Unified Batch All Triplet Loss for Visible-Infrared Person Re-identification

统一批次的全三元组损失多模态行人重识别

Wenkang Li, Ke Qi, Wenbin Chen, Yicong Zhou

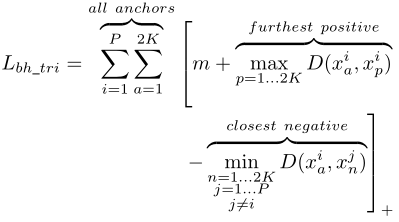
**摘要：**

批量硬三元组损失在行人重新识别任务中被广泛使用，但在可见红外行人重新识别中表现不佳。因为它只为小批次中的每个锚定图像优化最难的三元组，所以最难三元组中的样本可能都属于同一模态，这将导致模态优化的不平衡问题。为了解决这个问题，我们采用了批量全三元组选择策略，它从样本中选择所有可能的三元组进行优化，而不是最难的三元组。此外，我们引入了统一批量全三元组损失和余弦Softmax损失，以协同优化图像向量之间的余弦距离。类似地，我们将针对VI-ReID任务提出的异质中心三重损失重写。

**1.方法介绍**

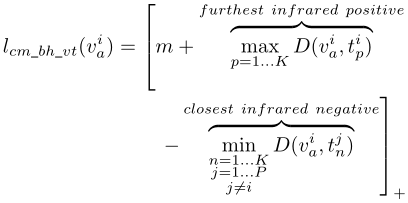
**1.1统一批次全三元组损失**

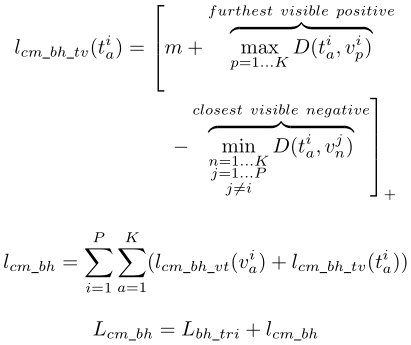
硬三元组损失（Batch Hard Triplet loss）：



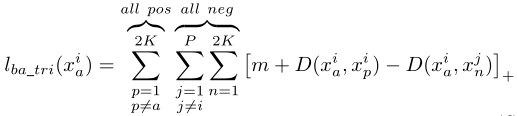
VI-ReID任务中硬三元组损失的问题是，它选择的用于形成三元组损失进行优化的样本可能都属于同一模态。这意味着，在每次迭代中，一个模态被优化得更多，另一个模态优化得更少。

跨模态硬三元组损失（Cross Modality Batch Hard Triplet loss）：

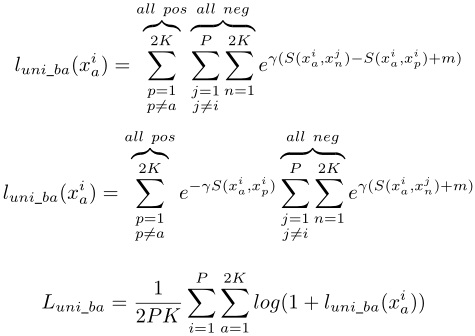




统一的全三元组损失（Batch All Triplet loss）：



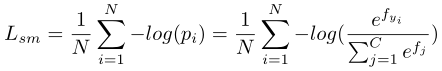
每个anchor有(2K−1)个正样本和2(P−1)K个负样本，因此有2(P−1)K(2K−1) 三元组。由于考虑了所有样本，模态优化不平衡问题不再存在，但是计算量太高。



这种方法将计算复杂度从O(2(P − 1)K(2K − 1))变成了O(2(P − 1)K + (2K − 1))，这样所有的三元组都能造成适当的损失，而不只有大于margin的三元组。γ是比例因子，S是余弦距离。

**1.2具有余弦Softmax损失的协同优化**

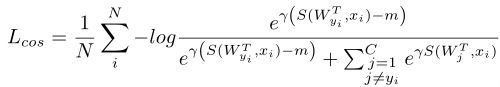
分类损失：





其中是完全连接层的权重。Wj表示W中的第j列，Wj可以被视为类j的中心。因此，Softmax损失优化了每个图像向量及其类中心的内积。

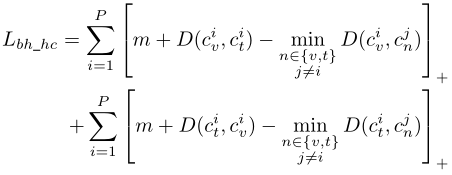
Softmax损失通常与三元组损失一起使用。然而，Softmax损失优化了向量内积，而三元组损失优化了欧几里德距离，它们不一致。更糟糕的是，在推理阶段，通常使用L2归一化向量，这意味着无论在推理阶段使用向量内积还是欧几里德距离作为度量函数，它都等效于余弦距离。这与训练期间使用的度量函数不一致。为了解决这些问题，我们引入Cosine Softmax损失，以使用统一批次全三重损失协同优化余弦距离。余弦Softmax损失可公式化为：

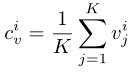


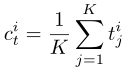
其中γ是比例因子，m是margin。S(x, y)计算x和y之间的余弦相似性。余弦Softmax损失和Softmax丢失之间的差异在于，前者优化样本向量和权重向量之间的余弦距离，而后者优化向量内积。此外，余弦Softmax损失增加了一个margin，使属于同一类别的点数更加集中。

**1.3 Batch All Hetero Center Triplet Loss**

Vi-ReID任务旨在匹配不同模态之间的人物图像，因此关心属于同一类但不同模态的样本之间的距离。因此，提出了批量硬异质中心三元组损失（Batch Hard Hetero Center Triplet loss）：

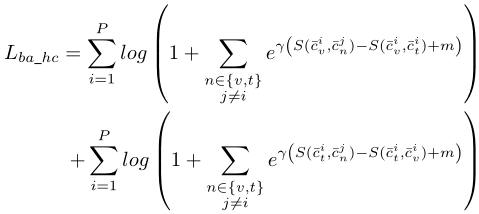






选择同一类不同的模态的中心作为正样本，不同类的距离最近的中心作为负样本，该损失缩短了同类的模态中心距离，增加了不同类之间的中心距离。

Batch All Hetero Center Triplet loss：



1.4模型框架

首先，引入随机灰度作为数据增强方法来减少模态差异，该方法以0.5的概率随机选择部分可见图像并将其转换为灰度图像。然后，我们将红外图像和处理过的可见图像连接起来，形成一个小批量，并将它们输入到Resnet50主干模型中，该模型的权重由两个模态共享。Resnet50输出N\*2048\*H\*特征图。我们使用1x1卷积层将其维数减少到1024。ReLU用作激活函数，卷积层之后没有BN层。我们使用全局平均池化从输出特征图中获得1D向量。由于最后一个ReLU使输出的所有值不小于0，因此我们使用无偏置项的BN层使其分布在0附近。

在推理阶段，我们使用cos距离 作为度量函数：



