



U-Net：用于生物医学图像分割的卷积网络

Olaf Ronneberger、Philipp Fischer 和 Thomas Brox

计算机科学系和 BIOS 生物信号研究中心、
德国弗莱堡大学

ronneber@informatik.uni-freiburg.de、

万维网主页：<http://lmb.informatik.uni-freiburg.de/>

摘要大量研究表明，成功的深度网络训练需要成千上万的注释训练样本。在本论文中，我们提出了一种网络和训练策略，该策略依赖于数据增强的强大功能，以更高效地利用可用的注释样本。该架构由一个用于捕捉上下文的收缩路径和一个用于精确定位的对称扩展路径组成。我们的研究表明，这种网络可以从极少的图像中进行端到端训练，并在电子显微镜堆栈神经元结构分割的 ISBI 挑战赛中表现优于之前的最佳方法（滑动窗口卷积网络）。使用在透射光显微镜图像（相衬和DIC）上训练的相同网络，我们在2015年ISBI细胞追踪挑战赛中以较大优势赢得了这些类别的比赛。此外，该网络速度很快。在最新的 GPU 上分割一幅 512x512 的图像不到一秒钟。完整的实现（基于 Caffe）和训练好的网络可在 <http://lmb.informatik.uni-freiburg.de/people/ronneber/u-net> 上获得。

1 引言

在过去两年中，深度卷积网络在许多视觉识别任务中的表现超过了目前的技术水平，例如[7,3]。虽然卷积网络早已存在[8]，但由于可用训练集的规模和所考虑网络的规模，它们的成功受到了限制。Krizhevsky 等人的研究[7]取得了突破性进展，他们在具有 100 万张训练图像的 ImageNet 数据集上对一个具有 8 层和数百万个参数的大型网络进行了监督训练。此后，更大、更深的网络也被训练出来 [12]。

卷积网络通常用于分类任务，在这种任务中，图像的输出是一个单一的类标签。然而，在许多视觉任务中，尤其是在生物医学图像处理中，所需的输出应该包括定位，即给每个像素分配一个类标签。此外，在生物医学任务中，通常无法获得数千张训练图像。因此，Ciresan 等人[1]在滑动窗口设置中训练了一个网络，通过提供像素周围的局部区域（补丁）来预测每

个像素的类别标签。

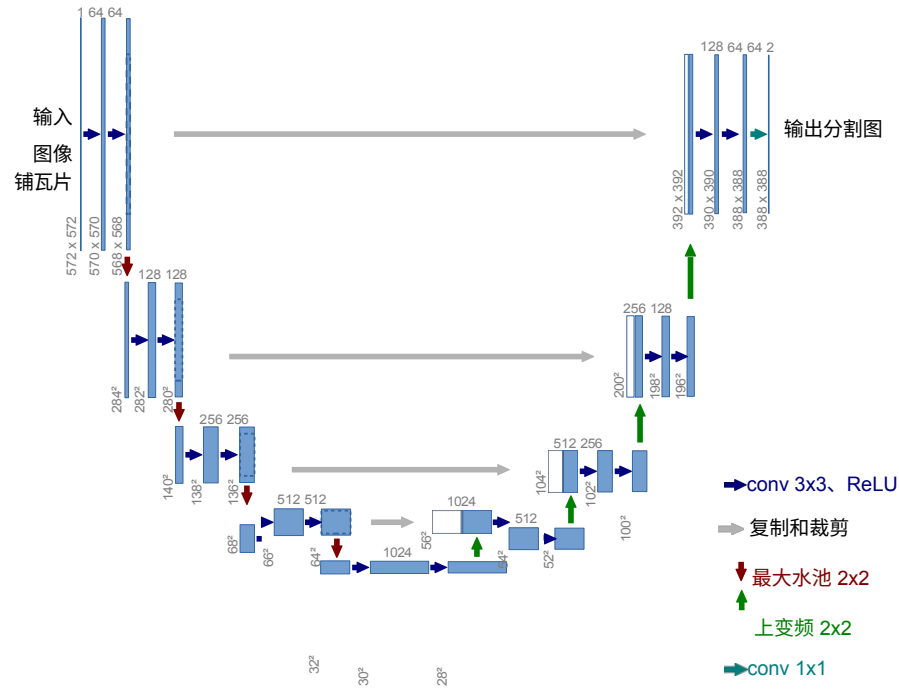


图 1.U-net 架构（以最低分辨率 32x32 像素为例）。每个蓝色方框对应一个多信道特征图。方框顶部表示通道数。x-y 大小位于方框左下边缘。白色方框代表复制的特征图。箭头表示不同的操作。

作为输入。首先，该网络可以定位。其次，以斑块为单位的训练数据远远大于训练图像的数量。由此产生的网络在 ISBI 2012 的电磁分割挑战赛中大获全胜。

显然，Ciresan 等人[1]的策略有两个缺点。首先，它的运行速度相当慢，因为必须对每个补丁分别运行网络，而且由于补丁重叠，存在大量冗余。其次，定位精度和上下文的使用之间需要权衡。较大的斑块需要更多的最大池化层，从而降低了定位精度，而较小的斑块则只允许网络看到很少的上下文。最近的方法[11,4]提出了一种考虑到多层特征的分类器输出。在实现良好定位的同时，还能利用上下文。

在本文中，我们采用了一种更优雅的架构，即所谓的“全卷积网络”[9]。我们对这一架构进行了修改和扩展，使其只需很少的训练图像就能工作，并产生更精确的分割结果；见图 1。[9]的主要思想是通过连续层来补充通常的卷积网络，其中池化算子被上采样算子所取代。因此，这些层提高了输出的分辨率。为了进行定位，将收缩路径中的高分辨率特征与上采样特征结合起来。

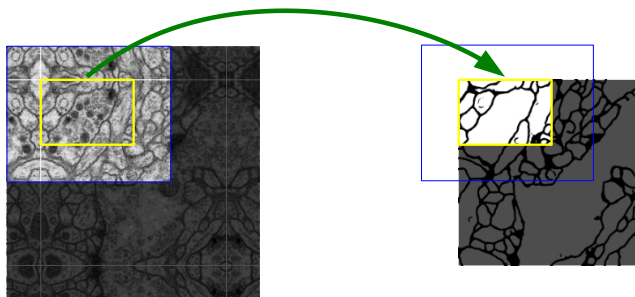


图 2.用于无缝分割任意大型图像的重叠-瓦片策略（此处分割电磁堆栈中的神经元结构）。黄色区域的分割预测需要蓝色区域内的图像数据作为输入。缺失的输入数据通过镜像推断出来。

输出。然后，连续的卷积层可以根据这些信息，学习组合出更精确的输出。

我们的架构有一个重要的变化，那就是在上采样部分，我们也有大量的特征通道，这使得网络可以将上下文信息传播到更高分辨率的层。因此，扩展路径与收缩路径基本对称，形成了一个 U 型结构。该网络没有任何全连接层，只使用每次卷积的有效部分，即分割图只包含输入图像中具有完整上下文的像素。这种策略允许通过重叠-瓦片策略对任意大的图像进行无缝分割（见图 2）。为了预测图像边界区域的像素，可以通过镜像输入图像来推断缺失的上下文。这种平铺策略对于将网络应用于大型图像非常重要，否则分辨率就会受到 GPU 内存的限制。

在我们的任务中，可用的训练数据非常少，因此我们通过对可用的训练图像进行弹性变形来过度扩充数据。这样，网络就能学习到这种变形的不变性，而无需在注释图像语料库中看到这些变换。这在生物医学分割中尤为重要，因为变形曾是组织中最常见的变化，而真实的变形可以有效地模拟出来。Dosovitskiy 等人[2]在无监督特征学习中展示了数据增强对学习不变性的价值。

在许多细胞分割任务中，另一个挑战是如何分离同一类别的触摸对象；见图 3。为此，我们建议使用加权损失，即在损失函数中，触摸细胞之间的分离背景标签获得较大权重。

由此产生的网络适用于各种生物医学分割问题。在本文中，我们展示了 EM 堆栈中神经元结构的分割结果（2012 年 ISBI 开始的一项正在进行的竞赛）。

Ciresan 等人的网络[1]。此外，我们还展示了 2015 年 ISBI 细胞追踪挑战赛中光镜图像的细胞分割结果。在最具挑战性的两个二维透射光数据集上，我们以较大优势胜出。

2 网络架构

网络结构如图 1 所示。它由收缩路径（左侧）和扩展路径（右侧）组成。收缩路径遵循卷积网络的典型架构。它由两个 3x3 卷积（无填充卷积）的重复应用组成，每个卷积之后都有一个整流线性单元（ReLU）和一个 2x2 最大池化操作（步长为 2）用于下采样。在每个下采样步骤中，我们都会将特征通道的数量增加一倍。扩展路径中的每一步都包括对特征图进行一次上采样，然后进行一次 2x2 卷积（“上卷积”），将特征通道数减半；与收缩路径中相应裁剪的特征图进行一次合并；以及两次 3x3 卷积，每次卷积后都进行一次 ReLU。由于每次卷积都会损失边界像素，因此必须进行裁剪。在最后一层，使用 1x1 卷积将每个 64 分量的特征向量映射到所需的类别数。整个网络共有 23 个卷积层。

为了实现输出分割图的无缝平铺（见图 2），必须选择输入平铺的大小，使所有 2x2 最大池化操作都应用于 x 和 y 大小相同的图层。

3 培训

输入图像及其相应的分割图被用于使用 Caffe [6] 的随机梯度下降实现来训练网络。由于未填充卷积，输出图像比输入图像小一个恒定的边界宽度。为了最大限度地减少开销并最大限度地利用 GPU 内存，我们倾向于使用大尺寸的输入图片，而不是大尺寸的批处理图片，从而将批处理图片减少为单张图片。因此，我们使用了较高的动量（0.99），从而使大量之前看过的训练样本决定了当前优化步骤中的更新。

能量函数的计算方法是对最终的特征图与交叉熵损失函数相结合。软最大值为

定义为 $p(\mathbf{x}) = \exp(a(\mathbf{x})) / \sum_{k=1}^K \exp(a_k(\mathbf{x}))$ 其中 $a_k(\mathbf{x})$ 表示

2K 是类别数， $p_k(\mathbf{x})$ 是近似最大函数。也就是说，对于具有最大激活 $a_k(\mathbf{x})$ 的 k ， $p_k(\mathbf{x}) \approx 1$ ；而对于所有其他 k ， $p_k(\mathbf{x}) \approx 0$ 。然后，交叉熵会在每个位置对 $p_{A(\mathbf{x})}(\mathbf{x})$ 与 1 的偏差进行惩罚，方法是

$$E = \sum_{\mathbf{x} \in \mathcal{E}} w(\mathbf{x}) \log(p_{A(\mathbf{x})}(\mathbf{x})) \quad (1)$$

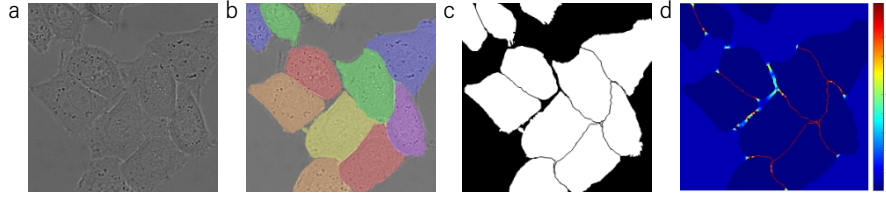


图 3.用 DIC（微分干涉对比）显微镜记录的玻璃上的 HeLa 细胞。(a) 原始图像。(b) 与地面实况分割叠加。不同颜色表示 HeLa 细胞的不同实例。(c) 生成的分割掩膜（白色：前景，黑色：背景）。(d) 带有像素损失权重的地图，以强制网络学习边界像素。

其中 $l: \Omega \rightarrow \{1, .K\}$ 是每个像素的真实标签，而 $w: \Omega \rightarrow \mathbb{R}$ 是我们引入的权重图，用于在训练中赋予某些像素更多的重要性。

我们预先计算了每个地面实况分割的权重图，以反映训练数据集中某个类别像素出现的不同频率，并迫使网络学习我们在相邻单元之间引入的小分隔边界（见图 3c 和 d）。

利用形态学运算计算分离边界。权重图的计算公式为

$$w(\mathbf{x}) = w_c(\mathbf{x}) + w_0 - \frac{(d_1(\mathbf{x}) + d_2(\mathbf{x}))^2}{\exp(2\sigma^2)} \quad (2)$$

其中， $w_c: \Omega \rightarrow \mathbb{R}$ 是平衡类频率的权重图， $d_1: \Omega \rightarrow \mathbb{R}$ 表示到最近单元边界的距离， $d_2: \Omega \rightarrow \mathbb{R}$ 表示到第二近单元边界的距离。在实验中，我们设置 $w_0 = 10$ 和 $\sigma \approx 5$ 像素。

在具有许多卷积层和不同网络路径的深度网络中，权重的好初始化极为重要。否则，部分网络可能会产生过多的激活，而其他部分则不会产生激活。理想情况下，初始权重的调整应使网络中的每个特征图都具有近似单位方差。对于采用我们的架构（交替卷积层和 ReLU 层）的网络，可以通过从标准偏差为 $\sqrt{2/N}$ 的高斯分布中提取初始权重来实现这一点，其中 N 表示一个神经元的输入节点数[5]。例如，对于前一层的 3×3 卷积和 64 个特征通道， $N = 9 \cdot 64 = 576$ 。

3.1 数据扩充

在只有少量训练样本的情况下，数据扩增对于网络获得所需的不变性和鲁棒性特性至关重要。对于

在显微图像中，我们主要需要移动和旋转不变性，以及对变形和灰度值变化的鲁棒性。尤其是训练样本的随机弹性变形，似乎是在注释图像极少的情况下训练分割网络的关键概念。我们在一个 3 乘 3 的粗网格上使用随机位移向量生成平滑变形。位移从标准偏差为 10 像素的高斯分布中采样。然后使用双三次插值法计算每个像素的位移。收缩路径末端的剔除层会执行进一步的隐式数据增强。

4 实验

我们展示了 u-net 在三种不同分割任务中的应用。第一个任务是分割电子显微记录中的神经元结构。图 2 展示了数据集和我们获得的分割结果。我们将结果全文作为补充材料提供。数据集由 EM 细分挑战赛 [14] 提供，该挑战赛始于 ISBI 2012，目前仍在接受新的贡献。训练数据是一组 30 幅图像（512x512 像素），来自果蝇初生幼虫腹侧神经索（VNC）的序列切片透射电子显微镜。每张图像都附有相应的完整注释的细胞（白色）和膜（黑色）地面实况分割图。测试集是公开的，但其分割图是保密的。将预测的膜概率图发送给组织者即可获得评估结果。评估是通过在 10 个不同级别上对地图进行阈值处理，并计算 "翘曲误差"、"兰德误差" 和 "像素误差" 来完成的[14]。

u-网络（输入数据 7 个旋转版本的平均值）在没有任何进一步预处理或后处理的情况下，翘曲误差为 0.0003529（新的最佳得分，见表 1），运行误差为 0.0382。

这明显优于 Ciresan 等人[1]的滑动窗口卷积网络结果，Ciresan 等人提交的最佳结果的翘曲误差为 0.000420，兰德误差为 0.0504。就 rand 误差而言，唯一表现较好的是 Ciresan 等人[2]的结果。

表 1.按扭曲误差排序的电磁分割挑战赛排名[14]（2015 年 3 月 6 日）。

等级	组名	翘曲错误	兰德错误	像素误差
	** 人类价值观	0.000005	0.0021	0.0010
1.	u-net	0.000353	0.0382	0.0611
2.	DIVE-SCI	0.000355	0.0305	0.0584
3.	IDSIA [1]	0.000420	0.0504	0.0613
4.	潜水	0.000430	0.0545	0.0582
.
10.	IDSIA-SCI	0.000653	0.0189	0.1027

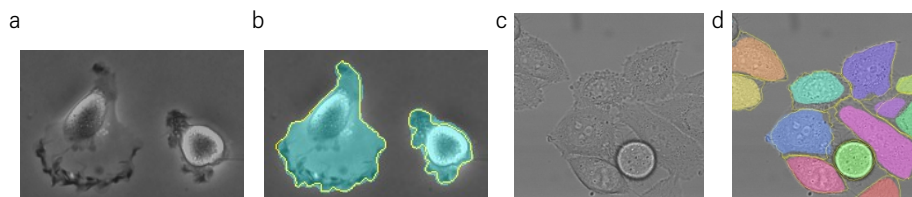


图 4. ISBI 细胞追踪挑战赛的结果。(a) "PhC-U373 "数据集输入图像的一部分。(c) "DIC-HeLa "数据集的输入图像。(d) 带有人工地面实况（黄色边框）的分割结果（随机彩色掩膜）。

表 2. 2015 年 ISBI 细胞追踪挑战赛的分割结果（IOU）。

名称	PhC-U373	DIC-HeLa
IMCB-SG (2014)	0.2669	0.2935
KTH-SE (2014)	0.7953	0.4607
美国住房委员会 (2014 年)	0.5323	-
2015 年第二名	0.83	0.46
u-net (2015)	0.9203	0.7756

该数据集上的算法使用了高度特定于数据集的后处理方法¹适用于 Ciresan 等人[1]的概率图。

我们还将 u-net 应用于光镜图像中的细胞分割任务。该分割任务是 2014 年和 2015 年 ISBI 细胞追踪挑战赛的一部分 [10,13]。第一个数据集 "PhC-U373"²第一组数据集 "PhC-U373 "包含通过相衬显微镜记录的聚丙烯酰亚胺基底上的胶质母细胞瘤-胃细胞瘤 U373 细胞（见图 4a,b 和补充材料）。它包含 35 幅部分注释的训练图像。在这里，我们的平均 IOU（"交集大于联合"）率达到 92%，明显优于第二名算法的 83%（见图 2）。第二组数据 "DIC-HeLa"³是用微分干涉对比（DIC）显微镜记录的平面玻璃上的 HeLa 细胞（见图 3、图 4c,d 和补充材料）。它包含 20 幅部分注释的训练图像。我们的平均 IOU 达到 77.5%，明显优于第二名的 46%。

5 结论

u-net 架构在不同的生物医学分割应用中都取得了非常出色的性能。由于采用了弹性变形的数据增强技术，u-net 架构在不同的生物医学分割应用中都取得了非常好的性能。

¹该算法的作者提交了 78 种不同的解决方案，以实现这一结果。

²数据集由 Sanjay Kumar 博士提供。加州大学伯克利分校生物工程系。加州伯克利（美国）

我们提供了基于 Caffe[6]的完整实现和训练后的网络。我们提供了基于 Caffe[6]的完整实现和训练后的网络。⁴我们确信，u-net 架构可以轻松应用于更多任务。

致谢

本研究得到了德国联邦政府和州政府卓越计划（EXC 294）以及联邦基金（Fkz 0316185B）的支持。

参考资料

1. Ciresan, D.C., Gambardella, L.M., Giusti, A., Schmidhuber, J.: Deep neural networks segment neuronal membranes in electron microscopy images. NIPS: 2852-2860 (2012)
2. Dosovitskiy, A., Springenberg, J.T., Riedmiller, M., Brox, T.: 利用卷积神经网络进行判别式无监督特征学习。In: NIPS (2014)
3. Girshick, R., Donahue, J., Darrell, T., Malik, J.: 用于对象检测和语义分割的丰富特征层次。In: IEEE 计算机视觉与模式识别（CVPR）会议论文集（2014 年）
4. Hariharan, B., Arbelaz, P., Girshick, R., Malik, J.: Hypercolumns for object segmentation and fine-grained localization (2014), arXiv:1411.5752 [cs.CV].
5. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification (2015), arXiv:1502.01852 [cs.CV].
6. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: 用于快速特征嵌入的卷积架构（2014 年），arXiv:1408.5093 [cs.CV].
7. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. NIPS: pp.
8. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: 反向传播应用于手写邮政编码识别。神经计算 1（4），541-551（1989 年）
9. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation (2014), arXiv:1411.4038 [cs.CV].
10. Maska, M., (...), de Solorzano, C.O.: 细胞追踪算法比较基准。生物信息学 30, 1609-1617 (2014)
11. Seyedhosseini, M., Sajjadi, M., Tasdizen, T.: 使用级联层次模型和逻辑互断正态网络进行图像分割。In: pp.
12. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition (2014), arXiv:1409.1556 [cs.CV]
13. WWW: 细胞追踪挑战赛网页, http://www.codesolorzano.com/celltrackingchallenge/Cell_Tracking_Challenge/Welcome.html
14. WWW: 电子邮件分段挑战赛网页, <http://brainiac2.mit.edu/>

⁴U-net 的实施、训练有素的网络和补充材料可查阅 <http://lmb.informatik.uni-freiburg.de/people/ronneber/u-net>