Discover Cross-Modality Nuances for Visible-Infrared Person Re-Identification

发现可见光-红外行人重识别的跨模态细微差别

Qiong Wu, Pingyang Dai, Jie Chen, Chia-Wen Lin, Yongjian Wu, Feiyue Huang, Bineng Zhong, Rongrong Ji

**摘要：**

现有的研究主要集中在通过调整不同模态的特征分布来缓解模态差异。但是细微有区别的信息，如眼镜、鞋子和衣服的长度，还没有得到充分的研究，特别是在红外模式中。在没有发现细微差别的情况下，仅使用形态对齐来匹配不同形态的行人是一项挑战，这不可避免地会降低特征的独特性。本文我们首先提出了一种模态缓解模块，以从提取的特征图中去除模态信息。然后，我们设计了一个模式对齐模块，它为一个人的不同模式生成多个模式图，以发现细微差别。最后，我们引入了一种互均值学习方式来缓解模态差异，并提出了一种中心簇损失。

**1.引言**

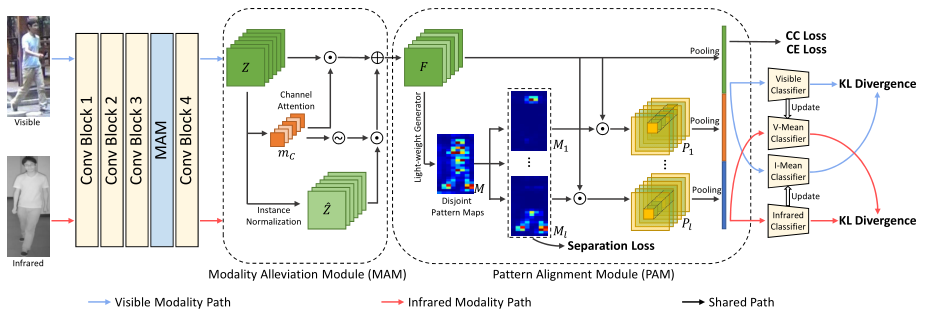
MPANet t框架含有两个模态缓解模块（MAM），用于缓解模态差异，一个模式对齐模块（PAM）用于发现不同模式中的细微差别，以及一种相互平均的学习方式，用于训练具有中心簇损失和交叉熵损失的模型，以进行身份识别。具体来说，MAM使用实例规范化来缓解模态差异，同时尽可能保持区分性。通过一个轻量级生成器，模式对齐模块生成一组模式图，这些图关注不同的模式以发现细微差别。该模块的输出是通过连接模式特征和全局特征获得的。为了以无监督的方式发现细微差别，设计了一个区域分离约束，以确保每个模式图都关注不同的模式。然后，提出了一种中心聚类损失，以减少相同身份的某些图案特征之间的距离，同时增加不同身份的特征中心之间的距离。我们进一步应用两个特定于模态的分类器从每个模态中学习特征的身份，并预测与它们相同的特征的分类结果。此外，通过以相互平均学习方式减少由不同模态特定分类器生成的同一图像的预测之间的分布差异，可以减轻模态差异。最后，这两个模块以端到端的方式级联并联合优化。通过以上工作，MPANet提取的特征是模态不变的，可以表示不同模式中的细微差别**。**

**2.主要贡献**

1. 为了发现细微差别并提取鉴别特征，提出了模式对齐模块（PAM），以无监督的方式发现不同模式中的细微差别，提出了中心簇损失和分离损失。
2. 为了在保持身份信息的同时缓解模态差异，提出了模态缓解模块（MAM），该模块在相互平均学习方式的指导下选择性地应用实例归一化。

**3.方法介绍**

**3.1整体框架**



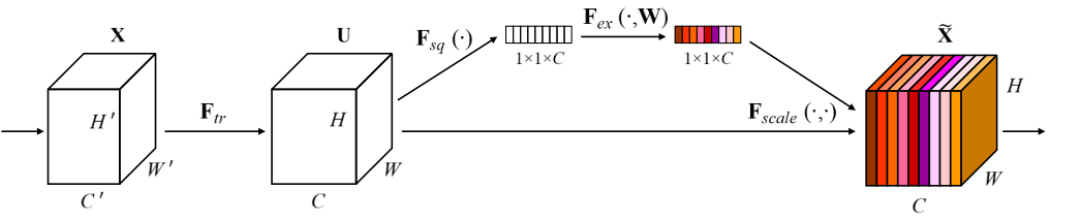
MPANet采用预处理的单流CNN从可见光和红外模式中提取特征地图。卷积块3和卷积块4提取的特征图分别送入模态缓解模块（MAM），该模块对特征图进行细化，以缓解模态差异，同时保持特征图的识别能力。

**3.2 Modality Alleviation Module (MAM) 模态缓解模块**

对于输入图像x，特征图作为MAM的输入，应用实例规范化（IN），它可以减少实例之间的差异。直接应用IN可能会损坏识别信息，我们应用通道注意引导IN来缓解模态差异，同时保留身份信息：

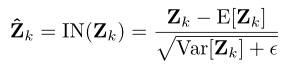


根据SE-Net，我们通过以下方法生成维度掩码mc：





其中g(·)表示全局平均池化，和是两个FC层中的可学习参数，δ(·)是ReLU激活函数，σ(·)是sigmoid激活函数，r=16。



其中是特征地图Z的第k个通道，用于避免被零除，平均值E[·]和标准偏差Var[·]按每个通道计算。

**3.3Pattern Alignment Module (PAM) 模式对齐模块**

PAM旨在发现不同身份之间不同模式的细微差别。我们将特征图分成l个模式，模式图M=[M1，M2，…，Ml] 。



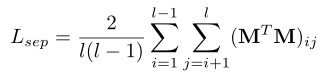
A(·)是轻量级生成器，是一个kernel size 为1的卷积层，σ(·)是sigmoid激活函数。每个模式图都应该注意不同的模式，以发现其中包含的细微差别。我们可以将特征图F拆分为l个模式，如下所示：





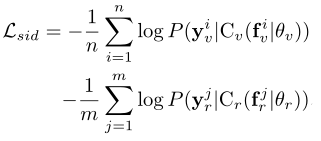


，为了确保模式映射可以捕获不同的模式，通过最小化每两个掩模M之间的重叠区域，分离损失可以监督模式图，以从不同的模式中学习特征。我们把模式特征图想成one-hot向量，细粒度特征的位置就为1，那么两个one-hot相乘，当有重复的地方时，那个地方的值就为1，其余为0。所以一旦两个M相乘，出现1，就证明有重复，否则全为0向量，所以我们让相乘的结果最小，也就是让1不出现，也就是不重复。



**3.4Modality Learning(ML)模态学习**

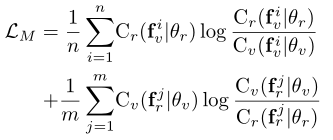
对于每个模态提取的特征f，分别使用一个特定模态的分类器，去计算交叉熵损失:



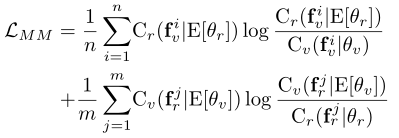
由于输入到每个分类器的训练图像来自特定的模态，分类器只能从其对应的模态中学习知识。因此，给定一个特征f，无论它来自哪个模态，如果两个模态特定分类器提供相同的预测，这意味着该特征可以被视为来自两个模态。换句话说，模态差异被消除。

KL散度又可称为相对熵，描述两个概率分布 P 和 Q 的差异或相似性

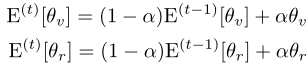




这样去训练分类器，可能分类器的学习并不是因为特征已经对齐了，而是一味的去模仿另一个分类器的参数。为了解决这个问题，一个好方法就是我不直接更新参数，而是使用加权的更新，这样保留了之前的参数，就不会直接模仿另一个分类器的参数了，这样的话，就可以强迫特征去对齐。

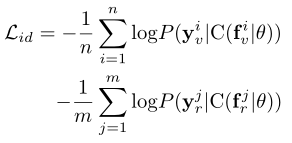


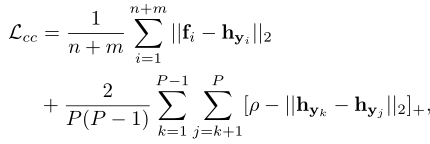
其中和分别表示两个均值分类器的参数，这些参数以时间平均方式更新。



和，参数。

因为两个模态的特征已经对齐，所以我们就可以把两个特征平等来看，进行id loss和中心损失的训练，这里的id loss两个模态用同一个分类器。





其中是当前批次中标记为的特征的平均值，P是当前批次的身份数，ρ是中心中的最小边距。中心簇丢失旨在将特征聚集到其中心，首先是各个特征和当前批次的特征平均值靠近，然后是每个身份的平均特征靠近，代表的该Id身份的平均特征。

MPANet的总损失L定义为:

