

# 论文题目：Person Transfer GAN to Bridge Domain Gap for Person Re-Identification----from cvpr2018

## 一、创新点

- 1、提出一种针对于 ReID 的生成对抗网络 PTGAN，可以实现不同 ReID 数据集的行人图片迁移，在保证行人本体前景不变的情况下，将背景转换成期望的数据集 style。
- 2、提出新的数据集 MSMT17，该数据集包括了多个时间段多个场景，包括室内外场景（呈现复杂灯光变化，包含最多的标注）。

## 二、MSMT17 数据集

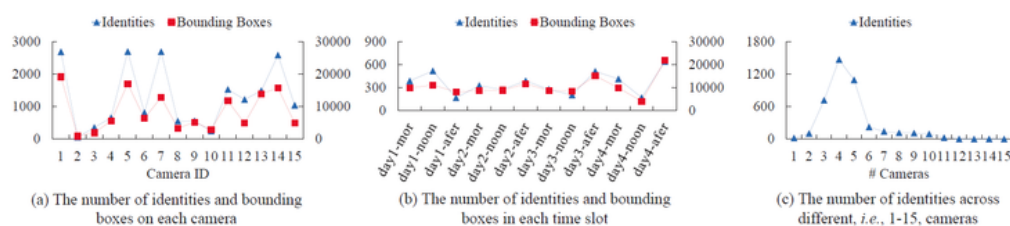
MSMT17 是一个大型的 ReID 数据集，该数据集有以下特性

- 数据采集时长约为 180 小时
- 总共有 15 个相机，其中 12 个室外相机，3 个室内相机
- 行人框由 Faster RCNN 机标完成
- 最后总共有 4101 个行人的 126441 个 bounding boxes

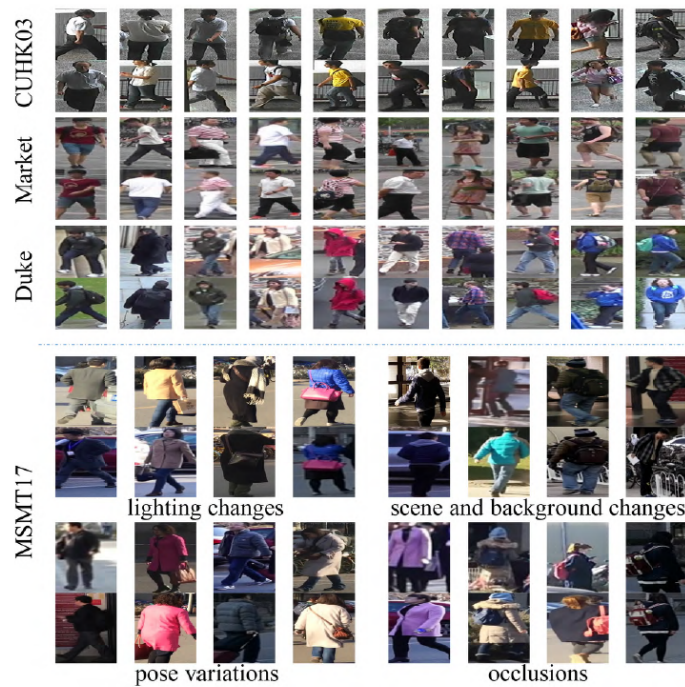
与现有的 ReID 数据集对比

Table 1: Comparison between *MSMT17* and other person ReID datasets.

Dataset	<i>MSMT17</i>	<i>Duke</i> [40]	<i>Market</i> [38]	<i>CUHK03</i> [20]	<i>CUHK01</i> [19]	<i>ViPeR</i> [8]	<i>PRID</i> [10]	<i>CAVIAR</i> [3]
BBoxes	<b>126,441</b>	36,411	32,668	28,192	3,884	1,264	1,134	610
Identities	<b>4,101</b>	1,812	1,501	1,467	971	632	934	72
Cameras	<b>15</b>	8	6	2	10	2	2	2
Detector	<b>Faster RCNN</b>	hand	DPM	DPM, hand	hand	hand	hand	hand
Scene	<b>outdoor, indoor</b>	outdoor	outdoor	indoor	indoor	outdoor	outdoor	indoor



数据集图片和其他数据集图片之间对比



### 三、PTGAN

Person Transfer GAN (PTGAN)是作者提出的一个针对于 ReID 问题的 GAN。这个 GAN 最大的特点就是在尽可能保证行人前景不变的前提下实现背景 domain 的迁移。

1、PTGAN 损失函数(loss)：

$$L_{PTGAN} = L_{style} + \lambda_1 L_{ID}$$

其中  $L_{style}$  是生成的风格损失(生成图像是否像新的数据风格),  $L_{ID}$  代表图像的 ID 损失(生成图像是否和原始图像是同一个人),  $\lambda_1$  平衡两个损失的权重。

2、作者使用 Cycle-GAN 去学习 style maping function.

PTGAN 的判别 loss：

$$L_{Style} = L_{GAN}(G, D_B, A, B) + L_{GAN}(\bar{G}, D_A, B, A) + \lambda_2 L_{Cyc}(G, \bar{G})$$

其中  $G$  代表从 A 到 B 的样式映射函数,  $\bar{G}$ 代表从 B 到 A 的样式映射函数,  $D_A$  和  $D_B$  分别是 AB 的样式鉴别器,  $L_{GAN}$  表示对抗性损失,  $L_{cyc}$  表示循环一致损失。

论文提出 LID 损失，用 PSPNet 提取的前景，这个前景就是一个 mask，最后 ID 损失为：

$$L_{ID} = \mathbb{E}_{a \sim p_{data}(a)} [|| (G(a) - a) \odot M(a) ||_2] + \mathbb{E}_{b \sim p_{data}(b)} [|| (\bar{G}(b) - b) \odot M(b) ||_2]$$

其中  $M(a), M(b)$  是两个分割出来的前景 mask，ID loss 将会约束行人前景在迁移过程中尽可能的保持不变。转换效果如下图所示：

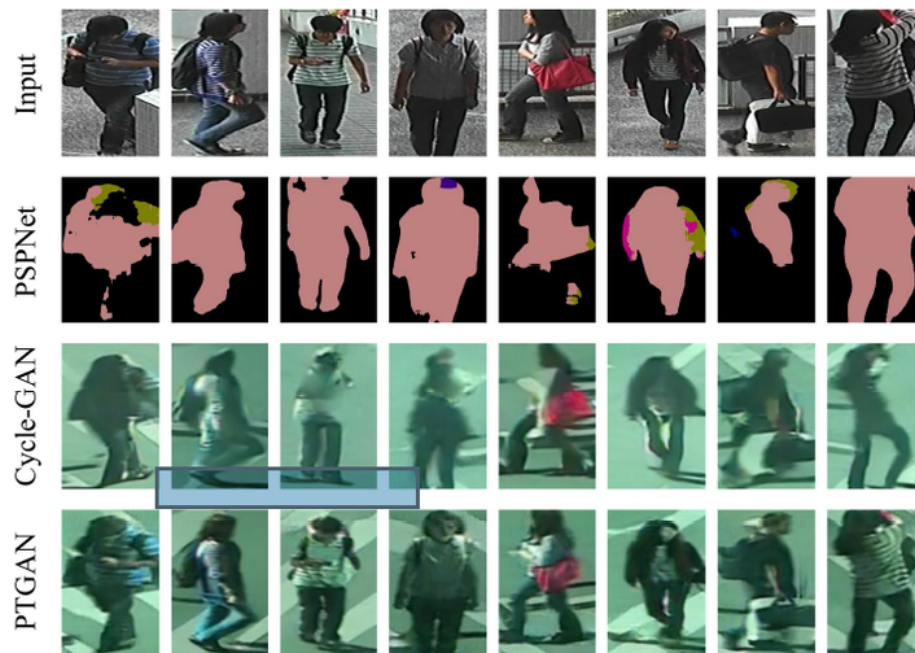


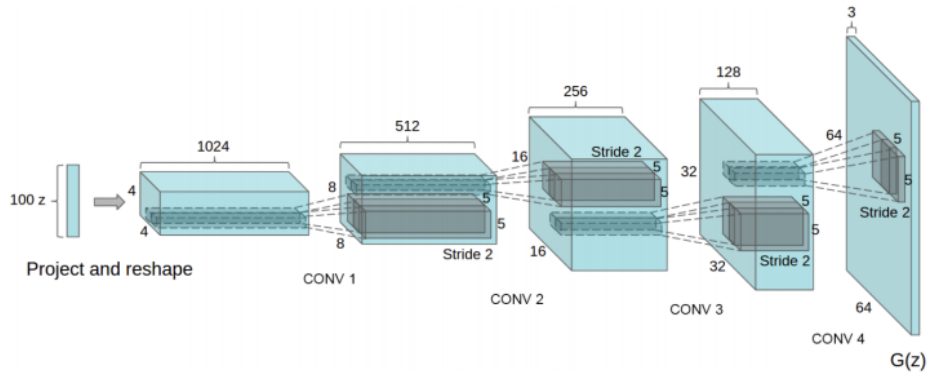
Figure 4: Comparison of the transferred images by PTGAN and Cycle-GAN from *CUHK03* to *PRID-cam1*. The second row shows the segmentation results by PSPNet. The pink regions are segmented as person body regions.

#### 四、实验

软件环境：

- tensorflow r1.1
- numpy 1.11.0
- scipy 0.17.0
- pillow 3.3.0

下载了官方的源代码，代码中主要由四个生成器(generatorA2B 和 generatorB2A，分别有 real 和 fake)，两个判别器(discriminatorA 和 discriminatorB)组成。网络框架跟原始的 GAN 网络相似，如下图所示



实验结果：

## PTGAN: from Duke to MSMT17



论文给出的实验结果：

Table 5: The performance of GoogLeNet for weakly supervised learning on *MSMT17*.

Training Set	R-1	R-10	mAP
<i>MSMT</i> (1%)	0.9	3.6	0.2
<i>MSMT</i> (2.5%)	2.0	7.4	0.5
<i>MSMT</i> (5%)	6.3	18.1	1.9
<i>MSMT</i> (10%)	11.5	26.9	3.7
<i>Duke</i> + <i>MSMT17</i> (10%)	16.1	33.1	5.5
<i>Duke</i> <sub><i>MS</i></sub> <sup>*</sup> + <i>MSMT17</i> (10%)	<b>18.0</b>	<b>36.4</b>	<b>6.2</b>
<i>Market</i> + <i>MSMT17</i> (10%)	12.6	28.5	4.4
<i>Market</i> <sub><i>MS</i></sub> <sup>*</sup> + <i>MSMT17</i> (10%)	<b>17.7</b>	<b>35.9</b>	<b>6.0</b>
<i>CUHK03</i> + <i>MSMT17</i> (10%)	11.9	28.3	4.1
<i>CUHK03</i> <sub><i>MS</i></sub> <sup>*</sup> + <i>MSMT17</i> (10%)	<b>14.3</b>	<b>31.7</b>	<b>4.6</b>

虽然论文没有用特别复杂的网络来训练，但是将另外一个数据集通过 PTGAN 迁移到 MSMT 上都能增加 MSMT17 数据集上 performance。并且从准确度上看 MSMT17 还是一个非常难的数据集。