

# 华中科技大学

## 本科生毕业设计（论文）开题报告

题    目：

院    系\_\_\_\_\_

专业班级\_\_\_\_\_

姓    名\_\_\_\_\_

学    号\_\_\_\_\_

指导教师\_\_\_\_\_

年    月

## 开题报告填写要求

### 一、 开题报告主要内容：

1. 课题来源、目的、意义。
2. 国内外研究现况及发展趋势。
3. 预计达到的目标、关键理论和技术、主要研究内容、完成课题的方案及主要措施。
4. 课题研究进度安排。
5. 主要参考文献。

### 二、 报告内容用小四号宋体字编辑，采用A4号纸双面打印，封面与封底采用浅蓝色封面纸（卡纸）打印。要求内容明确，语句通顺。

### 三、 指导教师评语、教研室（系、所）或开题报告答辩小组审核意见用蓝、黑钢笔手写或小四号宋体字编辑，签名必须手写。

### 四、 理、工、医类要求字数在3000字左右，文、管类要求字数在2000字左右。

### 五、 开题报告应在第八学期第二周之前完成。

## 一、课题的来源、目的、意义

行人重识别（Person Re-identification）也称行人再识别，是利用计算机视觉技术判断图像或视频序列中是否存在特定的行人的技术，被广泛认为是一个图像检索的子问题，给定一个行人图像，跨摄像头检索该行人图像。该研究在视频监控和安防领域都有广阔的应用场景。

最近，基于深度学习方法的行人重识别方法迅速兴起，并取得较大成就。然而，基于深度学习方法的行人重识别在实际应用中由于域间差异的存在，在已标注源域训练好的深度学习模型在无标注目标域性能不佳，但标注目标域需要大量人力物力。一些不依靠标签或依靠已标注的少量数据集的行人重识别方法被提出，包括：（1）直接对无标注数据集使用聚类等无监督学习方法，虽然完全不需要任何标签，但性能表现不佳；（2）利用已标注数据的源域有监督地训练深度模型，然后通过各种无监督域适应方法将模型迁移到无标注的目标域，域适应方法形式多样；（3）通过改进有监督的行人重识别模型，使得模型更加泛化，提取到具有域不变特征。

近年来，基于视频行人重识别也逐渐兴起，相比基于图像的重识别，视频行人重识别能发掘更丰富的时空行人特征。但是视频行人重识别的数据量相比以前也更加庞大，标注数据的成本的也大大增加。

本课题旨在探索出一种能够从图像到视频的无监督域适应方法，使得在具有较小数据量的源域训练的行人重识别模型在目标域取得较好的行人重识别性能。

## 二、国内外研究现状

无监督域适应方法的关键在于如何消除源域和目标域之间差异。

一些研究<sup>[1][2][3][9]</sup>使用无结对的GAN生成将源域的训练样本风格迁移为目标域图像，组成新的训练集训练在源域上预训练好的模型，以此让模型学得具有域不变性质的特征，增强模型的泛化能力。这种风格迁移是可视化的图像到图像模式。这种方法的缺陷是风格迁移的图片往往保留在源域的身份信息进行训练，由于源域图像存在噪声，风格迁移后可能会引起噪声的放大使得模型性能下降。

另一种域适应方法是使用伪标签评估<sup>[4][5][6][7][8][10][11][12]</sup>，该类方法通过已经在源域训练好的模型提取出目标域的特征，然后采用无监督聚类的方法为每个目标域行人赋予一个伪身份标签，也有论文采用了其他方式赋予伪标签，如Ye

等人<sup>[12]</sup>使用动态图匹配机制来进行两个视频序列间无监督的度量学习，本质上也是一种聚类方法。随后使用伪身份标签对源域预训练模型进行再训练。由于域间差异的存在，初始的假标签的错误较多，因此这种方法往往采取渐进式的训练方法。

还有一些方法利用了一些辅助信息进行域适应，Li等人<sup>[13]</sup>利用姿势评估模型提取目标域的姿势信息，提出一种姿势解离迁移模型。Huang等人<sup>[5]</sup>利用分部语义分割模型，在对源域进行行人重识别训练同时对目标域进行分部语义训练。Wang等人<sup>[14]</sup>利用行人属性信息，使模型学得一个同时具有属性语义和身份分类性能的特征表示空间。

### 三、关键技术和理论

从前文所述的各种跨域行人重识别方法中，我们可以发现无监督域适应方法的目的是消除源域和目标域间的差异，关键是引入目标域的信息重新对模型进行再训练，但难点在于目标域由于没有身份标签，缺少有效的监督学习手段，无监督学习性能低下。为解决此问题，我们可以利用目标域不需要标注但可以利用的已知信息来对模型进行有监督学习，来使模型迁移到目标域。

首先，非配对的GAN方法可无需标注便可获得目标域风格，将风格转化图像与源域图像混合训练有助于模型习得目标域的风格信息。

第二，在缺少目标域身份标签的情况，为了使模型学得目标域的一些特征表达，我们将采取将为目标域设计人工识别任务，预测图像的旋转，将图像水平切割几块，预测图像块的位置。这些任务与在源域的行人重识别任务一起构成多任务训练结构。由于训练时模型同时受到在源域的行人重识别损失函数的约束，我们相信，这种特征表达对在目标域的行人重识别有益。并且，在随后的目标域伪标签再训练时，这些特征表达可以是使模型收敛的很好的初始点。利用辅助信息的三种方法<sup>[13][5][14]</sup>开发了目标域不需标注的其他信息，如姿势，属性和语义，这些信息的标签由已经训练好的模型识别得到，随后与在源域的行人重识别任务一起构成多任务有监督地训练深度神经网络。但是，我们应注意使用这种方法时，目标域的辅助信息是由训练好的各种识别模型得到，识别模型的训练集和目标域也存在域间差异，且识别错误会引入噪声造成行人重识别模型性能下降。为目标域设计人工识别任务则可避免域间差异和引入噪声。

最后，由于我们的任务是在视频目标域做行人重识别，视频行人重识别数据集往往默认一个视频序列的图像属于同一人，因为一个视频序列的图像可以由目标跟踪算法得到。因此我们在最初时便已经知道了一个视频序列的图像一定属于同一类，同时由源域风格转化到目标域的图像一定与目标域的任何图像都不属于同一类。因此，对于一个目标域图像，我们可将其与在同一视频序列的图像做正样本，由源域风格转化的图像做负样本，构成三元损失函数对初始模型进行有监督训练。另外，如果使用无监督聚类方法对目标域赋伪标签时，我们可将同一序列的图像为一类作为聚类算法的初始解，或者聚类的种子起点，可很大提高聚类赋予的伪标签的准确性。同理如果使用度量学习等方法，也可作为这些图像也可作为初始的训练集。这一步可约束模型提取目标域与行人重识别相关的特征表达。

#### 四、完成课题的方案及主要措施

根据以上的分析，为了充分把模型迁移到目标域，我们把整个域适应过程分为三个阶段：风格迁移图像生成阶段、多任务预训练阶段以及跨域再训练阶段。

1，2，3阶段的训练框架如图1所示，三个阶段介绍如下：

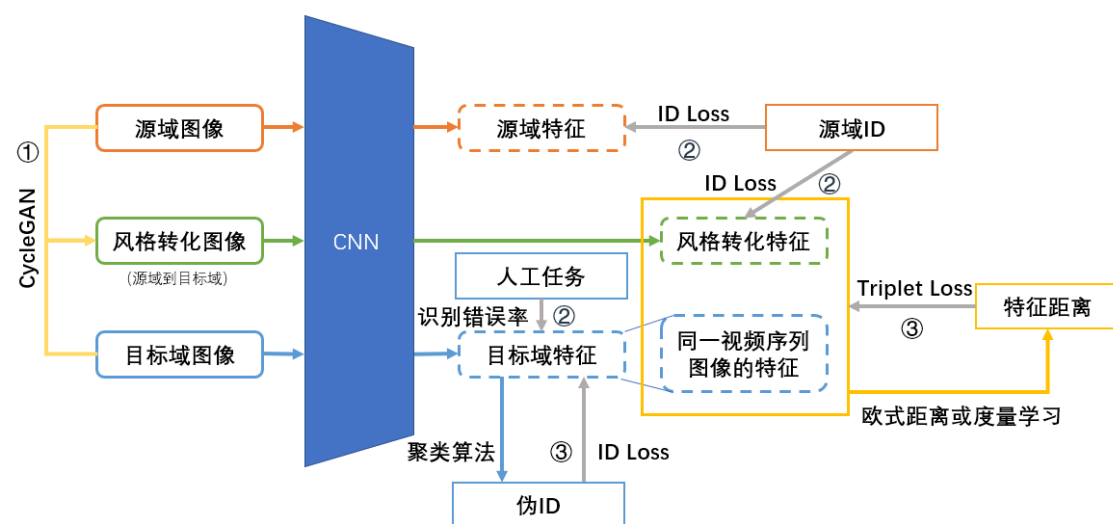


图1 三阶段训练框架，标号代表所处训练阶段

（1）**风格迁移图像生成阶段**。风格迁移图像生成的方法我采取主流的非配对的CycleGAN<sup>[16]</sup>模型。生成的样本将继承原图像的身份标签进行下面的训练。生成样本的个数将视源域和目标域的大小决定。

（2）**多任务预训练阶段**。我们决定采用主流的ResNet<sup>[17]</sup>作为特征提取器，

也作为行人重识别任务的基准模型。在具体进行训练时，考虑风格迁移同样会带来噪声影响，因此先完全使用源域图像进行训练几轮，再逐步加入风格迁移图像。对目标域的人工任务训练，我们将在分割图像块位置预测，旋转图像角度预测选择其中一种，如果实验效果不理想，我们将选择姿势评估<sup>[19]</sup>作为任务对目标域图像进行监督学习。

（3）**跨域再训练阶段**。目标域的聚类算法预期采用HDBSCAN<sup>[18]</sup>，这种聚类方法不需要设定聚类数量，只需设定聚类最小大小。考虑视频的数据量，进行聚类计算量巨大，因此有以下几种替代方案：（1）去除伪标签ID损失函数，仅采用三元损失函数；（2）考虑视频序列的特征融合问题，以一个视频序列为单位进行聚类；（3）直接使用源域的ID分类器为目标域产生伪标签；（4）每轮训练抽取目标域部分图像进行聚；（5）重新设计网络架构，使用孪生网络模型<sup>[20]</sup>，如图2所示。此种架构下，2阶段的种种举措可依旧实行。在3阶段时去除需要聚类的伪标签损失函数，目标域可根据两个图像的相似度概率为其打上伪标签，此外三元损失函数可用相似度损失函数代替。

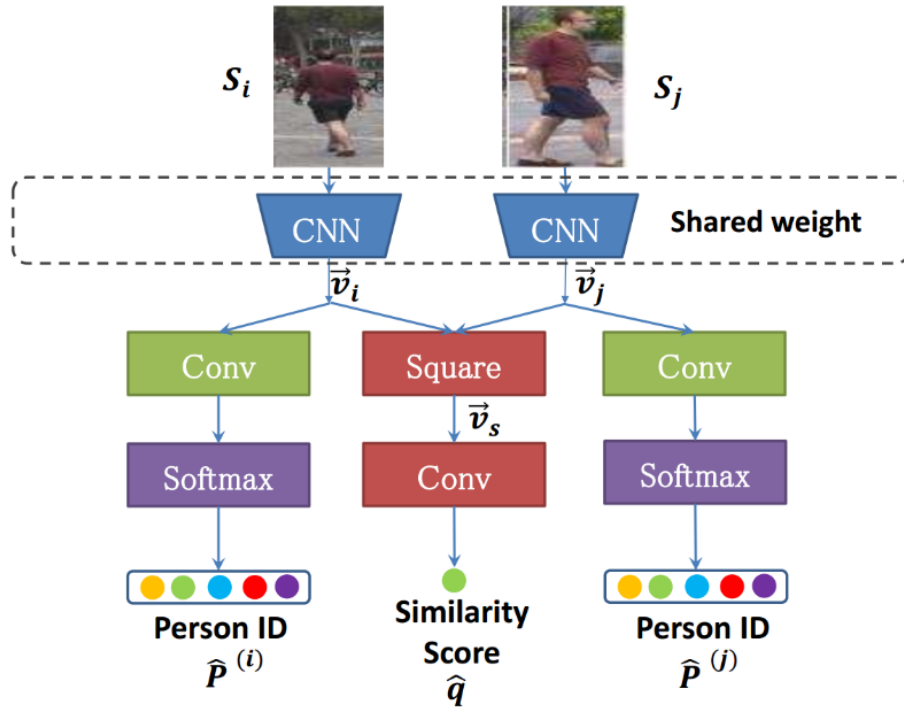


图2 孪生网络架构

由于三个阶段的训练框架过于复杂，在实际应用中可适度简化，如去除其中对目标域的人工任务部分。另外，尽管生成风格图像是必不可少的一部分，但仍

可免去在第二阶段的在源域训练时加入风格迁移图片的做法。

## 五、预期达到的目标

本课题将选用常用的行人重识别数据集DukeMTMC-reID<sup>[21]</sup>作为源域，视频行人重识别数据集MARS<sup>[22]</sup>作为目标域就所研究方案进行实验。目标是使用该方案训练出的模型在要比在源域训练好的模型直接在目标域进行行人重识别的效果要好。了解各阶段训练对模型训练的实际影响，并深入研究。

## 六、课题研究进度安排

学期	周次	工作任务
2019-2020	17周——19周	完成任务书
第一学期	寒假	外文翻译
	1周——2周	开题报告
2019-2020	3周——7周	代码编写和整理实验数据
第二学期	7周——11周	方案改进
	12周——14周	论文撰写，论文修改、论文查重

## 七、参考文献

- [1] Zhong Z, Zheng L, Zheng Z, et al. Camera Style Adaptation for Person Re-identification[C]. computer vision and pattern recognition, 2018: 5157-5166.
- [2] Wei L, Zhang S, Gao W, et al. Person Transfer GAN to Bridge Domain Gap for Person Re-identification[C]. computer vision and pattern recognition, 2018: 79-88.
- [3] Deng W, Zheng L, Ye Q, et al. Image-Image Domain Adaptation with Preserved Self-Similarity and Domain-Dissimilarity for Person Re-identification[C]. computer vision and pattern recognition, 2018: 994-1003.
- [4] Lv J, Chen W, Li Q, et al. Unsupervised Cross-Dataset Person Re-identification by Transfer Learning of Spatial-Temporal Patterns[C]. computer vision and pattern recognition, 2018: 7948-7956.
- [5] Huang H, Yang W, Chen X, et al. EANet: Enhancing Alignment for Cross-Domain Person Re-identification[J]. arXiv: Computer Vision and Pattern Recognition, 2018.
- [6] Fan H, Zheng L, Yan C, et al. Unsupervised Person Re-identification: Clustering

and Fine-tuning[J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2018, 14(4).

[7] Zhang X, Cao J, Shen C, et al. Self-training with progressive augmentation for unsupervised cross-domain person re-identification.[J]. arXiv: Computer Vision and Pattern Recognition, 2019.

[8] Yang F, Li K, Zhong Z, et al. Asymmetric Co-Teaching for Unsupervised Cross Domain Person Re-Identification.[J]. arXiv: Computer Vision and Pattern Recognition, 2019.

[9] Chen Y, Zhu X, Gong S, et al. Instance-Guided Context Rendering for Cross-Domain Person Re-Identification[C]. international conference on computer vision, 2019: 232-242.

[10] Lv J, Chen W, Li Q, et al. Unsupervised Cross-Dataset Person Re-identification by Transfer Learning of Spatial-Temporal Patterns[C]. computer vision and pattern recognition, 2018: 7948-7956

[11] Song L, Wang C, Zhang L, et al. Unsupervised Domain Adaptive Re-Identification: Theory and Practice.[J]. arXiv: Computer Vision and Pattern Recognition, 2018.

[12] Ye M, Li J, Ma A J, et al. Dynamic Graph Co-Matching for Unsupervised Video-Based Person Re-Identification[J]. IEEE Transactions on Image Processing, 2019, 28(6): 2976-2990.

[13] Li Y J, Lin C S, Lin Y B, et al. Cross-dataset person re-identification via unsupervised pose disentanglement and adaptation[C]//Proceedings of the IEEE International Conference on Computer Vision. 2019: 7919-7929.

[14] Wang J, Zhu X, Gong S, et al. Transferable Joint Attribute-Identity Deep Learning for Unsupervised Person Re-identification[C]. computer vision and pattern recognition, 2018: 2275-2284.

[15] Wang J, Zhu X, Gong S, et al. Transferable joint attribute-identity deep learning for unsupervised person re-identification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 2275-2284.

[16] Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. Unpaired image-to-image



translation using cycle-consistent adversarial networks. arXiv preprint arXiv:1703.10593, 2017a.

[17] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[18] Ricardo JGB Campello, Davoud Moulavi, and Jorg Sander. "Density-based clustering based on hierarchical density estimates. In Proc. Pacific-Asia. Conf. Knowledge discovery & data mining, pages 160–172, 2013.

[19] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[20] Z. Zheng, L. Zheng, and Y. Yang. A discriminatively learned cnn embedding for person re-identification. TOMM, 2017.

[21] Zhedong Zheng, Liang Zheng, and Yi Yang. Unlabeled samples generated by gan improve the person re-identification baseline in vitro. In Proc. IEEE Int. Conf. Comp. Vis., pages 3754–3762, 2017

[22] Zheng L, Bie Z, Sun Y, et al. Mars: A video benchmark for large-scale person re-identification[C]//European Conference on Computer Vision. Springer, Cham, 2016: 868-884.

华中科技大学本科生毕业设计（论文）开题报告评审表

姓名		学号		指导教师	
院（系）专业					
指导教师评语					
1. 学生前期表现情况。					
2. 是否具备开始设计（论文）条件？是否同意开始设计（论文）？					
3. 不足及建议。					
<div>(用蓝、黑钢笔手写或小4号宋体字编辑，签名必须手写。可加页，A4纸双面打印)</div>					
<div>指导教师（签名）：</div> <div>年 月</div> <div>日</div>					
教研室（系、所）或开题报告答辩小组审核意见					

教研室（系、所）或开题报告答辩小组负责人（签名）：

年

月 日

