

行人重识别研究综述(二)

2, 度量学习(Metric learning)是广泛用于图像检索领域的一种方法。

不同于表征学习, 度量学习旨在通过网络学习出两张图片的相似度。在行人重识别问题上, 具体为同一行人的不同图片相似度大于不同行人的不同图片。最后网络的损失函数使得相同行人图片(正样本对)的距离尽可能小, 不同行人图片(负样本对)的距离尽可能大。常用的度量学习损失方法有对比损失(Contrastive loss)、三元组损失(Triplet loss)、四元组损失(Quadruplet loss)、难样本采样三元组损失(Triplet hard loss with batch hard mining, TriHard loss)、边界挖掘损失(Margin sample mining loss, MSML)。首先, 假如有两张输入图片, 通过网络的前馈我们可以得到它们归一化后的特征向量。

Base model	Methods	Market1501			MARS			CUHK-SYSU			CUHK03		
		mAP	r = 1	r = 5	mAP	r = 1	r = 5	mAP	r = 1	r = 5	r=1	r=5	r = 10
Resnet50	Cls	41.3	65.8	83.5	43.3	59.3	75.2	70.7	75.0	88.1	51.2	72.6	81.8
	Tri	54.8	75.9	89.6	62.1	76.1	89.6	82.6	85.1	94.1	73.0	92.0	96.0
	Quad	61.1	80.0	91.8	62.1	74.9	88.9	85.6	87.8	95.7	79.1	95.3	97.9
	TriHard	68.0	83.8	93.1	71.3	82.5	92.1	82.4	85.1	94.7	79.5	95.0	98.0
	MSML	69.6	85.2	93.7	72.0	83.0	92.6	87.2	89.3	96.4	84.0	96.7	98.2
Inception-v2	Cls	40.7	66.3	84.1	45.0	62.6	77.9	74.2	78.2	89.7	50.5	68.8	77.4
	Tri	57.9	78.3	91.8	55.5	70.7	85.2	87.7	89.7	96.6	76.9	93.7	97.2
	Quad	66.2	83.9	93.6	65.3	77.8	89.9	88.3	90.2	96.6	81.9	96.1	98.3
	TriHard	73.2	86.8	95.4	74.3	84.1	93.5	83.5	86.1	95.2	85.5	97.2	98.7
	MSML	73.4	87.7	95.2	74.6	84.2	95.1	88.4	90.4	96.8	86.3	97.5	98.7
Resnet50-X	Cls	46.5	70.8	87.0	48.0	63.8	80.2	74.2	78.2	89.7	57.2	77.7	85.6
	Tri	69.2	86.2	94.7	68.2	79.5	91.7	89.6	91.4	97.0	82.0	96.3	98.4
	Quad	64.8	83.3	93.8	63.6	77.7	89.4	87.3	89.6	96.2	80.7	94.9	97.9
	TriHard	71.6	86.9	94.7	69.9	82.5	92.4	86.4	88.8	96.3	82.8	96.1	98.1
	MSML	76.7	88.9	95.6	72.0	83.4	93.3	89.6	90.9	97.4	87.5	97.7	98.9