**Inter-Task Association Critic for Cross-Resolution Person Re-Identifification**

**Abstract**

由不受约束的监视摄像机捕获的人像通常具有低分辨率（LR）。当与高分辨率（HR）画廊图像匹配时，这会导致分辨率不匹配问题，从而对人员重新识别（re-id）的性能产生负面影响。一种有效的方法是以**联合学习**的方式利用图像超分辨率**（SR）以及re-id**。但是，由于**在训练过程中梯度反向传播要困难得多**，因此该方案受到限制。在本文中，我们介绍了一种新颖的模型训练正则化方法，称为**任务间关联批评家（INTACT）**，以解决这一基本问题。具体而言，**INTACT发现了图像SR与reId之间的潜在关联知识**，并将其作为额外的学习约束，以增强SR-Reid模型在HR图像空间中的兼容性。这是通过参数化关联约束来实现的，该关联约束使它可以从训练数据中自动学习。广泛的实验证明了INTACT在使用五个标准人员re-id数据集的交叉分辨率re-id任务上相对于最新方法的优越性。

**Introduction**

人员重新识别（re-id）的目的是将不相干的监视相机视图所捕获的图像中的身份信息进行匹配[8]。 大多数现有方法都假定探测器图像和画廊图像具有相似且足够高的分辨率。 但是，由于摄像机和行人之间的距离不受限制，因此以各种分辨率捕获的人像只有十张。 该解决方案不匹配的问题给re-id带来了重大挑战。 与高分辨率（HR）图像相比，低分辨率（LR）图像包含的身份信息少得多，因此在分辨率之间直接匹配它们会导致性能大幅下降[11，18]。 例如，将标准人员身份模型[7]应用于交叉分辨率人员身份[18]时，可能会遭受高达19.2％的Rank-1下降。

为了解决分辨率不匹配问题[5、11、18、27]，已经开发了许多交叉分辨率重新识别方法。它们通常分为两类：（1）学习分辨率不变表示[5]和（2）激发图像超分辨率（SR）[11，27]。第一个类别旨在学习由LR和HR图像共享的特征表示空间，但由于LR图像中不存在，往往会丢失细粒度的判别性细节。第二类通常可以通过采用**级联SR**和**re-id**的多任务联合学习框架来解决此限制。但是，由于通过这种**级联的较重模型反向传播梯度的难度大大提高**，因此**该设计的模型训练效率很低**[2]。结果，**SR模型与Re-ID的兼容性较差**。最近，李等人。 [18]在一个统一的框架中结合了这两种方法来提高跨分辨率的re-id性能，但是仍然没有解决上述问题。

在这项工作中，我们通过引入一种新颖的正则化来解决此问题，该正则化称为任务间关联批评家（INTACT）。**INTACT是一种任务间关联机制**，可简化联合学习中的两个独特任务。 实际上，它由级联的多任务（SR＆re-id）网络和**关系评论者网络组成**。目的是增强SR和re-id之间的兼容性，即以**使得超分图像适合于re-id模型以在HR图像空间中执行身份匹配的方式来超分辨LR人物图像**。

这是通过INTACT分为两部分实现的：（

I）我们**使用专用网络对（未知）任务间关联约束进行参数化**，从而可以从HR训练数据中直接学习它。

（II）一旦学会了，以评论员的角色，**则将关联约束应用于监督SR模型**。 这意味着，SR模型训练会进一步受到约束以满足学习到的任务间关联。

我们做出三点贡献：

（1）我们提出了一种利用图像SR和re-id任务之间的关联来解决研究不足但仍很重要的交叉分辨率re-id问题的想法。

（2）我们制定了一种新颖的正则化方法，称为任务间关联批判（INTACT），以实现所提出的任务间关联。 INTACT建立在对关联进行参数化并且端到端可训练的基础上。

（3）广泛的实验显示了我们的INTACT在跨分辨率人员重新识别问题中基于五人重新识别基准的广泛的最新方法的性能优势。

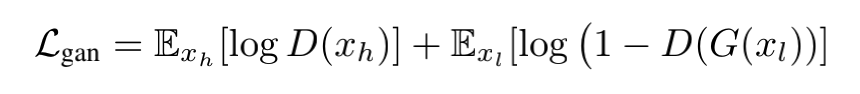
**Methodology**

**Joint Multi-Task Learning**

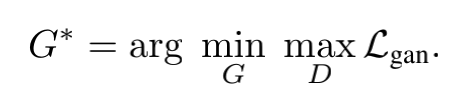
**Image super-resolution model ：**

为了训练SR模型，我们通常使用一组像素对齐的LR-HR图像对{{xl，xh}}。通常，我们通过对HR训练图像进行下采样来形成这样的对。在这项研究中，我们选择用于SR的生成对抗网络（GAN）模型[9]，因为它具有良好的性能[14]。

GAN解决了一个最小-最大优化问题，其中鉴别器D旨在从超分辨图像中区分出真实的HR，而生成器G则旨在生成可能使鉴别器蒙蔽的超分辨图像。 目标函数可以定义为：

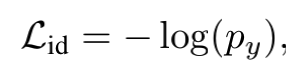


更具体地，生成器G试图相对于对抗鉴别器D使目标值最小化，而对手鉴别器D反而试图使该值最大化。最佳解决方案为：



**Person re-id model**

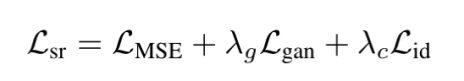
利用训练数据D，可以通过softmax交叉熵损失函数来训练任何现有的re-id模型（例如[33]）：



其中y是xl的真实身份，py是y类的预测概率。

**Joint multi-task learning**

为了建立联合的多任务学习管道，我们可以简单地通过使用SR的输出G（xl）作为re-id模型的输入来级联SR和re-id。 然后将总体目标函数表述为：



其中LMSE是逐像素内容损失，定义为LMSE = ||xh - G（xl）||22。 λg和λc是权重参数。

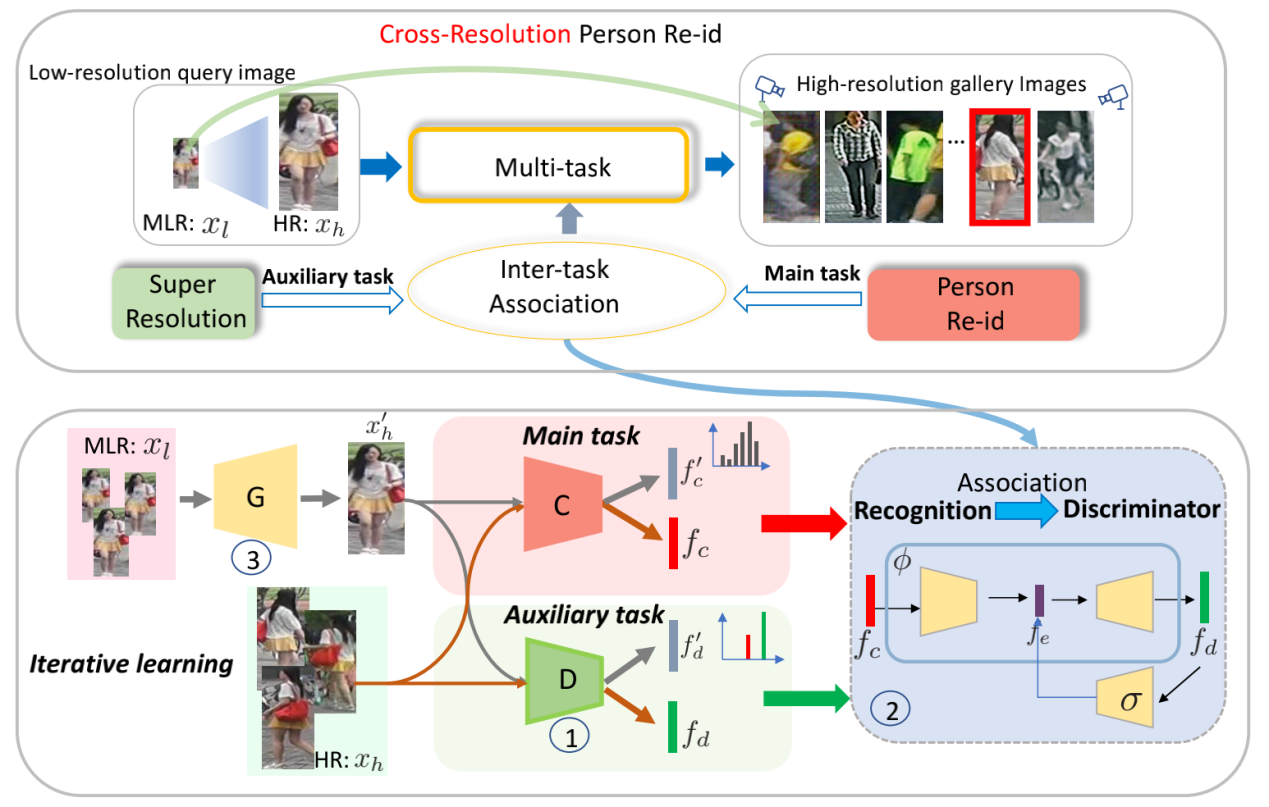
**Limitation**

尽管有一个很好的跨分辨率ReId解决方案，但该管道本质上受到限制。这是**由于通过两个级联模型反向传播梯度的难度大大提高**。结果，**SR模型训练没有被适当地约束以最大化re-id性能**，即，所得的SR模型与re-id模型不是很好地兼容。

**Inter-Task Association Critic**

为了解决这一基本局限性，我们引入了一种新的正则化，任务间关系评论家（INTACT）。 INTACT的**关键思想是利用SR和re-id任务之间的内在联系作为额外的优化约束**，以促进他们的联合学习和增强他们的兼容性。然而，**量化这样的任务间关联**并不是很简单的，这些任务间关联通常是复杂的并且是先验未知的。为了解决这个问题，我们**建议参数化该关联**。

具体来说，我们**利用专用网络来表示从主要任务（即Re-Id）到辅助任务（即SR）的关联**。这构成了INTACT的核心要素。在模型训练期间，INTACT由两部分组成。在第一部分中，它**使用原始的HR图像发现关联**。具体地，它**学习位于HR训练图像{xh}上的判别器和身份分类表示之间的关联网络**。在第二部分中，然后将所学习的关联作为正则项应用于SR模型训练中。具体地，鼓励从SR图像中提取的**判别器和分类表示**满足从真实HR数据中预先学习的关联约束。我们的INTACT方法的概述如图2所示。

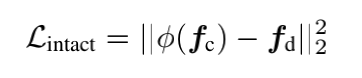


**Part I: Association Learning**

利用GAN模型，我们通过**判别器的特征激活**fd表示真实-错误的判断任务。对于身份分类任务，可以采用大量现成的re-id模型。 我们利用[33]中提出的一种最新方法来提取身份分类特征fc来表示身份。我们**独立地使用HR图像xh训练re-id模型**，以获得最佳身份表示。 它是经过一次性培训，冻结参数的，并充当以下模型优化的身份验证者。

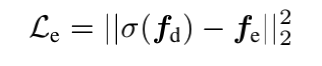
给定输入的LR图像xl，我们**希望生成器（SR模型）输出具有高度身份判别性的超分辨HR图像**G（xl）。为了实现这一点，我们建议**在**图像xh的真实-错误**鉴别表示表示fd和身份分类fc表示之间设计一个关联约束φ**。然后，考虑到**它们是SR图像G（xl）在训练期间需要接近的目标**，我们**用一个小型网络在HR训练图像xh上表示并学习φ**。

形式上，我们学习关联，这是从身份识别fc到判别器fd表示的转换。这是基于一个假设，即从HR训练图像中学习到的身份识别表示包含用于一般高分辨率分布的信息（真-假鉴别器会尝试学习）；尽管鉴别特征比身份特征的信息量相对较少，这是由于它们源自更简单的二进制分类任务。因此，学习这样的映射更加明智。特别地，我们将关系正则化推导为：

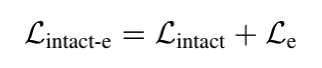


其目的是使用从**re-id模型提取的fc**和从鉴别器在HR图像xh上提取的fd来优化关联网络的参数φ。

为了促进学习φ，我们进一步施加了附加的桥梁约束来操纵优化方向。具体来说，我们将中间潜在特征空间fe与φ隔离开来，以便可以用目标fd的转换σ定义桥接操作：

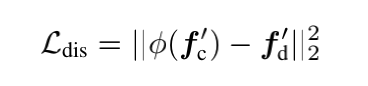


其中，fe是在φ的中间潜在空间中以fc为输入获得的。桥接模块σ与关联模块φ共同学习，组合为：



**Part II: Association Regularisation**

一旦如上所述学习了任务间关联网络φ，我们就将其视为评判者，**以正则化（规范？）**基于GAN的多任务学习网络中**SR模型（生成器）和鉴别器的学习**。我们通过相似地结合鉴别器和身份识别信息来蒸馏学习到的联系。特别地，该蒸馏损失具有等式5的形式。但应用于SR图像G（xl）：



其中f‘c和f’d是与上述xh类似的单个SR图像G（xl）的对应身份和判别器表示。

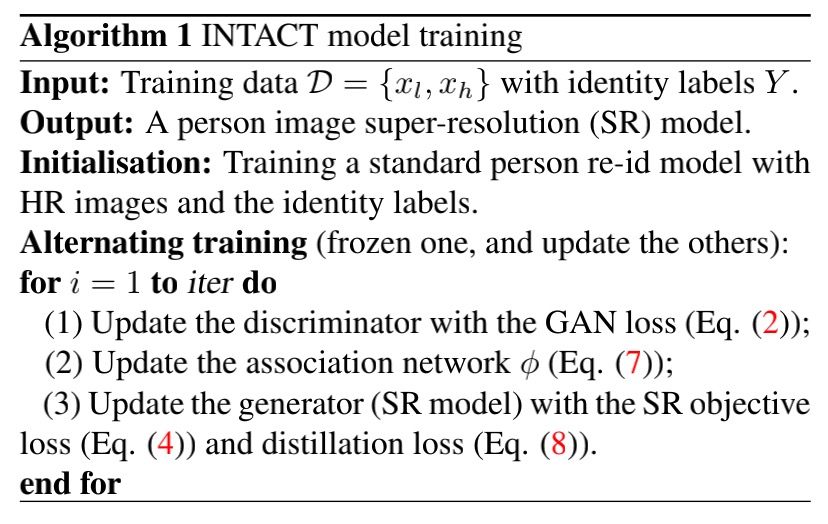
值得一提的是，与等式（5）不同。 在这里，我们冻结了在此步骤中功能上充当外部评论家的关联网络φ。 在精神上，此角色与ImageNet预训练的知觉损失VGG模型相似[13]。 结合GAN训练使用Ldis，我们实质上鼓励合成这样的HR图像，这些图像应尊重真实HR图像上身份和忠诚之间的相同关联关系。 这是我们INTACT模型背后的关键驱动力，该模型在单个公式中同时施加了监督信号和重要的是它们的相互作用。

**Remarks**

与使用加权损失总和进行任务间交互学习和交流的事实上的多任务推理不同，我们发现**两个任务之间的潜在关联是一种额外的学习约束**。重要的是，一旦参数化，该关联就可以以数据驱动的方式从原始训练数据本身中自动学习，而无需任何手工操作，也不需要临时知识。 因此，可以有效地缓解两个不同任务之间的内在冲突，从而**有利于整体模型学习过程朝着每个人的身份匹配**。此外，我们还可以考虑到INTACT采用了一种软集成设计，旨在通过最大程度地提高训练过程中的正相关性，将两个不同任务之间的基本目标联系起来。因此，两个学习目标可以在统一的学习过程中自适应地协作，并在个人追求和共同追求之间取得平衡。

**Model Training**

在模型训练中，我们的INTACT损失项与标准GAN优化又无缝集成了一步。 整个模型保持端到端的可训练性。 算法1总结了整个训练过程。

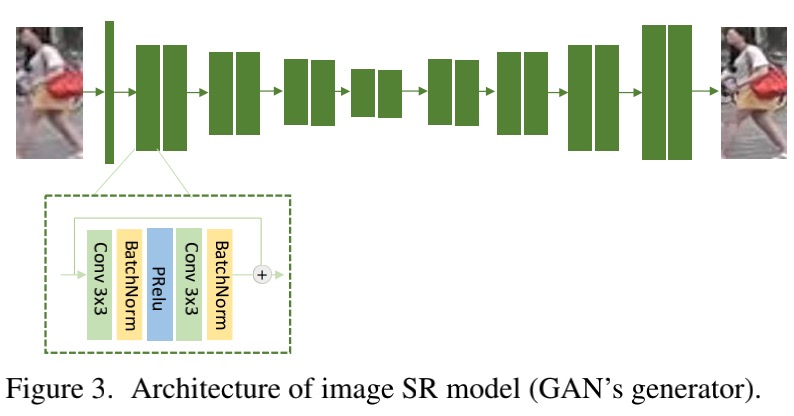


**4. Experiments**

**4.1. Datasets**

**Implementation Details**

在训练期间，通过对HR图像进行随机下采样r∈{2，3，4}次来生成变化的LR图像。 然后将所有LR图像的大小调整为256×128×3，以进行模型训练和部署。 我们使用残差块[10]作为模型的主干。对于SR生成器，我们采用了编码器解码器体系结构。具体来说，它由16个残差块组成，平均分布在8个组中。 分辨率从256×128像素下降到16×8像素，下降了16倍（由于前4个残差块组每个都有一个最大池化层），然后又增加回256×128，最后四个残差块组每个都有像素。生成器的结构如图3所示。判别器与[14]相似。ReId网络[33]已在HR训练数据上进行了预训练。 一旦训练，它在INTACT训练期间就被冻结了。



**Comparisons to State-of-the-Art Methods**

我们将INTACT与广泛的最新re-id方法进行了比较，其中包括：（1）常规人员re-id模型：CamStyle [32]和FD-GAN [7]； （2）基于超分辨率的模型：SLD2L [12]，SING [11]，CSR GAN [27]； （3）基于分辨率不变表示学习的模型：JUDEA [17]，SDF [26]，RAIN [5]； （4）结合了SR和分辨率不变表示学习的混合方法CAD [18]。

