

论文整理

2018年8月12日 20:52

概要	题目	要点	结论
车辆搜索 代码提及 论文	LEARNING A REPRESSION NETWORK FOR PRECISE VEHICLE SEARCH	<div>1. 论文提出的架构和STAGE1的几乎相同</div> <div>2. 核心思想: 把网络分成两个流, 分别学习粗粒度的attribute和细粒度的detail这两种表示</div> <div>3. 在第一个FC层后分成两个流:<div>a. 基于label的分类 - softmax (ACS), 粗粒度</div><div>b. 基于detail的相似性学习 - triplet loss (SLS), 细粒度</div></div> <div>4. 两个流通过Representation Layer连接, 论文给出的三种REP:<div>a. Product Repression Layer (PRL). 对两个向量进行element-wise乘积</div><div>b. Subtractive Repression Layer (SRL) element-wise减</div><div>c. Concatenated Repression Layer (CRL) 拼接向量</div></div> <div>5. 通过Representation Layer早期提取的ACS特征用于为随后的SLS学习过程提供某些反馈, 以便减少甚至消除有关嵌入到SLS中的那些粗粒度属性的信息, 并使其更多地关注那些潜在的细节。</div> <div>6. 提出Bucket Search, 减少retrieval time. (把有特定的特征组合的图片放当一个bucket, 搜索的时候只搜索对应的bucket)</div> <td><div>1. REP层可以抑制一定的从SLS-1转递到SLS-2的信息</div><div>2. MAP大于与之对比的四个模型: 其中两个是1 stream, 一个是2 steram+CCL, 一个是2 stream(无REP层)</div><div>3. REP+CRL取得最大MAP</div></td>	<div>1. REP层可以抑制一定的从SLS-1转递到SLS-2的信息</div> <div>2. MAP大于与之对比的四个模型: 其中两个是1 stream, 一个是2 steram+CCL, 一个是2 stream(无REP层)</div> <div>3. REP+CRL取得最大MAP</div>
车辆搜索	<div>1. RAM A REGION-AWARE DEEP MODEL FOR VEHICLE RE-IDENTIFICATION</div> <div>2. MobileNets Efficient Convolutional Neural Networks for Mobile Vision (补充第一篇)</div>	<div>1. 主要分为两个branch<div>1. global branch<div>a. Conv branch<div>i. attribute branch</div></div><div>b. BN branch (complementary global branch)</div></div><div>2. regional branch (learns regional features from three overlapping local regions)</div></div> <div>2. 六个branch分别计算loss, 最后加权平均得到这个网络的loss</div> <div>3. 第二篇论文补充内容:<div>1. weight prediction sub-network to predict feature weights (看不懂)</div><div>2. 新的distance loss<div>a. 替换了sample-wise triplet loss</div><div>b. 把相同汽车的样本图片作为一个集合, 训练时最大化图片集合之间的距离, 同时最小化图片集合内部样本之间的距离</div></div></div> <td>和另外几个网络比较, MAP有比较大的提高 (没有仔细看其它的网络)</td>	和另外几个网络比较, MAP有比较大的提高 (没有仔细看其它的网络)
distill	Distilling the Knowledge in a Neural Network	<div>a. soft targets : teacher模型输出</div> <div>b. soft target的优势: 使每个训练case包含更多的信息<ul style="list-style-type: none">▪ 例如MNIST中, teacher模型对输入数字"2"的输出中, 在"3", "7"分类会包含小分量, 包含了一些数字相似性的信息▪ 因此在small model中, 即使输入很少量的数据去训练, 也能获得一个比较好的结果(减少overfit)</div> <div>c. small model(student)的输入可以是unlabeled data</div>	<div>1. 有效压缩模型</div> <div>2. 数据集小的情况下仍然能有很好的效果.</div>
Multi-Level	A Deep Multi-Level Network for Saliency Prediction	<div>• 这篇论文提出一个新的结构, 把CNN不同层的特征提取并且组合起来.</div> <div>• a CNN extracts low, medium and high level features;<ul style="list-style-type: none">• 第三个pooling层的output(256 feature maps)• 最后一个pooling层的Output(512 feature maps)• 最后一个卷积层的output(512 feature maps)</div> <div>• 它们的spatial size是一致的, 然后被连接起来形成一个有1280 channels的张量</div>	<div>1. outperforms under all evaluation metrics on the SALICON dataset</div> <div>2. 提出了一个feature importance analysis, 分析结果表明提取的三个层都对最终的结果有影响.</div>
hashing	DEEP HASHING WITH MULTI-TASK LEARNING FOR LARGE-SCALE INSTANCE-LEVEL	<div>• test的时候, hash layer把大于0的变成1, 其它的变成0.</div> <div>• trian的时候应该就是hash layer就是用ReLU激活</div>	<div>1. 用ReLU激活比用Sigmoid好</div> <div>2. 用自己的车辆数据集pretrain效果比用ImageNet来pretrain好</div> <div>3. multi-task network里面ID+model+color的效果最好</div>

2. 六个branch分别计算loss, 最后加权平均得到这个网络的loss

3. 第二篇论文补充内容:

1. weight prediction sub-network to predict feature weights (看不懂)

2. 新的distance loss

a. 替换了sample-wise triplet loss

b. 把相同汽车的样本图片作为一个集合, 训练时最大化图片集合之间的距离, 同时最小化图片集合内部样本之间的距离