

基于 ResGRUU-Net 网络的图像分割算法

丁璇[†]

(陕西警察学院, 陕西 西安 710021)

摘要:为提高密集图像分割训练效率,在深度残差网络、GRU、U-Net 等网络模型基础上,设计了一种改进的 ResGRUU-Net 完成密集图像分割任务。首先,ResGRUU-Net 中卷积块由两个或三个 ResGRU 块组成,可通过多条连通通路从原始信息中提取出不同的特征,从而记忆和消化先前特征中包含的规律;其次为提高数据使用效率,采用随机切片和 Mosaic 切片方法对数据进行扩增;最后,分别在 DRIVE、STARE 和 CHASE_DB1 测试数据集上对训练模型进行评估。通过仿真分析,改进后的模型在测试数据集上的准确率分别提升了 0.28%、0.17% 和 0.46%,表明改进后的模型相较于改进前的模型在精度和训练速度上有所提高。

关键词: 图像分割;深度学习;残差网络;门控循环单元;数据增强

中图分类号: TP393

文献标识码: A

Image Segmentation Algorithm Based on ResGRUU-Net

DING Xuan[†]

(Shaanxi Police College, Xi'an, Shaanxi 710021, China)

Abstract: In order to improve the training efficiency of dense image segmentation, this paper designs an improved ResGRUU-Net based on the research of ResNet, GRU, U-net and other network models. Firstly, the convolution block in resgruu net is composed of two or three ResGRU blocks, which can extract different features from the original information through multiple connected paths, so as to memorize and digest the rules contained in the previous features; secondly, in order to improve the efficiency of data utilization, this paper uses random slicing and Mosaic slicing methods to amplify the data; finally, the training model was evaluated in data sets of DRIVE, STARE and CHASE_DB1, respectively. Through the simulation analysis, the accuracy rate of the improved model on the test data set is increased by 0.28%, 0.17% and 0.46%, respectively, which shows that the improved model has higher accuracy and training time than the original model.

Key words: image segmentation; deep learning; residual network; gated recursive unit; data enhancement

随着深度学习(Deep Learning, DL)技术的广泛应用,图像分类、分割、检测和跟踪等领域迎来了飞速发展时期,并取得了许多突破性的进展^[1]。基于深度学习的方法为提高模型性能而提出的改进技巧性方法,如激活函数^[2]、改进的正则化技术^[3]和深度残差网络^[4]等方法,为图像分类、分割和检

测等领域的飞速发展带来了契机。

基于全卷积网络(Fully Convolutional Networks, FCN)的卷积神经网络(Convolutional Neural Network, CNN)分割方法^[5]为自然图像分割的发展发挥了重要作用。另一种常用的语义图像分割方法为 SegNet^[6]。SegNet 由编码和解码网络

两部分组成, 编码网络是一个 13 层的 VGG-16 网络, 相应的解码网络使用像素分类层。SegNet 后来也产生了许多衍生网络, 如 Bayesian SegNet^[7]。这些体系结构中的大多数都是使用在 ImageNet 等大规模数据集中。然而, 对于密集的输出预测任务, 如医学图像分割, 上述算法效果不是很理想。因为密集图像分割任务需考虑小样本和数据不平衡等问题。

此外, 最早也是最流行的语义医学图像分割方法之一是 U-Net^[8]。根据 U-Net 体系结构, 网络主要由两部分组成: 卷积编码和解码单元。需注意, 每个卷积运算后接 ReLU 激活函数。对于编码单元中的下采样, 执行 2×2 最大池化操作。在解码阶段, 执行反卷积操作以向上采样得到特征映射。U-Net 模型为分段任务具备几个优点: 首先, 该模型允许同时使用全局位置和上下文信息; 其次, U-Net 只需要很少的训练样本即可实现分割任务; 最后, U-Net 可实现端到端的前向过程, 即输入原始图像并直接输出分割图像。然而 U-Net 也有许多可改进之处, 如训练时间长, 随机切片裁剪的方法会丢失部分全局信息, 导致网络无法记忆或学习特征之间的关系等。为此, 在 U-Net 网络基础上, 许多学者提出了其衍生网络模型, 如 ResU-Net^[9], U-Net++^[10], R2U-Net^[11] 等。

为提高密集图像分割训练效率, 受深度残差模型、GRU、U-Net 等模型的启发, 本文提出了一种改进后的 ResGRUU-Net 完成密集图像分割任务技术方法。为了验证所提出的模型性能, 在不同的医学数据集上进行了分析与评估。

1 相关理论

残差神经网络 (Residual Neural Network, ResNet) 可以解决由网络深度引起的模型性能退化和不收敛的问题。在 ResNet 的结构中 [如图 1 (a) 所示], 假设输入为 x , 网络学习的特征为 $H(x)$ 。通过学习残差 $F(x) = H(x) - x$, 使网络学习特征增加到 $F(x) + x$ 。当残差为零时, 累加层仅为恒等映射的作用, 可用来避免冗余网络层产生冗余。对于梯度下降的情况, 该网络可以有效地处理性能下降的问题。然而, 在实际应用中, 残差往往不等于零。它将使积累层根据输入的特征学习新的特征, 从而具有更好的性能。

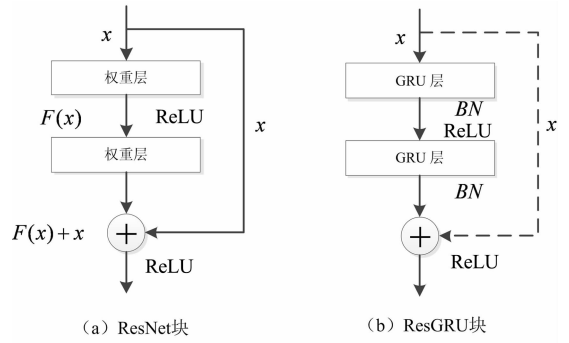


图 1 ResNet 块和 ResGRU 块

普通 GRU 可以在很大程度上解决梯度爆炸问题。然而, 一旦输入数据量增加, GRU 也会在一定程度上引起网络性能的退化, 导致输入信息的某些特性的丢失。为了解决这个问题, 本文提出了一个 ResGRU 残差块, 采用 GRU 模块提取特征之间的关系。ResGRU 的结构如图 1 (b) 所示, 虚线表示输入和输出的尺寸匹配。在结构上, 残差块的输出等于 GRU 最后一层的输出与输入 x 之和。如图 1 (b) 所示, 假设 GRU 的最后一层输出为 y , 则残差块的输出 y_R 可以表示为:

$$y_R = \text{ReLU}(\text{BN}_{\gamma, \beta}(y) + g(x_t, h_t)) \quad (1)$$

$$h_t \in H$$

其中, $\text{ReLU}(\cdot)$ 表示 ReLU 激活函数; $\text{BN}(\cdot)$ 表示 batch 规范化函数; γ 和 β 表示函数中的两个可学习变量; $g(\cdot)$ 是调节函数, 使得输入 x_t 和残差 h_t 具有相同的维数。通过残差连接, ResGRU 网络可以更好地记忆数据前后信息的相关性, 在保留原始数据特征信息的同时, 提高网络的性能。

2 网络结构

网络体系结构如图 2 所示, ResGRUU-Net 继承了 U-Net 网络模型中 U 形结构的特点, 包含下采样和上采样两部分。每个绿色/蓝色框对应一个多通道特征图。白色框表示复制的特征图。箭头与框图构成了不同的操作。与 U-Net 网络模型不同之处总结如下:

(1) ResGRUU-Net 中卷积块由两个或三个 ResGRU 块组成, 其中应用 3×3 个卷积核, 再经过 ReLU 激活层, 然后由一个批量归一化层将数据归一化处理;

(2) 对于下采样, 在卷积块之间使用一个 2×2 的最大池化层, 然后是 1×1 的卷积层;

(3)对于下采样,每个卷积块由一个反卷积层、两个卷积层和一个融合层(Concatenation 操作)组成;

(4)使用 1×1 卷积核与 Sigmoid 激活函数将特征映射为单个输出。

(5)最后,网络输出时将阈值 T 设为 0.5,生成

分割区域。

ResGRU 块由两个 GRU 模块组成,如图 1(b)所示。每个 GRU 模块完成后,进行批量归一化处理,使每个 GRU 模块具有独立的参数调整能力,从而加快模型收敛速度。此外,第一模块的批量归一化以 ReLU 为激活函数连接到激活层。

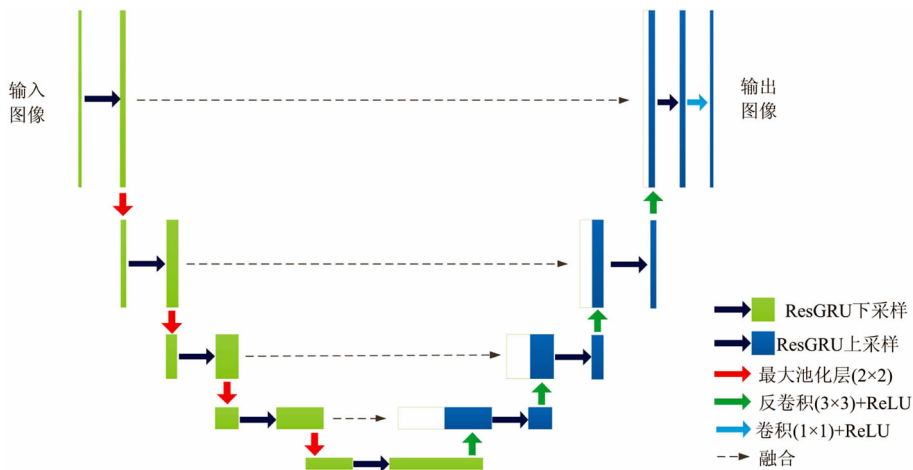


图 2 ResGRUU-Net 网络模型

与普通 GRU 单元相比,ResGRU 不仅具有独立的残差学习功能,而且可通过多条连通通路以原始信息中提取出不同的特征,从而记忆和消化先前特征中包含的规律。

3 仿真与分析

为了验证 ResGRUU-Net 网络模型的性能,本文在不同的视网膜图像血管医学影像数据集上进行了测试,测试数据集包括 DRIVE、STARE 和 CHASE_DB1。模型运行环境为:软件由 Keras 和 TensorFlow 框架搭建;硬件包括 ThinkStation 图行工作站, RAM 为 32 G,显卡为 NVIDIA GTX 1060,显存 6 G。

3.1 训练过程

3.1.1 数据集

DRIVE 数据集由 40 幅彩色视网膜图像组成,其中 20 幅用于训练,其余 20 幅用于测试。每个原始图像的大小为 565×584 像素。STARE 数据集包含 20 个彩色图像,每个图像的大小为 700×605 像素。CHASE_DB1 数据集包含 28 幅彩色视网膜图像,每个图像的大小为 999×960 像素。与其他另外两种数据集相比,DRIVE 数据集多了眼部轮

廓图像集。各数据集部分数据展示如图 3 所示。

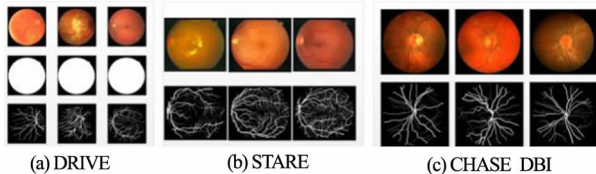


图 3 不同数据集数据展示

3.1.2 数据预处理

数据预处理包括灰度变换、标准化、对比度受限的自适应直方图均衡化(CLAHE)、伽马变换、高斯滤波、随机噪声等。不同数据集数据预处理效果如图 4 所示。

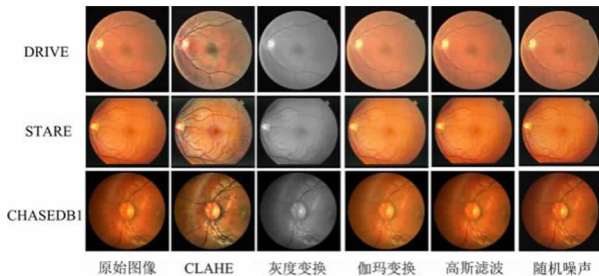


图 4 不同数据集数据预处理效果

3.1.3 数据增强

从数据集的相关介绍可知,3 种数据集本质上

为小样本数据,为提高数据使用效率,本文采用两种切片方法对数据进行扩增。其一为随机切片,另一个为 Mosaic 切片。两种切片按概率随机选取,最终共同完成数据增强处理操作。接下来,以 DRIVE 数据集中的 20 个图像为例,介绍本文所提数据增强方法。

(1) 随机切片

每个图像内以随机中心随机切片,且每个切片尺寸为 48×48 像素。

(2) Mosaic 切片

Mosaic 切片方法参考了 YOLO-v4 中 Mosaic 数据增强^[13]方法。不同之处在于 Mosaic 数据增强主要处理目标检测,且需要对目标真实框进行处理。本文中的 Mosaic 切片方法应用于图像分割,且需要对分割后 ground_truth 标签图像进行处理。本文中的 Mosaic 切片方法应用于图像分割,且需要对分割后 ground_truth 标签图像进行处理。具体过程总结如下:1) 随机选取 4 个 48×48 像素切片图像;2) 对四个图像按概率进行翻转、缩放操作,然后将 4 张图像按顺序摆放;3) 将 ground_truth 标签图像与原图进行相同处理;4) 将原图切片和 ground_truth 标签图像进行组合,形成最终增强图像。

DRIVE 数据集下的数据增强效果如图 5 所示。因每幅图像选取 9500 个切片,这样 20 幅图像总共包含 190000 个切片。训练时随机选取其中 80% 切片用于训练,剩下的 20% 切片用于验证。

本文使用的切片方法的优点总结如下:1) 这些切片允许网络访问有关像素的本地信息,该信息对整体预测具有潜在影响;2) 确保输入数据平衡;3) 在

整个图像上随机地进行采样,切片很可能完全在视野(FOV)之外,因此网络模型可以有效学习分辨 FOV 边界与血管的差异;4) 利用 Mosaic 切片可有效丰富数据样本,且网络计算时一次性可输入 4 张图像,在一定程度上能够提升训练效率。

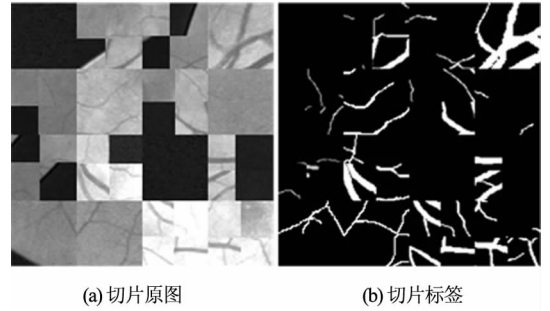
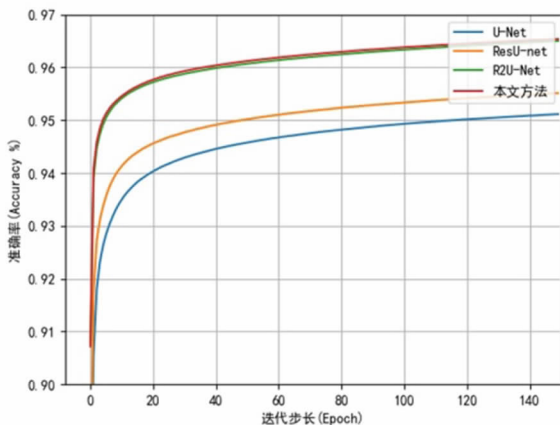


图 5 数据增强效果

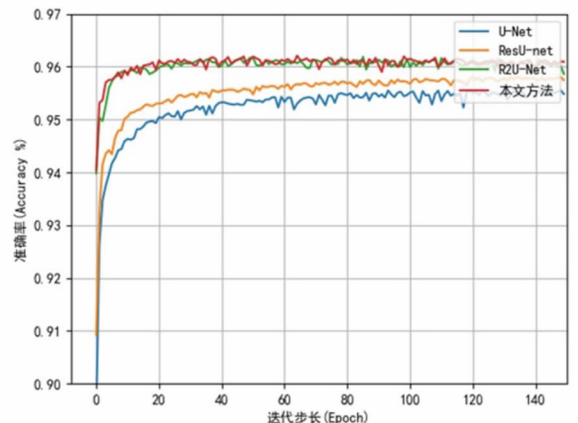
3.2 仿真结果分析

本文在训练时使用随机初始化方法和随机梯度下降优化方法,并采用分类交叉熵计算损失。

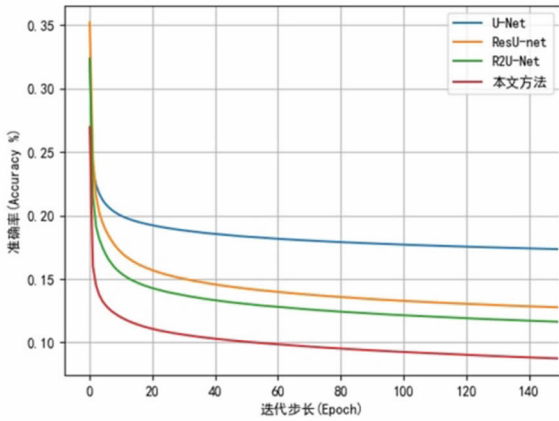
令训练时批处理量为 32,迭代步长为 150。图 6 所示为使用 U-Net、ResU-Net、R2U-Net 及本文方法在 DRIVE 数据集训练/验证准确率及 Loss 变化曲线,其中图 6(a)、(b) 分别为模型训练/验证准确率变化曲线,图 6(c)、(d) 为模型训练/验证 Loss 变化曲线。可以看出,在验证准确率和 Loss 变化曲线上,所提方法可以很快收敛,其他方法波动时间较长,表明本文方法在准确率及训练时间上均有所改善。



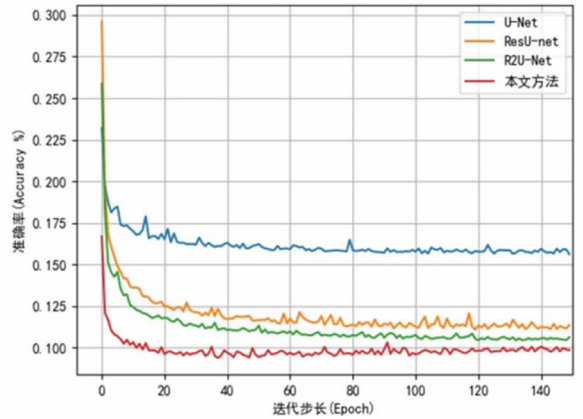
(a) DRIVE 数据集训练准确率



(b) DRIVE 数据集验证准确率



(c) DRIVE 数据集训练 Loss



(d) DRIVE 数据集验证 Loss

图 6 DRIVE 数据集下训练/验证曲线

图 7 所示为模型测试结果展示效果,第一行为输入,第二行为标签,最后一行为预测分割结果。

进一步,在数据集 DRIVE、STARE 和 CHASE_DB1 下对四种不同的模型进行了定量分析,结果如表 1 所示。四个指标分别表示如下:SE 为敏感度,SP 为特异性,AC 为准确率,AUC 为 ROC 曲线下的面积。综合比较四个指标,其中,U-Net 性能最低,其次为 ResU-Net 和 R2U-Net,本文提出的模型在 AC 和 AUC 性能更优。与 U-Net 相比,准确率分别提升了 0.28%、0.17% 和 0.46%;与 R2U-Net 相比性能提升不明显,但本文模型训练时间明显缩短。

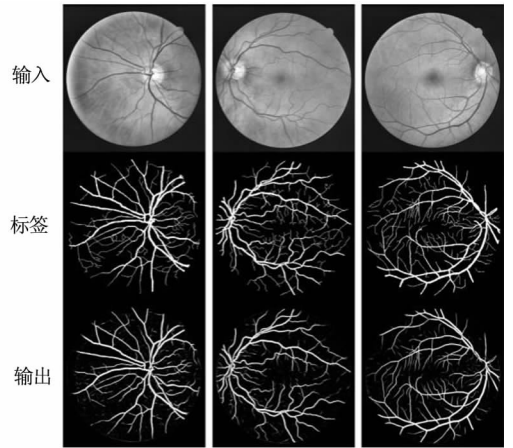


图 7 模型测试结果展示效果

表 1 不同数据集下各模型对比结果

数据集	模型	SE	SP	AC	AUC
DRIVE	R2U-Net	0.7792	0.9904	0.9556	0.9764
	ResU-Net	0.7726	0.9820	0.9553	0.9779
	U-Net	0.7537	0.9820	0.9531	0.9755
	本文方法	0.7664	0.9927	0.9559	0.9784
STARE	R2U-Net	0.8108	0.9871	0.9706	0.9909
	ResU-Net	0.8203	0.9856	0.9700	0.9904
	U-Net	0.8270	0.9842	0.9690	0.9898
	本文方法	0.8079	0.9869	0.9707	0.9908
CHASE_DB1	R2U-Net	0.7459	0.9836	0.9622	0.9803
	ResU-Net	0.7726	0.9820	0.9553	0.9779
	U-Net	0.8288	0.9701	0.9578	0.9772
	本文方法	0.7462	0.9830	0.9624	0.9805

4 结 论

本文对密集图像分割中常用的卷积神经网络、U-Net 等模型进行研究,提出了将 ResGRU 作为网络块嵌入 U-Net 模型,从而设计了加快网络训练速度的 ResGRUU-Net 网络模型。该模型继承了 U-Net 网络模型中 U 形结构的特点,包含下采样和上采样两部分。ResGRU 块由两个 GRU 模块组成,每个 GRU 模块具有独立的参数调整能力,从而加快模型收敛速度。

本文在模型验证时选取的数据集标签图主要为黑白图像,分类为简单的二分类任务。未来研究的方向可为多分类任务及迁移学习,即将模型应用到更多其他类别数据集,同时有效减少新数据集训练所用时间。

参考文献

[1] 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报, 2017, 40(6): 1229-1251.
[2] 蒋昂波, 王维维. ReLU 激活函数优化研究[J]. 传感器与微系

统, 2018, 37(2): 50-52.

- [3] 陈杰. 深度学习正则化技术研究及其在驾驶安全风险领域的应用[D]. 合肥: 中国科学技术大学, 2019.
- [4] 陆永帅, 李元祥, 刘波, 等. 基于深度残差网络的高光谱遥感数据霾监测[J]. 光学学报, 2017, 37(11): 314-324.
- [5] 田萱, 王亮, 丁琪. 基于深度学习的图像语义分割方法综述[J]. 软件学报, 2019, 30(0): 440-468.
- [6] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Seg-Net: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.
- [7] PETROV I, SHAKHURO V, KOUNSHIN A. Deep probabilistic human pose estimation[J]. IET Computer Vision, 2018, 12(5): 578-585.
- [8] ZHANG Z, LIU Q, WANG Y. Road extraction by deep residual U-Net[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(5): 749-753.
- [9] 杨云, 张立泽清, 齐勇. 结合优化 U-Net 和残差学习的细胞膜分割[J]. 计算机工程与设计, 2019, 40(11): 3313-3318.
- [10] PENG D, ZHANG Y, GUAN H. End-to-end change detection for high resolution satellite images using improved UNet++[J]. Remote Sensing, 2019, 11(11): 1382.
- [11] 刘楷东, 谢斌, 翟志强, 等. 基于 R2U-Net 和空洞卷积的羊后腿分割目标肌肉区识别[J]. 农业机械学报, 2020, 51(S2): 507-514.