

Daily report 2-3 邓瑞峰

1. 配置好了语义分割网络，用 SUNRGBD 数据集训练一个基于 Deeplab-Resnet101 的语义 label 的分割网络。这个语义分割网络的输入是 rgb image, depth map, normal map, 按照我们做室内布局估计的网络那样进行融合。输出是包含 38 类 label 的语义分割结果。

训练集包含 10335 组{rgb,depth,normals}, 尝试了几个初始学习率之后, 采用 $2.5e-4$ 。为了尽快得到用于初始化的模型, 测试时网络每隔 10000 次迭代, 仅在训练集的前 1000 张做测试。

2. LSUN 第二个版本的结果提交了接近两天, 对方还没回复, 又发了封邮件催。

TODO

1. 用上面训练得到的模型作为初始化, 训我们的室内布局估计网络, 如果效果有提升, 接着做后处理得到 LSUN_v3 的结果;
2. 再看看论文, 规范一下 reference, 考虑后处理怎么写。

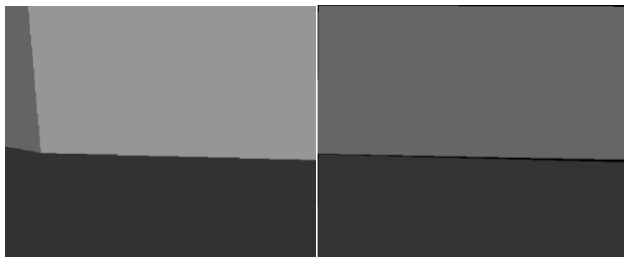
Daily report 2-2 邓瑞峰

1. 主要在处理 SUNRGBD 的数据, 因为希望得到 external data 的信息, 所以干脆把原训练集和测试集融合起来作为新的训练集, 使得训练数据足够多, 共 10335 张。

对新的训练集图片生成深度, 法向, 并且全部 resize 到 321x321. 由于使用的是另一个人分好的小块数据, 而且我又把训练集和测试集做了整合, 原来的 ground truth 和这些训练集图片的一一对应关系就被打乱了。按照之前划分的规则, 重新建立了一一对应。然后把 image, depth, normal, label 四组数据做成我需要的 lmdb 格式, 并检查。

做完数据后修改 DeepLab-resnet101 的网络结构, 去除 multi-scale 相关层, 还有分类之后的几层, 还没改完。

2. 按照昨天清华的作者回复给我的邮件, 我写了一个脚本计算忽略边界时的错误率, 在未剔除由于拓扑错误带来的额外误差时, 错误率为 7.40%.



result

ground truth

在调整以上这种结果时，收到了清华另一个作者的邮件（昨天给共一作的两个作者都发了邮件），出现了一个问题。今天收到的邮件说他们在评估 Hedau 数据集用的是 LSUN toolkit，然后说“I guess it would automatically ignore the semantic boundaries in all cases”。

但实际上，LSUN toolkit 并不会忽略边界，会把边界当做新的一类 label 计算错误率。于是我又发了一封邮件和他们确认，最后他们回我说他们的确用的是 LSUN toolkit。所以实际上是他们没有意识到这个问题，这也解释了为什么他们在 Hedau 数据集上的结果会比 LSUN 数据集差，而之前其他方法恰恰相反。所以我就没有继续改我的脚本了，打算就按 roomnet 里的那种方法计算错误率（6.63%）。

TODO

1. 完成语义分割网络配置，开始训练网络。

Daily report 2-1 邓瑞峰

1. 生成了第二个版本的 LSUN 测试集结果，已经提交；
2. 昨天发的邮件里，roomnet 的作者回复了我，他们处理边界的方式是“I perform boundary thinning to one pixel width and assign them to the closest segmentation label.”即把边界像素用最近的语义面像素 label 做填充，我之前其实也按照这种思路做过一组 ground truth，但是因为看了[10]即清华那篇论文之后认为他们算 error 时考虑了边界像素，所以放弃了这种 ground truth。我按照这种 ground truth 对我在 Hedau 数据集上的结果重新做评估，错误率是 6.63%（我们之前是 8.49）。



Method	ϵ_{pixel} (%)
Hedau et al. [1]	21.20
Mallya et al. [5]	12.83
Zhang et al. [15]	12.70
Dasgupta et al. [8]	9.73
Ren et al. [6]	8.67
Lee et al. [9]	8.36
Zhao et al. [10]	6.60

这样一来就非常接近清华的结果了。下午又给[10]的作者发了邮件，确认他们是如何处理 `ground truth` 的，结果晚上他们回复说，他们的方法是直接忽视边界像素。这个他们论文里的图没有描述，加上论文里说“the ground truth mask annotated by Hedau dataset is more strict”所以之前我产生了误解。打算明天按照清华的方法计算一下正确率，按照他们的方法就不能直接用 LSUN 的 toolkit 了。

3. 为了使用 `external data`，我打算选用 SUNRGBD 数据集先训一个语义分割网络（同[10]），但是数据下载下来之后发现他们的分割 `ground truth` 文件太大了，matlab 在 Load 的时候 RAM 不够会报错，吴建提供了一个 github 链接，是另一个人分类好的数据，把大文件拆分成了小文件。看了下数据集和使用方法。

TODO

1. 按清华的计算方法算一下在 Hedau 数据集上我们的结果；
2. 处理 SUNRGBD 数据，生成深度和法向，用一个 Deeplab-ResNet101 改一个语义分割的网络，尽快开始训练。

Daily report 1-31 邓瑞峰

1. 打算搞清楚在 Hedau 数据集上算 `pixel error` 的时候是否需要考虑边界这种 label。查了一下 Hedau 的算法 code，没有算 `error` 的代码。重新查了一下包括 roomnet 在内的其他几篇基于 FCN 方法的论文，都没有提到。Hedau 论文里的原句是 `pixel error: computed as the percentage of pixels on the box faces that disagree with ground truth`. 感觉还是比较模糊，论文里没提 `ground truth` 中有边界这件事。于是给 CFILE, roomnet, 还有 Delay..三篇文章的作者发了邮件问他们是怎么处理这个问题的。

个人觉得他们应该是不考虑边界的，否则应该像清华那样，在 Hedau 数据集的结果相比 LSUN 要差。

2. 修改论文剩余部分，主要是 **results** 部分补充说明。另外关于算法部分，后处理里定义的 **score function** 不知道是否保留，就没改。然后在 **Hedau** 数据集上的结果可能会变。

TODO

1. 生成第二个版本的 **LSUN** 测试集结果；
2. 尝试使用 **external data** 提高网络识别正确率，提高在两个数据集上的最终结果。对此要在室内语义分割的数据做预训练，训一个基于 **Deeplab-ResNet101** 的语义分割网络，需要做数据准备。
3. 视邮件回复情况修改 **Hedau** 数据集上的评估标准。

Daily report 1-30 邓瑞峰

1. 在大集群上配置完成 **multi-scale** 的 **Deeplab-ResNet101** 实验；
2. 为了对比 **external data** 的影响，配置了一组不用初始化，基于 **Deeplab-ResNet101** 的实验，但是网络没有收敛趋势，考虑放弃这组实验，直接按照清华和 **roomnet** 里的做法，先用室内语义分割的数据做预训练。
3. 在后处理的语义分割结果上提取关键点坐标，程序还在跑。
4. 修改论文：

先讲 3.3，加结果分析 (f)(g)多加注释 换了一下(g)的图 **done**

改表格， ϵ_{corner} 去掉年份 字体按原来字体 **done**

Check **roomnet** 在 **HEDAU** 数据集结果 **todo**

Tabel3 多加一个 **resnet** 的结果 我们的网络结构名字要改 **done**

训练部分讲 **pretrain** 的几句话，融合 **done**

Joint training 不讲了 **done**

后处理部分的 **score function** 该怎么讲 **todo**

两个数据集上定量的结果分析 **todo**

TODO

1. 修改论文剩余部分;
2. 生成第二个版本的 LSUN 测试集结果;
3. 尝试使用 **external data** 提高网络识别正确率, 提高在两个数据集上的最终结果。对此要在室内语义分割的数据做预训练, 训一个基于 Deeplab-ResNet101 的语义分割网络, 需要做数据准备。

Daily report 1-29 邓瑞峰

1. 对昨天说的网络输出结果进行后处理, 用两台电脑分别跑, 预计明天能跑完;
2. 修改论文, 把 introduction 和 relatedwork 融合, 去掉了 fig1, 修改了 fig2, 缩减成只有一个例子, 已上传到 github。
3. 按照 DeepLab-ResNet101 的原版网络结构, 即包含 multi-scale 相关的层, 配置实验。用来验证论文里说的, 单一尺度有利于提取全局信息, 且时间显存消耗低。修改网络配置文件时出了一点问题, 还在解决。

TODO

1. 修改论文的 results 部分。
2. 把 1 月 28 日统计的, 效果较差的图, 和目前后处理之后的结果对比, 找到较好的结果替换原来的图, 并生成相应的关键点坐标。然后生成第二个版本的 LSUN 测试结果。
3. 完成 multi-scale 的 DeepLab-ResNet101 实验配置;
4. 考虑做一组对比实验用来说明 **external data** 的作用。

Daily report 1-28 邓瑞峰

1. 之前尝试了基于 vgg16 的两种网络结构, 以 **rgbd** 为输入和 **rgbn** (法向) 为输入进行训练, 实际情况没有 **rgbdn** 效果好。其中加深度会使得结果与 **rgbdn** 非常接近 (百分之 0.05)。今天设置了下基于 **deplab-resnet101** 的结构, 挂了几组实验。
2. LSUN 反馈的结果只有总的 **pixel error** 和 **corner error**, 不知道每张的 **error**。今天对照原图, MC-FCN 输出结果, 和最终结果, 统计了一下测试集 1000 张中效果较差且可以通过对比进行提高的图。
3. 用扩大迭代次数、进一步 finetune 过的模型在测试集上跑了一组结果, 虽然

pixel-accuracy 还是 87%，但是对个别的图效果会有改善，这些个别的图就是 2 中的统计结果。

TODO:

1. 修改论文：
 - introduction & relatedwork 融合，问题定义的部分可压缩，fig1 可不要，fig2 例子可以少一点。
 - results 加数据结果分析（和其他方法对比）；加入 roomnet 的结果；说明我们对比清华方法主要优点，或者他们的缺点，external data 会带来哪些问题。
2. 在目前 LSUN 数据集的测试结果上尝试提高正确率，对上面 3 中得到的网络输出结果进行后处理，后期对统计的图进行调整。
3. 在 DeepLab-ResNet101 的原始版本进行实验，即包含多尺度相关的层。之前认为布局属于全局信息，去掉了 multi-scale 相关的层（也是为了减少显存占用，训练更快）。准备补一组加 multi-scale 的实验。
4. 视以上实验的结果，考虑是否需要变更算法。