# 基于深度学习的MRI脑卒中病灶分割方法综述

余唯一[1]，陈涛[1]，张军平[2]，单洪明[1]
（1. 复旦大学类脑智能科学与技术研究院，上海 200433；
2. 复旦大学计算机科学技术学院，上海 200433）

**摘　要：**脑卒中病灶自动分割方法成为近几年的研究热点。为了全面研究用于MRI脑卒中病灶分割的深度学习方法的现状，针对脑卒中治疗的临床问题，进一步阐述了基于深度学习的病灶分割的研究背景及其挑战性，并介绍脑卒中病灶分割的常用公共数据集（ISLES和ATLAS）。然后，重点阐述了基于深度学习的脑卒中病灶分割方法的创新与进展，从网络结构、训练策略、损失函数这3个角度对研究进展进行了归纳，并且对比了各种方法的优缺点。最后，讨论了该研究存在的困难和挑战以及未来的发展趋势。

**关键词：**脑卒中；医学图像分割；计算机视觉；深度学习；神经网络
**中图分类号：**TP39
**文献标志码：**A
**doi:** 10.11959/j.issn.2096−6652.202328

# A survey of deep learning-based MRI stroke lesion segmentation methods

YU Weiyi[1], CHEN Tao[1], ZHANG Junping[2], SHAN Hongming[1]
1. Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai 200433, China
2. School of Computer Science, Fudan University, Shanghai 200433, China

**Abstract:** Automatic stroke lesion segmentation has become a research hotspot in recent years. In order to comprehensively review current progress of deep learning-based MRI stroke lesion segmentation methods, start with the clinical problems of stroke treatment, we further elaborate the research background and challenges of deep learning-based lesion segmentation, and introduce common public datasets (ISLES and ATLAS) for stroke lesion segmentation. Then, we focus on the innovation and progress of deep learning-based stroke lesion segmentation methods, and summarize the research progress from three perspectives: network structure, training strategy, and loss function, and compare the advantages and disadvantages of various methods. Finally, we discusse the difficulties and challenges in this research and its future development trend.

**Key words:** stroke, medical image segmentation, computer vision, deep learning, neural network

## 0　引言

脑卒中是人类特别常见的脑血管疾病之一，也是全球第二大致死原因[1]。根据2015—2018年调查数据显示，全球大约有760万的脑卒中患者。对于老年人来说，第一次发生中风后1~5年内的死亡概率超过50%[2]。根据脑供血障碍的不同，中风可分为缺血性中风和出血性中风[3]。前者由输送血液到

**收稿日期：**2023−06−02；**修回日期：**2023−07−12
**通信作者：**单洪明，hmshan@fudan.edu.cn
**基金项目：**国家自然科学基金项目（No.62101136，No.62176059）；上海市自然科学基金项目（No.21ZR1403600）；上海市科技计划项目（No.20JC1419500）
**Foundation Items:** The National Natural Science Foundation of China (No.62101136, No.62176059), Natural Science Foundation of Shanghai (No.21ZR1403600), Shanghai Municipal Science and Technology Project (No.20JC1419500)

大脑的血管堵塞引起，占所有中风病例的 87%[4]；后者由脆弱的血管破裂引起，如动脉瘤、动静脉畸形、高血压等[5]。治疗脑卒中是一项极具挑战性的工作。以常见的缺血性中风为例，在中风发作后的几个小时内，如果没有及时向大脑重新供血，可能发生组织梗死，并且该损伤过程是不可逆的[6]。因此，观察脑卒中患者的脑部结构并发现病灶，是治疗脑卒中患者的关键。

神经影像学是检测、表征和预测急性中风（包括缺血性中风和出血性中风）的重要工具。其中，电子计算机断层扫描（computer tomography，CT）和磁共振成像（magnetic resonance imaging，MRI）是常见的成像手段。CT 可以用来识别中风类型，常用于中风后、实施紧急手术之前采集脑部信息[7]。虽然 CT 成本较低，成像速度快，受噪声影响较小，但异常病变在 CT 上不是清晰可见的。如果需要观察小脑、脑干和大脑半球内部的微小梗死，它也会受到限制[8-9]。虽然 MRI 成本高，扫描时间长，要求较高，仅部分医疗中心有条件配备，但它能够克服 CT 的局限性，并且在症状出现之前比 CT 更早发现梗死，所以 MRI 更常用于发现脑卒中病灶[10]。因此，本文仅探讨在 MRI 上对病灶进行分割的工作，如何与在 CT 上的分割方法[11-12]结合实现多模态分割，仍需进一步研究。

放射科医生通常会基于大脑结构和病灶特点手动划分病灶边界，这对医生的临床经验的要求很高，且耗时耗力。因此，研究者们希望通过深度学习实现对脑卒中患者磁共振图像病灶的全自动分割[13-16]。然而，用深度学习实现全自动病灶分割是非常具有挑战性的工作。首先，深度学习训练需要尽可能多的高分辨率磁共振（magnetic resonance，MR）图像，而患者 MR 图像的获取难度较大，MR 设备也比较昂贵。其次，不同患者的病灶形状、位置、大小和数目差异很大，如图 1 所示，病灶组织的多样性增大了分割难度。最后，病灶组织的边界不够清晰[17]，甚至同一位医生对同一位患者前后两次的划分结果都可能有差异，需要反复检查以确保数据标注的质量。

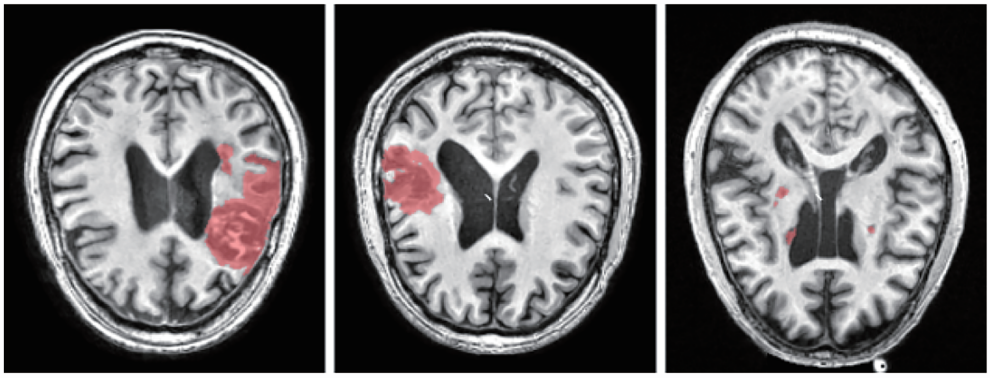


图 1　脑卒中患者的 T1 加权图像示例

国外已有一些综述论文整理和分析了利用深度学习技术检测脑卒中的方法[18-20]，但没有专门针对脑卒中病灶分割的详细总结。本文的主要贡献如下：介绍了脑卒中病灶分割的常用公共数据集；对深度学习在脑卒中病灶分割的应用进行了全面梳理，从网络框架、训练策略、损失函数这 3 个角度出发，对研究进展进行了归纳，并且对比了各种方法的优缺点；阐述了当前面临的困难和挑战，展望了未来的发展方向。

## 1　公开数据集

为了方便比较各类方法的效果，本节将介绍脑卒中病灶分割的公开数据集。

研究人员通常会招募脑卒中患者参与数据收集活动，这需要耗费大量的时间和资金成本。所以，脑卒中病灶分割的公开数据集较少，已有公开数据集如表 1 所示。接下来，将从图像预处理和图像标注这两个方面对表 1 中的数据集进行描述。

### 1.1　缺血性脑卒中病灶分割数据集

缺血性脑卒中病灶分割（ischemic stroke le-

表 1　脑卒中病灶分割的公开数据集

| 名称 | 训练样本数量/个 | 测试样本数量/个 |
|---|---|---|
| ISLES 2015 - SISS | 28 | 36 |
| ATLAS（v2.0） | 655 | 300 |

sion segmentation，ISLES）数据集有多个版本，但只有ISLES 2015 - SISS[21]包含临床采集的亚急性缺血性脑卒中患者的影像数据，涉及的MR模态有T2、T1、液体衰减反转恢复（fluid attenuated inversion recovery，FLAIR）和弥散加权图像（diffusion weighted image，DWI）。

（1）图像采集。所有数据都来源于德国，石勒苏益格-荷尔斯泰因大学医疗中心提供了56例病例，慕尼黑工业大学临床教学中心的神经放射科提供了8例病例。这两个中心均配备了3T Phillips扫描仪。

（2）图像预处理。所有图像均做了去头骨处理，并视情况进行位置校正。

（3）图像标注。由经验丰富的医学博士进行病灶标注，主要基于FLAIR图像完成标注，其他模态图像提供辅助信息。

### 1.2　脑卒中后病灶示踪数据集

脑卒中后病灶示踪（anatomical tracings of lesions after stroke，ATLAS）数据集的MR模态为T1。它有两个版本，首先发布的版本是ATLAS v1.2[22]，包含全球11个研究中心的304例病例。然而，基于ATLAS v1.2的大多数方法都没有开源代码，数据划分方式也不一致，训练容易过拟合，泛化性较差。

虽然当前没有公认标准，但ATLAS数据集的作者根据实践经验推断：对于深度学习方法，只有分割结果的Dice系数达到0.85以上，才能证明它是一种可取代人工的全自动分割方法，当前的实验结果与该阈值相差甚远。为了促进基于深度学习的脑卒中病灶分割方法的发展，新版本数据集ATLAS v2.0[23]发布。跟ATLAS v1.2相比，它的数据量更大，并且划分了训练集（公开人工标注）和测试集（不公开人工标注），这样可以帮助研究者开发出鲁棒性和泛化性更强的方法。所有人都可以在网页上用测试集进行性能评估，公平公开地对比各类方法。由于ATLAS v2.0已覆盖ATLAS v1.2，接下来的介绍将围绕ATLAS v2.0展开。

（1）图像采集。数据集由全球33个研究中心的955例病例组成。绝大多数中心都配备了3T扫描仪，涉及的品牌有Siemens、Phillips和GE，型号多种多样（公开文件中有包含数据集信息的excel表格）。

（2）图像预处理。预处理流程包括了去头骨、归一化处理、配准至MNI-152模板空间等，这样的操作便于提升深度学习方法的性能。

（3）图像标注。由8位专家对病灶进行手动分割，共同检查，以控制数据标注质量。

## 2　深度学习在脑卒中病灶分割的研究

近年来，随着深度学习的快速发展，越来越多的研究者关注深度学习在医学影像上的应用[24-28]。而医学影像分割就是其中一项基础且关键的任务，主要目的是针对医学影像，从像素水平识别出目标器官、组织或病变区域[29-31]。深度学习作为临床辅助诊断工具，能够实现病灶的全自动分割，从而减少医生在病灶划分上耗费的时间和精力[32-34]。

笔者按时间顺序整理了深度学习应用于脑卒中病灶分割的文献，如表2所示。各类方法在公共数据集上的性能将在第2.4节展示。

本节对应用于脑卒中病灶分割的深度学习方法进行了全面梳理（如图2所示），从网络结构、训练策略、损失函数这3个角度出发，对研究进展进行归纳。

### 2.1　网络结构

随着大规模数据和计算机硬件的飞速发展，深度学习在计算机视觉中的表现受到广泛关注。卷积神经网络（convolutional neural network，CNN）[70]是深度学习的经典结构之一，主要由卷积层、池化层和全连接层组成，具有较好的特征提取、分类能力，常用于目标检测、图像分割等领域。后来，基于语义分割的全卷积神经网络（fully convolutional network，FCN）[71]，把CNN的全连接层换成卷积层。

UNet[72]在FCN的基础上提出了编码器—解码器结构的神经网络，编码器用于获取上下文信息、提取特征，解码器用于精确定位，跳跃连接则实现浅层特征和深层特征的融合。目前，UNet是医学影像分割中性能最好、应用最广泛的网络结构[73]，大多数用于脑卒中病灶分割的深度学习方法是基于UNet结构展开研究的。

#### 2.1.1　注意力机制

对于不同的中风患者，脑卒中病灶的形状、位置、大小和数目差别较大。经验丰富的医生可以结合图像的多尺度信息综合考虑，但基于FCN的深度学习方法感受野小、缺乏长距离信息，从而忽略了这些差异。注意力机制最初源于对人类视觉的研

**表2 深度学习用于脑卒中病灶分割的现有方法**

| 年份 | 方法 | 特点 | 是否公开代码 |
|---|---|---|---|
| 2017年 | EDD Net and MUSCLE Net[34] | 提出双阶段训练策略，先获得初步分割结果，再进行细化 | 否 |
| 2018年 | uResNet[35] | 利用残差连接优化卷积模型 | 否 |
| 2018年 | Res-FCN[36] | 利用残差连接优化卷积模型 | 否 |
| 2019年 | X-Net[37] | 提出特征相似度模块以捕获上下文信息 | 是 |
| 2019年 | D-UNet[38] | 设计维度转换模块以融合3D和2D图像特征 | 否 |
| 2019年 | Multi-path 2.5D CNN[39] | 在轴状面、矢状面和冠状面方向上做2D分割后再融合，并进行3D后处理 | 是 |
| 2019年 | MSDF-Net[40] | 将空洞空间金字塔池化用于医学影像分割框架 | 否 |
| 2019年 | CLCI-Net[41] | 1.在UNet结构的编码器和解码器之间嵌入了空洞空间金字塔池化；<br>2.提出跨层特征融合策略以整合各层特征 | 是 |
| 2020年 | MI-UNet[42] | 利用大变形微分同胚度量映射算法[43]获得组织解剖分割，以作为额外输入 | 否 |
| 2020年 | A 3D+2D CNN approach[44] | 先用3D U-net获取病灶概率图，再用2D UNet获得最终预测结果 | 否 |
| 2020年 | 3D residual CNN[45] | 采用“放大&缩小训练策略” | 否 |
| 2020年 | MUDCap3[46] | 用胶囊网络代替卷积层和池化层 | 否 |
| 2020年 | Multi-task[47] | 图像重建和图像分割的多任务学习 | 否 |
| 2020年 | Hybrid Dice loss[48] | 结合Dice损失与BCE损失，应对正负样本不平衡问题 | 否 |
| 2021年 | DFENet[49] | 提出特征融合边界，引导网络进行2D-3D特征融合 | 否 |
| 2021年 | Variational mode decomposition[50] | 利用变分模态分解做预处理 | 否 |
| 2021年 | ME-Net[51] | 1.在UNet结构的编码器和解码器之间嵌入空洞空间金字塔池化；<br>2.提出残差编码器，用于提取高维度特征 | 否 |
| 2021年 | Priors from healthy subjects[52] | 将健康被试图像作为先验辅助分割 | 否 |
| 2021年 | MDAN[53] | 针对大脑“伪对称”的形状特点设计专门的网络结构 | 否 |
| 2022年 | 3D UNet[54] | 3D医学影像分割 | 是 |
| 2022年 | e-UNets[55] | 设计全局特征注意力模块，以获得全局特征 | 否 |
| 2022年 | STHarDNet[56] | 用HarDNet[57]模块代替卷积模块 | 否 |
| 2022年 | AGMR-Net[58] | 1.在解码器阶段设计粗糙粒度块注意力模块，应对类内不一致问题；<br>2.设计维度特征融合模块，应对类内不确定问题 | 否 |
| 2022年 | METranse[59] | 利用4种不同尺度的编码器提取多尺度特征，再利用Transformer进行全局特征建模 | 否 |
| 2022年 | MAPPING [60] | 提出一种新的分割后处理策略 | 是 |
| 2022年 | MLiRA-Net[61] | 设计多尺度长距离区域关注模块，在多尺度层面提取全局信息 | 否 |
| 2023年 | CarveMix[62] | 随机组合两张医学影像作为数据增强 | 否 |
| 2023年 | ICI loss[63] | 将Dice损失与实例级和实例中心损失结合，以减轻实例不平衡问题 | 是 |
| 2023年 | TSRL-Net[64] | 1.设计粗糙粒度残差学习模块修正残差连接，令解码器更关注于未预测区域；<br>2.提出一种新的目标感知损失，通过调节聚焦因子扩展聚焦样本区域 | 是 |
| 2023年 | Fine-grained segmentation[65] | 将中风影响区域定义为常规定义病变的详细子区域，并据此定义细粒度的分割任务 | 否 |
| 2023年 | SQMLP-net[66] | 同时进行脑卒中病变分割和TICI评级 | 否 |
| 2023年 | W-Net[67] | 提出一种新颖的双阶段病灶分割方法，利用CNN和Transformer作为骨架网络：<br>1.一阶段利用CNN的局部信息提取能力获得大致病变分割图，第二个阶段利用Transformer的强大全局上下文建模能力对边界进行细致分割；<br>2.设计边界变形模块和边界约束模块来应对模糊边界的苛刻条件 | 否 |
| 2023年 | SAN-Net[68] | 针对不同中心影像分布差异设计域适应分割算法：<br>1.提出自适应实例归一化（masked adaptive instance normalization，MAIN）算法；<br>2.利用梯度反转层迫使UNet编码器使用域分类器学习域不变表示 | 是 |
| 2023年 | FAN-Net[69] | 针对不同中心影像分布差异，设计域适应分割算法，提出基于傅里叶的自适应标准化模块 | 否 |

究。在认知科学中，由于信息处理的瓶颈，人类会选择性地关注一部分信息，忽略其他可见的信息。随着神经网络的快速发展，注意力机制[27,74]也得到了广泛的应用。简要来说，它是一种通过计算注意力权重、对特征重加权，以强化有效特征、抑制无效特征的嵌入式结构[75]。

根据注意力机制应用位置的不同，注意力机制可以分为通道注意力[76]与空间注意力[77]。前者对

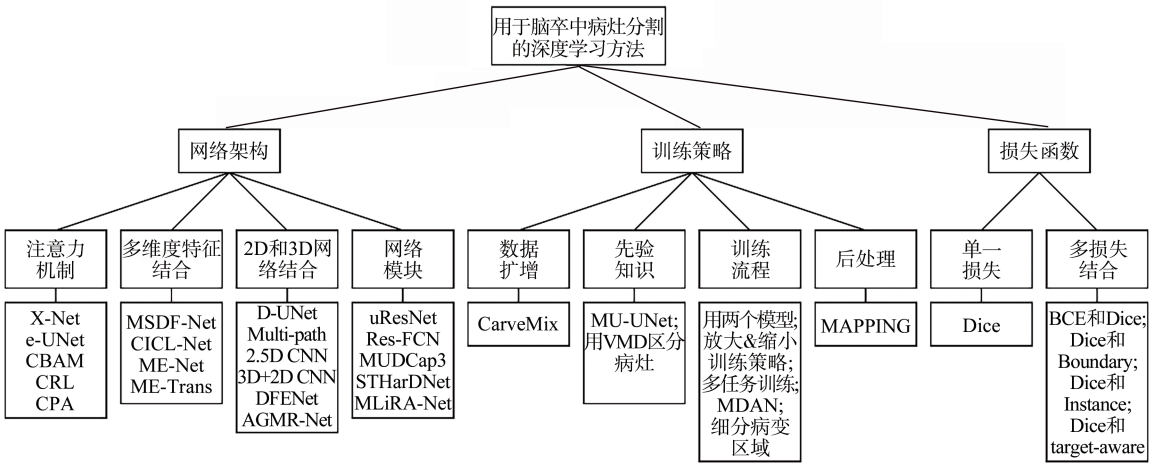



图2　深度学习在脑卒中病灶分割中的研究方法

特征图的各个通道进行加权；后者对特征图的各个像素进行加权，有效弥补了由卷积操作感受野不足导致的全局特征提取能力不足的缺陷。因此，为了充分利用所有像素之间的上下文信息，研究者们根据特定的医学图像分割任务，从注意力机制的角度优化网络结构。

为了在特征图中捕捉更丰富的上下文关系，参照非局部信息统计的注意力机制[78]，特征相似度模块（feature similarity module，FSM）[37]被提出。如图3所示，$H$、$W$和$C$分别代表特征图的高度、宽度和通道数，$HW$表示*reshape*操作后，原始的高度和宽度维度被展平为一个单一的维度，卷积（*conv*）后的“3×3”或“1×1”代表卷积核尺寸，圆圈内的“+”和“×”分别表示逐像素相加和矩阵相乘。特征相似度模块利用3×3卷积把通道数$C_0$压缩为$C$后，先计算输入特征像素间的特征相似度，由此得到每个像素的全局注意力，然后进行特征重加权。除此之外，还加入了残差连接以保证训练过程的稳定性。

与此类似的思路，全局特征注意力模块（global-feature attention block，GA-block）[55]被提出。如图4所示，利用1×1卷积把通道数从$C_0$压缩到$C$后，先计算输入特征通道间的权重，由此得到每个通道的全局注意力，然后进行特征重加权。除此之外，还加入了残差连接以保证训练过程的稳定性。全局特征注意力模块侧重的是不同通道间特征图的关系，而特征相似度模块侧重的是特征图上各像素点的关系。

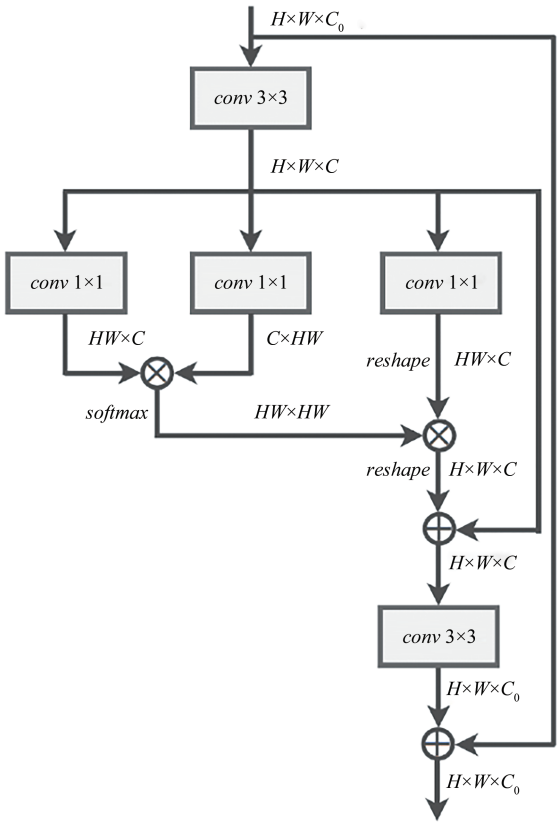



图3　特征相似度模块的结构[37]

卷积注意力模块（convolutional block attention module，CBAM）[79]结合了空间和通道的注意力机制，通过两个级联的空间和通道注意力模块，提升了全局特征能力，CBAM已在脑卒中分割网络中得到应用[59,65]。

粗糙粒度残差学习（coarse-grained residual learning，CRL）模块[64]修正残差连接，令解码器

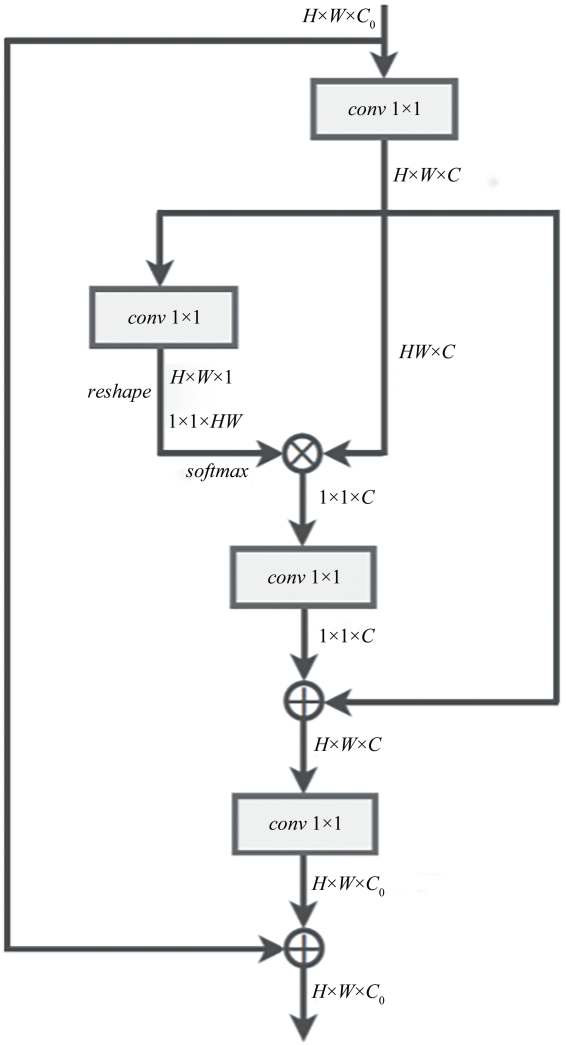



图4　全局特征注意力模块的结构[55]

更关注于未预测区域，并抑制大量背景噪声。该模块包含正向注意力（positive attention，PA）单元和逆向注意力（reverse attention，RA）单元两个部分。RA单元抑制已识别的目标信息，减少冗余信息，更有利于网络识别目标的残差特征，从而提高召回率。PA单元首先利用解码器中的粗糙粒度块注意力（coarse-grained patch attention，CPA）[58]，获得粗糙粒度的注意力图，并据此增强目标所在区域的特征表示，抑制大量不在目标区域的背景噪声特征，避免了噪声对网络的干扰，降低了错误正样本率。

图5展示了注意力机制和UNet框架的结合形式，将注意力模块嵌入UNet的编码器和解码器之间，便于提取高维特征图像素之间的上下文关系。总的来说，注意力机制通过计算注意力权重、对特征重加权，以达到强化有效特征、抑制无效特征的目的，帮助模型捕捉长距离上下文信息。

### 2.1.2　多尺度特征融合

对于不同的中风患者，病灶大小和形状的差异较大，由此增加了深度学习模型对病灶准确识别、精准定位的难度。针对该问题，利用多尺度特征融合提升模型对病灶分割的预测能力，该思路在医学影像领域应用广泛[80-84]。

空洞空间金字塔池化（atrous spatial pyramid pooling，ASPP）[85]以不同采样率对特征图进行空洞卷积，然后将结果拼接到一起，最后通过1×1卷积改变通道数，利用多尺度特征来捕捉图像的上下文关系。为了对不同尺度的语义信息进行有效融合，多尺度深度融合网络（multi-scale deep fusion network，MSDF-Net）[40]将空洞空间金字塔池化应用于医学影像分割框架UNet，并将其嵌入UNet结构的编码器和解码器之间。

跨层融合和上下文预测网络（cross-level fusion and context inference network，CLCI-Net）[41]不仅在UNet结构的编码器和解码器之间嵌入了空洞空间金字塔池化，还提出了跨层特征融合（cross-level feature fusion，CLF）策略。跨层特征融合策略指的是在进行下采样操作之前，将每个下采样层的输出与之前所有的特征拼接在一起，该策略可以整合各层特征，增强跨尺度信息之间的联系和互补性。除此之外，还在UNet的跳跃连接环节加入了卷积长短时记忆（ConvLSTM），这样连接浅层位置信息和深层特征信息就能更好地推断上下文关系，从而得到更精细的分割结构。

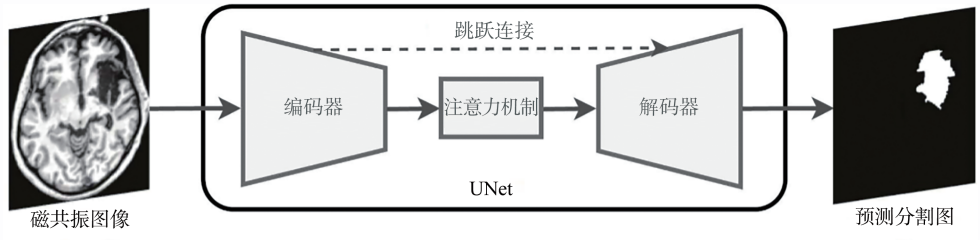



图5　注意力机制与UNet框架结合的示意

目前常用的编码器网络往往没有利用足够的深度来获取图像中更深层的语义信息，可能导致小病灶检测失败或病灶边界预测不准确。为了解决该问题，多编码器网络（multiple encoders network，ME-Net）[51]不仅在UNet结构的编码器和解码器之间嵌入了空洞空间金字塔池化，同时提出了残差编码器（如图6所示）用于提取高维度特征。图7展示了ME-Net的整体框架，残差编码器会生成4个特定尺寸的特征图，通过跳跃连接组合不同尺度的特征图，拼接后输入解码器进行上采样。与此类似，METrans[59]利用4种不同尺度的编码器来提取多尺度特征，再将特征图馈送到Transformer中进行全局特征建模，以克服UNet无法建模长距离上下文信息的缺点。

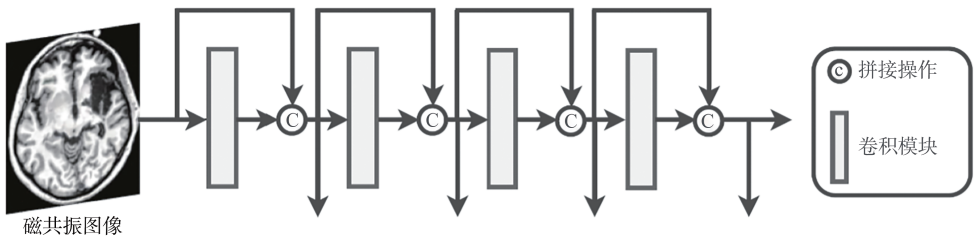



图6　残差编码器的结构[51]

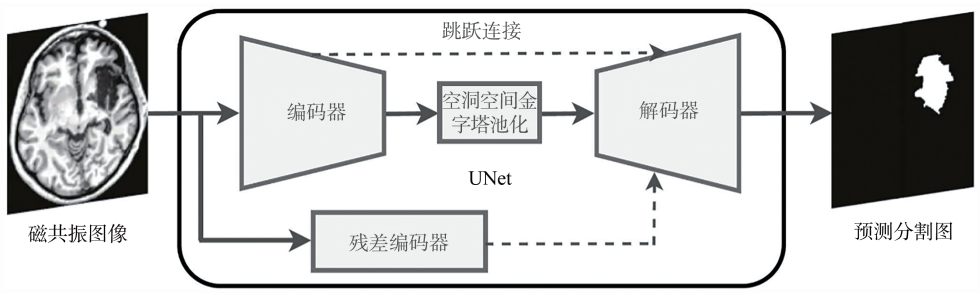



图7　多编码器网络的结构[51]

多尺度特征融合的优点是在不增加过多模型参数的前提下扩大了感受野，提升模型对病灶分割的预测能力。

### 2.1.3　2D和3D网络相结合

在同等数量的医学影像条件下，跟3D网络相比，2D网络的模型参数较少，但特征提取过程中缺乏3D空间信息；而3D网络能够捕捉图像的3D空间信息，但模型参数较多，样本数量较少[54,86]。为了同时利用2D和3D网络的优点，研究者们考虑将两者结合起来，共同处理医学影像。

一种常见思路是使用2D网络和3D网络共同构成2.5D网络[87-90]。为了在编码器阶段融合2D和3D网络的特征图，特征融合U型网络（dimension-fusion U shape network，D-UNet）[38]被提出。图8展示了D-UNet的整体框架，主干仍为2D UNet，在此基础上增加3D编码器，还设计了维度转换（dimension transform）模块，专门用于融合2D编码器和3D编码器输出的特征图。与此类似的方法还有特征融合边界引导网络（dimension fusion edge guided network，DFENet）[49]以及AGMR-Net[58]中的维度特征融合（cross-dimensional feature fusion，CFF）模块。

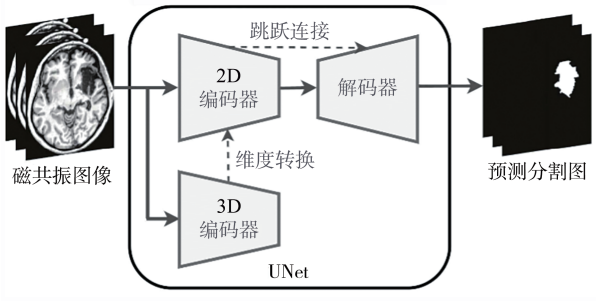



图8　特征融合U型网络的结构[38]

另一种思路是把2D和3D网络分别用于预测过程的不同阶段[39,44]。比如，为了丰富模型的输入信息，多路径2.5D卷积神经网络[39]先对切片做9种不同的预处理（在轴状面、矢状面和冠状面方向上的3种预处理方法），分别输入9个2D UNet，然后对9个2D UNet输出的2D特征图做方向校正，拼接为

3D特征图，最后用3D CNN做最终预测，如图9所示。除此之外，还有人提出2D+3D CNN方法[44]，先用3D UNet提取空间信息来获取病灶概率图，然后和2D切片一起输入2D UNet，最终得到预测结果。

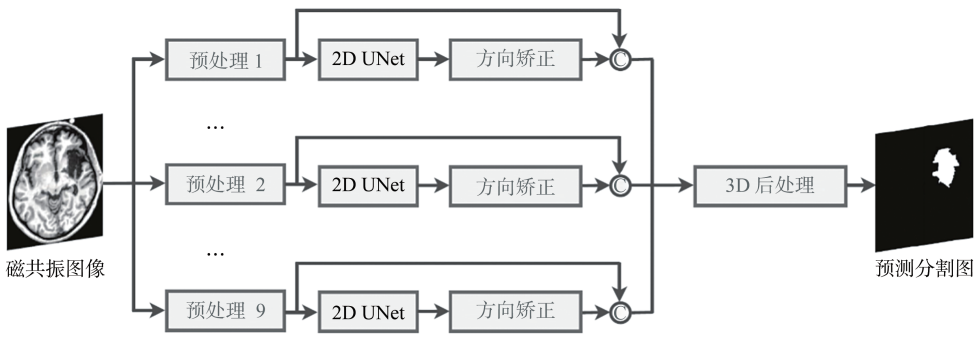



图9 多路径2.5D卷积神经网络的结构[39]

总之，2D和3D网络结合的目的是利用二者的优点，在模型参数尽量少的前提下捕捉3D空间信息。

**2.1.4** 优化网络模块

除此之外，还有人通过优化网络模块来提升模型的性能，比如残差连接[91]能够解决深层网络中梯度消失、梯度爆炸等问题，提升网络的表征能力[92]。uResNet[35]和Res-FCN[36]都是利用残差连接来优化卷积模型。以uResNet[35]为例，它以UNet为基础框架，通过残差连接（如图10所示）改善卷积模块。

CNN虽然通过卷积和池化解决了平移不变性问题，在一定程度上解决了旋转、放缩不变性问题，但是池化操作后多个规律分布的特征值变为一个特征值，不仅特征表达量减少，而且原特征值之间的相对位置关系被打破。如果对象发生旋转，则原已训练好的权重矩阵将不再完美适用。针对该问题，MUDCap3[46]以3D UNet为框架，用胶囊网络[93]替代卷积层和池化层。胶囊网络需要的训练数据较少，能提供等变映射使位置和形状信息得以保存，有利于图像分割。

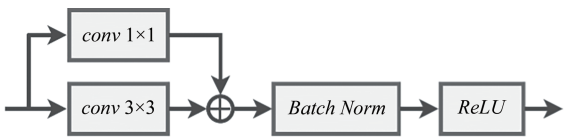



图10 残差模块的结构[35]

STHarDNet网络[56]（如图11所示）以UNet为整体框架。为了降低内存占用，用HarDNet[57]模块代替卷积模块。HarDNet基于DenseNet[94]做了连接稀疏化处理。同时，为了提取层次特征图和上下文关系，在跳跃连接的基础上加入了包含滑窗操作、具有层级设计的Swin Transformer[95]。

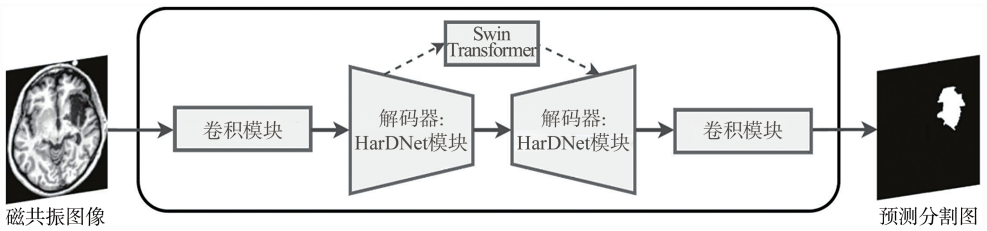



图11 STHarDNet的结构[56]

MLiRA-Net[61]将Transformer中的自注意力模块做了适当变形，设计多尺度长距离区域关注（multi-scale long-range and regional attention，MLiRA）模块，并将其引入分割网络的编码器中，以在多尺度层面提取全局信息。

**2.1.5** 小结

一是注意力机制分为通道注意力和空间注意力，它通过计算注意力权重、对特征重加权，达到强化有效特征、抑制无效特征的目的，从而捕捉长距离上下文信息。不过，注意力机制对样本数据量的要求比CNN更高，不适用于小样本数据。

二是多尺度特征融合利用空洞空间金字塔池化等模块嵌入UNet结构，扩大模型的感受野，从而提高预测能力。然而，空洞卷积对分割小病灶的提

升效果不明显，甚至造成信息冗余。

三是将2D和3D网络结合，便于利用二者的优点，不过结合方式可以朝着模型轻量化的方向进一步优化。

四是借鉴ResNet[91]、Swin Transformer[95]等算法来优化UNet网络模块。在以增加模型参数为代价换取性能提升的过程中，需要考虑设备要求和计算量。

### 2.2　训练策略

除了网络结构，研究人员在训练策略方面也做了许多努力。在不改变网络结构的情况下，数据扩增、增加先验知识、优化训练流程和后处理也能提升模型的预测性能。

#### 2.2.1　数据扩增

医学影像数据不如自然图像那么容易获取，尤其是患者的数据，而样本数据量不足会直接导致模型的性能不佳，所以数据扩增是有效的方法之一。常见的数据扩增方式有平移、旋转等，这类方法对医学影像不一定有效。由于大脑组织结构的复杂性和边界模糊性，对MR脑影像做数据扩增并不是一件容易的事[96-97]。

在自然图像领域，Mixup[98]是一种对图像进行混类增强的算法，它用两张图像和对应的标注做线性组合，并且组合是随机的。Cutmix[99]是基于Mixup改进的非线性组合方法，它不是对两张图像进行样本插值，而是从图像的空间角度出发，把一类图片上的某个区域做随机矩形框裁剪，然后拼接到另一类图片上，标注的处理和Mixup一致，按照合成图像中两张图像所占的比例进行混合标注计算。

根据这样的思路，CarveMix[62]被提出，用于脑卒中患者MR脑影像的数据扩增（如图12所示）。第一步，根据带标注样本a的病灶位置和形状，用概率分布采样得到感兴趣区域（region of interest，ROI）。第二步，取另一带标注样本b，ROI二值图、ROI二值互补图分别与样本a、样本b的影像做像素级别的点乘运算，最后相加，生成新样本c。第三步，ROI二值图、ROI二值互补图分别与样本a、样本b的病灶标注做像素级别的点乘运算，最后相加，生成新样本c的病灶标注。

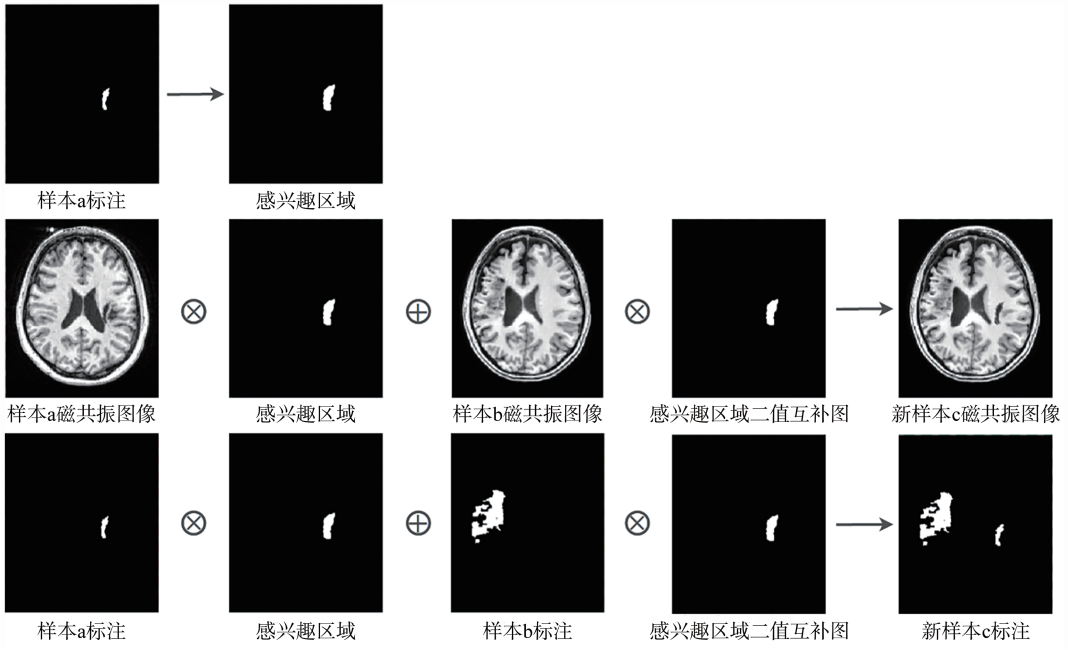



图12　CarveMix的流程[62]

受到人类大脑“伪对称”的形状特点的启发，文献[68]提出了一种简单而有效的数据增强方法——对称启发数据增强（symmetry-inspired data augmentation，SIDA），可以在减半内存消耗的同时，将样本大小增加一倍。

#### 2.2.2　增加先验知识

由于病灶分割跟脑组织结构有较强关联，病灶区域的信号强度值也显示异常，研究人员基于这些特点给网络输入增加先验知识，降低病灶分割的难度，从而提升性能。

MI-UNet[42]利用大变形微分同胚度量映射（large deformation diffeomorphic metric mapping，LDDMM）算法[43]对MR脑影像做组织解剖分割，将其跟磁共振脑影像作为UNet的输入，能够显著提升模型性能。类似地，文献[52]将一组健康被试的图像配准至患者的图像空间，为每个体素提供正常组织的参考强度值，由此得到异常分数图谱，这些图谱可以作为先验知识输入分割网络，帮助网络预测病灶区域。

有一种常用于分析MRI的预处理方法[100-102]——变分模态分解（variational mode decomposition，VMD），能够通过求解频域变分优化问题估计各个信号分量。文献[50]提出用VMD做预处理，区分可能的病灶区和非病灶区，帮助去除不必要的信息，然后用3D UNet做病灶分割。

总之，先验知识能够在不增加模型复杂度的前提下，降低模型的预测难度。不过，准确的先验知识获取难度较大。

### 2.2.3 优化训练流程

除此之外，也可以优化训练流程，充分发挥网络模型的作用，达到更好的预测效果。如图13所示，文献[34]利用两个不同的网络模型EDD Net和MUSCLE Net进行脑卒中病灶分割，先将图像输入EDD Net得到初步的分割结果，再将原图像和初步的分割结果输入MUSCLE Net得到进一步细化的分割结果，最后对EDD Net和MUSCLE Net输出的分割结果进行增强，得到最终预测结果。

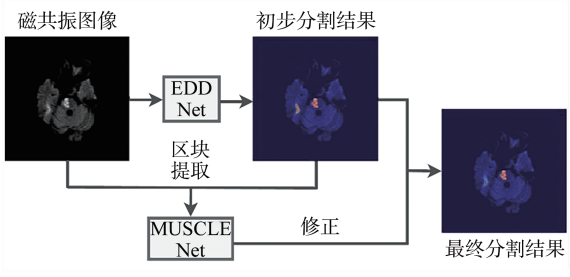



图13 卷积神经网络系统的流程[34]

文献[45]使用的网络模型是3D残差卷积网络，采用了“放大&缩小训练策略”，即先在小体积数据集上训练模型，然后在大体积数据集上对模型进行微调。该训练策略的优点是在放大阶段从原始立体影像中随机提取不同的小体积数据，相当于数据扩增，这样小体积数据可以产生正则化效果；缩小阶段的模型输入是大体积数据，便于模型学习到更广泛的背景，同时提高模型的鲁棒性。

多任务学习也可以在中风病灶分割中发挥作用[47]，主要任务是图像重建和图像分割，所用的网络模型分别是D5C5[103]和UNet。在此基础上，又设计了动态重加权损失约束，以保持重建任务和分割任务的平衡。多任务的目的是缓解分割模型在训练数据上过拟合的问题，从而提升模型性能。

传统方法将脑卒中病变定义为一个区域，而不是区分中风影响区域（stroke-affected region）和脑脊液（cerebrospinal fluid，CSF），这可能会影响游治疗效果。文献[65]将中风影响区域定义为常规定义病变的详细子区域，并根据大脑连通性将原始分割图细致化为8个子类：大脑灰质、非皮质区（大脑白质和皮层下）、小脑灰质、小脑白质、CSF（CSF+脑室+卒中后脑脊液空化）、颅骨、皮肤、中风影响区域。该文献将分割任务与是否发生病变这一分类任务结合，以提高检测准确率。

脑卒中分割和脑梗死溶栓（thrombolysis in cerebral infarction，TICI）评级是辅助诊断中风的两个重要但具有挑战性的先决条件。然而，大多数研究只关注两个任务中的一个，没有考虑它们之间的关系。SQMLP-net[66]同时进行脑卒中病变分割和TICI评级，用一个单输入双输出的网络处理这两个任务间的相关性和异质性。

镜像差异感知网络（mirror difference aware network，MDAN）[53]利用了大脑“伪对称”的形状特点。如图14所示，首先，基于UNet结构的Siamese编码路径把原始图像和翻转图像作为输入，然后，用Siamese对比监督损失和基于镜像位置的差异扩增（mirror position-based difference augmentation，MDA）模块进一步处理编码后的特征图，目的是通过增强特征的不对称性来检测异常区域。为了更好地将精细的对称性特征传递到解码器，每个阶段都使用了镜像特征融合（mirror feature fusion，MFF）模块，最后输出预测结果。

W-Net[67]拥有一个双阶段网路架构，第1个阶段利用CNN的局部信息提取能力获得大致的病变分割图，第2个阶段利用transformer的强大全局上下文建模能力对边界进行细致分割。为了更好地对边界进行分割，W-Net还设计了两个边界提升模块，即边界变形模块（boundary deformation module，BDM）和边界约束模块（boundary constraint module，BCM），来应对模糊边界的苛刻条件。

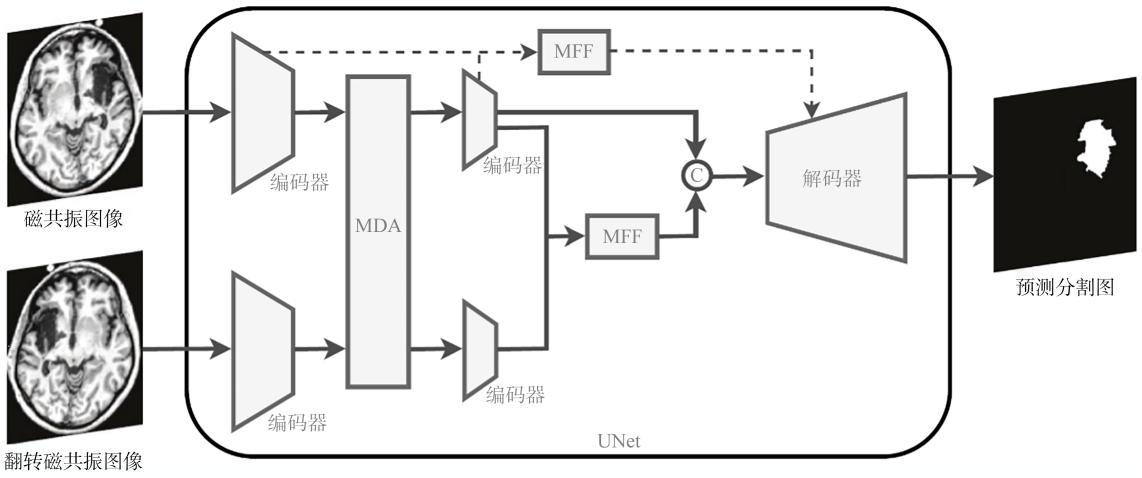



图 14　镜像差异感知网络的流程

BDM 利用循环卷积对初始边界进行校正，BCM 利用多尺度空洞卷积对目标边界进行动态约束。

优化训练流程虽然能够提升模型性能，但通常会引入更复杂的网络结构和训练步骤，如何尽量减少计算量和时间消耗也是需要考虑的问题。

#### 2.2.4　后处理

后处理的目的是使预测的连接部件数量与真实连接部件数量相匹配。考虑到ATLAS数据集中大病变的连接部件数量相较于小病变更多，MAPPING[60]提出的后处理策略，首先通过消除最大概率小于0.7的小连接组件来减少小病变的连接组件数量。然后，再通过对较大的病变应用高阈值（0.55）将连接组件分成几个部分。实验结果显示这种后处理策略能提高测试集上的评估指标。

#### 2.2.5　小结

一是数据扩增能够扩充训练样本，对模型训练有直接帮助。然而，要注意：跟自然图像相比，医学影像的组织结构更复杂，边界更模糊。

二是先验知识不会增大模型复杂度，又能降低预测难度。不过，如何保证先验知识的高质量是值得考虑的问题，否则会给模型带来错误的引导。

三是优化训练流程同样能提升模型的预测能力。然而，最佳实践是在训练流程变得复杂时，尽量保持测试阶段的简洁。

四是后处理对分割结果也有相当大的影响。

### 2.3　损失函数

除了网络结构、训练策略，所使用的损失函数对模型性能也有影响。本节将介绍脑卒中病灶分割中常用的损失函数及其组合。

#### 2.3.1　戴斯损失函数

戴斯（Dice）损失主要为了解决语义分割中正负样本极度不平衡的问题。然而，如果正样本为小目标，则会产生严重的震荡，因为在只有前景和背景的情况下，一旦对小目标的部分像素预测错误，那么损失函数的值将大幅变动，从而导致梯度变化剧烈。

Dice 损失计算的就是 1-Dice 系数，如式（1）所示：

$$\mathcal{L}_{\text{Dice}}=1-\frac{2\sum_{i=1}^{w}\sum_{j=1}^{h}p_{ij}\dot{p}_{ij}+\epsilon}{\sum_{i=1}^{w}\sum_{j=1}^{h}p_{ij}^{2}+\sum_{i=1}^{w}\sum_{j=1}^{h}\dot{p}_{ij}^{2}+\epsilon} \tag{1}$$

其中，$p_{ij}\in\{0,1\}$代表第 $i$ 行、第 $j$ 列像素的实际标注；$\dot{p}_{ij}\in[0,1]$代表第 $i$ 行、第 $j$ 列像素的病灶预测概率；$w\times h$ 代表输出尺寸；分母上的 $\epsilon=10^{-5}$ 用于防止分母为 0；当实际标注无病灶且预测完全正确时，分子上的 $\epsilon$ 用于使 $\mathcal{L}_{\text{Dice}}=0$。

#### 2.3.2　Dice损失和其他损失结合

（1）与二元交叉熵损失的结合

二元交叉熵（binary cross entropy，BCE）损失逐个检查每个像素，将每个像素类别的预测结果（概率分布向量）与独热编码标签向量进行比较，可用于大多数语义分割场景中。但它有一个明显的缺点：在只需要分割前景和背景的情况下，如果前景像素的数量远远小于背景像素（归属于正负样本的像素极度不平衡），就会使得模型严重偏向背景，导致分割效果不佳。BCE 损失计算式如下：

$$\mathcal{L}_{\text{BCE}}=-\left[p_{ij}\log\dot{p}_{ij}+\left(1-p_{ij}\right)\log\left(1-\dot{p}_{ij}\right)\right] \tag{2}$$

因此，研究人员考虑将Dice损失和BCE损失等权重相加，在正负样本不平衡的情况下，可以相互弥补局限性[42,48,53,55-56]。

（2）与边界损失的结合

针对Dice损失和BCE损失对语义分割中正负样本不平衡比较敏感的问题，边界损失函数[104]被提出，目的是最小化实际标注和预测分割的边界之间的距离。

用$G$表示二维空间$\Omega$中实际标注的二值图像，$\partial G$表示它的边界，$q_n$表示第$n$个像素在二维空间$\Omega$中的位置，水平集$\phi_G$: $\Omega \to R$，如式（3）所示。

$$\phi_G(q_n)=\begin{cases}-D_G(q_n), & q_n \in G\\ D_G(q_n), & \text{其他}\end{cases} \tag{3}$$

式中，$D_G$: $\Omega \to \mathrm{R}^+$用于评估$q_n$和$\partial G$上最近距离点$z_{\partial G}$的欧几里得距离。欧几里得距离的计算方式是$D_G(q_n)=||q_n-z_{\partial G}||_2$。

将边界损失函数视作$\partial G$的加权平均，其计算式如下：

$$\mathcal{L}_B=\frac{1}{N}\sum_{n=1}^{N}s_\theta(n)\phi_G(q_n) \tag{4}$$

式中，$N$表示二维空间$\Omega$中点的个数，$s_\theta(\cdot)$表示参数为$\theta$、输出为概率值的分割网络。

所以，有人将Dice损失和边界损失结合使用，以克服Dice损失对正负样本不平衡比较敏感的问题[44,67]。

（3）与实例级损失的结合

病灶分割任务存在严重的实例不平衡问题，即较大的实例支配较小的实例。文献[63]将Dice损失与实例级和实例中心（instance-wise and center-of-instance，ICI）损失结合，以减轻实例不平衡问题。

ICI损失受blob损失[105]与LesLoss[106]启发，由实例级损失与实例中心损失这两部分组成。用$y_c$表示二维空间$\Omega$中实际标注针对类别$c$的二值图像，$\dot{y}_c$表示分割网络针对类别$c$的预测输出。首先使用连接组件分析（connected component analysis，CCA）从$y_c$和$\dot{y}_c$中分别提取单个实例，提取到的连接组件分别表示为$I_{y_c,n}$与$I_{\dot{y}_c,n}$，其中$n$表示组件序号。

实例级损失通过比较每个真实实例和与该真实实例相交的预测实例，评估每个实例的分割质量，如式（5）所示：

$$\mathcal{L}_{\text{instance}}=\frac{1}{N}\sum_{n=1}^{N}\mathcal{L}_{\text{seg}}\left(\left\{I_{\dot{y}_c,m}\middle|\{I_{\dot{y}_c,m}\cap I_{y_c,n}\neq\varnothing\}\right\}_{m=1}^{M},I_{y_c,n}\right) \tag{5}$$

其中$N$表示图像中真实实例的总数，$M$表示预测输出中实例的总数，$\varnothing$表示空集。

实例中心损失度量标准化实例的分割质量，其中每个实例的大小和形状根据质心归一化为正方形（2D）或立方体（3D），记为$C(I,\delta)$。标准化尺寸由参数$\delta$控制，默认为1。具体计算式如下：

$$\mathcal{L}_{\text{CIS}}=\mathcal{L}_{\text{seg}}(C(\{I_{\dot{y}_c,m}\}_{m=1}^{M},\delta),C(\{I_{y_c,n}\}_{n=1}^{N},\delta)) \tag{6}$$

式（5）和式（6）中$\mathcal{L}_{seg}$可以是任意的分割损失，如Dice损失或焦点（focal）损失[107]。

（4）与目标感知损失的结合

为了减轻正负样本不平衡的问题，各种目标感知损失函数被提出。加权交叉熵[108]对正样本损失进行加权，当权重α>1时，可以减少错误负样本的数量；当权重α<1时，可以减少错误正样本。平衡交叉熵[109]额外对负样本损失给予权重，用于调整错误正样本率和错误负样本率。基于平衡交叉熵，focal损失[107]进一步提出了动态权重$\gamma$，使得简单样本的损失变小，而困难样本的损失变大，增加了对困难样本的关注，使模型更多地集中在困难样本的学习上。

然而，以上损失仅考虑正样本、不考虑负样本，会导致召回率和准确率不平衡。TSRL-Net[64]提出了一种新的目标感知损失，通过调节聚焦因子来扩展聚焦样本区域，其中原本只需要聚焦正样本的聚焦区域被扩展为包括部分负样本。聚焦样本区域$A$的计算式如下：

$$A_{x\times\beta+m,y\times\beta+n}=\text{sign}\left(\sum_{i=0}^{\beta}\sum_{j=0}^{\beta}g_{x\times\beta+i,y\times\beta+j}\right) \tag{7}$$

其中，$\beta$表示聚焦因子，$g$表示真实分割图，$m,n\in[0,\beta)$，$x,y\in[0,H/\beta)$，$H$表示图像高度。关注聚焦区域内所有样本的损失，可以在强烈关注正样本损失的同时，高度关注少量的负样本损失，以平衡准确率与召回率。该损失计算式如下：

$$loss_f=\frac{\sum_{i=0}^{N}-(A_i\alpha+|1-A_i|(1-\alpha))\left(1-p_ig_i-\left(1-p_i\right)\left(1-g_i\right)\right)^{\gamma}\log\left(p_ig_i+\left(1-p_i\right)\left(1-g_i\right)\right)}{N} \tag{8}$$

其中，$\alpha$表示针对聚焦样本的损失权重，设置成大于0.5以更关注聚焦区域，$\gamma\in[0,5)$调控对困难样品和简单样品的平衡程度，$N$表示图像中样本点的总数，$p$表示预测的分割概率图。

#### 2.3.3　小结

一是单一损失函数使用的是Dice损失，它是最常用的分割损失，不过在训练过程中，小病灶分割会影响损失函数的稳定性。

二是多损失结合能够帮助克服Dice损失的缺点，然而，各个损失函数的权重设置有待研究。对于不同的任务，最佳权重可能不一样。

### 2.4　各类方法性能对比

在介绍完各类方法后，笔者在表3中对比了主要方法在公共数据集上的性能。Multi-path 2.5D CNN[39]和Fine-grained segmentation[65]在实现过程中使用了外部数据，因此他们的受试者个数$n$由两部分组成：公共数据集受试者个数与外部数据集受试者个数，用“+”来分隔。METranse[59]的Dice系数远高于其他方法，可能是全局Dice系数。

由于那些基于ATLAS数据集的文献未明确指出训练集和测试集的划分方式，表3无法提供详细的数据划分信息。虽然各文献数据集划分方式不同，不同方法的结果不具备直接可比性，但是笔者还是给出了对应的实验结果，以提供一定的参考。可以看到，那些使用$k$折交叉验证的方法的实验结果差异较小，具有一定的比较价值。

对于定性结果，在调研了大量文献后，笔者发现很难找到覆盖度较广的实验结果，各文献所选的实例不一样，格式也不一样，难以总结，所以没有展示这个结果。

此外，基于深度学习的脑卒中病灶分割方法还有较大提升空间。第3节将介绍该任务的挑战以及潜在的解决方案。

## 3　挑战与展望

虽然深度学习在脑卒中病灶分割方向上具有广阔的应用前景和良好的潜力，但目前仍有许多未解决的困难。

### 3.1　病灶的多样性

对于不同患者，病灶的形状、位置、大小和数目各不相同（如图1所示），在样本稀缺的情况下，病灶的多样性会大幅提升模型的训练难度。

图15显示了ATLAS v1.2数据集中各中心病变大小的分布情况。从之前的研究[42,45,68]的实验结果可以发现，不管如何优化损失函数、网络结构和优化算法，模型在包含大量小病灶的中心（如中心8）上的表现总是较差。对于大病灶，模型更容易获得较高的Dice系数。

针对该问题，可以考虑数据扩增策略。旋转、翻转、平移、缩放或剪裁等传统数据扩增策略很难保持医学影像的形态特点，最好是设计更适合医学影像的扩增方式。比如，使用生成对抗网络[110-111]形成接近真实的样本数据：用生成模型捕捉真实样本的分布，根据分布由随机噪声生成假样本，判别器用于判别输入样本的真假；生成器和判别器通过不断地对抗训练，在提升判别器的辨别能力的同时，也帮助生成器产生更像真实样本的假样本。

### 3.2　病灶的边界模糊性

对于病灶分割而言，边界划分尤为重要。而医学影像中组织纹理复杂，磁共振图像容易产生伪影，所以病灶边界通常比较模糊，不易辨别。

针对该问题，可以设计跟边界相关的损失函数[104]，在损失优化过程中提高模型对边界的注意力，从而帮助模型更准确地划分病灶边界。除此之外，额外采集健康被试的医学影像，将其映射到标准空间作为正常组织的参照，模型预测出异常病灶后，还可以根据患者与健康被试的图像的强度差，估算预测病灶的边界不确定性，帮助医生更精确地发现病灶。

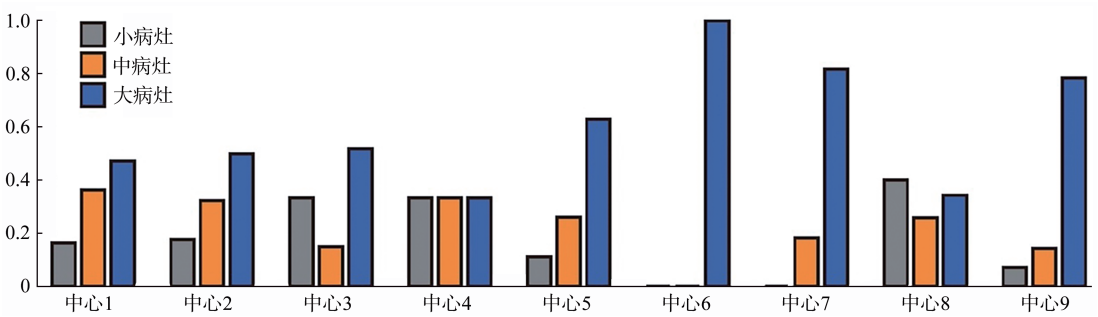



图15　9个中心的病灶大小分布（纵轴代表所占比例）。小病灶：小于1×10体素；中病灶：1×10至5×10体素；大病灶：大于5×10体素。图片来自文献［68］。

表3 深度学习用于脑卒中病灶分割的现有方法在公共数据集上的性能对比

| 数据集 | 受试者个数 $n$ | 验证方式 | 方法 | 实验结果（Dice 系数） |
|---|---|---|---|---|
| ATLAS | 229 | 留一域法交叉验证 | SAN-Net[68] | 0.5711 |
| | 229 | 留一域法交叉验证 | FAN-Net[69] | 0.5591 |
| | 239 | 六折交叉验证 | W-Net[67] | 0.6176 |
| | 229 | 五折交叉验证 | X-Net[37] | 0.4867 |
| | 54 + 45 | 五折交叉验证 | Multi-path 2.5D CNN[39] | 0.62 |
| | 229 | 五折交叉验证 | MI-UNet[42] | 0.5672 |
| | 229 | 五折交叉验证 | A 3D+2D CNN approach[44] | 0.5625 |
| | 239 | 五折交叉验证 | Hybrid Dice loss[48] | 0.5738 |
| | 229 | 五折交叉验证 | DFENet[49] | 0.5457 |
| | 229 | 五折交叉验证 | MDAN[53] | 0.5662 |
| | 220 | 五折交叉验证 | MLiRA-Net[61] | 0.6119 |
| | 304 | 五折交叉验证 | METranse[59] | 0.931 |
| | 229 | 80% 训练、20% 测试 | D-UNet[38] | 0.5349 |
| | 229 | 80% 训练、20% 测试 | e-UNets[55] | 0.592 |
| | 229 | 80% 训练、20% 测试 | AGMR-Net[58] | 0.594 |
| | 229 | 80% 训练、20% 测试 | TSRL-Net[64] | 0.634 |
| | 239 | 60% 训练、20% 验证、20% 测试 | Variational mode decomposition[50] | 0.6684 |
| | 229 | 70% 训练、30% 测试 | MSDF-Net[40] | 0.6875 |
| | 229 | 70% 训练、30% 测试 | MUDCap3[46] | 0.67 |
| | 229 | 52% 训练、18% 验证、30% 测试 | CLCI-Net[41] | 0.581 |
| | 229 | 52% 训练、18% 验证、30% 测试 | Multi-task[47] | 0.518 |
| | 239 | 76% 训练、11% 验证、13% 测试 | 3D residual CNN[45] | 0.64 |
| | 229 | 80% 训练、10% 验证、10% 测试 | ME-Net[51] | 0.6353 |
| | 220 | 77% 训练、23% 测试 | Priors from healthy subjects[52] | 0.6256 |
| | 229 | 77% 训练、23% 测试 | STHarDNet[56] | 0.5547 |
| | 220 | 77% 训练、23% 测试 | CarveMix[62] | 0.6391 |
| | 220 + 60 | 80% 训练、10% 验证、10% 测试 | Fine-grained segmentation[65] | 0.5596 |
| ATLAS R2.0 | 655 | 五折交叉验证 | 3D UNet[54] | 0.65 |
| | 955 | 655 训练、300 测试 | MAPPING [60] | 0.6667 |
| | 955 | 600 训练、55 验证、300 测试 | ICI loss[63] | 0.5817 |
| | 655 | 80% 训练、20% 测试 | SQMLP-net[66] | 0.7098 |
| ISLES | 64 | 五折交叉验证 | AGMR-Net[58] | 0.614 |
| | 64 | 五折交叉验证 | METranse[59] | 0.793 |
| | 28 | 五折交叉验证 | TSRL-Net[64] | 0.640 |

### 3.3 数据集标注难度大

目前，病灶分割的黄金标准是临床医生的手动标记，这项工作需要丰富的医学经验。而大多数专业医生需要在一线筛查和诊疗患者，这就导致了医学影像标记难度较大。

针对该问题，可以考虑结合半监督学习[112-113]和对比学习[114]，利用少量的标记数据和大量的未标记数据来训练模型。先利用未标记数据做预训练，把未标记原数据对应的两种扩增结果看作相似的正样本对，而将原数据和其他数据看作不相似的负样本对，用正负样本对帮助网络学习图像的高阶特征，再利用已标记数据做有监督学习，对网络进行微调，从而提升网络的分割能力。

在脑卒中病灶分割任务上，已有相关研究利用无监督学习[115]、弱监督学习[116]或半监督学习[117]来解决这一问题。

### 3.4 不同中心的影像分布差异大

机器学习模型表现好的前提假设是所有数据分布一致，而不同的MR扫描仪、扫描序列和种群都会导致影像数据分布存在差异。在这种情况下，提

升模型的泛化性也是一大难题。表4列出了ATLAS v1.2数据集的数据来源，可以看到各中心间存在巨大差异。首先，9个影像中心涵盖4个国家8个城市，不同人口的血管区域不同。其次，不同中心采用不同的3T磁共振扫描仪采集。第三，也是最重要的一点，即使使用同一台扫描仪，也可能使用不同的成像协议。这些因素导致在不同中心间进行泛化存在巨大难度。在ATLAS R2.0版本中，随着数据量的增加，引入更多的影像中心，使泛化问题变得更加困难。

**表4　ATLAS v1.2数据来源（来自文献［68］）**

| 中心信息 | 位置 | 扫描仪 | 患者人数 |
|---|---|---|---|
| Medical University General Hospital | 中国天津 | GE 750 Discovery | 55 |
| University of Tübingen | 德国图宾根 | GE Signa Excite | 34 |
| Sunnaas Rehabilitation Hospital | 挪威内索登 | Siemens Trio | 27 |
| NORMENT and KG Jebsen Centre for Psychosis Research | 挪威奥斯陆 | Siemens Trio | 12 |
| Department of Psychology | 挪威菲利普斯 | Phillips Achieva | 27 |
| Child Mind Institute | 美国纽约 | Siemens Trio | 14 |
| Nathan S. Kline Institute for Psychiatric Research | 美国奥兰治堡 | Siemens Trio | 11 |
| University of Texas Medical Branch | 美国加尔维斯顿 | GE 750 Discovery | 35 |
| University of Michigan | 美国安娜堡 | Siemens Trio | 14 |

针对该问题，可以将领域自适应[118-120]、元学习[121-122]与脑卒中影像分割结合。比如，在训练阶段加入判别器，源域数据和目标域数据被同时输入编码器，判别器对编码器输出的特征图进行分类；编码器和判别器通过不断地对抗训练，对齐源域和目标域的特征图，使得模型与目标测试数据的分布相匹配[123]。元学习使模型获取调整超参数的能力，可以在训练数据的基础上快速适应与训练数据分布不同的新数据[124-125]。

已有工作针对这一问题做出改进，可以利用留一域法[126]对模型在未知域上的泛化能力进行验证。受传统z-score归一化[127]和动态网络[128]的启发，SAN-Net[68]设计了一种掩码自适应实例归一化（masked adaptive instance normalization，MAIN）算法，通过动态学习输入的仿射参数，将来自不同域的输入MR图像标准化为与域无关的样式，以减少域间的差异。SAN-Net[68]还利用梯度反转层（gradient reversal layer，GRL）[129-130]迫使UNet编码器使用域分类器学习域不变表示，以进一步提高模型泛化能力。同样受传统z-score归一化[127]的启发，FAN-Net[69]提出基于傅里叶的自适应标准化（fourier-based adaptive normalization，FAN）模块，针对不同域的频域幅度分量进行自适应仿射参数学习，对源图像的风格信息进行动态归一化，以提高模型泛化能力。

### 3.5　保护患者的数据隐私

医学影像数据有限，常常需要结合多中心数据进行模型训练。由于严格的法律和伦理要求，患者隐私保护问题同样限制了医学影像的发展，阻碍了深度学习在脑卒中病灶分割中的广泛应用。

针对该问题，可以考虑将联邦学习用于医学影像[131-133]。联邦学习主要包含服务端和客户端两部分，客户端指的是各地医疗中心，客户端数据仅留在本地训练，训练得到的模型梯度或模型参数传输至服务端，这样就能在保护患者隐私的前提下生成全局模型。在此基础上，联邦学习与元学习[122]或领域自适应[134]的结合也是方向之一，既能保护患者隐私，也能提升病灶分割模型的泛化性。另外，也可以将联邦学习和多任务学习[135]结合，多任务可以包括重建、分割等，从而达到相互补充、提高网络特征提取能力的效果。

### 3.6　临床部署设备不足

一般来说，模型复杂度越高，其拟合能力越强，但较大模型的运行过程会占据更多的内存、显存，给计算设备带来过大的负担。大多数医院部署设备不足，大模型的计算效率低下，这是深度学习在脑卒中病灶分割领域进行临床应用时必须考虑的实际问题。

针对该问题，可以考虑模型压缩[136]，包含剪枝、知识蒸馏、结构重参数化等。以知识蒸

馏[137-138]为例：先构建一个轻量化小模型，利用性能更好的大模型的监督信息来训练这个小模型，使小模型达到更好的性能和精度。还可以借鉴RepVGG[139]，训练阶段使用较复杂的多分支模型，测试阶段把多分支模型的参数等价转换成单分支模型的参数，为临床部署减轻负担。

### 3.7 结合先进的分割框架

近年来，各种分割框架在医学图像分割任务中取得了令人满意的结果，如利用扩散模型产生既准确又多样的分割结果[140-142]；建立通用的医学图像分割模型，应对单一任务数据缺乏的问题[143-145]。如何将这些优秀的框架应用于脑卒中病灶分割的任务，需要进一步考虑。

## 4 结束语

深度学习对于脑卒中病灶分割的研究具有重要的意义和价值。本文先介绍了脑卒中病灶分割的公共数据集；然后，从网络框架、训练策略、损失函数这3个角度出发，分析了深度学习在脑卒中病灶分割领域的研究进展；最后，讨论了该领域面临的困难和挑战，展望了未来的发展趋势。

综上所述，深度学习为解决脑卒中病灶分割问题提供了全自动化的新技术，有很强的发展潜力，但要将其应用到医学临床应用中，仍存在很多问题。如何在保护患者隐私的前提下，提升模型的预测能力和泛化能力，值得我们思考和挖掘。

## 参考文献：


[1] KAMALAKANNAN S, GUDLAVALLETI A S V, GUDLAVALLETI V S M, et al. Incidence & prevalence of stroke in India: a systematic review[J]. The Indian Journal of Medical Research, 2017, 146(2): 175-185.

[2] VIRANI S S, ALONSO A, APARICIO H J, et al. Heart disease and stroke statistics-2021 update: a report from the American heart association[J]. Circulation, 2021, 143(8): e254-e743.

[3] DONKOR E S. Stroke in the 21st century: a snapshot of the burden, epidemiology and quality of life[J]. Stroke Research and Treatment, 2018, 2018: 1-10.

[4] WAJNGARTEN M, SILVA G S. Hypertension and stroke: update on treatment[J]. European Cardiology, 2019, 14(2): 111-115.

[5] YANG F, ZHAO J N, XU H D. Characteristics of hemorrhagic stroke following spine and joint surgeries[J]. BioMed Research International, 2017, 2017: 1-5.

[6] SAVER J L. Penumbral salvage and thrombolysis outcome: a drop of brain, a week of life[J]. Brain, 2017, 140(3): 519-522.

[7] EL-KOUSSY M, SCHROTH G, BREKENFELD C, et al. Imaging of acute ischemic stroke[J]. European Neurology, 2014, 72(5-6): 309-316.

[8] MAIR G, WARDLAW J M. Imaging of acute stroke prior to treatment: current practice and evolving techniques[J]. The British Journal of Radiology, 2014, 87(1040): 20140216.

[9] KIM J, CHHOUR P, HSU J, et al. Use of nanoparticle contrast agents for cell tracking with computed tomography[J]. Bioconjugate Chemistry, 2017, 28(6): 1581-1597.

[10] VILELA P, ROWLEY H A. Brain ischemia: CT and MRI techniques in acute ischemic stroke[J]. European Journal of Radiology, 2017, 96: 162-172.

[11] QIU W, KUANG H, TELEG E, et al. Machine learning for detecting early infarction in acute stroke with non-contrast-enhanced CT[J]. Radiology, 2020, 294(3): 638-644.

[12] LIANG K, HAN K, LI X, et al. Symmetry-enhanced attention network for acute ischemic infarct segmentation with non-contrast CT images [C]//Proceedings of the 24th International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2021: 432-441.

[13] 施俊, 汪琳琳, 王珊珊, 等. 深度学习在医学影像中的应用综述[J]. 中国图象图形学报, 2020, 25(10): 1953-1981.
SHI J, WANG L L, WANG S S, et al. Applications of deep learning in medical imaging: a survey[J]. Journal of Image and Graphics, 2020, 25 (10): 1953-1981.

[14] YEDAVALLI V S, TONG E, MARTIN D, et al. Artificial intelligence in stroke imaging: current and future perspectives[J]. Clinical Imaging, 2021, 69: 246-254.

[15] 秦海强, 张亚清, 张婧, 等. 人工智能在缺血性卒中诊断与治疗中的应用[J]. 中国现代神经疾病杂志, 2021, 21(1): 21-24.
QIN H Q, ZHANG Y Q, ZHANG J, et al. Application of artificial intelligence in diagnosis and treatment of ischemic stroke[J].Chinese Journal of Contemporary Neurology and Neurosurgery, 2021, 21(1): 21-24.

[16] 于长申, 巫嘉陵. 人工智能在脑卒中管理中的研究进展[J]. 中国现代神经疾病杂志, 2021, 21(1): 14-20.
YU C S, WU J L. Research progress of artificial intelligence in stroke management[J]. Chinese Journal of Contemporary Neurology and Neurosurgery, 2021, 21(1): 14-20.

[17] TYAN Y S, WU M C, CHIN C L, et al. Ischemic stroke detection system with a computer-aided diagnostic ability using an unsupervised feature perception enhancement method[J]. International Journal of Biomedical Imaging, 2014, 2014: 947539.

[18] KARTHIK R, MENAKA R, JOHNSON A, et al. Neuroimaging and deep learning for brain stroke detection-a review of recent advancements and future prospects[J]. Computer Methods and Programs in Biomedicine, 2020, 197: 105728.

[19] CASTILLO D, LAKSHMINARAYANAN V, RODRIGUEZ-ALVAREZ M J. MR images, brain lesions and deep learning[J]. Applied Sciences, 2021, 11(4): 1675.

[20] SOUN J E, CHOW D S, NAGAMINE M, et al. Artificial intelligence and acute stroke imaging[J]. American Journal of Neuroradiology, 2021, 42(1): 2-11.

[21] MAIER O, MENZE B H, VON DER GABLENTZ J, et al. ISLES 2015-a public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI[J]. Medical Image Analysis, 2017, 35: 250-269.

[22] LIEW S L, ANGLIN J M, BANKS N W, et al. A large, open source dataset of stroke anatomical brain images and manual lesion segmentations[J]. Scientific Data, 2018, 5(1): 1-11.

[23] LIEW S L, LO B P, DONNELLY M R, et al. A large, curated, open-source stroke neuroimaging dataset to improve lesion segmentation al-

gorithms[J]. Scientific Data, 2022, 9(1): 320.

[24] LIU X, FAES L, KALE A U, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis[J]. The Lancet Digital Health, 2019, 1(6): e271-e297.

[25] HUANG S C, PAREEK A, SEYYEDI S, et al. Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines[J]. NPJ Digital Medicine, 2020, 3 (1): 136.

[26] WANG T, LEI Y, FU Y, et al. A review on medical imaging synthesis using deep learning and its clinical applications[J]. Journal of Applied Clinical Medical Physics, 2021, 22(1): 11-36.

[27] 陈朝一, 许波, 吴英, 等. 医学图像处理中的注意力机制研究综述[J]. 计算机工程与应用, 2021, 58(5): 23-3.
CHEN Z Y, XU B, WU Y, et al. Overview of research on attention mechanism in medical image processing[J]. Computer Engineering and Applications, 2021, 58(5): 23-3.

[28] 陈弘扬, 高敬阳, 赵地, 等. 深度学习与生物医学图像分析2020年综述[J]. 中国图象图形学报, 2021, 26(3): 475-486.
CHEN H Y, GAO J Y, ZHAO D, et al. Review of the research progress in deep learning and biomedical image analysis till 2020[J]. Journal of Image and Graphics, 2021, 26(3): 475-486.

[29] TAJBAKHSH N, JEYASEELAN L, LI Q, et al. Embracing imperfect datasets: a review of deep learning solutions for medical image segmentation[J]. Medical Image Analysis, 2020, 63: 101693.

[30] LIU X, SONG L, LIU S, et al. A review of deep-learning-based medical image segmentation methods[J]. Sustainability, 2021, 13(3): 1224.

[31] FU Y, LEI Y, WANG T, et al. A review of deep learning based methods for medical image multi-organ segmentation[J]. Physica Medica, 2021, 85: 107-122.

[32] 冯龙锋, 陈英, 周滔辉, 等. CT图像肺及肺病变区域分割方法综述[J]. 中国图象图形学报, 2022, 27(3): 722-749.
FENG L F, CHEN Y, ZHOU T H, et al. Review of human lung and lung lesion regions segmentation methods based on CT images[J]. Journal of Image and Graphics, 2022, 27(3): 722-749.

[33] 周涛, 董雅丽, 霍兵强, 等. U-Net网络医学图像分割应用综述[J]. 中国图象图形学报, 2021, 26(9): 2058-2077.
ZHOU T, DONG Y L, HUO B Q, et al. U-Net and its applications in medical image segmentation: a review[J]. Journal of Image and Graphics, 2021, 26(9): 2058-2077.

[34] CHEN L, BENTLEY P, RUECKERT D. Fully automatic acute ischemic lesion segmentation in DWI using convolutional neural networks [J]. NeuroImage: Clinical, 2017, 15: 633-643.

[35] GUERRERO R, QIN C, OKTAY O, et al. White matter hyperintensity and stroke lesion segmentation and differentiation using convolutional neural networks[J]. NeuroImage: Clinical, 2018, 17: 918-934.

[36] LIU Z, CAO C, DING S, et al. Towards clinical diagnosis: automated stroke lesion segmentation on multi-spectral MR image using convolutional neural network[J]. IEEE Access, 2018, 6: 57006-57016.

[37] QI K, YANG H, LI C, et al. X-Net: brain stroke lesion segmentation based on depthwise separable convolution and long-range dependencies [C]//Proceedings of the 22th International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2019: 247-255.

[38] ZHOU Y, HUANG W, DONG P, et al. D-UNet: a dimension-fusion U shape network for chronic stroke lesion segmentation[J]. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2019, 18 (3): 940-950.

[39] XUE Y, FARHAT F G, BOUKRINA O, et al. A multi-path 2.5 dimensional convolutional neural network system for segmenting stroke lesions in brain MRI images[J]. NeuroImage: Clinical, 2020, 25: 102118.

[40] LIU X, YANG H, QI K, et al. MSDF-Net: multi-scale deep fusion network for stroke lesion segmentation[J]. IEEE Access, 2019, 7: 178486-178495.

[41] YANG H, HUANG W, QI K, et al. CLCI-Net: cross-level fusion and context inference networks for lesion segmentation of chronic stroke [C]// Proceedings of the 22th International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2019: 266-274.

[42] ZHANG Y, WU J, LIU Y, et al. MI-UNet: multi-inputs U-net incorporating brain parcellation for stroke lesion segmentation from T1-weighted magnetic resonance images[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 25(2): 526-535.

[43] WU J, TANG X. A large deformation diffeomorphic framework for fast brain image registration via parallel computing and optimization [J]. Neuroinformatics, 2020, 18: 251-266.

[44] ZHANG Y, WU J, LIU Y, et al. A 3d+ 2d cnn approach incorporating boundary loss for stroke lesion segmentation[C]//Proceedings of the 11th International Workshop on Machine Learning in Medical Imaging. Cham: Springer, 2020: 101-110.

[45] TOMITA N, JIANG S, MAEDER M E, et al. Automatic post-stroke lesion segmentation on MR images using 3D residual convolutional neural network[J]. NeuroImage: Clinical, 2020, 27: 102276.

[46] SAHAYAM S, ABIRAMI A, JAYARAMAN U. A novel modified U-shaped 3D capsule network (MUDCap3) for stroke lesion segmentation from brain MRI[C]//Proceedings of the 2020 IEEE 4th International Conference on Information and Communication Technology. Piscataway: IEEE, 2020: 1-6.

[47] QI K, GONG Y, LIU X, et al. Multi-task MR imaging with iterative teacher forcing and re-weighted deep learning[EB]. arXiv preprint, 2020, arXiv:2011.13614.

[48] LU Y, ZHOU J H, GUAN C. Minimizing hybrid Dice loss for highly imbalanced 3D neuroimage segmentation[C]//Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society. Piscataway: IEEE, 2020: 1059-1062.

[49] BASAK H, HUSSAIN R, RANA A. DFENet: a novel dimension fusion edge guided network for brain mri segmentation[J]. SN Computer Science, 2021, 2: 1-11.

[50] PAING M P, TUNGJITKUSOLMUN S, BUI T H, et al. Automated segmentation of infarct lesions in T1-weighted MRI scans using variational mode decomposition and deep learning[J]. Sensors, 2021, 21(6): 1952.

[51] ZHANG X, XU H, LIU Y, et al. A multiple encoders network for stroke lesion segmentation[C]//Proceedings of the 4th Chinese Conference on Pattern Recognition and Computer Vision. Cham: Springer, 2021: 524-535.

[52] LIU C, ZENG X, LIANG K, et al. Improved brain lesion segmentation with anatomical priors from healthy subjects[C]//Proceedings of the 24th International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2021: 186-195.

[53] BAO Q, MI S, GANG B, et al. MDAN: mirror difference aware network for brain stroke lesion segmentation[J]. IEEE Journal of Biomedical and Health Informatics, 2021, 26(4): 1628-1639.

[54] VERMA K, KUMAR S, PAYDARFAR D. Automatic segmentation

and quantitative assessment of stroke lesions on MR images[J]. Diagnostics, 2022, 12(9): 2055.

[55] SHIN H, AGYEMAN R, RAFIQ M, et al. Automated segmentation of chronic stroke lesion using efficient U-Net architecture[J]. Biocybernetics and Biomedical Engineering, 2022, 42(1): 285-294.

[56] GU Y, PIAO Z, YOO S J. STHarDNet: swin transformer with hardNet for MRI segmentation[J]. Applied Sciences, 2022, 12(1): 468.

[57] CHAO P, KAO C Y, RUAN Y S, et al. Hardnet: a low memory traffic network[C]//Proceedings of the 2019 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2019: 3552-3561.

[58] DU X, MA K, SONG Y. AGMR-Net: attention-guided multiscale recovery framework for stroke segmentation[J]. Computerized Medical Imaging and Graphics, 2022, 101: 102120.

[59] WANG J, WANG S, LIANG W. METrans: multi-encoder transformer for ischemic stroke segmentation[J]. Electronics Letters, 2022, 58(9): 340-342.

[60] HUO J, CHEN L, LIU Y, et al. MAPPING: model average with post-processing for stroke lesion segmentation[EB]. arXiv preprint, 2022, arXiv:2211.15486.

[61] WU Z, ZHANG X, LI F, et al. Multi-scale long-range interactive and regional attention network for stroke lesion segmentation[J]. Computers and Electrical Engineering, 2022, 103: 108345.

[62] ZHANG X, LIU C, OU N, et al. CarveMix: a simple data augmentation method for brain lesion segmentation[J]. Neuro Image, 2023, 271: 120041.

[63] RACHMADI M F, POON C, SKIBBE H. Improving segmentation of objects with varying sizes in biomedical images using instance-wise and center-of-instance segmentation loss function[J]. arXiv preprint, 2023, arXiv:2304.06229.

[64] LI L, MA K, SONG Y, et al. TSRL-Net: target-aware supervision residual learning for stroke segmentation[J]. Computers in Biology and Medicine, 2023, 159: 106840.

[65] LEE J, LEE M, LEE J, et al. Fine-grained brain tissue segmentation for brain modeling of stroke patient[J]. Computers in Biology and Medicine, 2023, 153: 106472.

[66] LIU L, CHANG J, LIANG G, et al. Simulated quantum mechanics-based joint learning network for stroke lesion segmentation and TICI grading[J]. IEEE Journal of Biomedical and Health Informatics, 2023, 27(7): 3372-3383.

[67] WU Z, ZHANG X, LI F, et al. W-Net: a boundary-enhanced segmentation network for stroke lesions[J]. Expert Systems with Applications, 2023: 120637.

[68] YU W, HUANG Z, ZHANG J, et al. SAN-Net: learning generalization to unseen sites for stroke lesion segmentation with self-adaptive normalization[J]. Computers in Biology and Medicine, 2023, 156: 106717.

[69] YU W, LEI Y, SHAN H. FAN-Net: fourier-based adaptive normalization for cross-domain stroke lesion segmentation[J]. arXiv preprint, 2023, arXiv:2304.11557.

[70] LECUN Y, BOSER B, DENKER J S, et al. Backpropagation applied to handwritten zip code recognition[J]. Neural Computation, 1989, 1 (4): 541-551.

[71] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 3431-3440.

[72] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2015: 234-241.

[73] LIU L, CHENG J, QUAN Q, et al. A survey on U-shaped networks in medical image segmentations[J]. Neurocomputing, 2020, 409: 244-258.

[74] 罗恺锴, 王婷, 叶芳芳. 引入注意力机制和多视角融合的脑肿瘤 MR 图像 U-Net 分割模型[J]. 中国图象图形学报, 2021, 26(9): 2208-2218.
LUO K K, WANG T , YE F F. U-Net segmentation model of brain tumor MR image based on attention mechanism and multi-view fusion [J]. Journal of Image and Graphics, 2021, 26(9): 2208-2218.

[75] NIU Z, ZHONG G, YU H. A review on the attention mechanism of deep learning[J]. Neurocomputing, 2021, 452: 48-62.

[76] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132-7141.

[77] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks [C]//Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7794-7803.

[78] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]//Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 3146-3154.

[79] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]//Proceedings of the 2018 European Conference on Computer Vision. Cham: Springer, 2018: 3-19.

[80] WANG R, FAN J, LI Y. Deep multi-scale fusion neural network for multi-class arrhythmia detection[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 24(9): 2461-2472.

[81] SINHA A, DOLZ J. Multi-scale self-guided attention for medical image segmentation[J]. IEEE Journal of Biomedical and Health Informatics, 2020, 25(1): 121-130.

[82] LIANG X, LI N, ZHANG Z, et al. Incorporating the hybrid deformable model for improving the performance of abdominal CT segmentation via multi-scale feature fusion network[J]. Medical Image Analysis, 2021, 73: 102156.

[83] LIU X, YANG L, CHEN J, et al. Region-to-boundary deep learning model with multi-scale feature fusion for medical image segmentation [J]. Biomedical Signal Processing and Control, 2022, 71: 103165.

[84] 陈舞, 孙军梅, 李秀梅. 融合多尺度残差和注意力机制的特发性肺纤维化进展预测[J]. 中国图象图形学报, 2022, 27(3):812-826.
CHEN W, SUN J M, LI X M. Multi-scale residual and attention mechanism fusion based prediction for the progression of idiopathic pulmonary fibrosis[J]. Journal of Image and Graphics, 2022, 27(3): 812-826.

[85] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the 2018 European Conference on Computer Vision. Cham: Springer, 2018: 801-818.

[86] ABULNAGA S M, RUBIN J. Ischemic stroke lesion segmentation in CT perfusion scans using pyramid pooling and focal loss[C]//Proceedings of the 4th International Workshop on Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. Cham, Springer, 2019: 352-363.

[87] LIAO H, TANG Y, FUNKA-LEA G, et al. More knowledge is better: cross-modality volume completion and 3D+ 2D segmentation for intracardiac echocardiography contouring[C]//Proceedings of the 21th International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2018: 535-543.

[88] TRAN D, WANG H, TORRESANI L, et al. A closer look at spatiotem-

poral convolutions for action recognition[C]//Proceedings of the 2018 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 6450-6459.

[89] GAN W, WANG H, GU H, et al. Automatic segmentation of lung tumors on CT images based on a 2D & 3D hybrid convolutional neural network[J]. The British Journal of Radiology, 2021, 94: 20210038.

[90] USHINSKY A, BARDIS M, GLAVIS-BLOOM J, et al. A 3D-2D hybrid U-net convolutional neural network approach to prostate organ segmentation of multiparametric MRI[J]. American Journal of Roentgenology, 2021, 216(1): 111-116.

[91] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770-778.

[92] SHANG W, SOHN K, ALMEIDA D, et al. Understanding and improving convolutional neural networks via concatenated rectified linear units [C]//Proceedings of the 2016 International Conference on Machine Learning. New York: PMLR, 2016: 2217-2225.

[93] SABOUR S, FROSST N, HINTON G E. Dynamic routing between capsules[J]. Advances in Neural Information Processing Systems, 2017, 30.

[94] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 4700-4708.

[95] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the 2021 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2021: 10012-10022.

[96] CHLAP P, MIN H, VANDENBERG N, et al. A review of medical image data augmentation techniques for deep learning applications[J]. Journal of Medical Imaging and Radiation Oncology, 2021, 65(5): 545-563.

[97] NALEPA J, MARCINKIEWICZ M, KAWULOK M. Data augmentation for brain-tumor segmentation: a review[J]. Frontiers in Computational Neuroscience, 2019, 13: 83.

[98] ZHANG H, CISSE M, DAUPHIN Y N, et al. Mixup: beyond empirical risk minimization[EB]. arXiv preprint, 2017, arXiv:1710.09412.

[99] YUN S, HAN D, OH S J, et al. Cutmix: regularization strategy to train strong classifiers with localizable features[C]//Proceedings of the 2019 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2019: 6023-6032.

[100] LAHMIRI S. Image characterization by fractal descriptors in variational mode decomposition domain: application to brain magnetic resonance[J]. Physica A: Statistical Mechanics and its Applications, 2016, 456: 235-243.

[101] GUDIGAR A, RAGHAVENDRA U, CIACCIO E J, et al. Automated categorization of multi-class brain abnormalities using decomposition techniques with MRI images: a comparative study[J]. IEEE Access, 2019, 7: 28498-28509.

[102] ACHARYA U R, MEIBURGER K M, FAUST O, et al. Automatic detection of ischemic stroke using higher order spectra features in brain MRI images[J]. Cognitive Systems Research, 2019, 58: 134-142.

[103] CHLEMPER J S, CABALLERO J, HAJNAL J, et al. A deep cascade of convolutional neural networks for dynamic MR image reconstructio [J]. IEEE Transactions on Medical Imaging, 2017, 37: 491-503.

[104] KERVADEC H, BOUCHTIBA J, DESROSIERS C, et al. Boundary loss for highly unbalanced segmentation[C]//Proceedings of 2019 International Conference on Medical Imaging with Deep Learning. New York: PMLR, 2019: 285-296.

[105] KOFLER F, SHIT S, EZHOV I, et al. Blob loss: instance imbalance aware loss functions for semantic segmentation[EB]. arXiv preprint, 2022, arXiv:2205.08209.

[106] ZHANG H, ZHANG J, LI C, et al. ALL-Net: anatomical information lesion-wise loss function integrated into neural network for multiple sclerosis lesion segmentation[J]. NeuroImage: Clinical, 2021, 32: 102854.

[107] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2017: 2980-2988.

[108] PIHUR V, DATTA S, DATTA S. Weighted rank aggregation of cluster validation measures: a monte carlo cross-entropy approach[J]. Bioinformatics, 2007, 23(13): 1607-1615.

[109] XIE S, TU Z. Holistically-nested edge detection[C]//Proceedings of the 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2015: 1395-1403.

[110] HUANG S W, LIN C T, CHEN S P, et al. Auggan: cross domain adaptation with gan-based data augmentation[C]//Proceedings of the 2018 European Conference on Computer Vision. Cham: Springer, 2018: 718-731.

[111] BISSOTO A, VALLE E, AVILA S. Gan-based data augmentation and anonymization for skin-lesion analysis: a critical review[C]//Proceedings of the 2021 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 1847-1856.

[112] ZHOU Y, HE X, HUANG L, et al. Collaborative learning of semi-supervised segmentation and classification for medical images[C]//Proceedings of the 2019 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 2079-2088.

[113] OLSSON V, TRANHEDEN W, PINTO J, et al. Classmix: segmentation-based data augmentation for semi-supervised learning [C]//Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2021: 1369-1378.

[114] SINGH A, CHAKRABORTY O, VARSHNEY A, et al. Semi-supervised action recognition with temporal contrastive learning[C]//Proceedings of the 2021 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 10389-10399.

[115] ZHANG X, OU N, LIU C, et al. Unsupervised brain tumor segmentation with image-based prompts[EB]. arXiv preprint, 2023, arXiv: 2304.01472.

[116] OU Y, HUANG S X, WONG K K, et al. BBox-guided segmentor: leveraging expert knowledge for accurate stroke lesion segmentation using weakly supervised bounding box prior[J]. Computerized Medical Imaging and Graphics, 2023: 102236.

[117] WANG S, CHEN Z, YOU S, et al. Brain stroke lesion segmentation using consistent perception generative adversarial network[J]. Neural Computing and Applications, 2022, 34(11): 8657-8669.

[118] CHEN C, DOU Q, CHEN H, et al. Synergistic image and feature adaptation: towards cross-modality domain adaptation for medical image segmentation[C]//Proceedings of the 2019 AAAI Conference on Artificial Intelligence. CA: AAAI, 2019, 33(01): 865-872.

[119] LIU L, ZHANG Z, LI S, et al. S-CUDA: self-cleansing unsupervised domain adaptation for medical image segmentation[J]. Medical Image Analysis, 2021, 74: 102214.

[120] WANG R, CHAUDHARI P, DAVATZIKOS C. Embracing the disharmony in medical imaging: a simple and effective framework for domain adaptation[J]. Medical Image Analysis, 2022, 76: 102309.

[121] LI X, YU L, JIN Y, et al. Difficulty-aware meta-learning for rare dis-

ease diagnosis[C]//Proceedings of the 23th International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2020: 357-366.

[122] LIU Q, CHEN C, QIN J, et al. Feddg: federated domain generalization on medical image segmentation via episodic learning in continuous frequency space[C]//Proceedings of the 2021 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 1013-1023.

[123] SANABRIA A R, ZAMBONELLI F, Ye J. Unsupervised domain adaptation in activity recognition: a GAN-based approach[J]. IEEE Access, 2021, 9: 19421-19438.

[124] LIU Q, DOU Q, HENG P A. Shape-aware meta-learning for generalizing prostate MRI segmentation to unseen domains[C]//Proceedings of the 23th International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2020: 475-485.

[125] ZHANG P, LI J, WANG Y, et al. Domain adaptation for medical image segmentation: a meta-learning method[J]. Journal of Imaging, 2021, 7(2): 31.

[126] ESTEBAN O, BIRMAN D, SCHAER M, et al. MRIQC: advancing the automatic prediction of image quality in MRI from unseen sites[J]. PLoS One, 2017, 12(9): e0184661.

[127] REINHOLD J C, DEWEY B E, CARASS A, et al. Evaluating the impact of intensity normalization on MR image synthesis[C]//Proceedings of the Medical Imaging 2019: Image Processing. Washington: SPIE, 2019, 10949: 890-898.

[128] HAN Y, HUANG G, SONG S, et al. Dynamic neural networks: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(11): 7436-7456.

[129] GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C]//Proceedings of 2015 International Conference on Machine Learning. New York: PMLR, 2015: 1180-1189.

[130] GANIN Y, USTINOVA E, AJAKAN H, et al. Domain-adversarial training of neural networks[J]. The Journal of Machine Learning Research, 2016, 17(1): 2096-2030.

[131] NG D, LAN X, YAO M M S, et al. Federated learning: a collaborative effort to achieve better medical imaging models for individual sites that have small labelled datasets[J]. Quantitative Imaging in Medicine and Surgery, 2021, 11(2): 852.

[132] SHELLER M J, EDWARDS B, REINA G A, et al. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data[J]. Scientific Reports, 2020, 10(1): 1-12.

[133] SARMA K V, HARMON S, SANFORD T, et al. Federated learning improves site performance in multicenter deep learning without data sharing[J]. Journal of the American Medical Informatics Association, 2021, 28(6): 1259-1264.

[134] WANG B, LI G, WU C, et al. A framework for self-supervised federated domain adaptation[J]. EURASIP Journal on Wireless Communications and Networking, 2022, 2022(1):1-17.

[135] YU T, KUMAR S, GUPTA A, et al. Gradient surgery for multi-task learning[C]//Proceedings of the Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2020, 33: 5824-5836.

[136] DENG L, LI G, HAN S, et al. Model compression and hardware acceleration for neural networks:a comprehensive survey[J]. Proceedings of the IEEE, 2020, 108(4): 485-532.

[137] CHO J H, HARIHARAN B. On the efficacy of knowledge distillation[C]// Proceedings of the 2019 IEEE International Conference on Computer Vision. Piscataway: IEEE, 2019: 4794-4802.

[138] BEYER L, ZHAI X, ROYER A, et al. Knowledge distillation: a good teacher is patient and consistent[C]//Proceedings of the 2022 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 10925-10934.

[139] DING X, ZHANG X, MA N, et al. Repvgg: making vgg-style convnets great again[C]//Proceedings of the 2021 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 13733-13742.

[140] WU J, FANG H, ZHANG Y, et al. Medsegdiff: medical image segmentation with diffusion probabilistic model[EB]. arXiv preprint, 2022, arXiv:2211.00611.

[141] WU J, FU R, FANG H, et al. Medsegdiff-v2: diffusion based medical image segmentation with transformer[EB]. arXiv preprint, 2023, arXiv: 2301.11798.

[142] CHEN T, WANG C, SHAN H. BerDiff: conditional bernoulli diffusion model for medical image segmentation[EB]. arXiv preprint, 2023, arXiv:2304.04429.

[143] HE S, BAO R, LI J, et al. Accuracy of segment-anything model (sam) in medical image segmentation tasks[J]. arXiv preprint, 2023, arXiv:2304.09324.

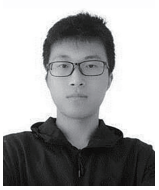
[144] MA J, WANG B. Segment anything in medical images[EB]. arXiv preprint, 2023, arXiv:2304.12306.

[145] BUTOI V I, ORTIZ J J G, Ma T, et al. Universeg: universal medical image segmentation[EB]. arXiv preprint, 2023, arXiv:2304.06131.

[作者简介]



**余唯一**（1998− ），女，复旦大学类脑智能科学与技术研究院硕士生，主要研究方向为深度学习和医学图像处理。



**陈涛**（1999− ），男，复旦大学类脑智能科学与技术研究院博士生，主要研究方向为深度学习和医学图像处理。



**张军平**（1970− ），男，复旦大学计算机科学技术学院教授、博士生导师，主要研究方向为机器学习、智能交通、生物认证与图像识别。



**单洪明**（1990− ），男，复旦大学类脑智能科学与技术研究院青年研究员、博士生导师，主要研究方向为机器学习和医学影像。