

基于迁移学习的伤口图像分割^①



陈志威^{1,2}, 赵奎¹, 曹吉龙³, 孙靖³, 马慧敏⁴

¹(中国科学院 沈阳计算技术研究所, 沈阳 110168)

²(中国科学院大学, 北京 100049)

³(中国医科大学附属第四医院, 沈阳 110032)

⁴(东软集团股份有限公司 医疗解决方案事业本部, 沈阳 110179)

通信作者: 赵奎, E-mail: 351282763@qq.com

摘要: 图像分割是计算机辅助阅片的基础, 伤口图像分割的准确率直接影响伤口分析的结果. 传统方法进行伤口分割步骤繁琐, 准确率低. 目前已有少部分人利用深度学习进行伤口图像分割, 但是他们都是基于小型数据集, 难以发挥深度神经网络的优势, 准确率难以进一步提高. 充分发挥深度学习在图像分割领域的优势需要大型数据集, 目前还没有关于伤口图像的大型公共数据集, 而制作大型伤口图像数据集需要人工标记, 耗费大量时间和精力. 本文提出基于迁移学习的伤口图像分割方法, 首先利用大型公共数据集训练 ResNet50 网络作为特征提取器, 再利用该特征提取器连接上两个并行的注意力机制后在利用小型伤口图像数据集进行再训练. 实验表明本方法的分割结果在平均交并比上有较大提高, 在某种程度上解决了缺乏大型伤口图像数据集而导致伤口图像分割准确率低的问题.

关键词: 迁移学习; 伤口图像; 图像分割; 预训练; 深度学习; 图像分割

引用格式: 陈志威, 赵奎, 曹吉龙, 孙靖, 马慧敏. 基于迁移学习的伤口图像分割. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/8640.html>

Wound Image Segmentation Based on Transfer Learning

CHEN Zhi-Wei^{1,2}, ZHAO Kui¹, CAO Ji-Long³, SUN Jing³, MA Hui-Min⁴

¹(Shenyang Institute of Computing Technology, Chinese Academy of Sciences, Shenyang 110168, China)

²(University of Chinese Academy of Sciences, Beijing 100049, China)

³(The Fourth Affiliated Hospital of China Medical University, Shenyang 110032, China)

⁴(Medical Solutions Business Headquarters, Neusoft Group Co. Ltd., Shenyang 110179, China)

Abstract: Image segmentation is the basis of computer-aided film reading, and the accuracy of wound image segmentation directly affects the results of wound analysis. However, the traditional method of wound image segmentation has cumbersome steps and low accuracy. At present, a few studies have applied deep learning to wound image segmentation, but they are all based on small data sets and can hardly give full play to the advantages of deep neural networks and further improve accuracy. Maximizing the advantages of deep learning in the field of image segmentation requires large data sets, but there is no large public data set on wound images as establishing large wound image data sets requires manual labeling, which consumes a lot of time and energy. In this study, a wound image segmentation method based on transfer learning is proposed. Specifically, the ResNet50 network is trained with a large public data set as a feature extractor, and then the feature extractor is connected with two parallel attention mechanisms for retraining with a small wound image data set. Experiments show that the segmentation results of this method are greatly improved in the average intersection over union (IoU), and this method solves the problem of low accuracy in wound image segmentation due to the lack of large wound image data sets to some extent.

Key words: transfer learning; wound image; image segmentation; pre-training; deep learning; image segmentation

① 基金项目: 国家水体污染控制与治理科技重大专项 (2018ZX07601001)

收稿时间: 2022-11-04; 修改时间: 2021-12-02, 2021-12-21, 2021-12-27; 采用时间: 2021-12-28; csa 在线出版时间: 2022-05-31

慢性伤口如静脉性溃疡、压力性溃疡、糖尿病性溃疡等具有高发病率、高患病率、高费用等特点,已成为医疗保健的一大难题.这类伤口的愈合时间长达数周甚至数百天,而住院条件无法满足患者所有治疗都在院内进行,因此院外延续性护理成为了慢性伤口患者治疗的重要组成部分.皮肤病变的常规诊断是通过活组织和周围皮肤的检查,伤口溃疡可以通过伤口部位特定类型的组织进行预测^[1-3].伤口观察与分析是伤口治疗的基础,所以在早些时候伤口护理通常是由患者和专业护理人员一对一进行.通常慢性伤口的患者行动非常不便利,近年来随着互联网的发展,伤口护理也从线下转到了线上,由患者拍摄伤口部位的像片通过网络传送给专业护理人员进行分析 and 诊疗.伤口护理人员严重依赖于伤口图像进行正确的诊断和治疗.近年来随着机器学习和深度学习的发展以及医疗设备的不断更新,部分人也开始利用机器学习和深度学习等算法辅助医生阅片.然而由患者拍摄的伤口图像通常包含大量背景,这对计算机阅片增加了许多干扰.为了提高计算机阅片的精度,对伤口图像进行分割,去除背景,提取出伤口区域成为了必不可少的环节.伤口图像的分割不仅可以帮助医生提取出伤口图像中感兴趣的区域,并且可以为伤口图像的分析提供形状和结构特征.伤口分割的准确度直接影响计算机对患者拍摄的伤口图片中伤口面积的测量和组织分析.但是由于缺乏带有标记的伤口图像,目前关于使用深度方法进行伤口分割的研究很少.在早些时候图像分割主要采用一些传统方法,如阈值分割法,区域生长法、区域分裂合并法等.国内利用传统方法对伤口图像进行分割的研究较少,国外有 Song 等^[4]分别采用阈值分割、K-means 聚类、边缘检测和区域增长 4 种方法对伤口图像进行分割,考虑到没有哪一种方法是对所有伤口图像有效的,他们将 4 种方法进行集成,对不同的伤口图像调节每个方法的参数,大大提升了分割系统的鲁棒性. Fauzi 等^[5]先在修正的 HSV 颜色空间进行概率图计算,然后再用区域生长法和最优阈值进行分割. Li 等^[6]利用神经网络对伤口图像进行分割,并且利用分水岭算法和动态阈值法分割结果进行校正,实验结果表明该方法在 950 张伤口图片的数据集上取得了不错的效果,但是该方法比较繁琐. Wang 等^[7]采用卷积神经网络 MobileNetV2^[8]进行分割,该网络比传统的卷积神经网络速度更快.由于缺乏大型的带有标记的伤口数据集,大多数研究人员都是在几百张伤口图片的小数据上进行训练,使用的网络不够深,难以发挥

深度神经网络的优势,在小数据集上的精度提高并无很大意义.由于人工标记的伤口图像需要大量的时间和精力,所以本文提出基于迁移学习的伤口图像分割方法.

本文利用 ResNet50^[9]在 ImageNet 数据集上的进行预训练,将训练好的 ResNet50 模型作为特征提取前端,后接 Danet 网络^[10]形成伤口图像分割网络,再在收集到的伤口图像数据集上进行再训练,得到伤口图像分割模型.该方法在平均交并比上达到了 87.92%,比之前的方法效果都要好.

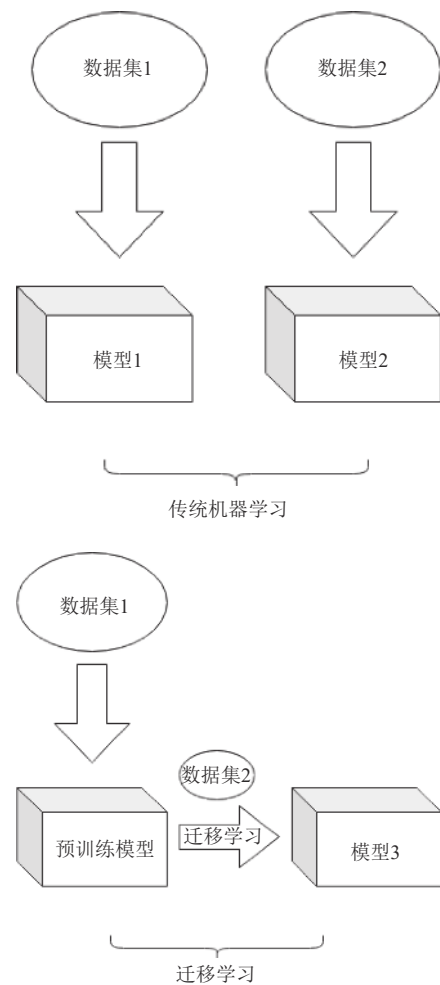


图1 传统机器学习与迁移学习

1 基本理论

1.1 迁移学习

迁移学习 (transfer learning)^[11,12],如图 1 所示,是指将已训练好的模型参数迁移到新的模型,从而达到加快模型的训练的目的.通过使用其他相关任务训练模型的参数,即使只有一个小数据集,也可能达到很好

的性能。

迁移学习按照迁移知识形式可以分为: 基于实例的迁移学习, 基于特征的迁移学习, 基于参数的迁移学习, 基于关系的迁移学习. 由于缺乏大型的伤口图像数据集, 可以利用迁移学习, 使用在其他大型数据集上的预训练模型作为目标任务模型的起点, 从而充分发挥深度神经网络的优势, 优化模型训练效率. 本文利用 ResNet50 网络在 ImageNet 数据集上的预训练模型作为伤口图像的特征提取前端.

1.2 ResNet50

通常, 在数据集足够大的情况下, 神经网络的深度越深, 模型的精度越高, 但是随着网络深度增加, 深度神经网络的退化问题也随之而来. 如果深层的网络后面的层是恒等映射, 那么模型就退化成一个浅层网络. ResNet 网络的提出正好解决了这个问题, 深度残差网

络不在拟合一个恒等映射 $H(X)=X$, 而是转换为学习一个残差函数 $F(X)=H(X)-X$, 只要 $F(X)=H(X)-X$. 只要 $F(X)=0$, 就构成了一个恒等映射. 而拟合残差比拟合恒等映射要容易的多. 本文利用一个在 ImageNet 数据集上预训练的 50 层的 ResNet 网络作为特征提取器, 来提取伤口图像的特征.

1.3 Danet

Danet 网络是一种双重注意力网络, 在传统的 FCN 网络^[11] 之上附加两种类型的注意力模块, 用来捕捉空间和通道维度中的全局特征依赖关系. 使用空间注意力模块^[13] 来学习特征的空间相互依赖性, 使用通道注意力模块^[14] 来模拟通道的相互依赖性. 通过添加这两个模块可以进一步提升模型的特征表示, 提升模型的分割性能, 图 2(a) 为空间注意力模块, 图 2(b) 为通道注意力模块.

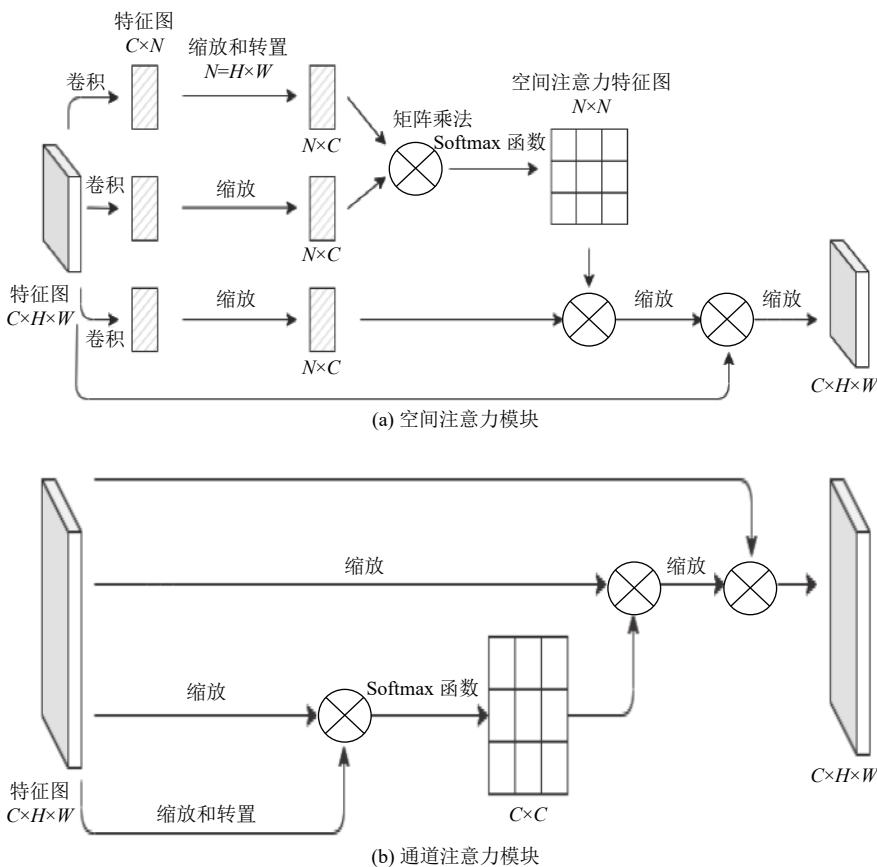


图 2 空间注意力模块和通道注意力模块

Danet 网络通过将这两种注意力并联, 将提取的特征图分别经过这两种注意力机制处理后形成的分别具有空间注意力和通道注意力的特征图, 再将这两种特

征图进行按元素求和完成特征图的聚合, 最后形成的特征图就融合了空间注意力和通道注意力. 整体结构如图 3 所示.

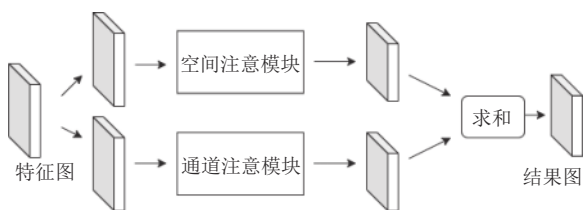


图3 Danet 结构

2 基于深度迁移学习的伤口图像分割

2.1 数据集

本文数据集包括预训练数据集 ImageNet 和再训练伤口图像数据集. ImageNet 数据集是一个大型公共的图像分类数据集, 含 120 万张, 1 000 个类别的图片, 该数据集足够大, 可以用来训练深度神经网络, 充分发挥深度神经网络的性能. 本文利用该数据训练的 ResNet50 网络, 作为图像分割的特征提取器. 伤口数据集总共有 810 张图片, 利用 labelme 图像标注软件进行手工标注. 部分伤口图像及标注图像如图 4 所示.

2.2 数据预处理

由于伤口图像数据集太小, 为了增加模型的泛化性,

本文对训练数据进行了增广操作, 对读取的每个小批量图像进行随机水平和垂直翻转. 通过对图像的随机翻转, 可以使感兴趣的物体出现在图像的不同位置, 减轻模型对物体出现位置的依赖性, 从而提高模型的泛化能力.



图4 部分伤口图像和标签

2.3 实验分析

ResNet50 整体结构如图 5 所示, 整体结构如下图所示, 整个网络层数为 50 层, 图中“7×7, 64, 2”意为通过尺寸为 7×7 卷积层处理, 输出通道为 64 个通道, 步长为 2. 预训练时对整个模型中的参数进行随机初始化, 利用该网络对 ImageNet 数据集进行图像分类, 由于 ImageNet 数据集中有 1 000 个类别, 所以最后的连接层的输出有 1 000 个类.

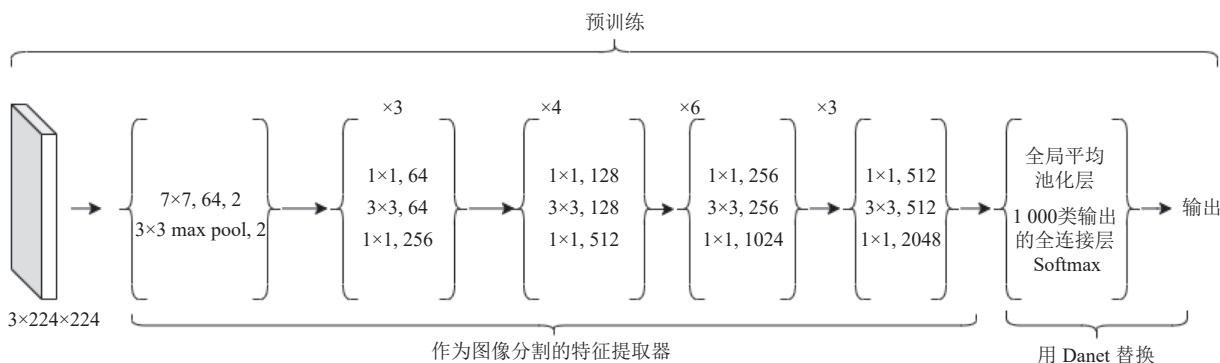


图5 ResNet50 结构

由于伤口数据集比较少, 本文不从头训练 ResNet50 网络, 而是直接将在 ImageNet 数据集上训练的 ResNet50 网络的参数迁移. 如图 5 所示, 本文将图像分类网络最后的全局平均池化层和全连接层替换成 Danet 网络进行图像分割. 训练时对特征提取器模块不进行随机初始化, 而是用预训练网络的参数权重. Danet 模块的权重进行随机初始化.

伤口图像与 ImageNet 网络中的图像差别比较大, ImageNet 数据集并不包含伤口图像, 所以对迁移的特征提取器中的网络权重进行微调. Danet 块中的参数需要从头开始学习所以设置较大学习率, 而作为特征提取器的

ResNet50 网络只需要微调所以设置较小学习率. 同时为了使网络更容易收敛, 本文对学习率还设置了衰减策略.

整个网络由经过预训练的 ResNet50 网络作为特征提取器, 后接具有空间注意力机制和通道注意力机制的 Danet 网络组成伤口图像分割网络再在伤口图像数据集上进行再训练得到伤口图像分割模型. 伤口图像首先由在 ImageNet 数据集上预训练的 ResNet50 网络作为特征提取器提取出特征图, 特征图分别通过空间注意力机制模块和通道注意力机制模块. 空间注意力模块学习特征的空间相互依赖性, 通道注意力模块来模拟通道相互依赖性. 将两个注意力模块处理后的

特征图进行求和以完成特征融合最后生成结果图像。

3 评价指标

伤口图像的分割结果最后将作为伤口图像分析的基础, 图像分割的精度越高, 能够为图像分析排除大量图像背景中无关信息的干扰, 图像分析的准确率也越高. 均交并比 (mean intersection over union, *MIoU*)^[15] 为图像分割的标准度量. 其计算真实值和预测值的交集和并集之比, 之后求平均值, 公式如下:

$$MIoU = \frac{1}{k+1} \frac{\sum_{i=0}^k P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}} \quad (1)$$

其中, k 表示共有 $k+1$ 个类别 (包含一个背景), P_{ij} 表示本属于类 i 但被预测为类 j 的像素数量, P_{ii} 为预测正确的像素数量.

Kappa 系数用于一致性检验, 也可以用于衡量分类精度. 通常 *Kappa* 系数落在 $[0, 1]$ 之间. 计算公式为:

$$Kappa = \frac{p_o - p_e}{1 - p_e} \quad (2)$$

其中, p_o 为准确率, 即每一类分类正确的样本数量之和除以总样本数. $p_e = \frac{a_1b_1 + a_2b_2 + \dots + a_ib_i}{n \times n}$, 其中 a_i 为每一类的真实样本个数, b_i 为预测出来的每一类样本个数. *Kappa* 值可分为 5 组来表示不同级别的一致性, 如表 1 所示.

表 1 *Kappa* 值等级

<i>Kappa</i> 值	等级
0.0-0.20	极低的一致性
0.21-0.40	一般的一致性
0.41-0.60	中等的一致性
0.61-0.80	高度的一致性
0.81-1.00	几乎完全一致

4 实验平台

本文实验使用的 PaddlePaddle 框架完成, 硬件信息和框架版本如表 2 所示.

5 结果与分析

本文进行了多组对比试验, 实验结果如表 3 所示. Unet^[16] 是 2015 年提出的图像分割网络, 被广泛运用于医学图像分割. Attention-Unet^[17] 是在 Unet 网络结构上加上 Attention 模块, ResNet50_Danet 是从零开始在伤口图像数据集进行训练的模型, 没有进行迁移学习.

Pretrain_ResNet50_Danet 是在将 ResNet50 在 ImageNet 进行预训练作为特征提取器和再在伤口图像数据集进行再训练的模型. 从表 3 实验结果可以看本文提出的基于迁移学习的 Pretrain_ResNet50_Danet 网络达到了最好的效果, 平均交并比达到了 0.8792, *Kappa* 系数也达到了最高的 0.8632, 达到了几乎完全一致. 比医学影像分割常用的网络 Unet 和加注意力机制的 Attention-Unet 网络在平均交并比上都有大幅提升. 同时, 基于迁移学习的 Pretrain_ResNet50_Danet 网络比用伤口数据集从零训练的 ResNet_50-Danet 网络平均交并比高 0.0476.

表 2 硬件信息和框架版本

属性	版本
Python版本	Python 3.7
框架版本	PaddlePaddle 2.1.0
CPU	4 Cores
RAM	32 GB
GPU	Tesla V100
显存	32 GB

表 3 实验结果

算法	<i>mIoU</i>	<i>Kappa</i> 系数
Unet	0.7723	0.7081
Attention-Unet	0.8168	0.7773
Resnet50_Danet	0.8316	0.7963
Pretrain_ResNet50_Danet	0.8792	0.8632

从实验结果可以看出, 本文基于迁移学习伤口图像分割算法在一定程度上解决了伤口图像数据集小而导致的训练的伤口分割模型精度低的问题, 通过在大数据集 ImageNet 上训练 ResNet50 作为特征提取器, 能有效提高伤口分割的精度. 部分分割结果如图 6 所示. 第 1 行是原图片, 第 2 行是手工利用 labelme 图像标注软件标注的图片, 第 3 行是利用训练好的 Pretrain_ResNet50_Danet 模型预测生成的伪彩色分割图片, 第 4 行是该模型预测的分割图与原图的叠加图.

6 总结

本文将在 ImageNet 数据集上训练的 ResNet50 网络去除最后的全局平均池化和全连接层作为特征提取器迁移至伤口图像分割任务中, 提取的特征图再通过 Danet 网络处理. 由于伤口图像数据集与 ImageNet 数据集存在较大差异, 我们对 ResNet50 网络设置较小的学习率, 并设置权重衰减进行微调, 同时 Danet 网络是从头开始训练的, 我们设置较大学习率. 实验表明本文得方法取得最好的效果, 我们还对比了相同的网络使

用迁移学习和不使用迁移学习得分割精度,实验表明在伤口图像数据集小的情况下,使用迁移学习能取得精度更高的分割结果。

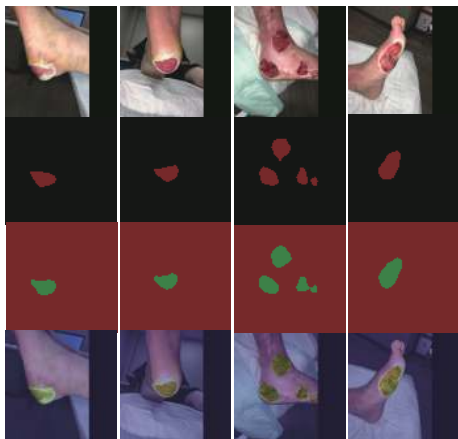


图6 分割结果图与标注图对照和原图对比

参考文献

- Blanco G, Bedo MVN, Cazzolato MT, *et al.* A label-scaled similarity measure for content-based image retrieval. Proceedings of 2016 IEEE International Symposium on Multimedia. San Jose: IEEE, 2016. 20–25. [doi: [10.1109/ISM.2016.0014](https://doi.org/10.1109/ISM.2016.0014)]
- Litjens G, Kooi T, Bejnordi BE, *et al.* A survey on deep learning in medical image analysis. Medical Image Analysis, 2017, 42: 60–88. [doi: [10.1016/j.media.2017.07.005](https://doi.org/10.1016/j.media.2017.07.005)]
- Seixas JL, Barbon S, Mantovani RG. Pattern recognition of lower member skin ulcers in medical images with machine learning algorithms. Proceedings of the 2015 IEEE 28th International Symposium on Computer-based Medical Systems. Sao Carlos: IEEE, 2015. 50–53.
- Song B, Sacan A. Automated wound identification system based on image segmentation and Artificial Neural Networks. Proceedings of 2012 IEEE International Conference on Bioinformatics and Biomedicine. Philadelphia: IEEE, 2012. 1–4.
- Fauzi MFA, Khansa I, Catignani K, *et al.* Computerized segmentation and measurement of chronic wound images. Computers in Biology and Medicine, 2015, 60: 74–85. [doi: [10.1016/j.combiomed.2015.02.015](https://doi.org/10.1016/j.combiomed.2015.02.015)]
- Li FZ, Wang CJ, Liu XH, *et al.* A composite model of wound segmentation based on traditional methods and deep neural networks. Computational Intelligence and Neuroscience, 2018, 2018: 4149103. [doi: [10.1155/2018/4149103](https://doi.org/10.1155/2018/4149103)]
- Wang CB, Anisuzzaman DM, Williamson V, *et al.* Fully automatic wound segmentation with deep convolutional neural networks. Scientific Reports, 2020, 10(1): 21897. [doi: [10.1038/s41598-020-78799-w](https://doi.org/10.1038/s41598-020-78799-w)]
- Sandler M, Howard A, Zhu ML, *et al.* MobileNetV2: Inverted residuals and linear bottlenecks. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 4510–4520. [doi: [10.1109/cvpr.2018.00474](https://doi.org/10.1109/cvpr.2018.00474)]
- He KM, Zhang XY, Ren SQ, *et al.* Deep residual learning for image recognition. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 770–778. [doi: [10.1109/cvpr.2016.90](https://doi.org/10.1109/cvpr.2016.90)]
- Fu J, Liu J, Tian HJ, *et al.* Dual attention network for scene segmentation. Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach: IEEE, 2019. 3141–3149. [doi: [10.1109/CVPR.2019.00326](https://doi.org/10.1109/CVPR.2019.00326)]
- Yosinski J, Clune J, Bengio Y, *et al.* How transferable are features in deep neural networks? Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal: NIPS, 2014. 3320–3328.
- Oquab M, Bottou L, Laptev I, *et al.* Learning and transferring mid-level image representations using convolutional neural networks. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 1717–1724. [doi: [10.1109/CVPR.2014.222](https://doi.org/10.1109/CVPR.2014.222)]
- Jaderberg M, Simonyan K, Zisserman A, *et al.* Spatial transformer networks. Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: NIPS, 2015. 2017–2025.
- Hu J, Shen L, Sun G. Squeeze-and-excitation networks. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 7132–7141.
- Garcia-Garcia A, Orts-Escolano S, Oprea S, *et al.* A review on deep learning techniques applied to semantic segmentation. arXiv: 1704.06857, 2017.
- Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer, 2015. 234–241.
- Oktay O, Schlemper J, Le Folgoc L, *et al.* Attention U-Net: Learning where to look for the pancreas. arXiv: 1804.03999, 2018.

(校对责编:牛欣悦)