SFANet: A Spectrum-aware Feature Augmentation Network for Visible-Infrared Person Re-Identification

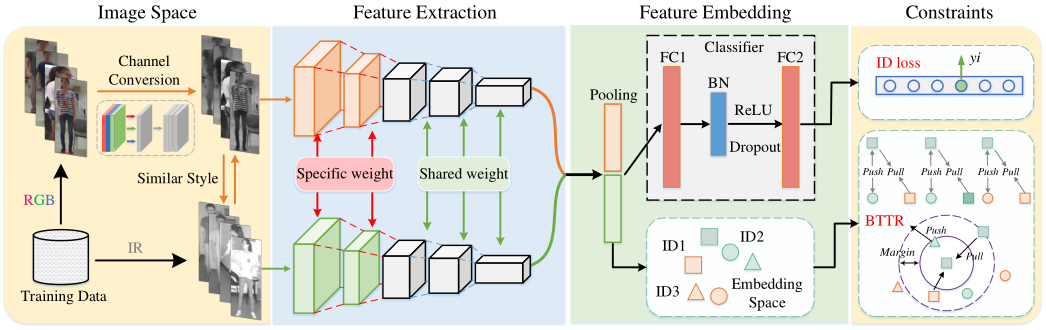
SFANet：一种用于可见-红外跨模态行人重识别的频谱感知特征增强网络

**1.主要贡献：**

1. 首先利用逐元素通道累加的方法，将给定RGB图像点对点转换为灰度图像，并通过简单的复制操作，将单通道图像扩展到三通道灰度图。
2. 计算ID损失的时候，将单FC层前新加了一个FC层和BN层，可以增加ID损失的效率。
3. 在ranking loss中设计了三种双向损失，跨模态损失，模态内损失，模态间损失。

**2.方法介绍**

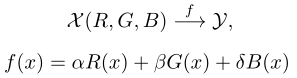
**2.1整体框架**



首先利用逐元素通道累加的方法，将给定RGB图像点对点转换为灰度图像，并通过简单的复制操作，将单通道图像扩展到三通道灰度图。将灰度图和IR图经过一个双流网络提取特征，池化后得到特征向量。一方面，特征向量经过双线性的FC层计算ID损失；另一方面计算ranking loss中的三种双向损失。

**2.2灰度图像生成**

给定一个RGB图像，对RGB图像的每个元素的三个通道进行累加，将得到的每个数值依次放到灰度图像中对应的像素位置上。然后通过简单的复制操作，将单通道图像扩展到三通道灰度图。生成的图像和IR图像风格相像而且保留了RGB图像的一些结构信息，计算公式如下：



其中α、β和δ的值分别为0.299、0.587和0.114。

**2.3双流网络**

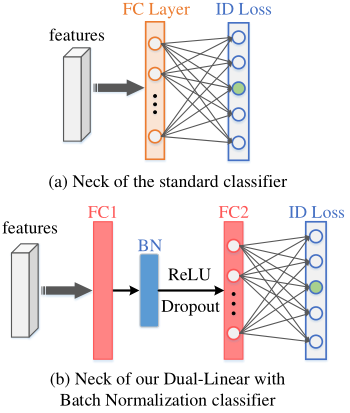
网络包括两个路径，灰度图路径和IR图路径。每个路径包括两个阶段，特定模态的特征提取（layer0和layer1）、共享特征的提取。resnet50作为backbone。

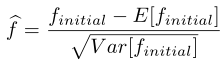
**2.4具有批量标准化的双线性ID嵌入 （Dual-Linear with Batch Normalization ID Embedding）**

**IDE网络（identity embedding network）**是ReID中一个基础的baseline，它的最后一层是全连接层，输出身份的预测。本文原始FC层之前增加了一个新的FC层和BN层。

**双全连接层：**添加了一个新的线性层（带有偏差项）来将不同模式的特征向量投影到一个更有区别的共同特征空间中。然后，利用最后一个线性层(没有偏差）以产生分类结果。这样，一个具有强先验偏差的全连接层以这种方式有助于身份损失。

**BN层：**通过批标准化，最后一个FC层的输入被转换成一种零均值和单位方差的形式，这有助于重新校准嵌入向量的通道。





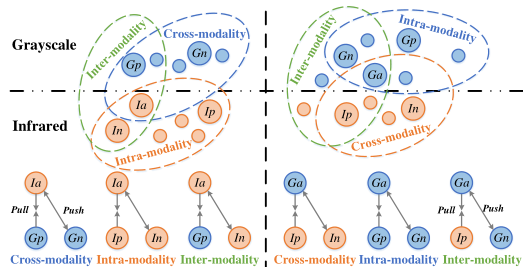
参数和是缩放和移位参数，将E去除后：



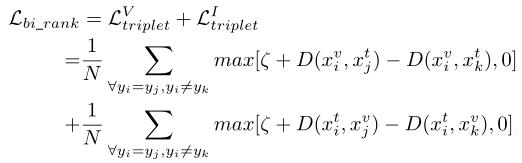
最后计算ID loss:



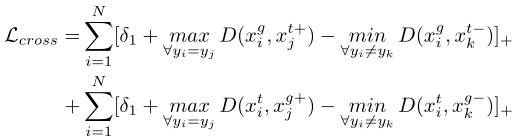
**2.5双向三约束损失**



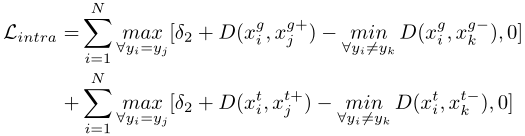
普通的基于三元组损失：



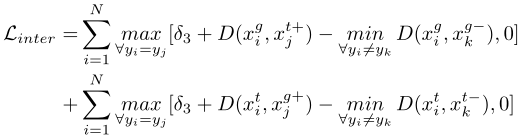
跨模态损失，anchor在一个模态，正负样本在另一个模态的双向三[元组](https://so.csdn.net/so/search?q=%E5%85%83%E7%BB%84&spm=1001.2101.3001.7020)损失。



模态内损失，anchor和正负样本都在一个模态的双向三元组损失。



模态间损失，anchor和负样本在一个模态，正样本在另一个模态的双向三元组损失。



总的损失：

