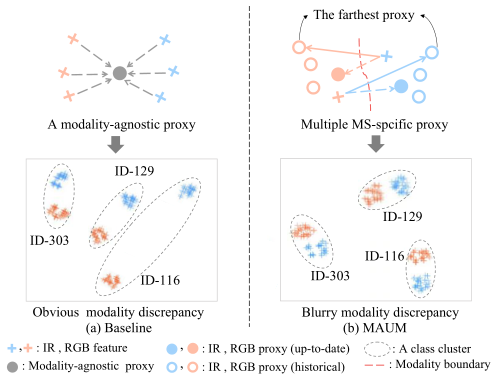
记忆增强单向度量在跨模态行人重识别中的应用

**摘要：**

本文通过抑制模态差异来解决跨模态行人重识别 (re-ID) 问题。在跨模态 re-ID 中，query 和 gallery 图像是不同的模态。给定一个训练 ID，流行的深度分类 baseline 是对两种模态共享相同的代理 (即最后一层的权重向量)。作者发现这样的做法对模态差异有相当大的容忍度，因为共享代理会作为两个模态之间的中间中继。为此，提出了一种记忆增强单向度量学习方法 (MAUM)，包括两种新的设计，即 单向度量 和 基于记忆的增强。具体来说，MAUM 首先在每个模态下独立学习特定模态代理 (MS-Proxies)，之后，MAUM 使用已经学习过的 MS-Proxies 作为静态引用，在对应的模态中关闭特征。这两个单向的指标 (IR图像到RGB代理 以及 RGB图像到IR代理) 共同缓解了中继效应，有利于跨模态联合。通过将 MS-Proxies 存储到 memory banks 以增加参考的多样性，进一步增强了跨模态关联。作者展示了 MAUM 在模态平衡情景下，改善了跨模态 re-ID 的效果，另外对于模态不平衡情景也具有很好的鲁棒性。



(a) 在基线中，每个标识对于两个模态都有一个通用的代理，充当IR和RGB特征之间的中继。(b) MAUM有两个特定于模式的代理(MS-Proxies，橙色实点表示RGB，蓝色实点表示IR)。每个MS-Proxy都是固定的静态引用，用于在对应的模态(虚线箭头)中拉近特性。此外，MAUM将历史MS-Proxies (空点)存储到两个内存库中，一个用于IR模态，一个用于RGB模态。相应地，每个标识都有多个IR和RGB代理。离模态边界最远的MS-Proxies成为硬正引用，因此具有更强的“拉近”效果(实箭头)。

**主要贡献：**

1. 提出了一种新的记忆增强单向度量学习方法，用于跨模态 re-ID 问题。它在两个单向上学习显式的跨模态度量，并通过基于内存的增强进一步增强它们；
2. 考虑了模态不平衡问题，这是跨模态 re-ID 中一个重要的现实问题。通过调整特定模态的增益，MAUM 对模态不平衡问题表现出较强的鲁棒性；

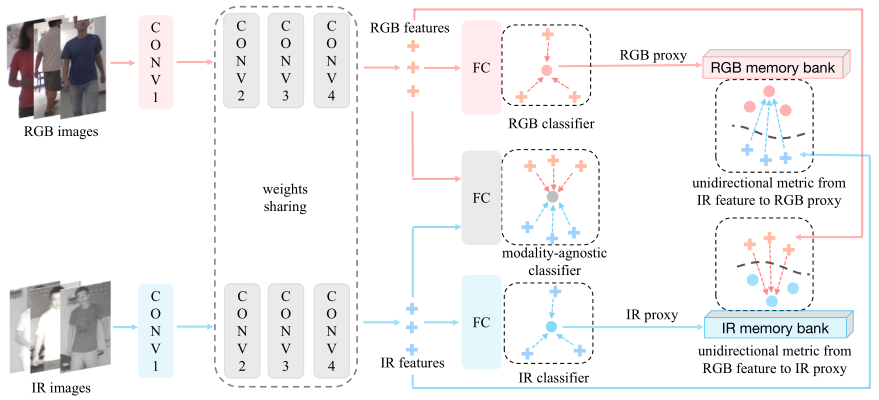
**相关内容：**

数据不平衡是深度学习的一个重要挑战。以往的研究大多关注类别失衡问题，主要介绍了两种方法，即重采样和重加权。在训练中对少数类(样本少)或频繁类(样本多)进行过采样或过采样，目的是在每次迭代中平衡头尾数据。重加权为损失函数中的不同类甚至不同样本分配自适应权重。

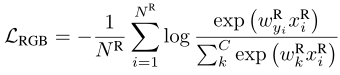
在MAUM中，特定于模式的增强是自然分离的，并允许对特定的模式进行独立的增强。它使MAUM对模态不平衡具有较强的鲁棒性。

**方法：**

MAUM 采用 ResNet50 作为 backbone，接受 RGB 和 IR 图像作为输入。MAUM 将第一个卷积块分成两个独立的分支，以适应特定模态的低级特征形式，一个用于 RGB，另一个用于 IR。为了计算效率，两种模态共享所有的卷积块。对于卷积特征映射，MAUM 使用全局平均池化 (GAP) 为每个输入图像生成深度嵌入。基于这种普遍采用的 backbone 设置，提出的 MAUM 着重于其新的记忆增强单向度量学习方法。



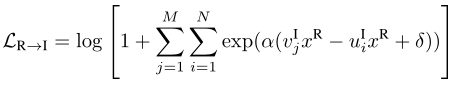
MAUM采用ResNet50作为骨干，两种模式共用“conv2”到“conv4”的参数。将RGB和IR图像映射到深度嵌入空间中，分别得到RGB和IR特征。MAUM有三个分ID类器，分别是RGB分类器，IR分类器和通用的分类器。RGB (IR)分类器只接受RGB (IR)特征，以便学习到的MS-Proxies具有高度特异性，并减轻中继效应。给定已经学习过的MS-Proxies， MAUM在每次迭代后将它们存储到两个相应的内存库中。该记忆库具有三个关键功能，即单向度量学习、通过漂移增强和抵抗模态失衡。分类器采用交叉熵损失：



其中上标“R”表示RGB模态，是当前小批量的RGB数量，C是类别数量，我们使用权重向量作为RGB模态中的代理。

在完全训练特定于模式的代理之后，MAUM将它们收集到两个相应的内存库中。具体来说，我们使用队列策略来更新内存库。我们将RGB模式和IR模式的内存大小分别设置为SRGB和SIR。在内存库达到其大小限制后，我们将最新的代理放入队列，并将最老的代理出队列。存储器对MAUM有三个关键功能。首先，他们冻结已经学习的MS-Proxies，并使用它们作为单向度量学习的静态参考。其次，通过累积历史MS-Proxies，利用模型漂移现象来增加MS-Proxies的多样性。第三，它们帮助MAUM获得针对模态失衡的额外鲁棒性，因为基于记忆的增强是特定于模态的，可以独立调整以重新平衡IR和RGB模态的增强。

在特定于模式的分类器中，每个标识只有一个IR和RGB代理，但将历史MS-Proxies存储到内存库中会逐渐增加其数量。因此，在RGB (IR)内存库中，每个标识都有多个RGB (IR)代理，为单个IR (RGB)特性提供多个正引用。具体来说，对于单个RGB特征xR，我们假设IR记忆库中有N个正引用和M个负引用 (上标“I”表示IR模态)。RGB图像到IR代理的单向度量的损失函数定义为:



特征和代理是l2 normalized，是比例因子，是边距参数，在实践中，损失函数是对当前小批处理中的所有RGB特性的平均值。IR图像到RGB代理的单向度量的损失函数和RGB图像到IR代理类似。