

毕业论文开题答辩

深度行人再识别学习

姓名学号: 王兴路 3140102282

指导老师: 李英明

年级专业: 2014 级信息工程

3 月 16 日下午 13:00 行政楼 108, 二次开题答辩, 每位同学答辩的时间不少于 15 分钟 (含提问), 学生陈述时间为 8-10 分。

各组答辩结束后, 由各答辩组组长录入开题答辩评语和成绩 (接学校最新要求开题成绩有变化), 成绩总分 30 分 (文献综述 10 分、开题报告 15 分、外文翻译 5 分), 开题不通过 (总分低于 18 分) 的学生将参加二次开题答辩。参加二次答辩的学生成绩不高于 24 分 (中等)。

大纲

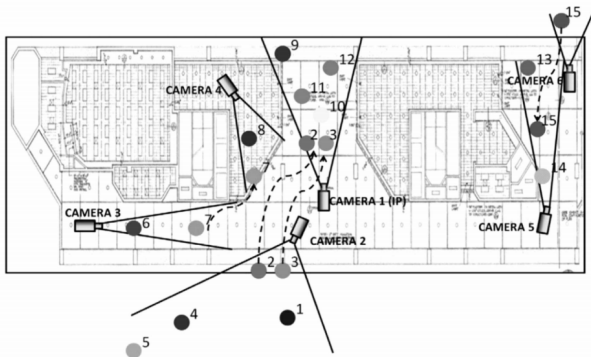
背景介绍与研究内容

技术路线与设计方案

进度安排与预期目标

背景介绍

- ▶ 行人再识别在智能视频监控、智能安防邻域应用广泛
- ▶ 摄像机网络已广泛布控于各种公共场合中，急需跨摄像头检索行人的技术
- ▶ 摄像头采集了**海量**数据，对海量数据进行压缩，建立索引并快速查找



毕设开题

└背景介绍与研究内容

└背景介绍

背景介绍

- 行人再识别在智能视频监控、智能安防领域应用广泛
- 摄像机网络已广泛布控于各种公共场合中，急需跨摄像头检索行人的技术
- 摄像头采集了海量数据，对海量数据进行压缩，建立索引并快速查找



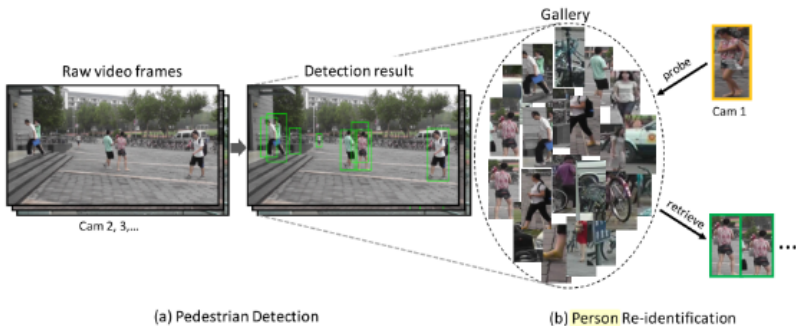
随着监控摄像机的大规模安装，摄像机网络已广泛布控于各种公共场合中。

利用摄像机网络，我们可以追踪同一行人的轨迹。但是面对多路监控和海量的视频数据，监控人员很容易产生疲惫和应接不暇的状况，我们需要智能视频分析技术弥补人类的不足。

行人再识别定义

前提假设

- ▶ 检测器已经成功检测
- ▶ 再识别**不包括**检测步骤



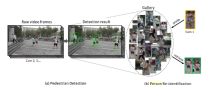
毕设开题

└ 背景介绍与研究内容

└ 行人再识别定义

前提假设

- 检测器已经成功检测
- 再识别**不包括**检测步骤



行人再识别定义

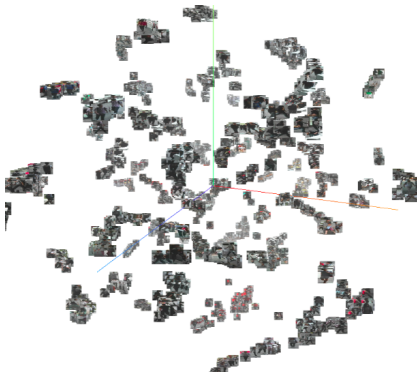
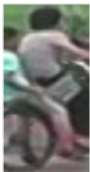
目前学者们普遍采用一个简化版本的定义. 这个版本的简化之处在于, 他的前提是检测器已经将从图像中检测出行人, 也就是说整个系统的输入是右图的检测后的图片

行人再识别定义

行人再识别 (Person Re-identification, *aka.* ReID)

输入某一行人的询问图片 (Probe), 在图像或者视频集合 (gallery) 中跨摄像头检索所有包含该行人的图片.

pid 425 cid 3



毕设开题

└ 背景介绍与研究内容

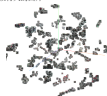
└ 行人再识别定义

目前从图像检索的角度来定义, 行人再识别是这样的一个过程... 比如左图是询问图片, 右图是测试阶段的图像集合

行人再识别定义

行人再识别 (Person Re-identification, aka. ReID)

输入某一行人的询问图片 (Probe), 在图像或者视频集合 (gallery) 中的摄像头检索所有包含该行人的图片.



行人再识别定义

行人再识别 (Person Re-identification, *aka.* ReID)

输入某一行人的询问图片 (Probe), 在图像或者视频集合 (gallery) 中跨摄像头检索所有包含该行人的图片。



毕设开题

└背景介绍与研究内容

└行人再识别定义

行人再识别定义

行人再识别 (Person Re-identification, aka. ReID)

输入某一行人的查询图片 (Probe), 在图像或者视频集合 (gallery) 中的摄像头检索所有包含该行人的图片。



右侧第一行是模型的预测结果, 第二行是预期的结果. 由于询问图片较难, 模型的 **top-5** 均预测错误. 除了图形检索, 行人再识别还有很多与之相关的领域, 细粒度物体识别 **fine-grained classification**/少示例大规模分类 **few shot extreme classification**.

细粒度物体识别与再识别的联系是, 想要区分不同的行人, 只能将注意力集中在某些有鉴别力的部位.

少实例大规模分类与再识别的联系是, 再识别有海量的 **id**, 每个 **id** 的训练样本很少, 图片质量通常不好

存在的挑战

Occlusion



Illumination



Pose



Misalign

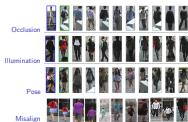


毕设开题

└ 背景介绍与研究内容

└ 存在的挑战

存在的挑战



the learning of comprehensive features of pedestrians for fine-grained tasks remains an open problem.

Desired features for open-set FR are expected to satisfy the criterion that the maximal intra-class distance is smaller than the minimal inter-class distance under a certain metric space. This 由于行人的非刚性运动、检测器的误差、摄像头的视角变化, 同一行人的不同图片往往存在严重的空间失配 (**Spatial Misalignment**); 行人没有可靠的生物特征, 只能从属性、语义层面的特征加以区分; 未标定的摄像机参数、巨大的时空跨度, 这些都进一步增加了再识别的难度; 同时现有的数据集规模相对较小, 不存在 ImageNet 或者 MegaFace 这样的大规模、可以泛化迁移 (**Transfer**) 到任意子领域 (**domain**) 的数据集。这导致数据集间存在较大偏差 (**domain bias/domain shift**), 从一个数据集到另一个数据集, 模型的性能通常都会下降。

大纲

背景介绍与研究内容

技术路线与设计方案

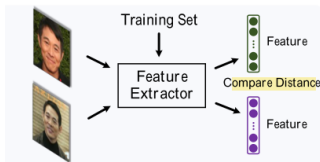
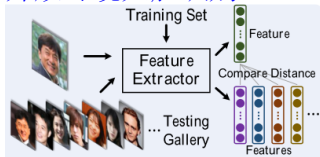
进度安排与预期目标

相关领域

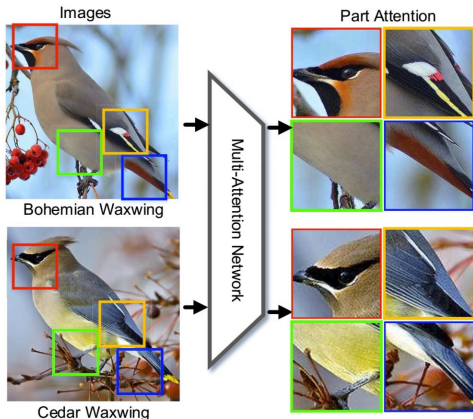
少样本大规模分类



开放环境人脸识别



细粒度物体识别



关键词： 图像检索、度量学习、表征学习、不平衡学习.....

相关工作

- ▶ **Identification** 转化为分类问题: 注意力机制提取特征 [2], [6], 全局 + 局部特征 [3]
- ▶ **Verification** 转化为两幅图片的验证问题/Verification: [4]
- ▶ **Embedding** 直接学习低维嵌入特征: 在三元损失的基础上强掉难样本的重要性 [1], 全局 + 局部特征 [5], 注意力机制 [7]

2018-03-15

毕设开题

└ 技术路线与设计方案

└ 相关工作

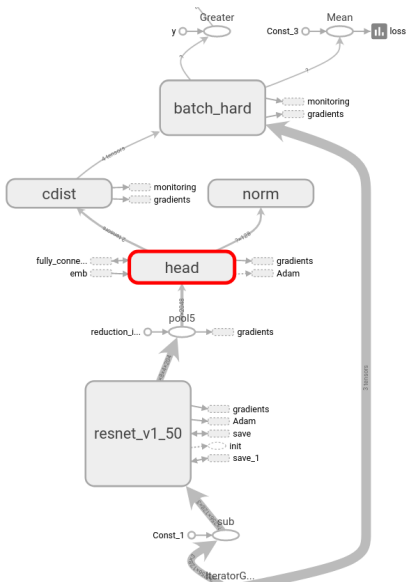
相关工作

- **Identification** 转化为分类问题: 注意力机制提取特征 [2], [6], 全局 + 局部特征 [3]
- **Verification** 转化为两幅图片的验证问题/Verification: [4]
- **Embedding** 直接学习低维嵌入特征: 在二元损失的基础上强调难样本的重要性 [1], 全局 + 局部特征 [5], 注意力机制 [7]

主流方法分为三类, 采用直接学习 **embedding** 的方法, 首先介绍采用的 **baseline** 方法

基本方法: 训练

- ▶ 特征提取
- ▶ 难样本挖掘
- ▶ 计算损失, 反向传播



2018-03-15

毕设开题

└ 技术路线与设计方案

└ 基本方法: 训练

resnet50

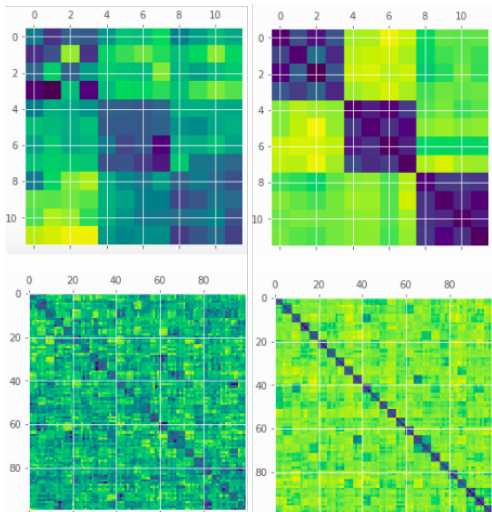
基本方法: 训练

- 特征提取
- 卷积本挖掘
- 计算损失, 反向传播



基本方法: 训练

- ▶ 特征提取
- ▶ 难样本挖掘
- ▶ 计算损失, 反向传播



2018-03-15

毕设开题

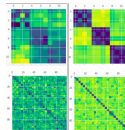
└ 技术路线与设计方案

└ 基本方法: 训练

分别是什么

基本方法: 训练

- ▶ 特征提取
- ▶ **神经网络**
- ▶ 计算损失, 反向传播



基本方法: 训练

- ▶ 特征提取
- ▶ 难样本挖掘
- ▶ 计算损失, 反向传播

$$L_{tri} = \frac{1}{PK} \sum_{a \in batch} \left(\max_{p \in A} d_{a,p} - \min_{n \in B} d_{a,n} + \alpha \right)_+ \quad (1)$$

毕设开题

└ 技术路线与设计方案

└ 基本方法: 训练

基本方法: 训练

- 特征提取
- 采样与挖掘
- 计算损失, 反向传播**

$$L_{\text{rel}} = \frac{1}{PK} \sum_{p \in \text{batch}} \left(\max_{p' \in \mathcal{A}} d_{LP} - \min_{q \in \mathcal{B}} d_{Lq} + \alpha \right)_+ \quad (1)$$

最后计算损失, 反向传播

基本方法: 测试

- ▶ 数据增强 (随机擦除/多次截取), 特征提取
- ▶ 计算欧式或余弦距离矩阵
- ▶ 重排序 [8], 预测最终排名

2018-03-15

毕设开题

└ 技术路线与设计方案

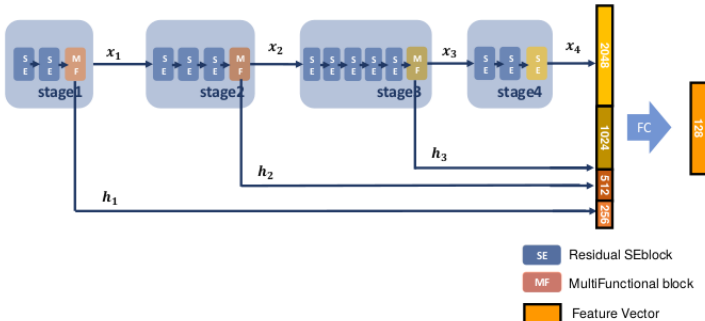
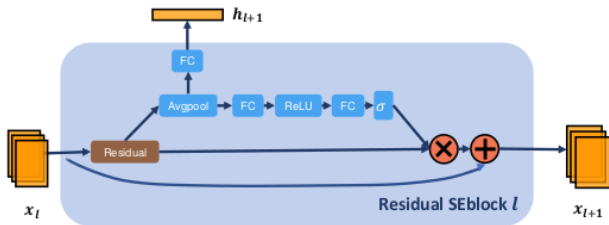
└ 基本方法: 测试

基本方法: 测试

- 数据增强 (**随机擦除** / **多次扰动**), 特征提取
- 计算欧氏或余弦距离矩阵
- **重排序** [8], 预测最终排名

随机擦除, random erasing, 重排序 rerank, 使用

多尺度注意力特征



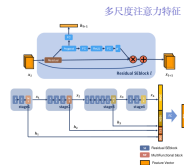
2018-03-15

毕设开题

└ 技术路线与设计方案

└ 多尺度注意力特征

也许是比较小的改进



根据关系图在全局挖掘难样本

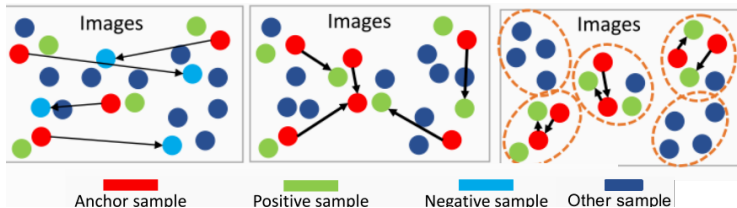


Figure: 不同的形成 batch 的方法:左图:一般的三元损失随机选取三元组,中图:[1] 在保证一个类至少两个样本的情况下随机选取,右图:我们希望在形成 batch 时难样本就不是随机选取的

2018-03-15

毕设开题

└ 技术路线与设计方案

└ 根据关系图在全局挖掘难样本

全局挖掘, 原来的图是讲什么, 我们的方法

根据关系图在全局挖掘难样本



Figure: 不同的形成 batch 的方法:左图一般的三元组随机选取三元组,中图[4]在保证一个类至少两个样本的情况下随机选取,右图我们希望在那成 batch 时难样本就不是随机选取的

大纲

背景介绍与研究内容

技术路线与设计方案

进度安排与预期目标

进度安排

2017/11/18–2017/12/18	调研文献, 制定规划
2017/12/18–2018/01/19	实现基准模型: TriHard [1]
2018/01/19–2018/03/30	注意力机制和难样本挖掘的初步探索
2018/03/30–2018/05/16	在难样本挖掘方面进一步探索
2018/05/16–2018/06/08	撰写毕业论文, 答辩

2018-03-15

毕设开题

└─ 进度安排与预期目标

└─ 进度安排

进度安排

2017/11/18-2017/12/18	调研文献, 制定规划
2017/12/18-2018/01/19	实现基准模型: TriHard [1]
2018/01/19-2018/03/30	注意力机制和难样本挖掘的初步探索
2018/03/30-2018/05/16	在难样本挖掘方面进一步探索
2018/05/16-2018/06/08	撰写毕业论文, 答辩

很难说方向一定可行

Table: 数据集与评估协议的统计信息

Dataset	CUHK03	Market1501	CUHK01	DukeMTMC	VIPeR
identities	1,467	1,501	971	1,812	632
images	13,164	32,668	3,884	36,411	1,264
views	2	6	2	8	2
train IDs	1,367	751	871/485	702	316
test IDs	100	751	100/486	1110	316

2018-03-15

毕设开题

└ 进度安排与预期目标

└ 现阶段实验

需要在各个数据集上验证

现阶段实验

Table: 数据集与评估协议的统计信息

Dataset	CUHK03	Market1501	CUHK01	DukeMTMC	VIPeR
identities	1,467	1,501	971	1,812	632
images	13,164	32,668	3,884	36,411	1,264
views	2	6	2	8	2
train IDs	1,367	751	871/485	702	316
test IDs	100	751	100/486	1110	316

name	top-1	top-1.rk	mAP
cuhk03label.res	86.28	93.88	82.49
cuhk03label.se.concat	89.33	96.34	86.42
cuhk03label.se.concat.dop	89.83	97.59	86.70
cuhk03label.res.dop	86.84	94.66	84.18
cuhk03label.se	88.28	95.51	84.74
cuhk03detect.res	82.79	91.11	80.08
cuhk03detect.se.concat	84.68	91.78	81.68
cuhk03detect.se.concat.dop	85.65	92.15	82.91
cuhk03detect.se	85.50	92.84	82.78
cuhk03detect.se.concat.35	85.03	92.09	82.65
cuhk03detect.se.concat.45	85.01	92.66	82.65
cuhk03detect.se.sum	84.75	91.22	82.06
cuhk03detect.se.sum.dop	85.93	91.53	83.21
cuhk03detect.res.new_avg	83.61	90.29	81.28
cuhk03detect.res.dop	83.70	91.85	80.78
market1501.res	85.61	86.12	60.52

毕设开题

└ 进度安排与预期目标





└ 现阶段实验

只进行了初步实验, 并不是都有效果的

现阶段实验

name	top-1	top-1 rk	mAP
cuhk03label.res	86.28	93.88	82.49
cuhk03label.se.concat	89.33	96.34	86.42
cuhk03label.se.concat.dop	89.83	97.59	86.70
cuhk03label.res.dop	86.84	94.66	84.18
cuhk03label.se	88.28	95.51	84.74
cuhk03detect.res	82.79	91.11	80.06
cuhk03detect.se.concat	84.68	91.78	81.68
cuhk03detect.se.concat.dop	85.65	92.15	82.91
cuhk03detect.se	85.50	92.84	82.78
cuhk03detect.se.concat.35	85.03	92.09	82.65
cuhk03detect.se.concat.45	85.01	92.66	82.65
cuhk03detect.se.sum	84.75	91.22	82.06
cuhk03detect.se.sum.dop	85.93	91.53	83.21
cuhk03detect.res.new_avg	83.61	90.29	81.28
cuhk03detect.res.dop	83.70	91.85	80.78

参考文献 I

-  Alexander Hermans, Lucas Beyer and Bastian Leibe. “In defense of the triplet loss for person re-identification”. In: *arXiv preprint arXiv:1703.07737* (2017).
-  Xihui Liu et al. “Hydraplus-net: Attentive deep features for pedestrian analysis”. In: *arXiv preprint arXiv:1709.09930* (2017).
-  Longhui Wei et al. “GLAD: Global-Local-Alignment Descriptor for Pedestrian Retrieval”. In: *CoRR* abs/1709.04329 (2017). arXiv: 1709.04329. URL: <http://arxiv.org/abs/1709.04329>.
-  Zhang Yaqing et al. “Semantics-aware deep correspondence structure learning for robust person re-identification”. In: *IJCAI Int. Jt. Conf. Artif. Intell.* 2016-Janua (2016), pp. 3545–3551. ISSN: 10450823.

参考文献 II



Xuan Zhang et al. “AlignedReID: Surpassing Human-Level Performance in Person Re-Identification”. In: *CoRR* abs/1711.08184 (2017). arXiv: 1711.08184. URL: <http://arxiv.org/abs/1711.08184>.



Haiyu Zhao et al. “Spindle net: Person re-identification with human body region guided feature decomposition and fusion”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 1077–1085.



Liming Zhao et al. “Deeply-Learned Part-Aligned Representations for Person Re-Identification”. In: *CoRR* abs/1707.07256 (2017). arXiv: 1707.07256. URL: <http://arxiv.org/abs/1707.07256>.

参考文献 III



Zhun Zhong et al. "Re-ranking Person Re-identification with k-reciprocal Encoding". In: *CoRR* abs/1701.08398 (2017).
arXiv: 1701.08398. URL:
<http://arxiv.org/abs/1701.08398>.

Thank you!