Deep Multi-Patch Matching Network for Visible Thermal Person Re-Identification

用于多模态行人重识别的深度多补丁匹配网络

Pingyu Wang, Zhicheng Zhao, Fei Su, Yanyun Zhao, Haiying Wang, Lei Yang, and Yang Li

**1.主要贡献**

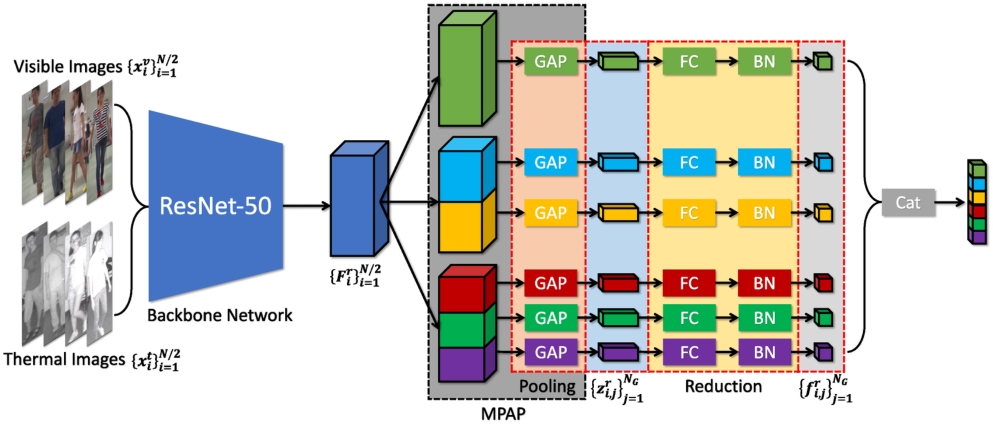
设计了MPMA的模态对齐损失函数，同时平衡和减少粗粒度和细粒度特征的模态差异。

提出了CPCD的相关蒸馏损失函数，以在不同的补丁之间传递语义知识，增强跨模态表示。

提出了PAPA的多任务方法，用于动态调整补丁级别的损失权重，以连续地对困难的补丁任务进行优先级排序。

**2.方法介绍**

**2.1整体框架**



Multi-Patch Matching Network(MPMN)框架由两个主要可学习模块组成。第一个模块是ResNet-50中几个卷积块的堆栈，第二个模块称为Multi-Patch Average Pooling（MPAP）。给定输入图像，该模块网络输出卷积特征图。

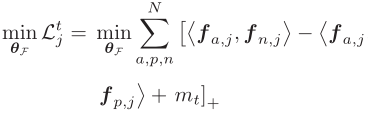
我们将划分为g个水平特征图，以保持局部细粒度信息。然后，对g个水平特征图进行全局平均池化（GAP）得到g个局部特征向量。



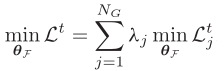
通过g从1变为g来提取多补丁特征以挖掘粗粒度和细粒度视觉语义，最终获得 NG =G(G + 1 )/2个局部特征向量。为了减少特征尺寸，将所有的块的特征向量与FC层和BN层进行连接，压缩为，在测试阶段，所有块的特征向量进行拼接，通道数合起来为D。

损失函数：

三元组损失函数：



<v1，v2＞表示两个向量v1和v2之间的余弦相似性, 为了学习多补丁特性，多补丁三元组损失公式优化为：



其中损失权重λj的值代表不同的块特征学习的权重值。在测试阶段，所有补丁特征都被连接起来作为最终的人物表示，因此，粗粒度和细粒度语义都被挖掘出来，以完善特征的全面性。

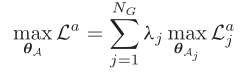
该部分是对N张图片的NG个块进行的三元组运算，计算的是不同图片的同一位置的块的余弦相似性。

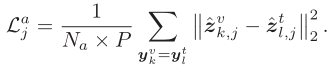
**2.2多补丁模态对齐MPMA**

最大化子空间差异：为了使模态分布更易于区分，我们为模态对齐器Aj使用FC层来对第j个块特征进行子空间投影。等式中的梯度只被反向传播到θA，主干网络θF是固定的。子空间特征的公式为：



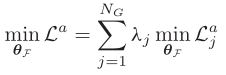
子空间维数P默认设置为P＝C/4。





Na是来自不同模态的正样本对的数量

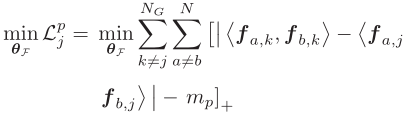
最小化子空间差异：在获得最优模态对准器之后，我们需要通过最小化以下目标来减少硬特征子空间的模态差异，等式中的梯度只被反向传播到主干网络θF，θA是固定的。只训练主干网络，从而获得模态不相关的特征。



**2.3交叉块的相关蒸馏CPCD**

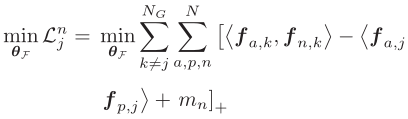
粗粒度特征是健壮的，但不那么有辨别力，而细粒度特征是有辨别力但不那么健壮。因此，如果有效利用不同块相关性，这两个特征可能是互补的。通过度挖掘NG块特征之间的信息相关性。根据来自不同块的两个特征对的相对相似性，我们构造了两种类型的相关性。

正交叉块的相关性：



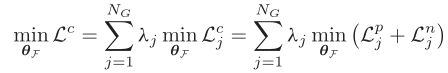
其中并且，来自相同图像对的不同块之间的相似性差异小于阈值。相同图像对的两个块之间的相似性应当接近，即。

负交叉块的相关性：



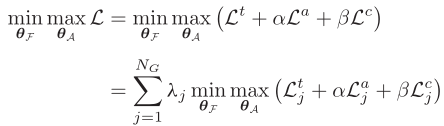
其中和，负样本对相同块之间的相似性和正样本对相同块之间的相似性之差应该小于阈值。

总体相关性：



**2.4补丁感知优先级注意**

将上述的损失函数加起来即：

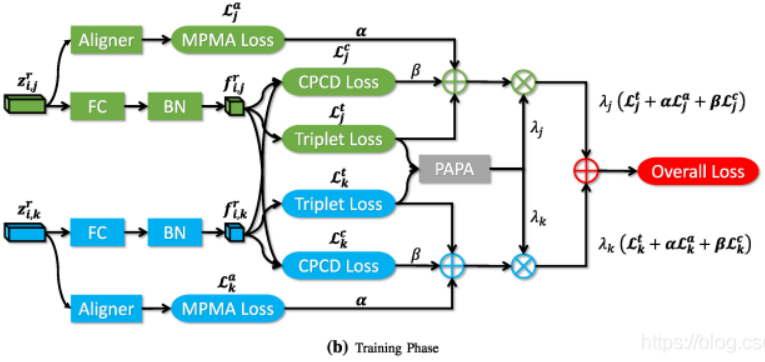


在实验中，=0.1且=0.1，对引入注意力机制：



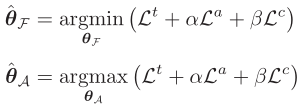
使用三元组损失的代表任务的难度，是训练的总轮数，，开始时，足够小，所有快的权重一样，随着训练轮数的增加，最后不同块任务之间的权重分配更不平衡，并且硬任务的权重分配比简单任务的权重更大。





**2.5训练优化**

通过联合优化最小和最大目标来执行学习最优特征表示的过程。由于这两个目标函数的优化目标相反：



可以使用随机梯度下降优化算法来实现该极大极小博弈。通过合并梯度反向层（GRL）可以有效地执行极大极小优化，梯度反向层在正向传播时是透明的，但其值乘以−1当反向传播时。如果在模态对准器之前添加GRL，则可以同时执行极大极小优化。

