进行加权。

将SEBlock添加在第一个7\*7的卷积层之后，构成的SEResNet如图5所示。

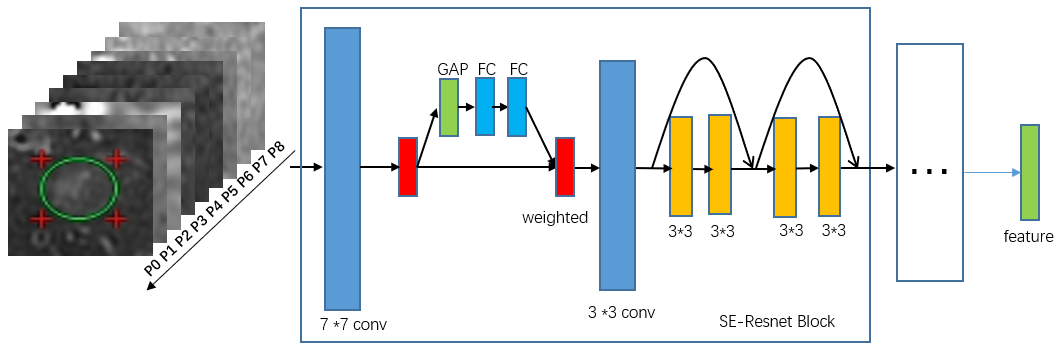


图5 SEResnet结构

首先通过7\*7的卷积层降低特征图大小，再经过一个SEBlock提取特征，并结合每个模态对最后分类任务的贡献程度，之后在经过一个3\*3的卷积层把通道数目与ResNet18同步。

**(3) 多尺度特征融合策略**

在多尺度特征融合方面，本课题在FPN的基础上加以改进，将其与（2）中的SEResnet结构融合在一起，用于对多尺度的特征的融合。SE-FPN具体的网络结构如下图6。

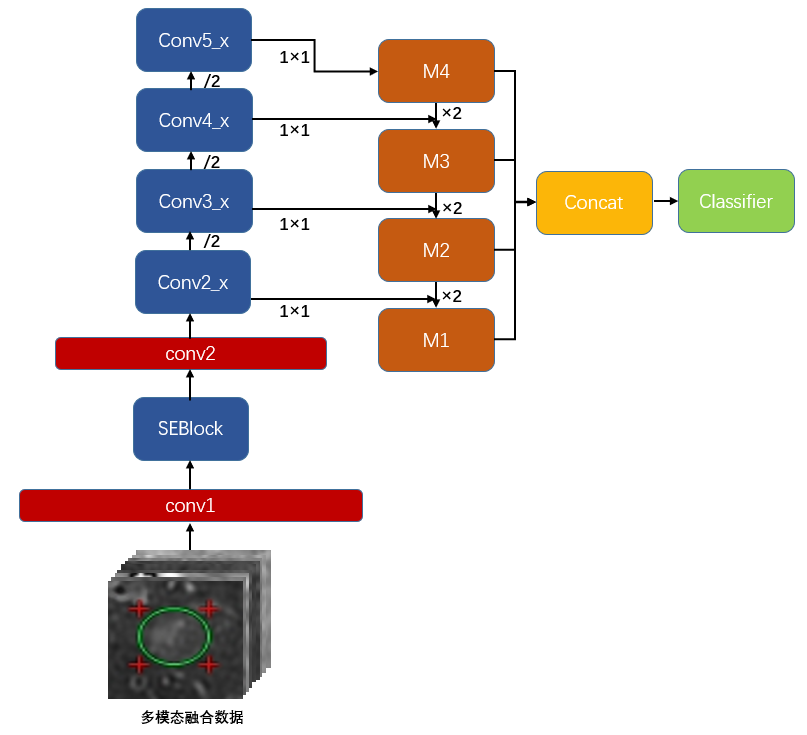


图6 多尺度特征融合网络

FPN的目标是利用卷积网络本身带有的层次性语义特征，来构建特征金字塔。FPN包含两个部分：第一部分是自底向上的过程，第二部分是自顶向下和侧向连接的融合过程。其中，自下至上的通路conv1和conv2分别代表7\*7和3\*3的卷积，SEBlock代表（2）中所提出的模型，在conv2\_x, conv3\_x, conv4\_x和conv5\_x中，底层特征可以保留更多的分辨率，但是特征语义较为低级，随着网络的深度越深，则特征的语义越高级，也因此，改进的FPN融合了低级和高级语义，即较低的层反映较浅层次的图片信息特征像边缘等；较高的层则反映较深层次的图片特征像物体轮廓、乃至类别等，并可以兼顾不同尺寸物体，同时控制了计算量。而自上至下的通路是在处理每一层信息时会把上一层的high level信息作为输入，即将上层的特征图等比例放大之后再与本层的特征图进行逐元素加。上层的特征输出一般其特征图更小，因此需要进行上采样之后才能与本层的特征融合。横向的1\*1的卷积用于使得不同的特征图有着相同的通道数目，便于特征融合与逐元素加，并且可以将CNN网络层特征与经过上采样所得的各个尺度的特征相关联。

最后Concat层将四种尺度的特征融合用于最后分类器的分类。

**(4) 小样本学习方法**

1. **迁移学习&度量学习**

针对医疗影像数据量较少的问题，本课题首先考虑引入在人脸识别[42]和行人再识别等任务**[41]**中表现比较好的Triplet损失[42]。在针对训练样本各类样本数目有限的问题中，可以引入人为定义的距离（例如欧式距离）用于度量不同样本编码之间的距离，并且通过损失的计算，训练一个合适的特征空间，在这个特征空间中，减少相同类之间的距离，增加不同类之间的距离。之后，再使用经过微调的分类器对这个特征空间所映射好的特征编码进行分类。Triplet损失函数如下：

在数据的采样方面Anchor，Positive和Negative为三种特征编码。其中Anchor和Positive为同类样本，Anchor和Negative为不同类样本，数据采样的方法如图8所示，所以当negative位于以Anchor为圆心，Anchor到Positive的距离为半径的圆中时，我们把三元组称为Hard Negative，而当Negative到Anchor的距离大于圆的半径，但小于圆的半径加上margin时，我们把三元组称为Semi-Hard Negative，而当Negative位于这两者之外时，我们把三元组称作Easy Negative。在数据的使用上，我们使用Hard Negative和Semi-Hard Negative对SE-FPN进行训练。

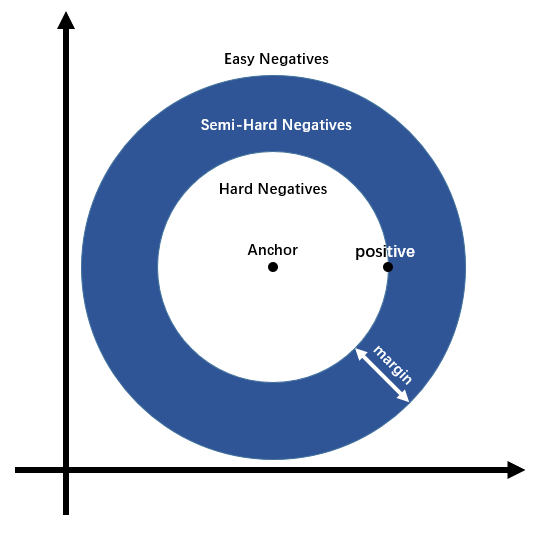


图7三元组采样方法

特征空间训练网络如图8所示。

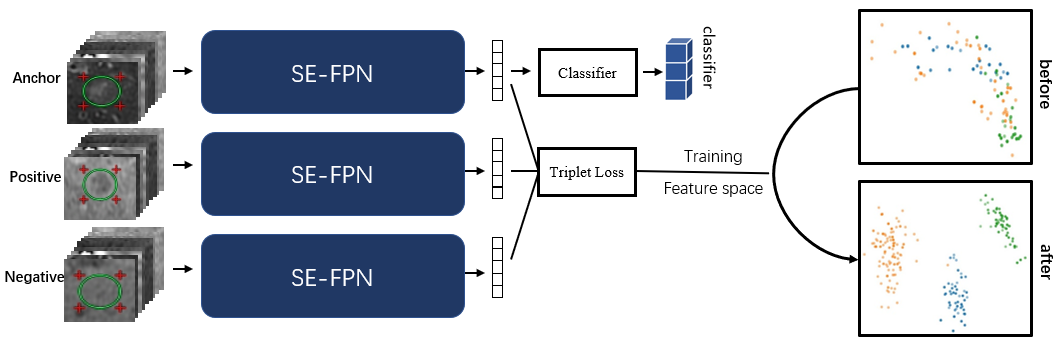


图8 特征空间训练网络

将3个权值共享的SE-FPN作为特征提取器，提取多模态影像的多尺度特征。第一步先用采样好的（Anchor，Positive，Negative）三元组作为特征空间的训练数据，待训练至网络收敛之后，用一个浅层分类器在剩余的训练数据中微调，之后在测试集中验证分类效果。

而针对现有的一大一小两个数据集，可以使用迁移学习的方式使用相对大的数据集用于SE-FPN的训练，而将浅层分类器在小数据集中进行微调，提升模型在小数据集中的分类表现。

1. **元学习&度量学习**

在目前的小样本学习领域，Relation Network[29]打开了使用神经网络来学习样本编码在特征空间的距离度量的大门。由于缺少足够的训练数据，Flood Sung等人使用了元学习的方式来训练关系网络，即通过大量子任务组成的Meta-train集合中抽取任务进行训练，通过学习大量的任务，从而学习到内在的元知识，从而能够快速且准确的处理新的同类任务。

而在当前的任务中，除了现有的HCC数据，还有另一批局灶性肝病数据，其中有囊肿95例、血管瘤108例、胆管细胞癌35例、转移瘤54例、肝脏局灶性结节增生28例、肝脓肿41例、瘤栓27例。

对于HCC分化程度的三分类任务，可以使用另外的七种局灶性肝病数据，每次随机采样三类的数据使用模型进行三分类任务的训练。具体的网络结构如图9所示。

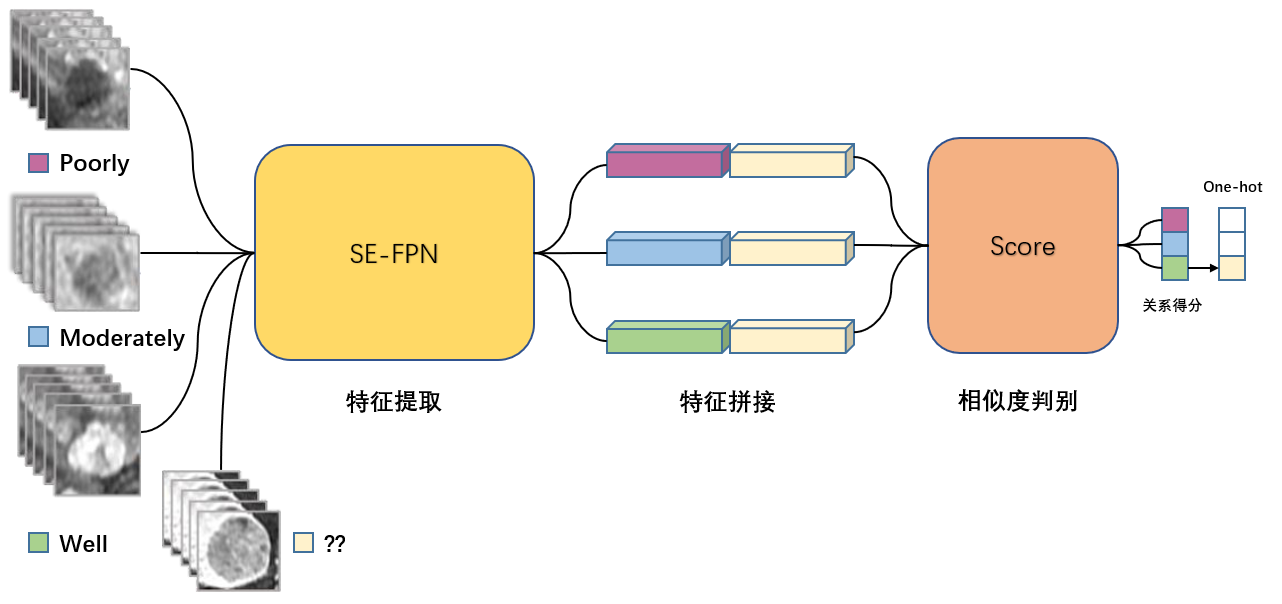


图9 基于元学习和度量学习的相似度判别网络

其中左侧的三个输入为有标签但样本较少的样本集，下方的输入为测试集。将一组样本集的数据和一个测试集的数据共同输入到特征提取网络之中，经过特征提取之后，将样本集的三类数据所产生的特征编码分别与测试集数据的特征编码拼接，之后送入Score模块进行相似度度量。最终判别结果是测试集数据的特征编码与样本集中相似度最高的一类特征编码的类别相同。

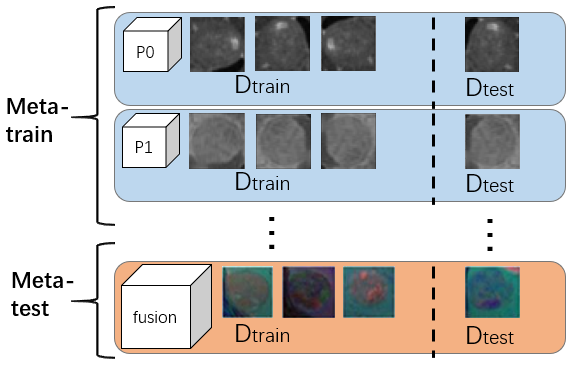


图10 元训练集与元测试集的划分

**3.3可行性分析**

本课题基于深度学习方法，提出以多模态融合，多尺度特征融合和小样本学习方法完成肝细胞癌分化程度的分类与诊断，并依靠前人的理论基础及实现工具，结合现有的图像处理、深度学习等方法的研究成果保证了课题的可实现性。以多模态融合，多尺度特征融合，小样本学习的方法处理医学影像分类问题，已经有一些学者在其他脏器使用另一些方法验证其有效性。综上，我们提出的引入多模态融合，多尺度特征融合和小样本学习方法分类模型具有一定的可行性。

1. **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两点：

1. 提出一种结合多模态医学影像和多尺度特征的医学影像分类模型。
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，提升模型分类效果。
3. 建立一个新的局灶性肝病的多模态MRI影像数据集。
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：

|  |  |
| --- | --- |
| 2019.9～2020.12 | 收集阅读相关文献资料，制定数据采集与标注方案 |
| 2020.1～2020.2 | 与医院沟通逐步开展数据标定工作并整理数据集 |
| 2020.2～2020.8 | 搭建深度学习模型，在自建数据集上进行模型效果验证 |
| 2020.9～2021.1 | 根据实际的肝病数据优化网络模型，不断研究性能更好的方案 |
| 2021.1～2021.6 | 整理相关资料，撰写毕业论文，完成论文答辩 |

预期研究结果：

构建基于深度学习的肝细胞癌分化程度分类模型，结合多模态影像，多尺度特征和小样本学习方法，优化模型，提升分类效果，并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. 作为负责人申请的第十七届研究生科技基金项目《DICOM影像的交互式标注系统和肝细胞癌的辅助诊断系统》已顺利结题，并获得软件著作权一项。
6. 在投论文一篇
7. **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）**
8. 设备条件：实验室电脑及相关软件；
9. 实验场地条件：北京工业大学信西304b实验室；
10. 可能遇到的问题：
11. 知识储备不足：需要不断地阅读领域相关文献，积累研究所需知识；
12. 代码编写困难：需要查阅各项资料，并向导师进行咨询；
13. 实验结果不好：通过调参和优化算法来提升实验结果；
14. 论文撰写问题：可参考优质论文，并向导师咨询请求协助。

**（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] Yang, Ju Dong, and Lewis R. Roberts. "Hepatocellular carcinoma: a global view." Nature reviews Gastroenterology & hepatology 7, no. 8 (2010): 448.

[2] Parikh, Sameer, and David Hyman. "Hepatocellular cancer: a guide for the internist." The American journal of medicine 120, no. 3 (2007): 194-202.

[3] Regimbeau, Jean Marc, Eddie K. Abdalla, Jean Nicolas Vauthey, Gregory Y. Lauwers, François Durand, David M. Nagorney, Iwao Ikai, Yoshio Yamaoka, and Jacques Belghiti. "Risk factors for early death due to recurrence after liver resection for hepatocellular carcinoma: results of a multicenter study." Journal of surgical oncology 85, no. 1 (2004): 36-41.

[4] Lecun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553): 436.

[5] Sun, Dongdong, Minghui Wang, and Ao Li. "A multimodal deep neural network for human breast cancer prognosis prediction by integrating multi-dimensional data." IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB) 16, no. 3 (2019): 841-850.

[6] Spasov, Simeon, Luca Passamonti, Andrea Duggento, Pietro Liò, Nicola Toschi, and Alzheimer's Disease Neuroimaging Initiative. "A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease." Neuroimage 189 (2019): 276-287.

[7] Xie, Yutong, Jianpeng Zhang, Yong Xia, Michael Fulham, and Yanning Zhang. "Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest CT." Information Fusion 42 (2018): 102-110.

[8] Burrel, Marta, Josep M. Llovet, Carmen Ayuso, Carmela Iglesias, Margarita Sala, Rosa Miquel, Teresa Caralt et al. "MRI angiography is superior to helical CT for detection of HCC prior to liver transplantation: an explant correlation." Hepatology 38, no. 4 (2003): 1034-1042.

[9] De, M. Santis, Renato Romagnoli, Alessandro Cristani, Giorgio Cioni, Alessandra Casolo, Francesco Folchi Vici, and Ezio Ventura. "MRI of small hepatocellular carcinoma: comparison with US, CT, DSA, and Lipiodol-CT." Journal of computer assisted.

[10] Nakanishi, Mitsuru, Makoto Chuma, Shuhei Hige, Tokuhiko Omatsu, Hideki Yokoo, Kazuaki Nakanishi, Toshiya Kamiyama et al. "Relationship between diffusion-weighted magnetic resonance imaging and histological tumor grading of hepatocellular carcinoma." Annals of surgical oncology 19, no. 4 (2012): 1302-1309.

[11] Burrel, Marta, Josep M. Llovet, Carmen Ayuso, Carmela Iglesias, Margarita Sala, Rosa Miquel, Teresa Caralt et al. "MRI angiography is superior to helical CT for detection of HCC prior to liver transplantation: an explant correlation." Hepatology 38, no. 4 (2003): 1034-1042.

[12] De, M. Santis, Renato Romagnoli, Alessandro Cristani, Giorgio Cioni, Alessandra Casolo, Francesco Folchi Vici, and Ezio Ventura. "MRI of small hepatocellular carcinoma: comparison with US, CT, DSA, and Lipiodol-CT." Journal of computer assisted tomography 16, no. 2 (1992): 189-197.

[13] Akai, Hiroyuki, Shigeru Kiryu, Izuru Matsuda, Jirou Satou, Hidemasa Takao, Taku Tajima, Yasushi Watanabe et al. "Detection of hepatocellular carcinoma by Gd-EOB-DTPA-enhanced liver MRI: comparison with triple phase 64 detector row helical CT." European journal of radiology 80, no. 2 (2011): 310-315.

[14] Wang, Q. and Que, D., 2016, July. Staging of Hepatocellular Carcinoma Using Deep Feature in Contrast-Enhanced MR Images. In 2nd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2017). Atlantis Press.

[15] Bevilacqua, V., Brunetti, A., Trotta, G.F., Carnimeo, L., Marino, F., Alberotanza, V. and Scardapane, A., 2017. A deep learning approach for hepatocellular carcinoma grading. International Journal of Computer Vision and Image Processing (IJCVIP), 7(2), pp.1-18.

[16] Jia, Xibin, Yujie Xiao, Dawei Yang, Zhenghan Yang, Xiaopei Wang, and Yunfeng Liu. "Temporal-Spatial Feature Learning of Dynamic Contrast Enhanced-MR Images via 3D Convolutional Neural Networks." In Chinese Conference on Image and Graphics Technologies, pp. 380-389. Springer, Singapore, 2018.

[17] Yang, Da-wei, Xi-bin Jia, Yu-jie Xiao, Xiao-pei Wang, Zhen-chang Wang, and Zheng-han Yang. "Noninvasive Evaluation of the Pathologic Grade of Hepatocellular Carcinoma Using MCF-3DCNN: A Pilot Study." BioMed research international 2019 (2019).

[18] Zhou, Qing, Zhiyong Zhou, Chunmiao Chen, Guohua Fan, Guangqiang Chen, Haiyan Heng, Jiansong Ji, and Yakang Dai. "Grading of hepatocellular carcinoma using 3D SE-DenseNet in dynamic enhanced MR images." Computers in biology and medicine 107 (2019): 47-57.

[19] Hu, Jie, Li Shen, and Gang Sun. "squeeze-and-excitation networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7132-7141. 2018.

[20] Huang G , Liu Z , Maaten L V D , et al. Densely Connected Convolutional Networks[C]// CVPR. IEEE Computer Society, 2017.

[21] Medela A, Picon A, Saratxaga C L, et al. Few shot learning in histopathological images: reducing the need of labeled data on biological datasets[C]//2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE, 2019: 1860-1864.

[22] Schroff, Florian, Dmitry Kalenichenko, and James Philbin. "Facenet: A unified embedding for face recognition and clustering." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 815-823. 2015.

[23] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[24] Shanis, Zahil, et al. "Intramodality Domain Adaptation Using Self Ensembling and Adversarial Training." Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data. Springer, Cham, 2019. 28-36.

[25] Li, Xiaomeng, et al. "Difficulty-aware Meta-Learning for Rare Disease Diagnosis." arXiv preprint arXiv:1907.00354 (2019).

[26] Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation

of deep networks. In: ICML. pp. 1126–1135. JMLR. org (2017)

[27] Jiang, Xiang, et al. "Task Adaptive Metric Space for Medium-Shot Medical Image Classification." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2019.

[28] Alex Nichol and John Schulman. Reptile: a scalable metalearning algorithm. arXiv, 2018.

[29] Sung, Flood, Yongxin Yang, Li Zhang, Tao Xiang, Philip HS Torr, and Timothy M. Hospedales. "Learning to compare: Relation network for few-shot learning." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1199-1208. 2018.

[30] Kerfoot, Eric, et al. "Synthesising Images and Labels Between MR Sequence Types with CycleGAN." Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data. Springer, Cham, 2019. 45-53.

[31] Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translationusing cycle-consistent adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV) (2017)

[32] 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.

[33] 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.

[34] Liu, Yucheng, et al. "Cross-Modality Knowledge Transfer for Prostate Segmentation from CT Scans." Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data. Springer, Cham, 2019. 63-71.

[35] Jia, Xibin, et al. "Multi-parametric MRIs based assessment of Hepatocellular Carcinoma Differentiation with Multi-scale ResNet." KSII Transactions on Internet & Information Systems 13.10 (2019).

[36] Wu, Yicheng, et al. "Multiscale network followed network model for retinal vessel segmentation." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2018.

[37] Burt P, Adelson E. The Laplacian Pyramid as a Compact Image Code[J]. IEEE Transactions on Communications, 1983, 31(4): 532–540.

[38] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden. Pyramid methods in image processing. RCA engineer, 1984

[39] He K, Zhang X, Ren S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 37(9): 1904–1916.

[40] Lin T-Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//International Conference on Computer Vision and Pattern Recognition, 2017: 2117–2125.

[41] Sun Y, Zheng L, Yang Y, et al. Beyond Part Models: Person Retrieval with Refined Part Pooling (and a Strong Convolutional Baseline)[C]//European Conference on Computer Vision, 2018.

[42] Schroff F, Kalenichenko D, Philbin J. FaceNet: A Unified Embedding for Face Recognition

and Clustering[C]//International Conference on Computer Vision and Pattern Recognition, 2015:

815–823.