

# Person Transfer GAN to Bridge Domain Gap for Person Re-Identification

2018/9/12 王诺

尽管行人再识别 (ReID) 的表现得到了显着提升, 但实际场景中的许多具有挑战性的问题尚未得到充分研究, 例如, 复杂的场景和灯光变化, 视点和姿势的变化, 以及摄像头网络中的大量人像数据。为了便于研究征服这些问题, 文章作者提供了一个名为 MSMT171 的新数据集, 其具有许多重要特征, 例如: 1) 原始视频由室内和室外场景中部署的 15 个摄像头网络拍摄, 2) 视频中包含了长时间和复杂的照明变化, 以及 3) 它包含当前最大数量的注释标识, 即 4,101 个标识和 126,441 个边界框。由于不同数据集之间通常存在“域间隙 (Domain Gap)”, 使得在不同数据集上进行训练和测试时会导致严重的性能下降。为了减轻注释新训练样本的昂贵成本, 本文作者提出了一个人像转移生成对抗网络 (PTGAN) 来弥补领域差距。综合实验表明, PTGAN 可以显着缩小域间隙。

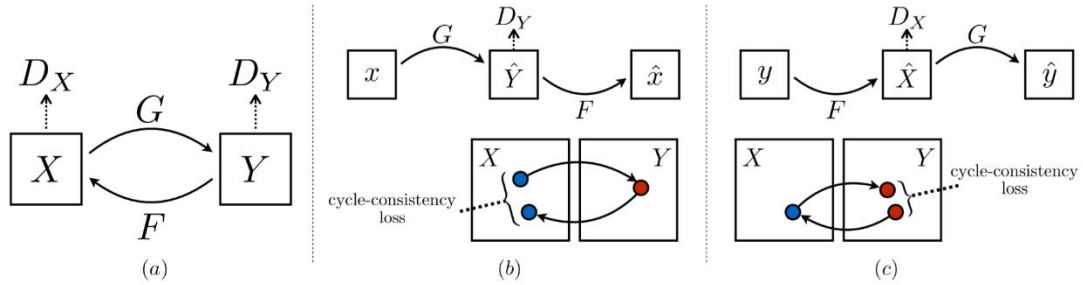


图 1

$$\mathcal{L}_{PTGAN} = \mathcal{L}_{Style} + \lambda_1 \mathcal{L}_{ID}, \quad (1)$$

$$\begin{aligned} \mathcal{L}_{Style} = & \mathcal{L}_{GAN}(G, D_B, A, B) \\ & + \mathcal{L}_{GAN}(\bar{G}, D_A, B, A) \\ & + \lambda_2 \mathcal{L}_{cyc}(G, \bar{G}), \end{aligned} \quad (2)$$

$$\begin{aligned} \mathcal{L}_{ID} = & \mathbb{E}_{a \sim p_{data}(a)} [\| (G(a) - a) \odot M(a) \|_2] \\ & + \mathbb{E}_{b \sim p_{data}(b)} [\| (\bar{G}(b) - b) \odot M(b) \|_2], \end{aligned} \quad (3)$$

图 2

图 1 为 CycleGAN 的结构图(参考文章 Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks)，本文提出的 PTGAN 即在其基础上，结合行人再识别这一应用，在损失函数方面进行了改进。行人再识别看重图像中的人物特征，为加强图像转换过程中对人像特征的保护，引入了 ID 损失函数（如图 2 中公式（3）所示）。其中  $G(a)$ 表示  $a$  经 PTGAN 转换后的图像， $M(a)$ 表示  $a$  的前景图像。

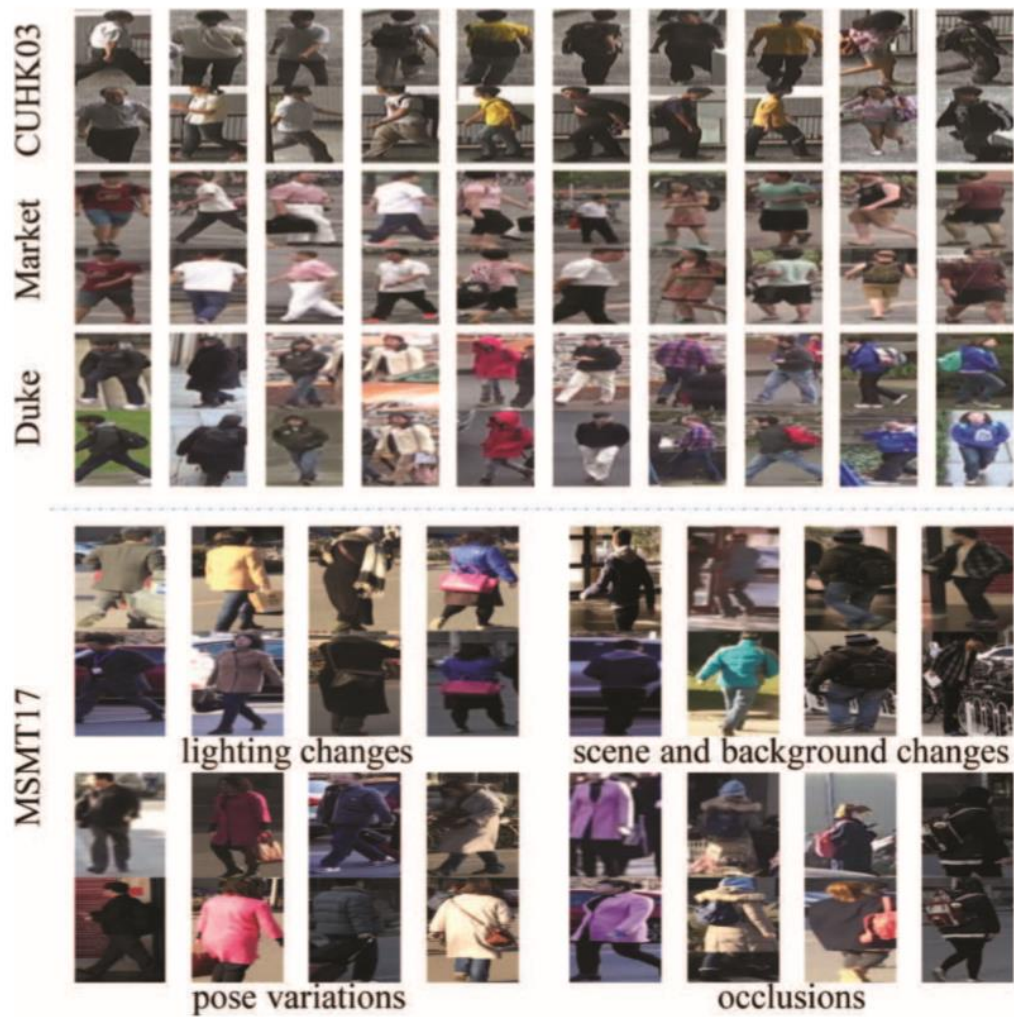


图 3

图 3 对本文中提供的新的行人再识别数据集 MSMT17 进行了部分展示，并与原有行人再识别数据集 CUHK03、Market、Duke 进行了比较。

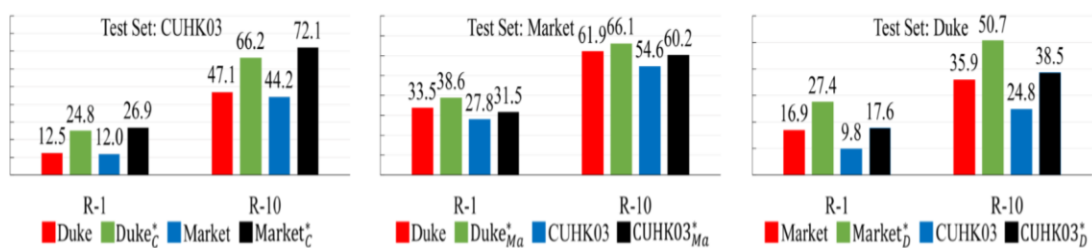


图 4

图 4 通过实验数据证明了 PTGAN 转换的有效性。其中  $Duke_c^*$  表示使用经 PTGAN 转化后的 Duke 数据集为训练集, CUHK03 为测试集进行的实验。由图中的数据可以看出, 从左到右, 每两个长方形, 都满足右侧高于左侧, 证明了本文所提出的方法的有效性。