

Richer Convolutional Features for Edge Detection 论文阅读

 blog.csdn.net/qq_25624231/article/details/79037964

Richer Convolutional Features for Edge Detection是2017年cvpr中一篇边缘检测文章，准备以此为基础，1) 了解深度学习中边缘检测的发展；2) 如何使用深度神经网络进行边缘检测，3) 思考应用到医学图像中的可能性。

介绍

边缘检测是计算机视觉中的经典问题。传统方法通过局部的亮度、颜色、梯度和纹理或者其他人工设计特征进行边缘或者非边缘的分类。Generalized Boundaries from Multiple Image Interpretations是2014年TPAMI上的文章，可以作为其中的代表。但是边缘通常是具有丰富的语义信息的，仅通过局部线索难以得到令人满意的结果。尽管存在传统方法想要尽可能结合全局信息。

卷积神经网络近来也被用于边缘检测，如DeepEdge，N4N4-Fields，CSCNN，DeepContour和HED。

本文通过对卷积神经网络中不同卷积层的输出进行可视化，观察到中间层含有大量有用的细节，因此想要充分的利用CNN所提取的特征。最终，文章提出一个框架来利用所有卷积层的特征来完成边缘检测。

相关工作

边缘检测作为最基本的计算机视觉问题，有50年左右的历史。其方法可被简单的分成三类：先驱，基于人工特征和基于深度学习。

先驱们主要集中在利用图像梯度上，如著名的Sobel和Canny等。由于其高效，目前仍有应用，但是其准确率很难达到现代应用的要求。

其后人们开始利用灰度，梯度，纹理和学习范式来实现边缘检测。Konishi et al. 通过学习两类边缘的响应的概率分布提出了第一个基于数据驱动的边缘检测方法。Martin et al. 将亮度，颜色和纹理结合起来作为Pb特征，并训练相应的分类器。

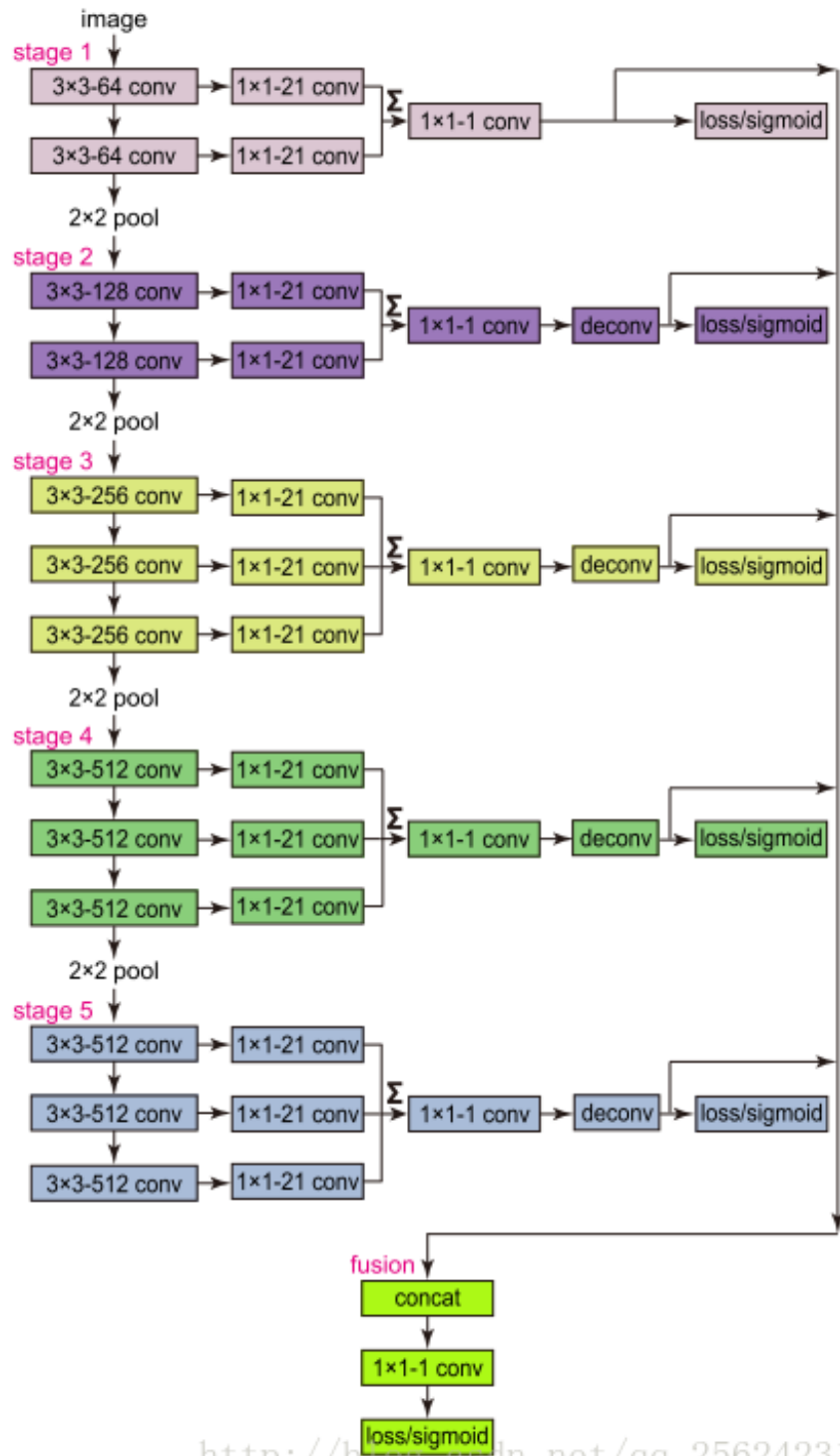
上述两种方法让我对概率分布到分类器这个过程有了一点理解。对于特征，指定概率分布是利用人的先验知识来限制其在特征空间的分布，因为不希望，或者说不是很了解，所以常常使用信息量最小的高斯分布。而分类器，比如SVM或者ANN，则是通过数据，隐式地学习概率分布，所以效果更好，同时对于数据的需求也更大。

Arbelaez et al 通过Normalized Cuts方法将局部信息和全局信息结合。Lim 使用了新的特征，Sketch tokens可以作为mid-level的特征信息。Dollar et al 使用随机森林来表示局部图像块中的结构。然而，上述方法都是基于人工设计的特征，因此缺乏足够的能力来表示high level信息。

深度学习方法中，Ganin et al 通过CNN和最近邻搜索，提出N4N4-Fields。Shen et al 将轮廓数据分成子类，并通过学习模型参数在子类上拟合。Hwang et al 将轮廓检测作为逐像素的分类问题，利用DenseNet 对每个像素点提取特征，然后用SVM进行分类。Xie et al 提出HED，实现图像到图像的训练和预测。其网络模型是以VGG16为基础，通过一个 $ks=1$ 卷积层，一个解卷积层和一个softmax层，实现多个side output，并将这些输出融合得到边缘检测结果。Li et al 提出一种无监督的边缘检测方法，目前表现不好。**综上**，这些基于CNN的方法往往只利用了每个卷积阶段(VGG16有5个阶段)最后一层的特征。因此本文提出一种全卷积网络来高效地利用每一个CNN层的特征。

Richer Convolutional Features

1. 网络结构



http://blog.csdn.net/qq_25624231

2. Annotator-robust损失函数

首先利用训练数据中不同标记者的结果生成gt的边缘概率图，根据其值的大小分成正类和负类以及模糊的结果。然后定义loss如下：

$$l(X_i; W) = \begin{cases} \alpha \cdot \log(1 - P(X_i; W)) & \text{if } y_i = 0 \\ 0 & \text{if } 0 < y_i \leq \eta \\ \beta \cdot \log P(X_i; W) & \text{otherwise,} \end{cases}$$

其中。超参数 λ 用于平衡正负类。

因此最终的loss为：

$$\alpha = \lambda \cdot \frac{|Y^+|}{|Y^+| + |Y^-|}$$
$$\beta = \frac{|Y^-|}{|Y^+| + |Y^-|}.$$

$$L(W) = \sum_{i=1}^{|I|} \left(\sum_{k=1}^K l(X_i^{(k)}; W) + l(X_i^{fuse}; W) \right),$$

。对于图像中每个点以及每个阶段的loss进行累加。

3. 多尺度边缘检测

为了提高精度，作者使用了1.5,1,0.5三个尺度的图像进行检测，然后通过平均得到最后的结果。

4. 与HED相比

1. HED只使用了每个阶段最后一个卷积层的信息；
2. 本文提出了新的损失函数，可以看做是对数据的一种清理；
3. 多尺度策略提高边缘检测效果。

实验

网络：VGG模型+pretrain_model；1*1的卷积层(stage1-5)用高斯分布

(N(0,0.012)N(0,0.012)) 初始化。1*1卷积层(fusion)使用0.2初始化。偏差全部0初始化。

训练：SGD：1e-6，每10k iter除以10，momentum=0.9, weight decay=0.0002. 训练40k次。

阈值选择：数据集优化(ODS)和图片优化(OIS)

Metrics: F-measure

1. BSDS500数据集

Method	ODS	OIS	FPS
Canny [6]	.611	.676	28
EGB [16]	.614	.658	10
MShift [10]	.598	.645	1/5
gPb-UCM [2]	.729	.755	1/240
Sketch Tokens [36]	.727	.746	1
MCG [3]	.744	.777	1/18
SE [14]	.743	.763	2.5
OEF [24]	.746	.770	2/3
DeepContour [47]	.757	.776	1/30 [†]
DeepEdge [4]	.753	.772	1/1000 [†]
HFL [5]	.767	.788	5/6 [†]
N ⁴ -Fields [19]	.753	.769	1/6 [†]
HED [58]	.788	.808	30[†]
RDS [37]	.792	.810	30[†]
CEDN [59]	.788	.804	10[†]
MIL+G-DSN+MS+NCuts [29]	.813	.831	1
RCF	.806	.823	30[†]
RCF-MS	.811	.830	8[†]

2. NYUD数据集

NYUD数据集中含有RGB图像和深度图像，模型针对这两个图像分别训练，最后联合起来预测。

3. Multicue数据集

Method	ODS	OIS	FPS
OEF [24]	.651	.667	1/2
gPb-UCM [2]	.631	.661	1/360
gPb+NG [22]	.687	.716	1/375
SE [14]	.695	.708	5
SE+NG+ [23]	.706	.734	1/15
HED-HHA [58]	.681	.695	20[†]
HED-RGB [58]	.717	.732	20[†]
HED-RGB-HHA [58]	.741	.757	10 [†]
RCF-HHA	.705	.715	20[†]
RCF-RGB	.729	.742	20[†]
RCF-HHA-RGB	.757	.771	10 [†]

Method	ODS	OIS
Human-Boundary [41]	.760 (.017)	–
Multicue-Boundary [41]	.720 (.014)	–
HED-Boundary [58]	.814 (.011)	.822 (.008)
RCF-Boundary	.817 (.004)	.825 (.005)
RCF-MS-Boundary	.825 (.008)	.836 (.007)
Human-Edge [41]	.750 (.024)	–
Multicue-Edge [41]	.830 (.002)	–
HED-Edge [58]	.851 (.014)	.864 (.011)
RCF-Edge	.857 (.004)	.862 (.004)
RCF-MS-Edge	.860 (.005)	.864 (.004)

作者最后提到，1) 测试了混合RCF和HED的输出，结论是RCF利用所有的卷积层信息是好的。
2) 在1*1卷积后面加上Relu会损害训练效果，甚至导致不能收敛。

关于一个网络的线性和非线性程度，有什么理论论证吗？

项目源代码

总结一下：

- 1) 边缘检测的发展是从局部信息开始，到结合高层信息。从结果对比来看，一些传统方法的效果差距并不是很大。
- 2) 神经网络用于边缘检测，目前来看主要是对卷积层输出信息的再利用。关于无监督神经网络边缘提取的文章可以再阅读一下。
- 3) 医学图像相比传统图像，只有灰度信息，因此已有人工设计的特征不够丰富。但是鉴于训练数据的问题，所以接下来可以尝试先用传统方法进行边缘检测，然后考虑无监督以及有监督训练的可能性。