

基于Focal Recall Loss的阿尔兹海默症病灶分割模型

陈子洋¹, 王卓薇¹, 邱俊豪²

¹广东工业大学, 计算机学院, 广东 广州

²广东工业大学, 机电学院, 广东 广州

收稿日期: 2021年12月21日; 录用日期: 2022年1月18日; 发布日期: 2022年1月25日

摘要

目前的阿尔兹海默症病灶分割算法主要以深度分割网络作为主流算法对组织病理区域进行分割, 但是这些方法在面临类别不均衡的情况时, 整体性能会受少数类的影响而陷入局部最优的情况。针对上述情况, 首先, 本文提出一种新的语义分割模型, 名为Global Attention UNet (GAUNet), 该模型嵌入了全局注意力模块以及组线性层模块对特征空间中的通道信息以及空间信息进行挖掘, 从而提高模型的特征表示能力。其次, 针对类别不均衡的问题, 本文提出一种局部回归损失函数(Focal Recall Loss), 针对每轮次召回情况动态调整各个类别的权重, 从而使得模型更关注少数类的类别特征信息。本文所提出的方法在Alzheimer's Disease Neuroimaging Initiative (ADNI)数据集中对6种组织类别区域(额叶、颞叶、顶叶、海马体、中脑、半卵圆中心)同时进行分割, 与当下的模型相比, 本文提出的方法在少数类别半卵圆中心的IOU比当前最新方法高出6.19%。

关键词

阿尔兹海默症, 病灶分割, 类别不均衡, 组线性层模块, 注意力机制

The Alzheimer's Disease Segmentation Model Based on Focal Recall Loss

Ziyang Chen¹, Zhuowei Wang¹, Junhao Qiu²

¹School of Computer, Guangdong University of Technology, Guangzhou Guangdong

²School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou Guangdong

Received: Dec. 21st, 2021; accepted: Jan. 18th, 2022; published: Jan. 25th, 2022

文章引用: 陈子洋, 王卓薇, 邱俊豪. 基于 Focal Recall Loss 的阿尔兹海默症病灶分割模型[J]. 计算机科学与应用, 2022, 12(1): 95-107. DOI: 10.12677/csa.2022.121011

Abstract

Currently, deep segmentation approaches become the main algorithms for lesion segmentation in Alzheimer's Disease. However, these methods hardly extract the edge information of regions that belonged to the minority category. In order to solve the above problem, firstly, we propose a novel segmentation model, called Global Attention UNet (GAUNet). Here, we embedded a Global Attention Block into our proposed model to explore the channel and spatial feature to improve the feature discriminative. Secondly, we propose the Focal Recall Loss function to alleviate the problem of the imbalance category. Specifically, our loss function can dynamically adjust the weights of each category based on each epoch of recall ratios. Finally, our proposed method applies to the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset to segment six regions including the Frontal Lobe, Temporal Lobe, Parietal Lobe, Hippocampus, Midbrain and Semi-oval center. The experimental results show that our proposed method achieves 6.19% higher IOU in the category of semi-oval center than several state-of-the-arts.

Keywords

Alzheimer's Disease, Lesion Segmentation, Imbalance Category, Group Linear Layer, Attention Mechanism

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

阿尔兹海默症是一种退行性脑部疾病。它是一种神经系统疾病，该疾病会诱发脑部组织萎缩，导致记忆丧失和认知能力下降，对人类的个人和社会生活产生了令人难以置信的负面影响。根据最近的统计数据，现在有超过 4680 万人患有痴呆症，4400 万人被诊断为阿尔茨海默病。这个数字在 2050 年将增加到 1.315 亿[1]。据研究表明，通过观察与 AD 相关的神经病理标志物可以有助于观察患者发病程度。在这些 AD 相关的标志物中，通常以海马体、中脑等组织区域最能表征阿尔兹海默症的病灶表征性。通过观察这些区域的萎缩程度来判断当前样本的患病严重程度。因此，能够准确分割出以上区域对于判定病情尤为重要。

一般地，对于这些区域的寻找通常是依赖人工经验对这些区域进行标注，但是这样的方式容易出现人为的失误，从而影响后续的诊断方案。因此一些研究认为，利用传统的图像处理算法应用在本场景中对特定组织区域进行分割可以准确地描绘这些组织区域的边界轮廓，从而能辅助判定这些区域的萎缩变化情况。比如，Qi 等人[2]借助区域生长技术对肿瘤区域进行区域分割。其中应对肿瘤区域中密度方差不一致的问题，结合贝叶斯决策规则以及模型匹配度原则，动态调整组织区域的密度方差。此外，Li 等人[3]新的统一水平集模型，整合图像梯度、区域竞争和先验信息，用于 CT 肝脏肿瘤的分割。通过无监督的模糊聚类来估计肝脏肿瘤的概率分布，并利用它来增强对象指示函数，定义方向性气球力和调节区域竞争。总的来说，上述提及的方法可以有效地分割出所需要的组织区域，应用在脑部 MRI 进行器官组织区域分割同样有效，但是这些方法处理速度较慢，在同时处理大批量样本时效率较低。

而随着深度学习技术的不断发展，越来越多的神经网络模型都应用在大脑肿瘤组织分割场景以及阿

尔兹海默症病灶区域分割场景,并且在速度以及精度上都得到提升。一些研究表明,利用 CNN 卷积神经网络作为编码器对脑肿瘤特征进行逐层编码,可以有效地提取出病灶区域特征[4] [5] [6]。以上分割模型主要以 FCN 作为主要的分割编码模型,通过改变不同卷积模块的数量以及卷积核的尺寸,从而提取出具有不同感受野的语义特征。但是 FCN 在结构设计中存在一定缺陷,即该模型在进行上采样的时候只是采用双线性插值的方式进行结果尺度还原,这样会导致在像素还原的过程中,组织轮廓区域的边缘信息很容易丢失。因此基于以上问题,Lopez 等人[7]提出利用 UNet 网络模型对 FCN 模型进行改进并应用在脑肿瘤组织分割的应用场景中,效果得到有效改进。该模型通过增加了拼接特征图的操作,将编码器的特征图信息与解码器的特征图信息进行结合,减少在上采样还原过程中的信息丢失。但是,这个结构在拼接信息时无法将其他层的特征信息进行融合,从而缺乏了不同尺度的特征信息,这样容易忽略小尺寸的病灶区域。因此,Islam 等人[8]提出一种基于 Attention 的 UNet 结构,在跳跃连接结构中嵌入了注意力机制模块来增强从编码器输出特征的表征性,通过强化这些特征的通道信息,从而使得增加不同区域特征的判别性。此外 Wang 等人[9]提出一种基于多尺度特征融合模块的病灶分割网络 GLF-Net 对不同尺度的脑肿瘤进行分割。其中所提出的多尺度融合模块主要以两部分组成。第一部分是全局多尺度注意力模块,加入该模块的目的是为了利用全局上下文信息,从空间和通道两方向指导低层次特征的重新校准,从而提高有效的多尺度特征的利用率,抑制低层次特征的噪声。而第二部分是低密度多尺度特征融合模块。该模块采用自下而上、自上而下的结构来捕捉背景信息,生成语义特征,并融合不同尺度的特征信息。这个模块可以在网络中逐层整合局部密集的多尺度背景特征,从而提高网络对边界像素之间相互依赖关系的编码能力。但是这些拼接操作会大幅度增加模型的计算量,从而降低模型的推理速度。因此,一些研究开始对模型结构进行简化。比如,Cheng 等人[10]提出一种级联式轻量级网络,名为 MECU-Net。这个模型可以实现在占用少量显存以及计算量的情况下达到较高的分割精度。Pendse 等人[11]在传统的 UNet 结构上进行改进,借鉴了 mobilenetV2 中逆残差结构作为编码器模块。其中在该结构中,利用深度可分离卷积,进一步地降低模型计算量。

总的来说,上述的分割算法在一定程度上解决了由于手工标注病灶区域所带来的失误,但是这些方法仍然存在以下问题。在本文所研究的阿尔兹海默症病灶分割中,由于不同组织区域分布极度不均衡,往往属于少数类的组织区域的特征表征能力要弱于多数类的组织区域。而这些少数类往往是较为重要的组织区域,比如海马体区域、半圆中心区域。这些区域之所以成为少数类因为整个大脑 MRI 中可观察到这些区域的切片数量不多,所以标注出来的区域会比额叶区域、颞叶区域要少很多。基于上述问题,本文首先将提出一个新的分割模型名为 Global Attention UNet (GAUNet),该模块中所使用的空间变换注意力模块(Global Transform Mechanism)中引入组线性层(Group Linear Layer)结构,用以扩增注意力模块中的高维映射空间。同时,引入用于均衡类别权重的损失函数,名为 Focal Recall loss,通过这个函数根据类别之间的比例来对动态生成不同类别权重,从而缓解上述的类别不均衡问题。

2. 阿尔兹海默症病灶分割算法介绍

2.1. GAUNet 分割模型

2.2.1. 整体结构

本文所提出的病灶分割模型 GAMUNet 主要分为编码器部分、Global Transform Mechanism 注意力模块以及解码器部分,整体流程图如图 1 所示。假设输入数据为 $X \in \mathbb{R}^{B \times C \times H \times W}$,其中 B 为样本批次, C 为通道数, H 为高度, W 为宽度。将这个数据输入到模型的编码器部分中进行特征提取。首先,该样本会经过卷积块、批归一化层以及 ReLU 三个模块,从而输出低级语义特征图。然后通过下采样的方式对该特征图进行维度扩增,扩大特征点的映射距离,从而使得模型获得较大的感受野。然后经过三层的下

采样后完成了整个编码过程。相比之下，在模型第三层所输出的特征图的语义信息会比第一层的语义信息要抽象，而这些语义信息表征着每个特征块所表示的语义响应强度。模型通过感受不同语义特征块之间的响应强度，来判别不同像素点之间的语义差别，继而可以判断该区域的语义差异性。然后，经过编码器输出的特征图经过本文所提出的全局注意力模块进行空间信息的扩增，以此丰富每一层编码器所产生的空间信息。最后，通过解码器部分对扩增后的特征图进行上采样，还原特征图的原有的尺度，并将每一层编码器所输出的特征图进行拼接，从而在解码部分可以弥补信息丢失的缺陷。

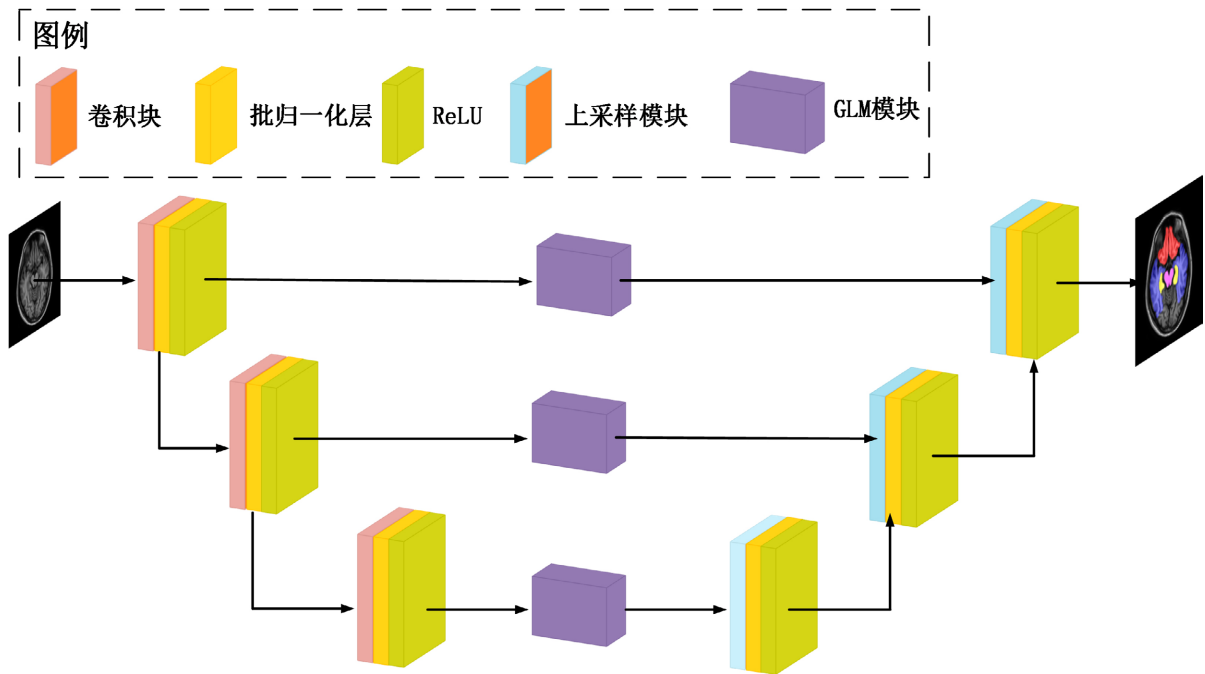


Figure 1. The total architecture of GAUNet

图 1. GAUNet 整体模型结构

具体来说，每层编码器的输出特征图可定义为 X_l ，其中 l 为当前的层数，如第一层编码器所输出的特征则可表示为：

$$X_1 = f_{bn} (f_{3 \times 3} (X)) \quad (1)$$

其中， $f_{3 \times 3}(\cdot)$ 表示具有 3×3 卷积核大小的卷积模块，而 $f_{bn}(\cdot)$ 表示批归一化模块。当达到第三层的编码器时，则输出的特征图就是 X_3 。通过这一编码过程，会得到具有不同通道特征的三组特征图，然后将通过注意力模块对这些特征图进行特征增强，以此增强特征的代表能力。

2.2.2. 全局注意力模块

利用上述编码器所输出的特征，本小节将介绍本文提出的全局注意力模块，这个模块受[12] [13]所启发，并对原有的模块进行了改进。这个模块主要是将通道注意力模块以及空间注意力模块进行串行结合，对输出特征图进行通道维度以及空间维度进行特征信息抽取。具体的注意力模块结构如图 2 所示。其中，先将输入特征通过通道变换的方式进行张量维度转换，然后将转换后的特征图进行通道维度进行通道信息，在此过程中，采用了 Group Linear Layer 模块特征空间扩展，这样可以增加通道信息的交互融合，然后将这些通道信息进行维度还原，并与原输入特征图进行加权，最终输出的具有通道信息的特征图。将

该特征图进行空间信息的提取，首先会经过卷积块进行特征映射，然后通过空间通道分解，将特征空间按照通道方向进行分组，将分组完成后的特征空间依次进行特征提取，最后再将这些特征组进行融合。具体来说，假设此时当前输出的特征图是位于第三层的编码器的输出位置，然后将输出后的特征图输出本小节提出的注意力模块中。首先，输出的特征图为 $X_3 \in \mathbb{R}^{B \times C \times H' \times W'}$ ，然后先将该特征图进行通道变换，其变换过程为：

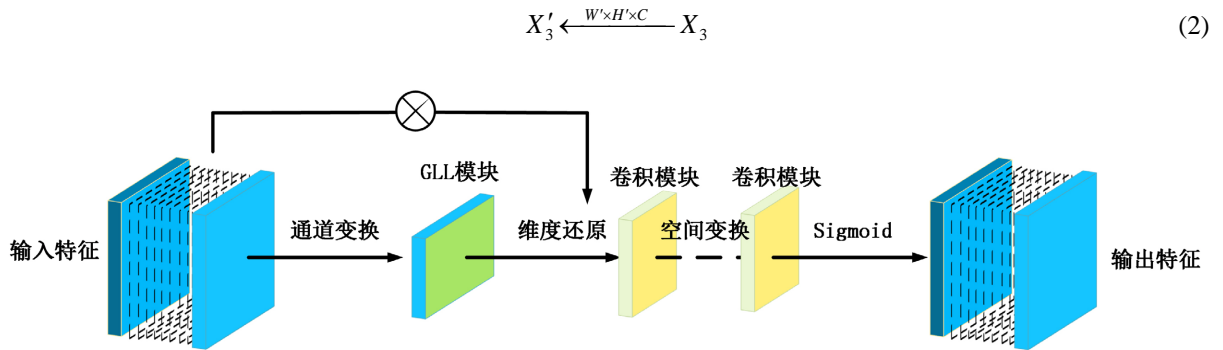


Figure 2. The structure of Global Attention Module
图 2. 全局注意力模块结构

然后，将变换后的特征图输入到 Group Linear Layer (GLL)模块中进行非线性映射。其中该 GLL 模块的流程图如图 3 所示。

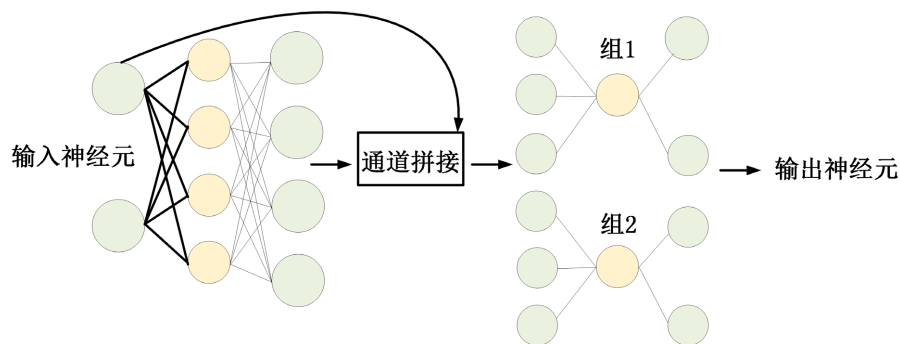


Figure 3. The structure of Group Linear Layer Module
图 3. 组线性层模块结构

由上图所示，该模块主要有两层不同结构的线性层构成。首先，输入神经元会接收由上述特征图，然后通过第一层线性层进行特征映射，然后为了防止某个神经元出现失活，则引入残差映射的结构，将输入的神经元特征与第一层线性层输出的特征图进行拼接。从而形成更高维度特征空间。然后将该特征空间进行分组，由上图 3 所示，将特征空间分成 2 组，然后对每组神经元分别进行特征映射，然后将映射完成后的特征图进行聚合，最后聚合输出特征图进行通道特征还原，最后输出特征图为 X_3'' 。此时输出的特征图则具有通道特征信息，从而增强模型的可解释性。然后再将上述的特征图 X_3'' 与原输入特征图 $X_3 \in \mathbb{R}^{B \times C \times H' \times W'}$ 进行加权组合，最后形成具有通道信息的特征图 X_c ：

$$X_c = X_3 \otimes X_3'' \quad (3)$$

其中， \otimes 表示加权组合操作。

在获取通道特征图后，则对该特征图进行空间信息挖掘。首先，将该特征图 X_c 通过卷积块 $f_{7 \times 7}$ 进行特征高维映射，然后将映射完成后的特征图进行通道分组，并且对每组特征图进行二次特征提取，然后再将提取完成的特征图进行聚合，最后利用 sigmoid 激活函数对聚合完成的特征图进行输出。最后得到具有空间注意力特征的输出特征图 X_S ：

$$X_S = \text{Sigmoid} \left(f_{7 \times 7} \left(\sum_r^C f_{7 \times 7} (X_S^r) \right) \right) \quad (4)$$

其中， r 表示通道分组数。

2.2. Focal Recall Loss 损失函数

本小节将介绍本文所提出的局部召回损失函数(Focal Recall Loss)。该损失函数结合每次训练后各个类别的召回情况，动态地调整每个类别的权重参数。在以往类别不平衡算法中，尽管也是对当前类别分布进行调整，但是这些调整方式会导致模型过度关注离群点样本，从而使得模型难以收敛到最优收敛位置[14]。而本小节提出的损失函数，则通过计算每轮次中每个类别的召回率，而评估当前模型找到每个类别的性能，从而将该召回率作为权重指导交叉熵进行梯度调整。具体来说，假设 $Y_{gt} \in \mathbb{R}^c$ ， $Y_{pre} \in \mathbb{R}^c$ 表示模型输出的预测值以及对应的数据标签，然后，统计出分割标签中的像素类别分布情况 $Cls(c)$ ：

$$Cls(c) = \sum_{n=1}^N X_n \quad (N = C \times W \times H) \quad (5)$$

其中 N 表示标签图片中所有像素点的总和。此时得到每个类别之间的比例分布情况，然后在训练过程中统计每个轮次中，每个类别的召回率 $F_{rec}^c \in (0,1)$ ：

$$F_{rec}^c = \frac{FN}{FN + TP} \quad (6)$$

其中， FN 表示负假样本， TP 表示正阳样本，此处统计真假阳样本的方式是通过统计当前预测的掩码与真实标签之间的交并比，并将阈值设置在 0.7 作为分界来进行划分。通过公式(6)所示，可以计算每个类别的召回情况，并将该召回情况作为权重与交叉熵进行结合，可以得到召回损失函数 G_{rec} ：

$$G_{rec} = - \sum_{c=1}^C (1 - F_{rec}^c) \times Cls(c) \times \log(Y_{pre}) \quad (7)$$

但是，由公式(7)则容易在每轮次中由于召回率波动而使得梯度更新不稳定的问题，因此，在公式(7)的基础上加入了难度系数权重，以此加大对难样本的权重比值，最后可以得到损失函数为 G_{total} ：

$$G_{total} = -\alpha \sum_{c=1}^C \sum_{Y_n=c} (1 - F_{rec}^c)^\beta \times Cls(c) \times \log(Y_{pre}) \quad (8)$$

其中， α 为正负样本的调整权重系数， β 为对难样本挖掘的关注程度系数。

3. 实验结果与分析

3.1. 数据集分析

本次实验是在 ADNI 数据集[15]对 MRI 图像进行病灶区域分割。本次收集的数据集总数为 2000 张，其中这些区域包括额叶、颞叶、顶叶、海马体、中脑以及半卵圆中心。对整体数据集中各个像素类别进行统计，统计结果如下图 4 所示。

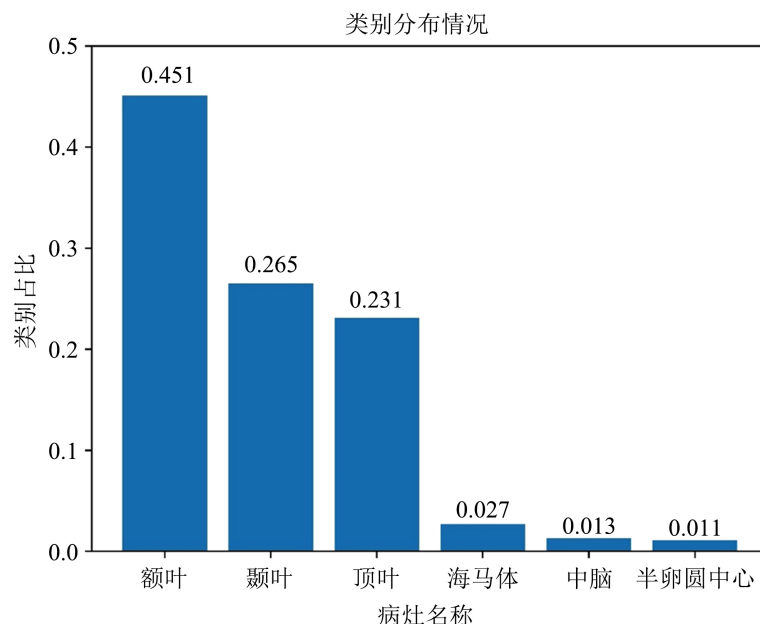


Figure 4. Distribution of Alzheimer's disease lesions by tissue category

图 4. 阿尔兹海默症病灶组织类别分布图

由上图所示，海马体、中脑、半卵圆中心这三个区域占比总数的最少。以上分布为典型的长尾分布，在这种分布中，模型会对长尾类的特征表征能力较弱，这样很容易让模型陷入局部最优，从而降低模型的特征提取性能。

3.2. 实验参数设置

基于上述类别分布，在本次实验中采取五折交叉验证的策略，对训练集以及验证集数据进行循环训练，确保每个像素类别都得到充分的采样。此外，本次训练环境是在 NVIDIA GTX3060Ti 的显卡环境进行实验，在参数设置上，batch 数量以 16 作为每个批次的吞吐数量。其次，在优化器的选择上，会选择 Adam 作为优化器，配合 StepLR 作为学习率调整方案对梯度收敛方向、学习率的下降速率进行调整。此外，在学习率(LR)初始化方面，LR 会初始化为 $1e-4$ 。总体训练轮次为 400 个 epoch。

3.3. 对比分析

为了验证本文所提出的模型综合性能，将选择以下对比模型参与评估：

1) UNet [16]: 该模型以 U 形网络模型构成了编码器结构以及解码器结构。该结构首次采用了拼接特征图的操作方式，弥补特征图上采样过程中所造成的特征信息丢失。

2) UNet++ [17]: 该模型在上述模型的基础上进行改进，在每层的编解码器之间加入了级联结构，并在每层结构中加入了深监督训练策略，其目的是利用各个尺度的特征信息来辅助主解码器对目标边缘进行分割。

3) Attention UNet (AttUNet) [18]: 该模型以 UNet [16]作为主要基线模型，通过嵌入通道注意力模块，从而缓解在上采样过程中所丢失的边缘信息的情况。

4) BiSeNet [19]: 该模型是一个轻量级分割网络，该网络与 UNet [16]的不同，它通过构建双路结构的方式，分别对空间位置信息以及语义信息同时提取，然后再通过构建特征融合模块将上述提取的两种特征进行融合，从而在提高推理速度的同时也可以确保一定的精确度。

5) Deeplabv3 [20]: 该模型通过堆叠不同尺度的空洞卷积, 并构建了并行空间金字塔池化模块(ASPP), 该模块在确保一定的图像分辨率的同时也扩大了感受野。

6) DFANet [21]: 该模型是针对浅层网络的特征表征能力较弱的问题, 采用分支组合空间细节以及上下文信息进行跨层特征聚合, 从而弥补语义特征以及结构细节特征之间的差距。

7) ENet [22]: 该结构也是一个轻量级分割网络, 在该网络中, 会利用 1×1 卷积模块来降低映射维度, 同时该模型分解卷积的方式来扩增模型的宽度, 并配合空洞卷积的结构, 以及极少的参数量来扩展模型感受野, 此外, 利用非对称的编解码结构降低了整理参数量。

8) FastSCNN [23]: 该网络为一个目前最轻量级的分割网络, 该模型摒弃使用 U 型结构, 而是借鉴了 BiSeNet [19] 构建了双分支结构, 其中较深的分支用以处理低分辨的特征图, 而另外一个分支则是处理高分辨率的特征图。与 BiSeNet 不同的是, 该网络并未采用特征融合模块方式进行合并特征, 而是通过分组卷积的方式, 让较深的分支特征部分充当特征升维模块, 然后将升维后的特征与另一个分支的特征进行结合。

在选择完成上述模型后, 则分别对 6 个类别进行分割, 以下见表 1~6 为分割后的结果。

Table 1. The comparison experiment of segmenting Frontal Lobe regions
表 1. 各个方法对额叶分割情况

模型	召回率(%)	精准度(%)	IOU (%)
UNet [16]	90.17	90.06	82.01
UNet++ [17]	90.74	92.06	84.16
AttUNet [18]	93.24	93.36	87.5
BiSeNet [19]	89.23	90.48	81.57
Deeplabv3 [20]	91.94	91.28	84.51
DFANet [21]	79.54	87.54	71.45
ENet [22]	90.14	87.55	79.90
FastSCNN [23]	29.05	59.26	27.22
Ours	92.80	93.70	87.35

Table 2. The comparison experiment of segmenting Temporal Lobe regions
表 2. 各个方法对颞叶分割情况

模型	召回率(%)	精准度(%)	IOU (%)
UNet [16]	86.16	89.09	77.94
UNet++ [17]	86.61	91.25	79.97
AttUNet [18]	89.02	93.26	83.64
BiSeNet [19]	83.02	91.77	77.27
Deeplabv3 [20]	88.91	91.04	81.75
DFANet [21]	71.29	84.83	63.23
ENet [22]	84.63	86.24	74.56
FastSCNN [23]	16.47	56.07	14.59
Ours	88.95	94.49	84.56

Table 3. The comparison experiment of segmenting Parietal Lobe regions
表 3. 各个方法对顶叶分割情况

模型	召回率(%)	精准度(%)	IOU (%)
UNet [16]	88.28	83.33	75.03
UNet++ [17]	87.55	86.11	76.71
AttUNet [18]	92.19	88.35	82.2
BiSeNet [19]	83.30	86.57	73.77
Deeplabv3 [20]	87.40	86.07	76.58
DFANet [21]	85.76	64.66	58.39
ENet [22]	83.28	79.89	68.85
FastSCNN [23]	8.06	32.06	6.89
Ours	89.90	90.63	82.25

Table 4. The comparison experiment of segmenting Hippocampus regions
表 4. 各个方法对海马体分割情况

模型	召回率(%)	精准度(%)	IOU (%)
UNet [16]	66.32	62.40	47.38
UNet++ [17]	64.06	67.53	48.97
AttUNet [18]	76.42	80.99	64.80
BiSeNet [19]	56.74	67.00	44.34
Deeplabv3 [20]	69.59	78.94	58.69
DFANet [21]	58.52	36.42	28.94
ENet [22]	53.07	52.09	35.66
FastSCNN [23]	0	0	0
Ours	77.69	81.41	65.98

Table 5. The comparison experiment of segmenting Midbrain regions
表 5. 各个方法对中脑分割情况

模型	召回率(%)	精准度(%)	IOU (%)
UNet [16]	77.61	93.08	73.38
UNet++ [17]	95.19	86.11	82.52
AttUNet [18]	94.97	92.67	88.34
BiSeNet [19]	85.84	76.41	67.86
Deeplabv3 [20]	89.07	90.46	81.42
DFANet [21]	55.70	77.85	48.08
ENet [22]	87.08	82.52	73.52
FastSCNN [23]	5.14	26.9	4.51
Ours	92.90	89.75	83.99

Table 6. The comparison experiment of segmenting Semi-oval Center regions
表 6. 各个方法对半卵圆中心分割情况

模型	召回率(%)	精准度(%)	IOU (%)
UNet [16]	59.80	76.93	54.89
UNet++ [17]	63.15	71.92	50.66
AttUNet [18]	65.70	76.93	54.89
BiSeNet [19]	58.75	69.48	46.7
DeepLabv3 [20]	65.71	78.08	55.48
DFANet [21]	40.47	55.13	30.44
ENet [22]	46.98	70.64	39.31
FastSCNN [23]	1.83	80.56	1.82
Ours	69.94	82.81	61.08

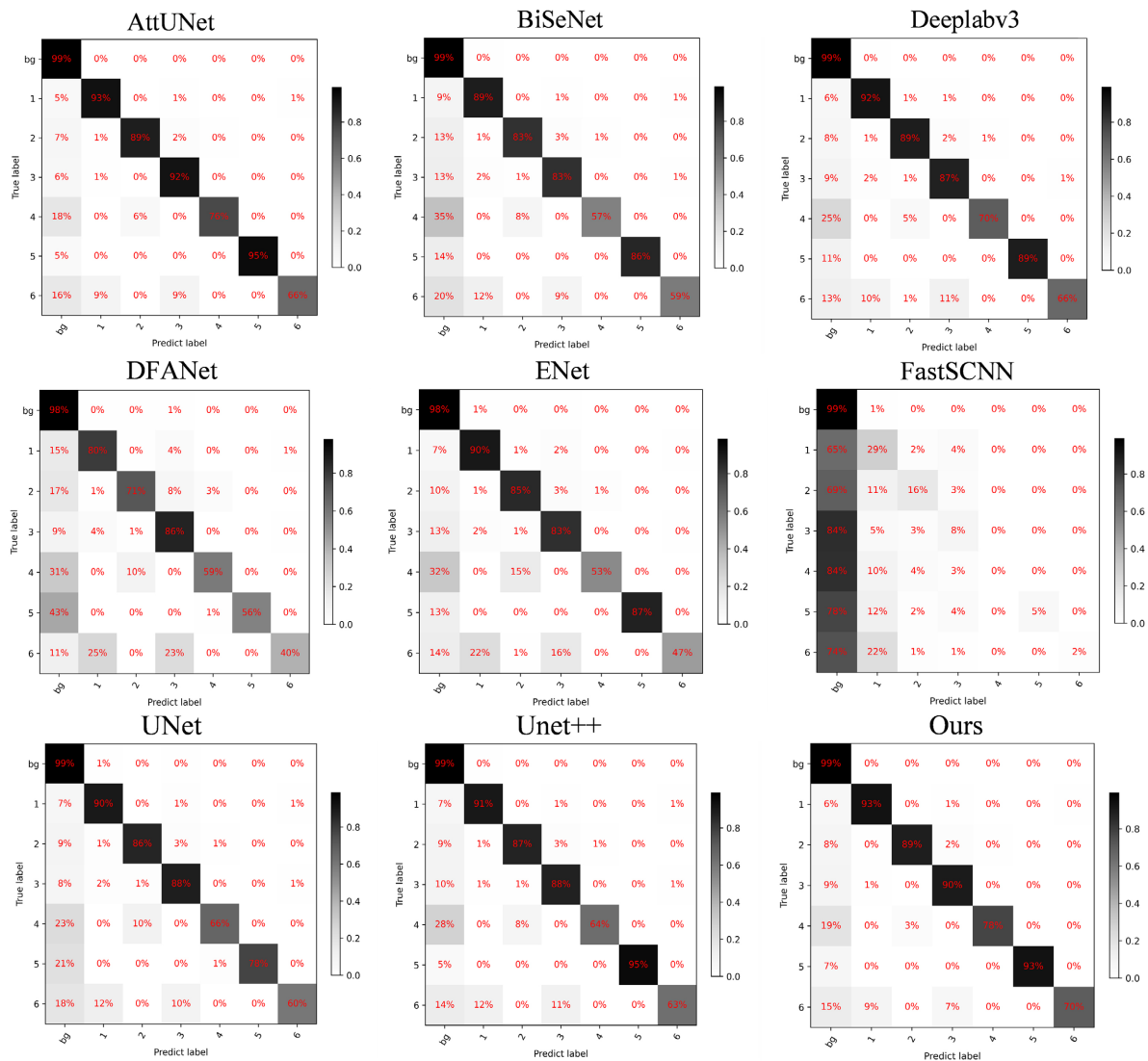


Figure 5. The confusion matrix results between each of methods
图 5. 各个方法对于 6 个类别的分类混淆矩阵结果

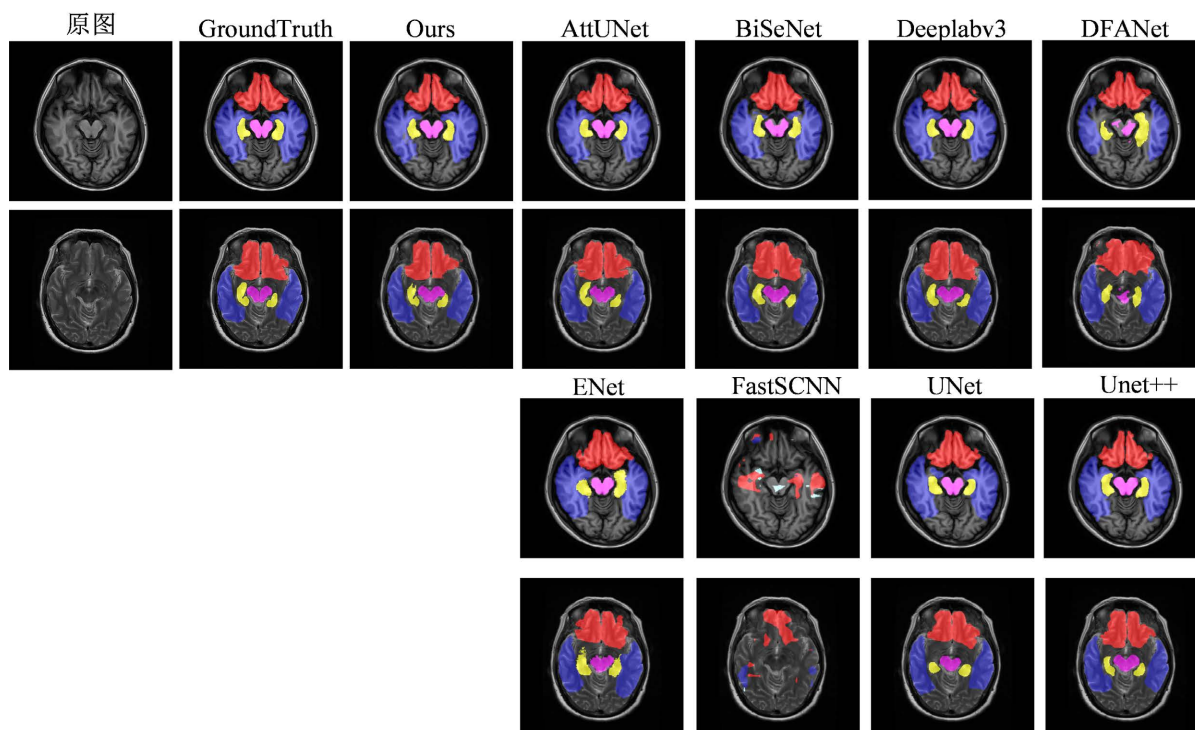


Figure 6. The inference results between each of methods

图 6. 各个方法对于 6 个类别的推理结果

由以上实验中得出，本模型在 6 个组织类别区域中性能表现良好，其中在少数类区域分割中对于海马体区域以及半卵圆中心的分割效果尤为明显，其中对海马体区域分割的 IOU 值为 65.98%，召回率为 77.69% 以及精准度为 81.41%，比目前最优的模型性能在 IOU 指标、召回率以及精准度上分别超过 1.18%，1.27% 以及 0.42%。而对于半卵圆中心的分割性能，本文提出的模型在 IOU 指标上达到 61.08%，召回率可达到 69.94% 以及精确度可达到 82.81%。其中上述两组类别的模型表现情况如图 5 所示。其中颜色越深表示模型对该类别的分类性能越高，同时上述所标注的类别序号分别表示为：1 号类别为额叶，2 号类别为颞叶，3 号类别为顶叶，4 号类别为海马体，5 号类别为中脑，6 号类别为半卵圆中心。从图 5 可知，类别 4 以及类别 6 的分类效果在各个模型的表现都较差。其中 FastSCNN 在本数据集中分割效果表现欠拟合。而本模型在类别 4 和类别 6 的准确率分别可以达到 78% 以及 93%。但是类别 5 的准确率并未达到最高水平，其原因是类别 5 的组织区域尺度较小，难以完全地提取其边缘特征。此外，各个模型在实际模型推理的情况如图 6 所示。其中红色表示额叶，蓝色表示颞叶，黄色表示半卵圆中心，紫色表示中脑。GroundTruth 为真实标注的图片。由图 6 可知，本模型对于第一份样本中各个区域的提取情况较为理想，相比之下，Attention UNet 在前额的提取边缘能力稍弱于本模型。其次，对于轻量级模型的提取性能来看，ENet、DFANet、BiSeNet 以及 FastSCNN 来看，除了 FastSCNN 在训练过程中出现欠拟合情况，其余模型性能对于颞叶区域提取性能较为稳定。其中性能最优的是 ENet，因为该网络对额叶区域边缘信息的丢失情况最少，而且该模型的参数量最低。

而观察第二张样本的时候，则由于该样本的模态与第一章的模态不同，使得该图像的光照情况偏暗，从而使得整体模型的性能表现较差。其中，本模型在提取额叶区域时出现部分区域的信息缺失，而对于该区域提取性能最为稳定的是 Attention UNet。从这个现象可看出，首先，U 型网络性能对于多类标签分割是最稳定的。其次，对编解码器之间加入注意力机制有效保留大部分显著区域的信息。而本模型则在

训练过程中加入类别不均衡策略,在对半卵圆中心的提取情况较 Attention UNet 较好,但同时额叶的提取情况较弱。这也是后续需要优化的方向。

总的来说,综合图 5 以及图 6 来说,U 型网络对阿尔兹海默症病灶组织区域提取性能较稳定,其次,加入类别不均衡策略可以有效保留少数类组织边缘信息。

4. 结论

在阿尔兹海默症语义分割过程中,针对分割组织区域类别不均衡问题,本文提出利用 GAUNet 进行组织病理区域分割,该模型通过全局注意力模块集成通道特征信息以及空间特征信息,提高模型对各个组织区域特征的代表能力,同时通过嵌入了组线性层模块(GLL),在进行扩大高维空间上增强了各个神经元之间的响应强度,避免在反向传播中由于某个神经元失活而导致整体模块性能下降。同时利用 Focal Recall Loss 损失函数缓解由于类别不均衡而使得少数类的特征表示能力弱的问题。在本文所使用的 ADNI 数据集中,本文提出的模型可以有效地将各个区域进行分割。未来的工作将着重将本文提出的模型的参数量进行优化,同时对于多尺度信息提取能力的增强同样需要继续优化。

基金项目

广东省重点领域研发计划——神经退行性疾病早期智能预警高级机器学习技术与示范应用 2019B010109001。

参考文献

- [1] Prince, M., Wimo, A., Guerchet, M., Ali, G., Wu, Y., Prina, M., *et al.* (2019) Alzheimer's Disease International: World Alzheimer Report 2015: The Global Impact of Dementia: An Analysis of Prevalence, Incidence, Cost and Trends 2015. Alzheimer's Disease International, London.
- [2] Qi, Y., Xiong, W., Leow, W.K., *et al.* (2008) Semi-Automatic Segmentation of Liver Tumors from CT Scans Using Bayesian Rule-Based 3D Region Growing. *The MIDAS Journal*. <https://doi.org/10.54294/rofibw>
- [3] Li, B.N., Chui, C.K.K., Chang, S.K.Y. and On, S.H. (2012) A New Unified Level Set Method for Semi-Automatic Liver Tumor Segmentation on Contrast-Enhanced CT Images. *Expert Systems with Applications*, **39**, 9661-9668. <https://doi.org/10.1016/j.eswa.2012.02.095>
- [4] Zikic, D., Ioannou, Y.A., Criminisi, A. and Brown, M. (2014) Segmentation of Brain Tumor Tissues with Convolutional Neural Networks.
- [5] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A.C., Bengio, Y., Pal, C.J., Jodoin, P.-M. and Larochelle, H. (2017) Brain Tumor Segmentation with Deep Neural Networks. *Medical Image Analysis*, **35**, 18-31. <https://doi.org/10.1016/j.media.2016.05.004>
- [6] Thaha, M.M., Kumar, K.P.M., Murugan, B.S., Dhanasekeran, S., Vijayakarhick, P. and Selvi, A.S. (2019) Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *Journal of Medical Systems*, **43**, 294. <https://doi.org/10.1007/s10916-019-1416-0>
- [7] Lopez, M.M. and Ventura, J. (2017) Dilated Convolutions for Brain Tumor Segmentation in MRI Scans. *International MICCAI Brainlesion Workshop*, Vol. 10670, 253-262.
- [8] Islam, M., Vibashan, V.S., Jose, V.J.M., Wijethilake, N., Utkarsh, U. and Ren, H.L. (2019) Brain Tumor Segmentation and Survival Prediction Using 3D Attention UNet. *International MICCAI Brainlesion Workshop*, Shenzhen, 17 October 2019, 262-272. https://doi.org/10.1007/978-3-030-46640-4_25
- [9] Wang, H., Wang, G.T., Liu, Z.J. and Zhang, S.T. (2019) Global and Local Multi-Scale Feature Fusion Enhancement for Brain Tumor Segmentation and Pancreas Segmentation. *International MICCAI Brainlesion Workshop*, Shenzhen, 13-17 October 2019, 80-88. https://doi.org/10.1007/978-3-030-46640-4_8
- [10] Cheng, X.C., Jiang, Z.K., Sun, Q.L. and Zhang, J.X. (2019) Memory-Efficient Cascade 3D U-Net for Brain Tumor Segmentation. *International MICCAI Brainlesion Workshop*, Shenzhen, 13-17 October 2019, 242-253. https://doi.org/10.1007/978-3-030-46640-4_23
- [11] Pendse, M., Thangarasa, V., Chiley, V., Holmdahl, R., Hestness, J. and DeCoste, D. (2020) Memory Efficient 3D U-Net with Reversible Mobile Inverted Bottlenecks for Brain Tumor Segmentation. *International MICCAI Brainlesion*

- Workshop 2020, Lima, 4 October 2020, 388-397. https://doi.org/10.1007/978-3-030-72087-2_34
- [12] Liu, Y.C., Shao, Z.R. and Hoffmann, N. (2021) Global Attention Mechanism: Retain Information to Enhance Channel-Spatial Interactions.
- [13] Mehta, S., Koncel-Kedziorski, R., Rastegari, M. and Hajishirzi, H. (2018) Pyramidal Recurrent Unit for Language Modeling. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Brussels, 31 October-4 November 2018, 4620-4630. <https://doi.org/10.18653/v1/D18-1491>
- [14] Tian, J.J., Mithun, N.C., *et al.* (2021) Recall Loss for Imbalanced Image Classification and Semantic Segmentation. *Proceedings of the 2021 International Conference on Learning Representation*, Vienna, 4 May 2021.
- [15] Sahumbaiev, I., Popov, A., Ramírez, J., Górriz, J.M. and Ortiz, A. (2018) 3D CNN HadNet Classification of MRI for Alzheimer's Disease Diagnosis. *Proceedings 2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*, Sydney, 10-17 November 2018, 1-4. <https://doi.org/10.1109/NSSMIC.2018.8824317>
- [16] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, 5-9 October 2015, 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [17] Zhou, Z.W., Siddiquee, M.M.R., Tajbakhsh, N. and Liang, J.M. (2018) UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018*, Granada, Vol. 11045, 3-11. https://doi.org/10.1007/978-3-030-00889-5_1
- [18] Li, S.Z., Zhang, J.W., Ruan, C.Y. and Zhang, Y.C. (2019) Multi-Stage Attention-Unet for Wireless Capsule Endoscopy Image Bleeding Area Segmentation. *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, San Diego, 18-21 November 2019, 818-825. <https://doi.org/10.1109/BIBM47256.2019.8983292>
- [19] Fan, M.Y., Lai, S.Q., Huang, J.S., Wei, X.M., Chai, Z.H., Luo, J.F. and Wei, X.L. (2021) Rethinking BiSeNet for Real-Time Semantic Segmentation. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 9711-9720. <https://doi.org/10.1109/CVPR46437.2021.00959>
- [20] Azad, R., Asadi-Aghbolaghi, M., Fathy, M. and Escalera, S. (2020) Attention Deeplabv3+: Multi-Level Context Attention Mechanism for Skin Lesion Segmentation. *European Conference on Computer Vision*, Glasgow, 23-28 August 2020, 251-266. https://doi.org/10.1007/978-3-030-66415-2_16
- [21] Li, H.C., Xiong, P.F., Fan, H.Q. and Sun, J. (2019) DFANet: Deep Feature Aggregation for Real-Time Semantic Segmentation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 9514-9523. <https://doi.org/10.1109/CVPR.2019.00975>
- [22] Karimov, A., Razumov, A., Manbatchurina, R., Simonova, K., Donets, I., Vlasova, A., Khrantsova, Yu.S. and Usheinin, K. (2019) Comparison of UNet, ENet, and BoxENet for Segmentation of Mast Cells in Scans of Histological Slices. *2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*, Novosibirsk, 21-27 October 2019, 544-547. <https://doi.org/10.1109/SIBIRCON48586.2019.8958121>
- [23] Poudel, R., Liwicki, S. and Cipolla, R. (2019) Fast-SCNN: Fast Semantic Segmentation Network.