**Learning Cross-Resolution Representations by**

**Cycle-Consistency for Person Re-Identification**

通过行人重识别的循环一致性学习跨分辨率的表示

**摘要**：

人物重识别(re-id)经常将低分辨率probe与高分辨率gallery图像匹配。许多方法能处理SR后的图像，但对原始的低分辨率图像泛化性能很差。作者希望生成的图像表示在分辨率变化的时候对id的识别是一致的，是具有鲁棒性的。因此作者提出了CCR模型来学习分辨率变化时identity不变的表示。给定训练图像，我们的bi-MetaSR模块可实现从低到高然后再到低的自适应分辨率周期，反之亦然。在解决方案循环中，我们强制其表示保持一致。这种周期一致性促使我们的模型在各种分辨率下充分利用关键的区分身份信息。为了进一步改善跨分辨率的表示，我们提出了一种与HR一致的损失，它将低分辨率的表现拉向更具信息量的高分辨率对应对象。由于增强了分辨率变化的鲁棒性，我们的模型可以在测试时直接处理各种分辨率，而无需额外的超分处理。

跨分辨率Re-id：匹配不同摄像头捕捉到的人是否是同一个人。现实世界中一般摄像头捕捉到的人的分辨率一般是不可控的，这里将HR人图像定义为gallery集，监控摄像头捕捉到的人定义为probe集合。

以前的方法一般将LR图像SR处理，这种情况下模型仅适应SR图像，在LR图像上泛化性能不好，对分辨率变化不鲁棒。SR方法可以在不改变关键信息的基础上引入更多纹理细节，从而便于识别，但是严重依赖SR模块会引入额外的计算成本，图片SR会引入噪声会影响识别结果。因此作者寻找高低分辨率一致的行人表示。可以直接将LR图像送入识别器识别而不用超分。

为了学习与超分辨率方法身份一致的表示，获得更丰富的信息，并且不引入噪声，提出了CCR模型，依赖循环一致性学习跨分辨率的表示。给定LR训练图像，我们的模型首先将其放大以进行恢复，然后进行缩小以重建原始的低分辨率。 在分辨率循环中，我们强制其表示身份表示保持一致。

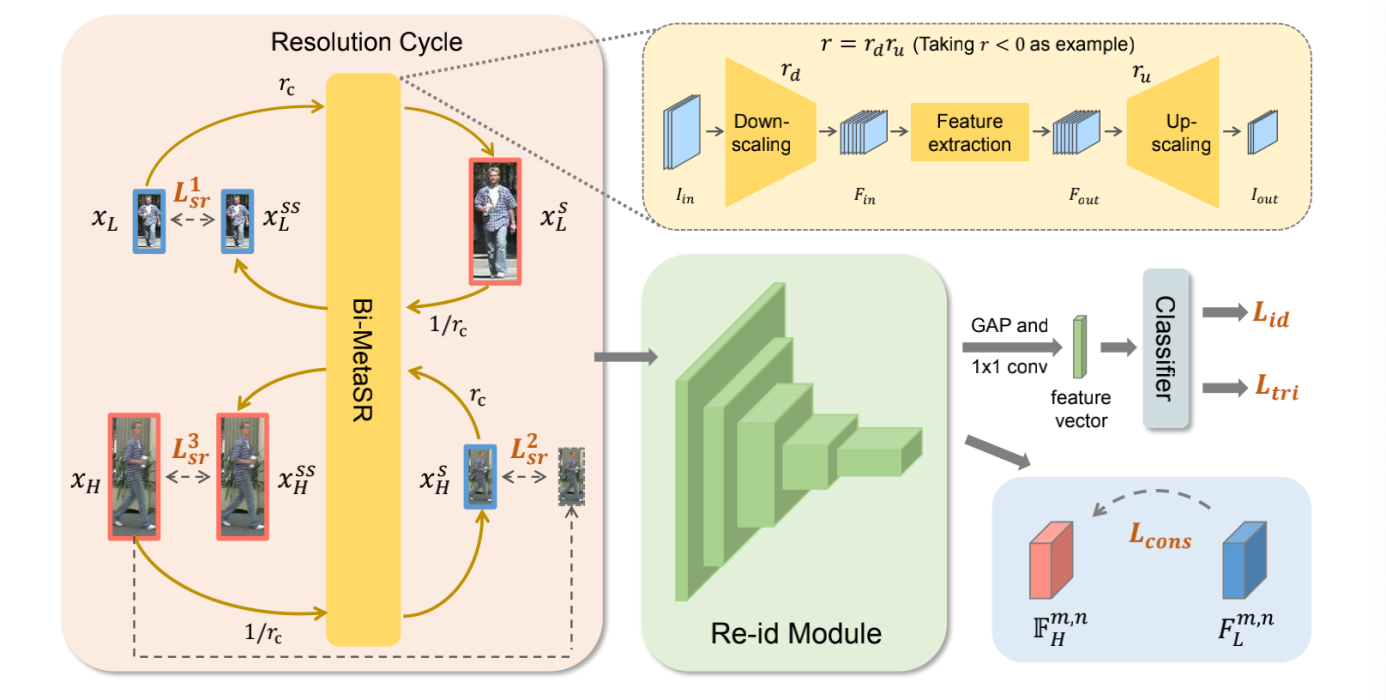
后向重构LR图像会限制SR过程产生更少的噪声，使上采样的特征表示更有意义。分辨率cycle仅仅对图片细节有影响，并不影响身份信息。bi-MetaSR模块可实现一个周期内的双向分辨率转换。

提出了一个损失，以HR作为锚点，使LR表示向HR移动。

与现有的针对单向放大的SR模块不同，作者改进的bi-MetaSR可以以任意比例因子自适应地实现双向放大和缩小。

循环一致性通过规范化从一种模态到另一种模态的映射，然后再回到原始模态，从而提高了特征学习的水平。

方法



CCR模型通过周期一致性来学习分辨率不同但是身份一致的表示。如图，模型由两个周期组成，其中图像分辨率有所变化，但它们的身份表示始终保持一致。

在训练期间，模型将一对具有相同身份但被不同摄像机捕获的LR和HR图像XH和XL作为输入。上循环通过将LR图像先超分再降级重构回LR图像。（rc是比例）。下循环相反。Bi-MetaSR受三个损失函数的和监督。reid模块提取所有的两个循环中的特征表示，通过身份损失和triplet损失保持身份一致性。HR一致的损失Lcons使得生成的LR表示更靠近信息丰富的HR表示。

测试时，只使用Re-id模块提取特征。

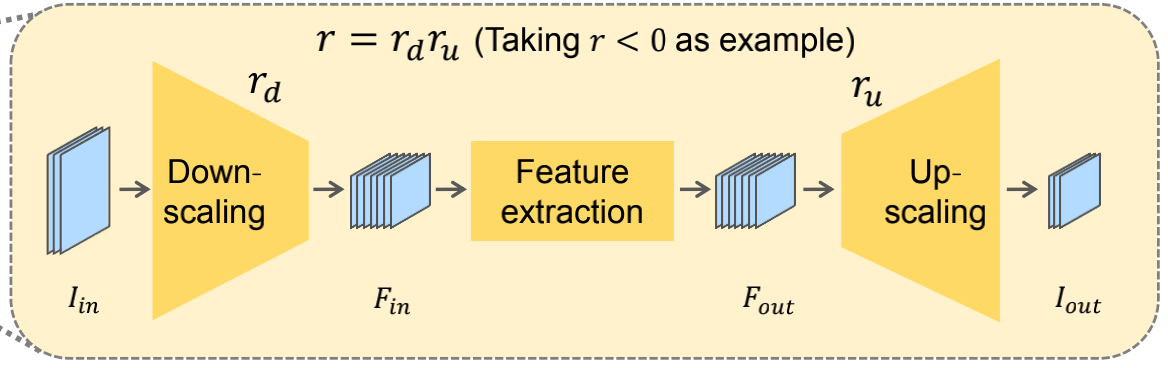
Bi-MetaSR：

包括上采样和降采样的双向操作。如图，Bi-MetaSR分为三部分，bi-MetaSR模块包括三个

部分，包括meta-downscale，深度提取和meta-upscale层。

将期望的比例因子r分解：r=rd\*ru 。满足rd= 1，ru =r进行放大或rd =r，ru= 1进行缩小。

给定输入图像I，meta-downscale层首先通过降采样因子rd进行降采样提取降级特征，深度提取层保持特征图大小不变提取更深的特征，最后通过meta-upscale层恢复图像。



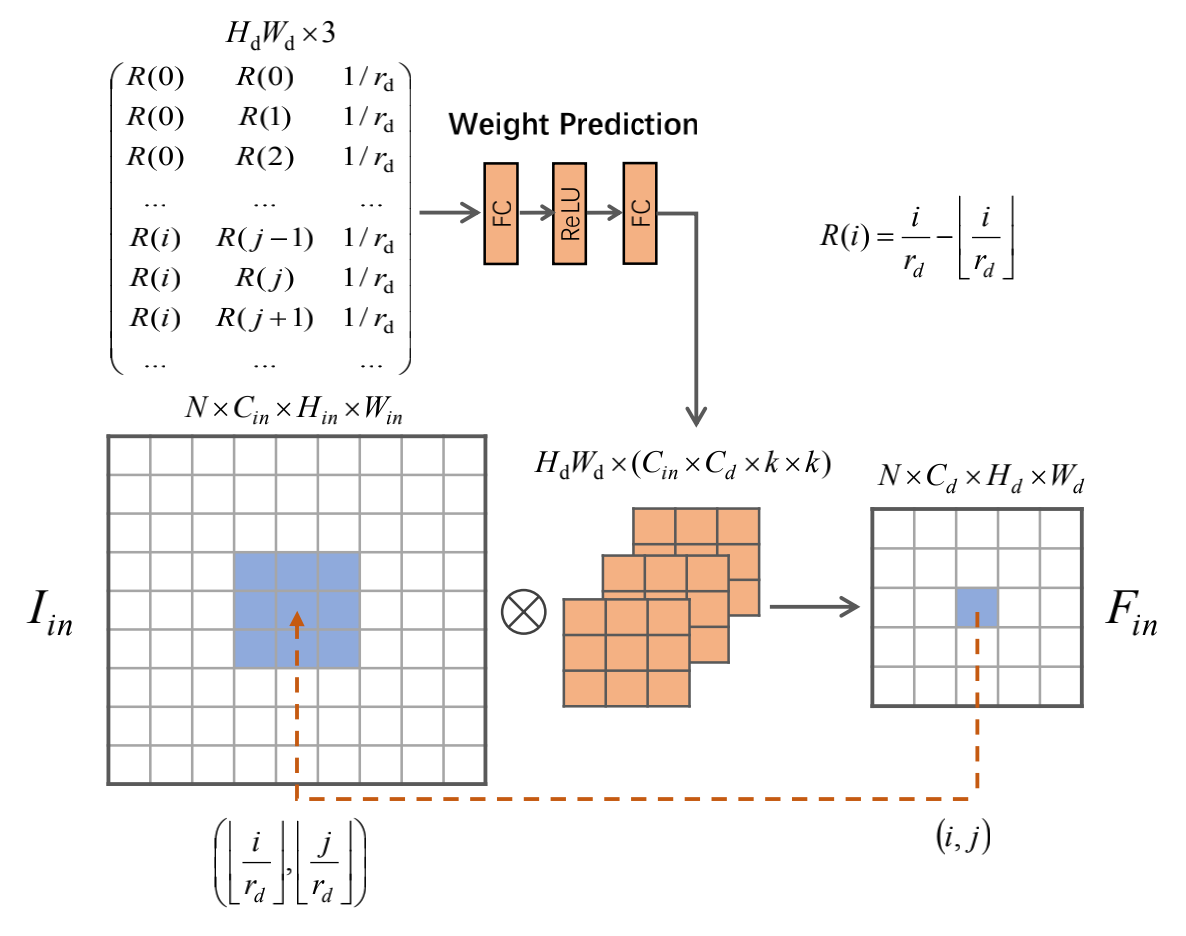
**Meta-downscaling**

图像的放大和缩小类似于像素点的扩展和压缩，在本文中，bi-MetaSR将所有分辨率转换公式化为首先压缩然后扩展信息的过程。基于此，设计了一个meta-downscaling层，以根据比例因子自适应地压缩输入信息。输入Iin，通过meta-downscaling层根据比例因子rd提取特征Fin：



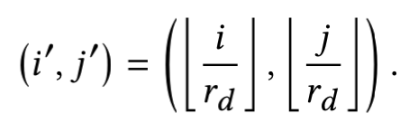
θd代表下采样函数，rd是下采样因子。如果r>1 rd=1，表示输出特征图与输入图片大小一致，但是通道增加。如果rd=r，表示特征图缩小到需要的尺寸。

**meta-downscaling层**：

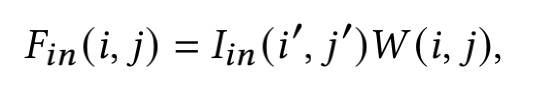


meta-downscaling层设计如下：

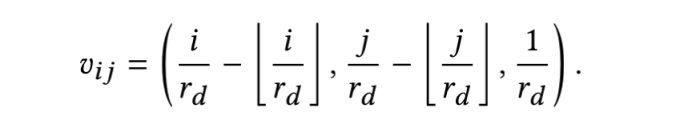
特征图Fin的特征向量在（i，j）处的值由输入图像Iin在（i'，j'）处的像素决定，满足：



因此，提取特征向量需要从（i'，j'）到（i，j）的映射。为了简化，假设映射是线性的，可以公式化为（W为权重）：



为了预测权重矩阵W(i，j)，使用权重预测层，这些层将与坐标（i，j）相关的向量Vij和缩减因子rd作为输入。Vij定义为：



此定义有两个考虑。首先，位于不同位置的像素应作为权重矩阵响应不同的卷积核，并且最好使用相对坐标而不是绝对坐标。因此，将前两项设计为xxxx和xxxx。其次，仅考虑坐标将使相同相对位置的像素始终具有相同的权重矩阵，而不管缩小因子是多少。 为了使权重矩阵适应不同的缩减因子，增加了第三项。

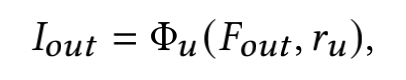
以Vij为输入，权重预测层由两个全连接层以及它们之间的ReLU激活函数组成。作为滤波器，将输出权重矩阵与输入图像上以（i'，j'）为中心的相应区域相乘，以生成特征图上（i，j）处的特征向量。 因此，我们实现了从输入图像到特征图的自适应缩小。

**Deep extraction and meta-upslcaing.**

使用深度提取层进一步提取缩小的特征图，但是在深度提取期间特征图的大小保持不变。

使用RDN作为深度提取层。

最后使用**meta-upslcaing重建图像到期望的大小。与meta-downscaling原理相同。**



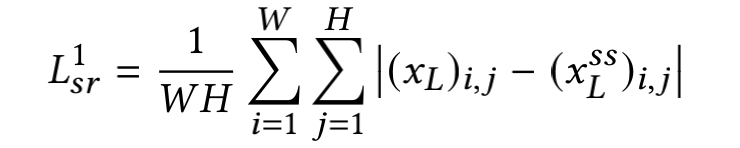
**循环一致性**：

**分辨率循环**：

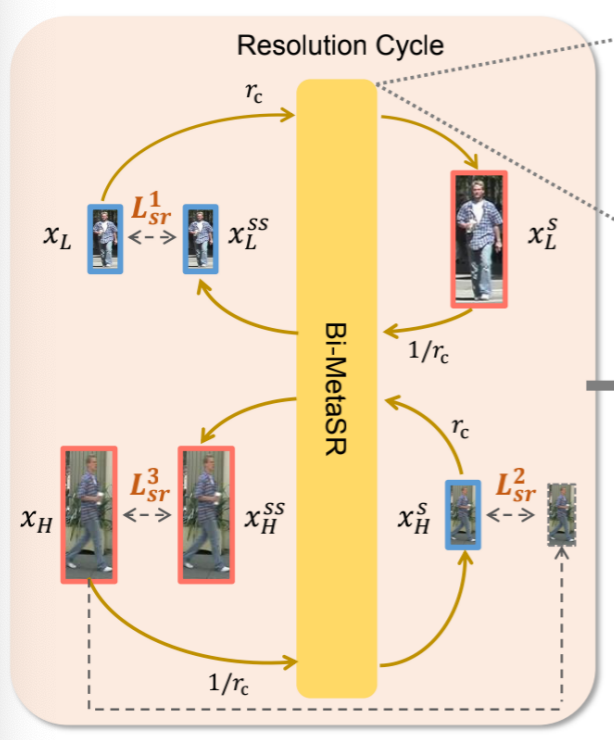
在本文中，我们旨在学习分辨率变化时保持身份同一性的表示。直观地，upscaled的表示信息越丰富，交叉分辨率表示就越具有竞争力。为了使upscaled的表示更具丰富的信息，我们从放大的对应项中向后重构了原始LR图像，这可以限制前向放大恢复更多有用的细节而不是噪声。

为了使biMetaSR能生成有效的图片，引入两个损失：

生成的LR图片与ground truth对比L1：

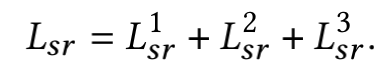


I，j代表像素坐标。该损失请注意也可以规范放大过程，因为放大和缩小共享bi-MetaSR模块并建立端到端循环。



关于循环比例因子rc，Wang建议自适应选择比例因子以统一所有图像的分辨率是一种有效的策略。受他们的发现启发，我们采用基于size的方式为XL选择循环比例因子rc。我们首先将可选的比例因子的集合预定义为{2，3，4}，并计算训练集中所有HR图像的平均分辨率RH（图像高度和宽度的乘积）。如果XL的分辨率为RL，则循环比例因子rc满足rc∈{2，3，4}，并且rc是RH与RL之比的最接近整数。 这样，训练期间的每个放大图像Xsl始终具有相似的分辨率。

同样设计了另外一个Cycle，以另外一个摄像头拍到的XH开始，与XL属于同一个人。先缩小，再重建HR。分别在上采样和下采样操作建立两个loss。通过双线性差值下采样得到的HR图片和原始图片作为ground truth。总的损失是三个SR损失的和：

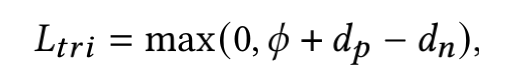


总的SR损失用于使bi-MetaSR模块生成接近ground truth的有效图像内容。

两个双循环共享bi-MetaSR并互相补充。 在第一个循环中进行的重构使放大操作可以恢复有用的身份细节并抑制噪声，而在第二个循环中进行重构则可以使缩小操作以自适应地保留关键身份信息。 它们共同促进bi-MetaSR学习双向分辨率转换。 此外，单个周期主要提高了图像内分辨率变化的鲁棒性。 从XL到XH开始联合训练两个双周期可以帮助将鲁棒性扩展到图像间分辨率变化。

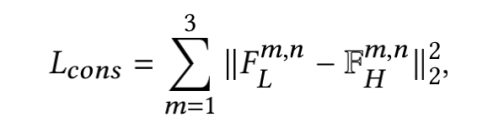
**Identity consistency.**

使用了re-id模块来提取两个双循环中六张图像的特征表示。选择ResNet50作为re-id模块。在全局平均池化层后面加一个1\*1的卷积，将维数从2048减小到512。因此，re-id模块中为每个图像获取512大小的特征向量。然后，由完全连接的层和softmax函数组成的分类器用于预测每个图像的身份分类。使用身份损失和triplet损失限制身份的一致性。



Dp dn代表正对和负对样本的距离。

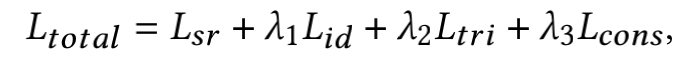
仅执行上面两个损失函数可能会将学习到的交叉分辨率表示引入HR和LR表示之间的中间状态。因为是将两种表示互相推进的。应该使得生成的LR表示向HR移动。因此提出了HR一致性损失：



F代表第m个对的HR和LR特征，从每个残差块的最后一层提取出来。n代表特征级别（五个残差块）。

M=1、2、3因为双循环只有三对LR、HR特征。（下面循环只取GT和下采样版本），以HR特征作为常量。

**总损失**：



优化总的CCR模型。通过端到端训练，bi-MetaSR可以实现面向re-id的分辨率增强或降低，并且re-id模块可以学习对分辨率变化具有鲁棒性的跨分辨率表示。 得益于这种鲁棒性，测试时我们不再需要bi-MetaSR更改分辨率。 re-id模块直接提取所有原始测试图像的特征，因此我们的模型高效，快速。

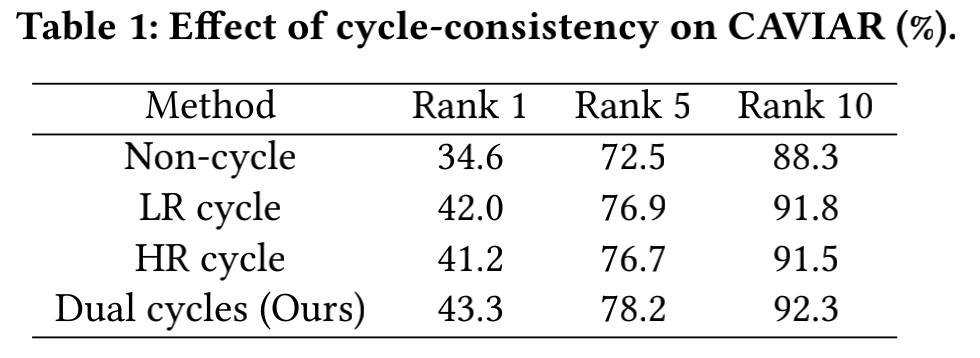
实验：

测试：  
为了进行测试，使用所有LR图像构建查询集，并使用每个人的一张随机选择的HR图像构建gallery。

预训练两个模块

消融研究：

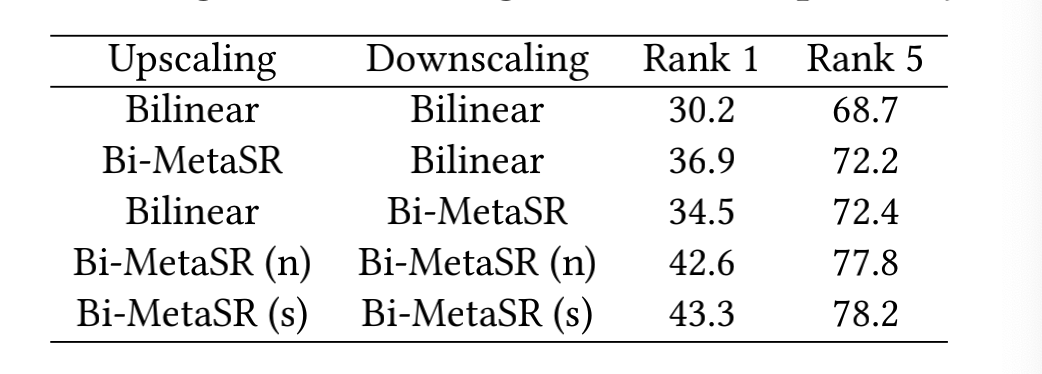
Non-Cycle：两个cycle仅保留上采样和下采样，仅有四张图片和部分损失作为训练。



首先，在没有重构信号反馈的情况下，单向放大会产生更多无用的噪声，而单向缩小会捕获更多琐碎的身份信息。

LR / HR周期分别从真实的LR / HR图像开始，并着重于推动原始LR / HR表示对其放大/缩小版本具有鲁棒性。 此外，两个对偶周期共享整个模型，并进一步提高了为bi-MetaSR转换分辨率和提取re-id的交叉分辨率表示的能力。

bi-MetaSR可以实现构建分辨率周期所需的自适应分辨率增强或降低。 我们通过将bi-MetaSR替换为用于放大或缩小的双线性插值来验证其有效性。



这很容易理解，因为插值是评估图像像素的固定公式，而bi-MetaSR是参数化的模块，可以对其进行优化以合成身份信息以进行re-id。通过使用re-id模块进行的端到端训练，bi-MetaSR可以在放大时恢复信息细节，而在缩小时可以保留重要信息。 因此，bi-MetaSR具有更好的自适应能力，可以实现面向re-id的分辨率转换。

**Losses**：

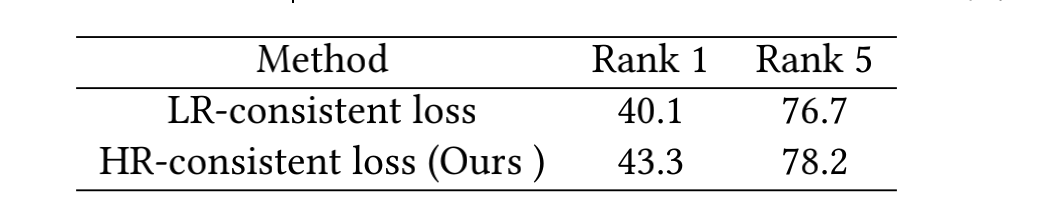
SR损失：用于学习从高到低、从低到高的分辨率转换。

Lid增强不同类别的判别性。

Ltri增强类间距离，减少类内距离

Lcons使得L特征向H特征靠近

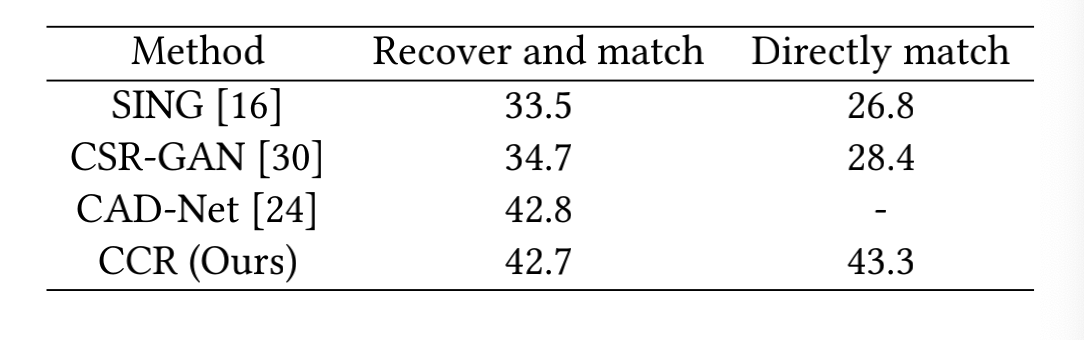
以LR为锚点代替HR：



因为HR包含更多的图像细节

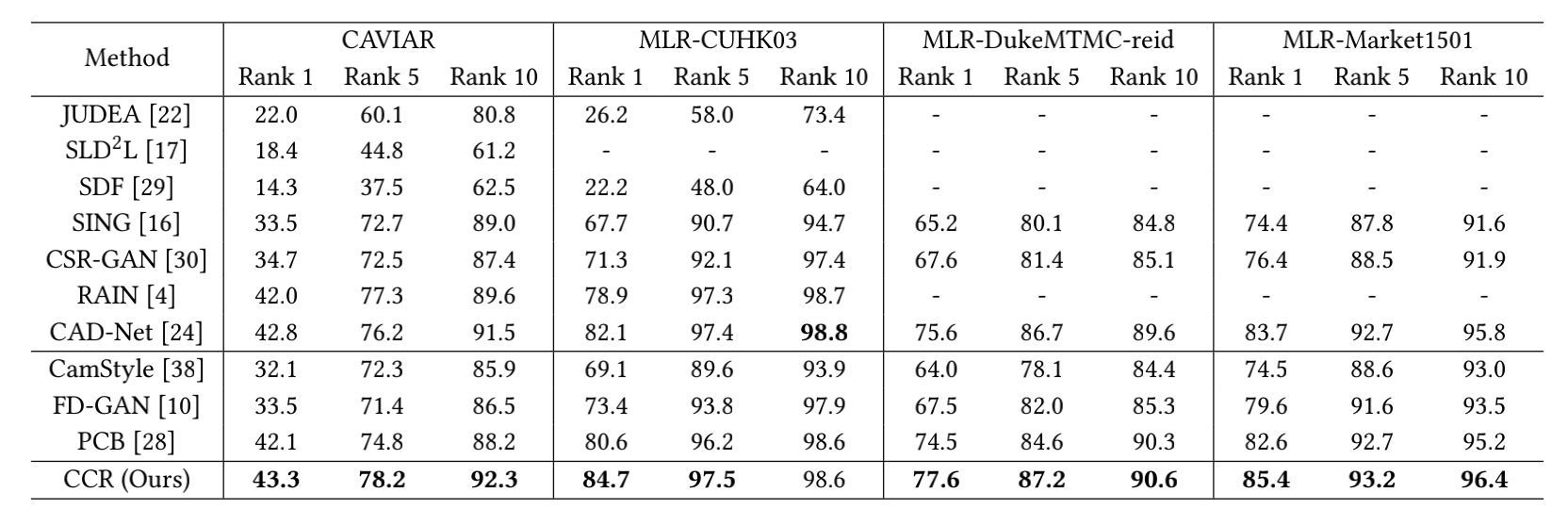
分辨率变化的鲁棒性：

其他低分辨率方法都是先超分再Reid，然而BiSR仅在训练阶段学习跨分辨率的表示，测试时不需要复原。其他方法测试时使用两种方式：先超分/直接复原。



第三种方法结果缺失作者因为找不到合适方法去测，可以看到分辨率变化，不鲁棒，因为移除了超分模块。作者的模型鲁棒，不超分也行。

与最新技术比较：



为了提高针对re-id的分辨率变化的鲁棒性，我们提出了一种新颖的CCR模型，该模型依赖于循环一致性来学习交叉分辨率表示。通过在分辨率的周期性变化过程中提取一致的身份表示，我们的模型学会了完全捕获与分辨率无关的身份属性。为了建立周期一致性，我们改进了bi-MetaSR模块，以自适应地实现分辨率的提高和降低。我们建议的HR一致性损失，可以通过指导交叉解决方案代表信息量丰富的HR表示而不是LR来使交叉分辨率表示更具竞争力。我们进行了广泛的实验，以表明我们的模型是有效的，并且在四个数据集上均达到了最新水平。此外，我们的模型可以摆脱超分过程，并在测试时直接提取各种分辨率的表示。这减少了推理时间和资源消耗，这使我们的模型在实际场景中可行。但是，可视化结果表明，我们的模型从视觉上恢复了竞争性较低的细节。我们将在未来的工作中尝试提高bi-MetaSR的恢复质量，以进一步改善交叉分辨率的表示。