

U-Net: Convolutional Networks for Biomedical Image Segmentation 论文分析报告

一、研究背景

生物医学图像处理中，许多任务需要像素级定位（将类别标签分配给每个像素），但标注数据往往稀缺（难以获取数千张训练图像）。传统卷积网络多用于分类任务（输出单标签），难以直接解决定位问题。此前 Ciresan 等人采用滑动窗口策略（输入像素周围补丁预测类别），虽能定位，但存在两大缺陷：1) 速度慢（重叠补丁冗余计算）；2) 上下文与定位的权衡（大补丁需更多池化层降低定位精度，小补丁上下文信息少）。为解决这些问题，论文提出 U-Net 架构，基于全卷积网络（Fully Convolutional Network, FCN）改进，实现精准定位与上下文利用的平衡，且仅需少量标注数据。

2

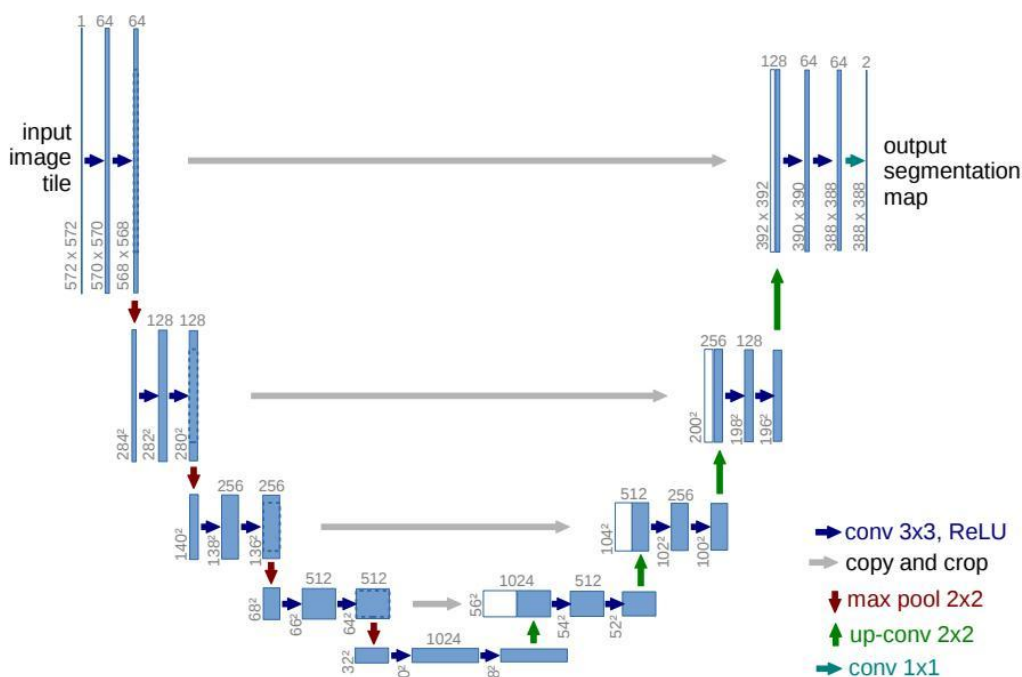


图 1: 生物医学图像像素级定位示例

二、核心架构：U 型网络设计

U-Net 架构以对称的 U 型结构为核心，分为收缩路径（左侧，提取上下文信息）和扩张路径（右侧，实现精准定位）两部分，无全连接层，仅通过卷积操作处理特征图：

1. 收缩路径：遵循传统卷积网络设计，通过重复的 3×3 卷积（无填充，Valid Convolution）+ ReLU 激活+ 2×2 最大池化（步长 2）逐步下采样。每下采样一次，特征通道数翻倍（如从 $64\rightarrow 128\rightarrow 256\rightarrow 512\rightarrow 1024$ ），以捕捉更广泛的上下文信息。
2. 扩张路径：通过上采样（Upsampling）+ 2×2 卷积（降低通道数）+ 跳跃连接（Skip Connection）+ 3×3 卷积+ReLU 逐步恢复分辨率。跳跃连接将收缩路径中对应层级的高分辨率特征图（裁剪后）与扩张路径的特征图拼接，融合局部细节（低层级）与全局上下文（高层级），实现精准定位。
3. 输出层：最后通过 1×1 卷积将特征图映射到目标类别数（如细胞/膜的二分类），生成像素级分割图。

此外，为处理大尺寸图像，论文提出重叠 Tile 策略：对输入图像分块，块间重叠区域通过镜像扩展补充上下文，确保分割图无缝衔接（避免 GPU 内存限制）。

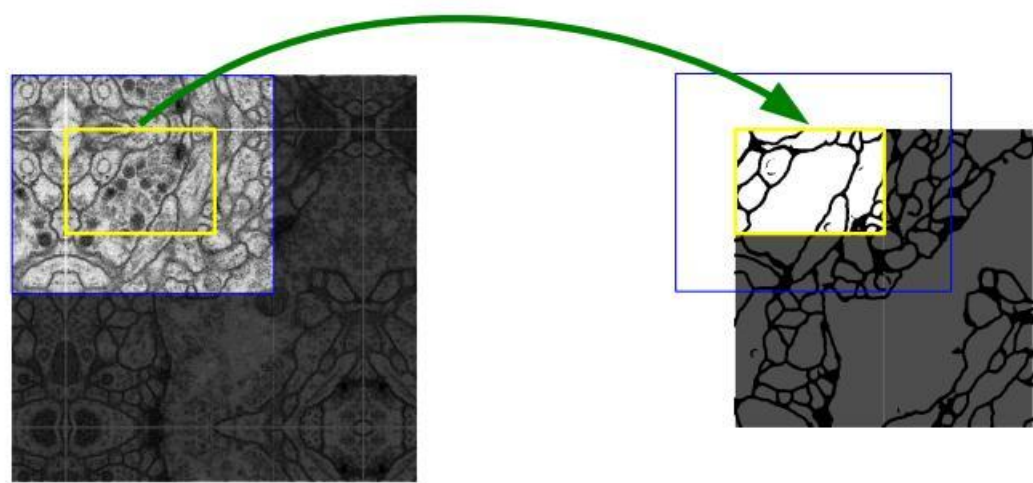


图 2: U-Net 架构示意图

三、训练策略：数据增强与损失设计

U-Net 的训练针对小样本场景优化，核心策略如下：

1. 损失函数：采用像素级交叉熵（结合 Softmax），并引入权重地图（Weight Map）解决类别不平衡与细胞边界分割问题。权重地图包含两部分：（1）平衡类别频率的权重（ w_c ）；（2）强化细胞边界的权重（基于像素到最近/次近细胞边界的距离，迫使网络学习细小区间），公式为：

$$w(\mathbf{x}) = w_c(\mathbf{x}) + w_0 \cdot \exp\left(-\frac{(d_1(\mathbf{x})+d_2(\mathbf{x}))^2}{2\sigma^2}\right) \quad (w_0 = 10, \sigma \approx 5 \text{ 像素})。$$

2. 数据增强：通过弹性变形（随机位移向量+双三次插值）、旋转、平移等增强训练数据，模拟生物组织的自然变形，减少对标注数据的依赖。此外，收缩路径末尾的 Dropout 层进一步隐含增强。

3. 权重初始化：采用高斯分布初始化权重（标准差 $\sqrt{2/N}$ ， N 为神经元输入节点数），确保各特征图方差近似为 1，避免梯度消失/爆炸。

4. 优化器：使用 Caffe 的 stochastic gradient descent (SGD)，动量 0.99（利用历史样本更新梯度），批次大小为 1（优先大输入块以利用 GPU 内存）。

四、实验结果

论文在三个生物医学分割任务上验证了 U-Net 的性能，均取得优异结果：

1. 电子显微镜 (EM) 神经元分割：使用果蝇幼虫腹神经索的 30 张训练图像（512×512 像素），U-Net（输入旋转 7 次平均）的翘曲误差（Warping Error）为 0.0003529（新纪录），Rand 误差为 0.0382，显著优于 Ciresan 等人的滑动窗口网络（翘曲误差 0.000420，Rand 误差 0.0504）。

2. 相衬显微镜 (PhC-U373) 细胞分割：35 张部分标注训练图像，U-Net 的交并比 (IOU) 达 92%，远超第二名的 83%。

3. 微分干涉对比 (DIC-HeLa) 细胞分割：20 张部分标注训练图像，U-Net 的 IOU 达 77.5%，显著优于第二名的 46%。

实验表明，U-Net 在小样本场景下仍能保持高精度，且推理速度快（512×512 图像在 GPU 上耗时<1 秒）。

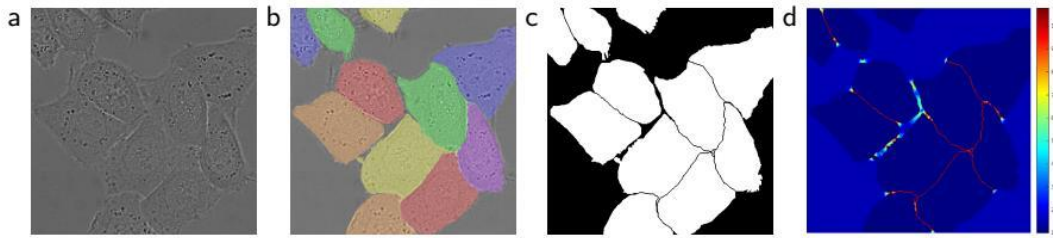


图 3：相衬显微镜细胞分割结果

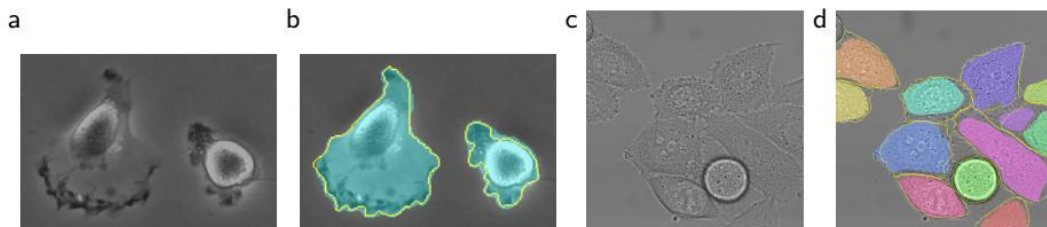


图 4：微分干涉对比细胞分割结果