

周报

冯浩哲

2018.6.17

周报

本周工作

下周的工作安排

文献阅读报告

参考文献

本周工作

1. 肺结节分类网络经过一系列微调，在验证集上的精度已经达到了92%，但是离原论文给出的97%结果还差一些
2. 按郝老师要求，将毕业论文改成一篇小论文，拟发表在软件学报上
3. 阅读CVPR2018的文献

下周的工作安排

1. 将肺结节分类网络再进行一些调整，如扩大模型容量，增加参数，采用Joint Learning方法等。如果仍然复现不出来，我将邮件联系原文作者，询问原网络所有模糊的地方。
2. 继续将毕业论文改成小论文的任务

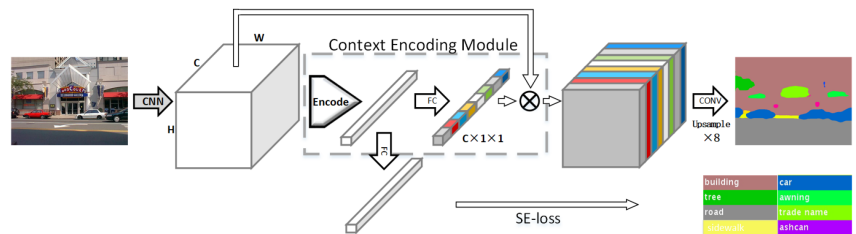
文献阅读报告

本周我看了4篇文献，都是CVPR2018的spotlight论文，我将简要介绍文献内容，并思考相应文献的思想与方法对我们现在的工作的启示。

1. Context Encoding for Semantic Segmentation[1]

- 文献内容
 - 针对的问题
如何在语义分割中利用语义信息
 - 拟解决的困难
用语义信息来提高神经网络语义分割效果
 - 解决困难的方法

作者采用了4阶段空洞卷积为基本流程，并在后面2阶段将输出的卷积特征与语义特征相联系，并提出了SE-Loss来利用文本信息生成损失。



如图所示，作者对我们卷积提出的特征将其编码为一个向量，并用SE-Loss计算其损失，用损失函数进行反馈调节，同时这个向量还可以作为每一个channel的权重来对输出结果进行处理。

- 达到的结果

作者在Cifar10数据集上获得了比state of art高3.45%的结果。

- 文献启示

该文献提出了如何利用语义文本信息来增强语义分割系统的操作，提出的方案简单易用，可以用在我们的系统中。

2. An Analysis of Scale Invariance in Object Detection – SNIP[2]

- 文献内容

- 针对的问题

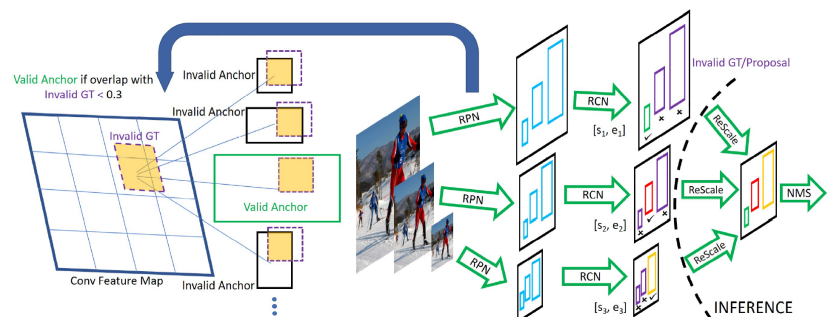
物体检测中的尺度不变量

- 拟解决的困难

如何设计一个网络，它可以更好利用多个尺度的图像输入进行物体检测

- 解决困难的方法

作者在原来的Feature Pyramid基础上，利用检测网络对于中等尺度的物体检测效果最好这一事实，对每一个尺度都只选取中等大小的RoI计算损失并进行反向传播，如图所示



这样作者认为可以减少冗余信息，更好利用多尺度特征

- 达到的成果

在COCO数据集上取得了state of art的结果

- 文献启示

这篇文献给出了如何利用多尺度数据进行物体检测的方法，很有借鉴意义。

3. Relational inductive biases, deep learning, and graph networks[3]

- 文献内容

- 针对的问题

- 深度学习解释性问题

- 拟解决的困难

- 深度学习中的因果关系问题

- 解决困难的方法

- 作者给出了一个基于图的网络结构 (Graph Network)，它的主要计算单元是 Graph Network Block，将图作为输入，对结构执行计算，并返回Graph作为输出。图的实体由节点，边的关系与全局属性进行定义。

- 文献启示

作者提出的图网络结构能将传统的贝叶斯因果网络和知识图谱，与深度强化学习融合，作者认为这是神经网络下一步的发展方向。

4. Tell Me Where to Look: Guided Attention Inference Network[4]

- 文献内容

- 针对的问题

如何用Image Level的标注来给出对图像的Attention

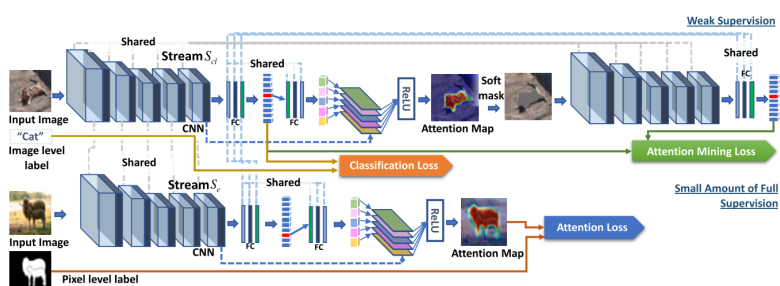
- 拟解决的困难

给出可以用来辅助进行语义分割的图像Attention

- 解决困难的方法

大体上的方法是，给定图像，分类网络输出对于c个类别上的分布，对于每一类可以利用Grad-CAM的方法得到一张Attention Map。根据Attention Map把重要的部分遮盖掉得到c张不同的图。然后同一个分类网络需要对每张处理过的图分类，使得对应的正确类别上的分数尽可能的低。

作者设计了如下给出Attention的流程：



首先，我们可以输入图像，然后用图像级标注进行分类。最后我们将用于分类的向量全连接生成对应的Attention Map，我们可以通过共享参数，用生成的Attention Map进行分割任务。

注意到图像右上角绿色部分作者提出了Attention Mining Loss，这个网络以Attention Map作为输入，目的是为了让网络能注意到图像上所有的区域，因此其设置为

$$L_{am} = \frac{1}{n} \sum_c s^c(I^{*c})$$

n代表n个类，而 $s^c(I^{*c})$ 代表第c类的得分。

- 文献启示

该文献提出的弱监督注意力机制融合了之前在这一方面的研究，非常值得借鉴，其中文献中提到的Grad-CAM,CAM方法都是神经网络可视化方面很好的思路。

参考文献

[1]Zhang H, Dana K, Shi J, et al. Context encoding for semantic segmentation[C]//The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018.

[2]Singh B, Davis L S. An Analysis of Scale Invariance in Object Detection-SNIP[J]. arXiv preprint arXiv:1711.08189, 2017.

[3]Battaglia P W, Hamrick J B, Bapst V, et al. Relational inductive biases, deep learning, and graph networks[J]. arXiv preprint arXiv:1806.01261, 2018.

[4]Li K, Wu Z, Peng K C, et al. Tell me where to look: Guided attention inference network[J]. arXiv preprint arXiv:1802.10171, 2018.