Cross-Modality Person Re-Identification via Modality Confusion

and Center Aggregation

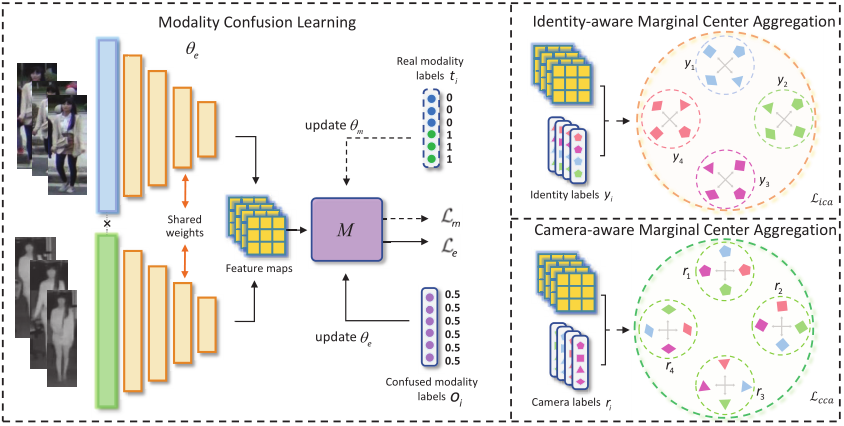
基于模态混淆和中心聚集的跨模态行人重识别

**1.主要贡献：**

1. 提出了一种用于cm-ReID的新型模态混淆学习网络（MCLNet）。
2. 引入了一种身份感知的边缘约束中心聚集策略，它提取了中心化特征，同时保持了多样性，以提高具有边缘约束的泛化能力。
3. 设计了一种相机感知学习方案，该方案应用相机标签监督，通过相机感知表示来丰富可辨别性。

**2.方法介绍**

**2.1整体框架**



采用AGW的VI-ReID通用框架作为基线。特征提取器是一个双流网络，连续提取模态特定和模态共享特征。具体而言，为了处理两种异构模态的差异，在第一个卷积块中，独立处理RGB和IR图像，以学习浅层特征。之后，两个流的以下四个块共享参数并共同提取高维特征。

2.2模态混淆学习(Modality Confusion Learning)

目标是实现模态分类器无法区分输入图像的模态的混淆。从形式上讲，对于每个样本图像，都有一个身份标签、一个实模态标签和一个混淆模态标签。使用二维向量来定义one-hot标签。对于每个输入样本，真实模态标签对于RGB图像设置为[1，0]，对于IR图像设置为[0，1]。对于混淆的模态标签，对于来自两个不同模态的所有样本，其设置为[0.5，0.5]。

模态混淆学习包含两个组件：特征提取器和模态混淆模块M。参数表示M作为模态混淆模块（MCM）。它本质上是一个两层分类器，其目的是将输入图像准确地区分为特定模态。的特征为，M输出模态预测概率，并将其与真实模态标签进行比较。M的损失函数可以表示为：



假设输入为RGB图像，则=[1，0]，其损失为：



表示模态混淆模块，当越接近1时，损失函数越小，即当模态预测准确时，损失最小。旨在最小化模态分类器的损失，以帮助网络区分模态。

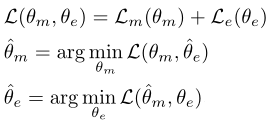
为了实现模态混淆，将特征提取器的预测概率与混淆的模态标签进行比较。损失函数可以表示为：



假设输入为RGB图像，则=[0.5，0.5]，其损失为：



表示特征提取器，当时，损失函数最小，即特征分布尽可能的相似。在训练阶段，交替更新和，直到它们达到平衡。和可以优化如下：



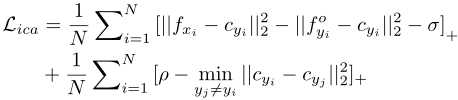
**2.3身份识别的边缘中心聚合（Identity-aware Marginal Center Aggregation）**

中心损失：

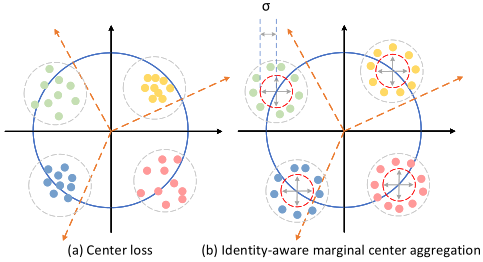


ICA将相同身份的特征约束在一定范围内。σ来确保属于同一类的样本不太靠近中心，防止特征过度拟合到特征中心，以适度保持身份特征的多样性。

用于身份预测的ICA损失可以表示为：

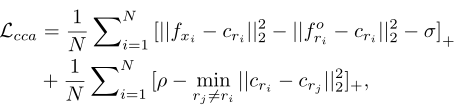


其中，表示第i个样本，是的身份标签，是第i个身份的类中心，是最接近中心的特征嵌入，N是该批次中的样本数，σ是超参数。第一项中的第一个元素是中心损失的一般形式，第二个元素是样本和中心之间的最小距离。两者相减以便将样本特征逐渐远离中心保持一个小的margin。超参数σ迫使样本与其类中心保持合理的距离。σ可视为图中内圈（红色）的半径。这种设计避免了过于严格的中心集中。第二项计算不同身份中心的最小距离。通过在不同身份中心之间应用多个约束，网络比较身份相似性而不是样本相似性。



**2.4摄像头感知边缘中心聚合（Camera-aware Marginal Center Aggregation）**

与ICA类似，损失函数如下：



**2.5总结**



