

Deep learning

Deep learning

U-Net: Ronneberger2015MICCAI

面临问题:

解决方法:

实验:

3D U-Net: Çiçek2016MICCAI

面临问题:

解决方法:

实验:

V-Net: Milletari20163DV

面临问题:

解决方法:

实验:

MobileNetV1: Howard2017arxiv

面临问题:

解决方法:

实验:

MobileNetV2: Sandler2018CVPR

面临问题:

解决方法:

实验:

MobileNetV3: Howard2019arxiv

面临问题:

解决方法:

实验:

Template: NAME: {authoname}+{year}+{conference/journal}

面临问题:

解决方法:

实验:

U-Net: Ronneberger2015MICCAI

架构类文章 Ronneberger O., et al. U-Net: Convolutional Networks for Biomedical Image Segmentation, MICCAI 2015

面临问题:

the success of CNN was limited due to the size of the **available training sets** and the size of the **considered networks**.

现有CNN的性能受限于数据集的大小以及网络架构的参数数量

In many visual tasks, especially in biomedical image processing, the desired output should include localization, i.e., a class label is supposed to be assigned to each pixel.

对于除分类以外的视觉任务，如医学图像处理的定位任务，需要对于图片中的每个像素生成对应的标签。

a network in a sliding-window setup to predict the class label of each pixel by providing a local region (patch) around that pixel ... has two drawbacks.

已有在ISBI数据集上成功的方法仍然存在两个主要弊端:

1. 由于使用滑窗方法来预测像素的label, 因此运算过程会有很多重复部分, 数据冗余造成运行缓慢;
2. 在上下文使用和定位精度需要进行取舍

解决方法:

- 采用了全卷积网络 fully convolutional network
- 使用 upsampling 取代pooling操作
- 并且结合contracting path, 将下采样阶段的feature map与上采样后的进行维度上的拼接, 以提高定位精度
- 并且使用 Overlap-tile strategy 用镜像样本, 来用于后续分割
- 使用基于“弹性变形”的数据增广方法来生成训练样本 use excessive data augmentation by applying elastic deformations to the available training images
- 使用带有权重的损失函数, 来进行相似物体的分离

-
- 训练阶段
 - 使用尽可能大的输入, 因此将batch减少至一张图片 (favor large input tiles over a large batch size and hence reduce the batch to a single image), 并且使用momentum=0.99, 用过去样本决定未来更新方向
 - 能量方程 = a pixel-wise soft-max + the cross entropy loss function
 - 预先计算权重图 weight map = 类出现频率 + 当前像素与最近两个细胞的靠近程度决定 (靠近两细胞边界越近, 权重越大)
 - 使用与层数和卷积核大小相关的参变量生成用于初始化的高斯分布 drawing the initial weights from a Gaussian distribution
 - 数据增广
 - 使用随机向量在3x3的网格内进行平滑变形操作, 然后结合插值生成新样本 smooth deformations using random displacement vectors

实验:

We demonstrate the application of the u-net to three different segmentation tasks, EM segmentation challenge, ISBI cell tracking challenge 2014 and 2015.

主要在三个数据集上进行测试, 在速度和精度都取得很大进步。

3D U-Net: C, i, cek2016MICCAI

架构类文章 C, i, cek O., et al. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation, MICCAI 2016

面临的问题:

- Volumetric data is abundant in biomedical data analysis. Annotation of such data with segmentation labels causes difficulty
- 在生物医疗图像中, 体素信息非常丰富, 但是标注异常困难。因此使用完全标注的三维图像得到训练数据, 非常不可行。

解决方法:

- 使用少量分割数据, 来生成整体稠密的分割结果
- 使用完全新颖的3D架构, 3维信息作为输入, 然后中间设计的操作也都是三维 the network proposed in this paper takes 3D volumes as input and processes them with corresponding

3D operations, in particular, 3D convolutions, 3D max pooling, and 3D up-convolutional layers.

- 与U-Net不同的是, 在3dmaxpooling丢失信息前的3dconv先进行维度扩增(扩增回路则是使用反操作, 先压缩维度卷积再3dmaxpooling), 来实现avoid bottlenecks in the network architecture
- 使用BN层来加速训练时的收敛 se batch normalization [4] for faster convergence
- 同样使用特别设计的 weighted loss function
- 结合特殊数据增广形式 (apply a smooth dense deformation field on both data and ground truth labels), 然后相关数据使用的插值方法为: B-spline interpolation

实验:

最后实验部分, 分别进行了Semi-Automated Segmentation半标注分割, 通过进行少数的标注, 实现整个三维结构的分割; 同时也进行了Full-Automated Segmentation全自动分割, 通过再小数据集上训练, 以实现在大数据集上的泛化。

V-Net: Milletari20163DV

架构类文章 Milletari F., et al. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, 3DV 2016

面临的问题:

- 目前在卷积神经网络上主要的进展来自于对于2D图像的处理, 而对于三维分割任务却缺乏关注
- 对于MRI分割问题, 主要的挑战性因素来自于较大的外观变化 (due to large appearance variation across different scans), 如较大形变或是密度分布变化; 并且由于城乡会带来伪影和失真 (often affected by artefacts and distortions due to field inhomogeneity.)
- MRI问题中, 因为样本不均, 神经网络可能会到达局部最小值, 导致预测强烈偏向占据较多位置的背景。(It is not uncommon that the anatomy of interest occupies only a very small region of the scan. This often causes the learning process to get trapped in local minima of the loss function yielding a network whose predictions are strongly biased towards background.)

解决方法:

- 相比于二维问题, 而是直接采用三维卷积 instead of processing the input volumes in a 2D slice-by-slice fashion, we propose to directly use 3D convolutions.
 - 使用5x5x5卷积为主要卷积核并使用2x2x2卷积(结合stride2)为下采样模块, 然后使用PReLU作为激活函数
 - 使用了U-Net的shortcut path, 来增加后续反卷积过程的细节 (In this way we gather fine grained detail that would be otherwise lost in the compression path and we improve the quality of the final contour prediction.)
 - 在分割领域, 引入了Res模块通路
- 提出一个新的目标损失函数, 基于dice overlap coefficient的loss: **dice loss**

$$D = \frac{2 \sum_i^N p_i g_i}{\sum_i^N p_i^2 + \sum_i^N g_i^2}; \text{ 对于第} j \text{ 个体素进行求导, 可得: } \frac{\partial D}{\partial p_j} = 2 \left[\frac{g_j \left(\sum_i^N p_i^2 + \sum_i^N g_i^2 \right) - 2 p_j \left(\sum_i^N p_i g_i \right)}{\left(\sum_i^N p_i^2 + \sum_i^N g_i^2 \right)^2} \right], \text{ 由分}$$

式求导法则 $(u/v)' = (u'v - uv')/v^2$ 容易导出

实验:

最后实验部分, 分别进行了Semi-Automated Segmentation半标注分割, 通过进行少数的标注, 实现整个三维结构的分割; 同时也进行了Full-Automated Segmentation全自动分割, 通过再小数据集上训练, 以实现在大数据集上的泛化。

MobileNetV1: Howard2017arxiv

架构类文章 Howard A., et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications, arxiv 2017

面临的问题:

提出了一种适用于移动端地神经网络架构，并且引入两个超参数用于实现在时间延迟与准确率上的权衡

解决方法:

- 相比使用缩小/分解/压缩已有网络或者使用知识蒸馏（使用大型网络协助训练小网络）的方达，本文借鉴了在Inception中所使用的 depthwise separable convolutions，提出一种全新的架构进行小网络训练。
- 主要包括depth-wise卷积（对应PyTorch中 `nn.Conv2d` 中 `groups=input_channels` 的概念）以及1x1的point-wise卷积操作进行组合
- 参数量对比，由于主要使用1x1的卷积代替之前一步的3x3卷积，总体参数计算量约为原来的1/8~1/9:

$$\frac{D_K \cdot D_K \cdot M \cdot D_F \cdot D_F + M \cdot N \cdot D_F \cdot D_F}{D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F} = \frac{1}{N} + \frac{1}{D_K^2}$$

- 引入宽度乘子 α 来实现更窄的网络以进行维度的压缩
- 引入分辨率乘子 ρ 来实现更窄的网络以进行输入图像分辨率缩减

实验:

- 自身实验：与常规卷积，浅层网络，以及不同大小的宽度/分辨率因子进行充分实验，证明采用更瘦的网络可以比更浅的网络取得更好的效果（making MobileNets thinner is 3% better than making them shallower.）
- 数据集实验：在几个应用数据集/网络架构上进一步测试：Fine Grained Recognition——Stanford Dogs dataset, Large Scale Geolocalization——PlaNet, Object Detection——COCO object detection, Face Embeddings——FaceNet

MobileNetV2: Sandler2018CVPR

架构类文章 Sandler M., et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks, CVPR 2018

面临的问题:

- 进一步拓展MobileNetV1
- 实验发现，过多使用ReLU激活函数引入的非线性可能造成对已有信息的破坏。Experimental evidence suggests that using linear layers is crucial as it prevents non-linearities from destroying too much information.

解决方法:

- 提出新的模块：the inverted residual with linear bottleneck
- 在模块的输出层将传统的ReLU层取消掉，直接线性输出，理由是：ReLU变换后保留非0区域对应于一个线性变换，仅当输入低维时ReLU能保留所有完整信息。
- 利用了res模块，但是与以往不同中间降低维度操作的是，inverted residual中间层结合扩增因子（t, expansion ratio）去更好捕获物体信息，并且可以实现精度与计算量的权衡

实验:

除了进行ablation studies，还针对检测和分割数据集进行后续的修改，在ImageNet, classification, COCO object detection, VOC image segmentation数据集进行实验分析。

MobileNetV3: Howard2019arxiv

架构类文章 Howard A., et al. Searching for MobileNetV3, arxiv 2019

面临的问题：

- 希望进一步部署到移动平台上，以及尽可能更好的进行精度和时间延迟上的权衡
- 近年来NAS，神经网络架构搜索将能够进一步提高神经网络性能，如基于MnasNet是基于MobileNetV2进行的NAS

解决方法：

- 将DW, PW, inverted residual, SE等作为基本block进行NAS
- 重新设计网络头尾开销巨大部分 (We redesign the computationally-expensive layers at the beginning and the end of the network.)
 - 经过搜索与验证，简化网络最后部分，移除原有the projection and filtering layers in the previous bottleneck layer
 - 经过搜索，使用16个卷积核代替了原有32个 (We were able to reduce the number of filters to 16 while maintaining the same accuracy as 32 filters using either ReLU or swish.)
- 结合现有swish非线性激活函数，并进一步优化为hard swish函数，在整体架构的后半部分使用 (We also introduce a new nonlinearity, h-swish, a modified version of the recent swish nonlinearity, which is faster to compute and more quantization-friendly)
 - 原有swish使用sigmoid函数：

$$\text{swish}x = x \cdot \sigma(x)$$

- 减少卡销，结合ReLU以及固定常数：

$$\text{h-swish}[x] = x \frac{\text{ReLU6}(x + 3)}{6}$$

实验：

- 使用手机端进行测试
- 进行了分类/检测/分割任务的实验，并提出了新型分割decoders LR-ASPP

Template: NAME: {authorname}+{year}+{conference/journal}

架构/综述类文章 {引用格式}

面临的问题：

解决方法：

实验：