

Cross-batch memory for embedding learning

跨越时空的难样本挖掘

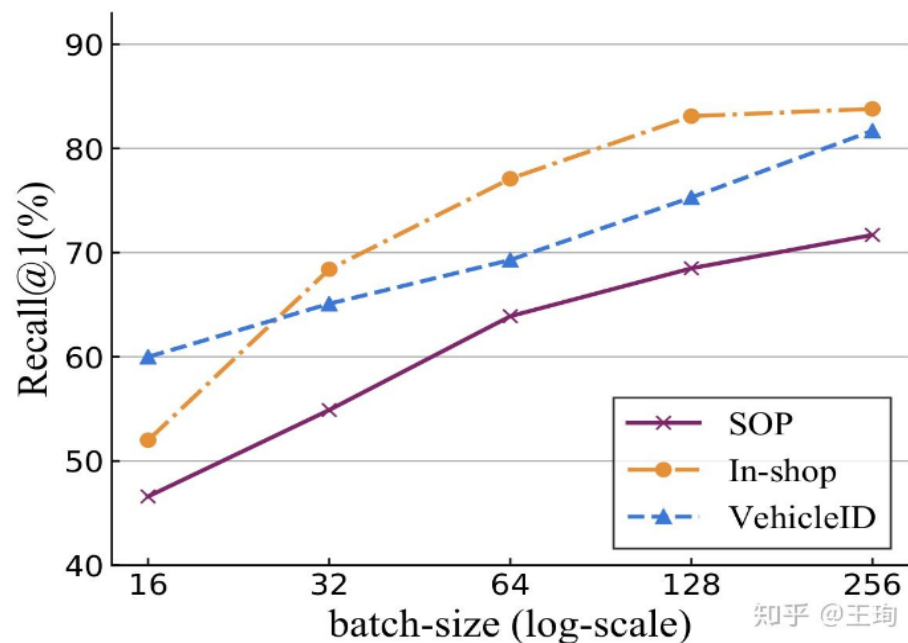
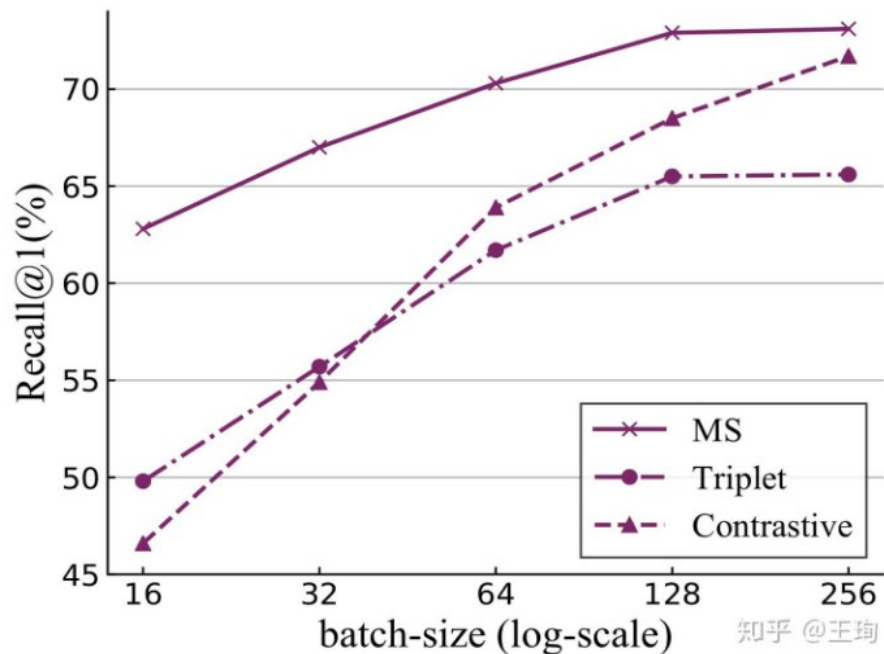
<https://arxiv.org/pdf/1912.06798.pdf>

背景和动机

- 难例挖掘是度量学习的核心问题
- 当前解决方案
 - 在一个mini-batch内，通过各种角度衡量难度，然后对难样本对给予更高的权重
 - 对整个数据集聚类，每次生成mini-batch不是从整个数据集上采样，而是从一个子集或子类采样，这样生成的难例的比例会提高
- 存在问题：都是从mini-batch中选取样本，单次训练可利用的数据总量有限

背景和动机

□ mini-batch对模型的影响



mini-batch越大，训练结果越好，是否可以无限制增加mini-batch?

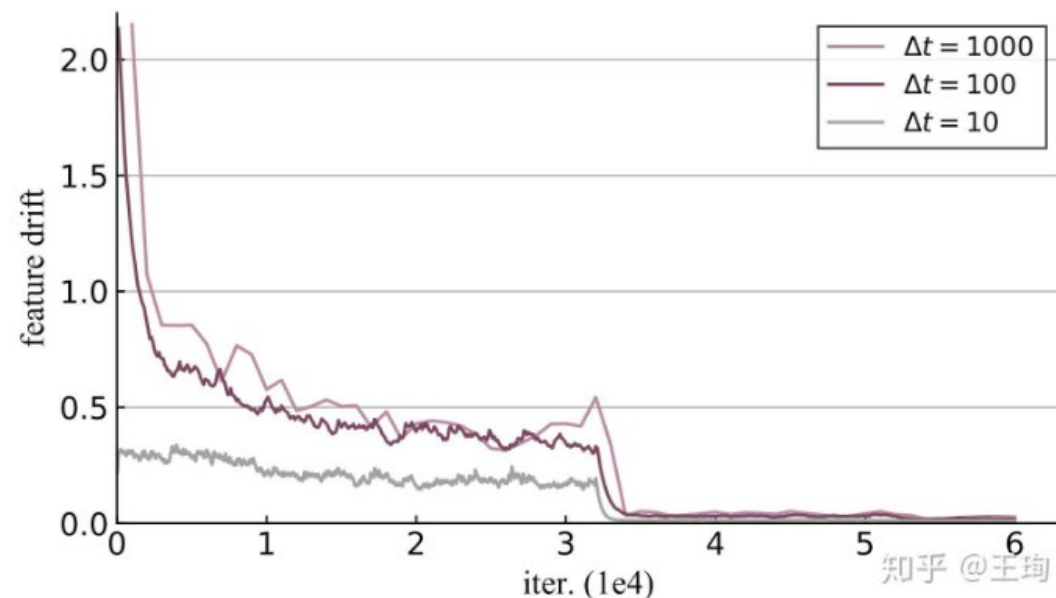
跨时空难例挖掘

□ XBM

XBM提出跨越时空进行难例挖掘，把过去mini-batch的样本特征也拿过来作比较，形成难例样本对

□ 特征偏移

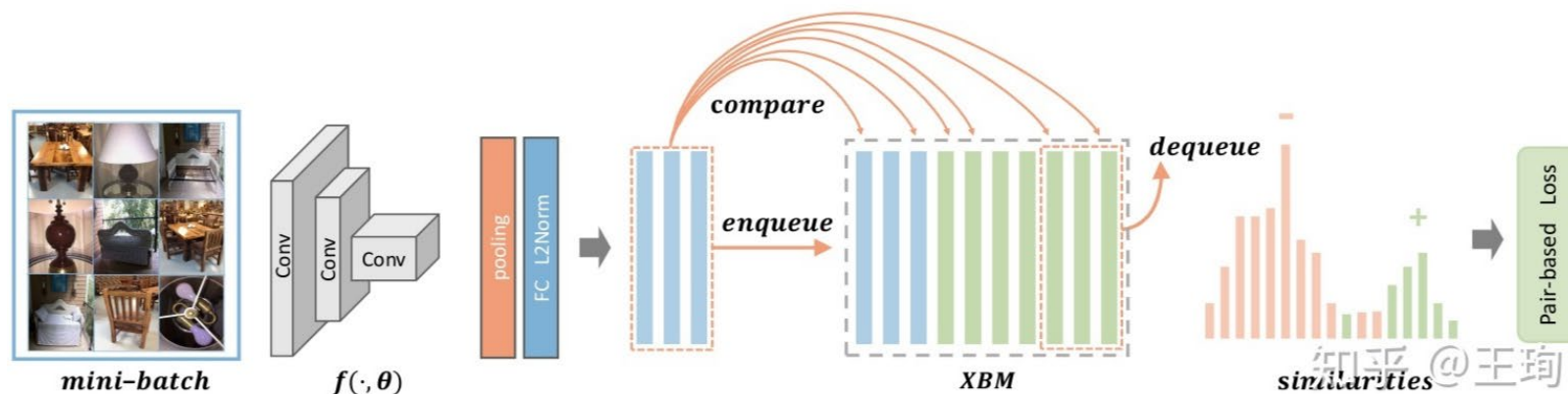
观察样本特征随着模型训练的偏移量，称之为feature drift(特征偏移)



随着训练的进行，模型逐步稳定，特征的偏移也变小

跨时空难例挖掘

□ 模型架构



□ 训练步骤

- 1、模型首先迭代N次，只在mini-batch上做数据挖掘
- 2、在迭代N次以后，认为特征偏移量很小，将特征数据入库存储
- 3、同时在mini-batch以及数据库上做数据挖掘

跨时空难例挖掘

□ 公开数据集实验结果

Method		Small		Medium		Large	
		1	5	1	5	1	5
GS-TRS [5]		75.0	83.0	74.1	82.6	73.2	81.9
BIER [22]	G^{512}	82.6	90.6	79.3	88.3	76.0	86.4
A-BIER [23]	G^{512}	86.3	92.7	83.3	88.7	81.9	88.7
VANet [4]	G^{2048}	83.3	95.9	81.1	94.7	77.2	92.9
MS [36]	B^{512}	91.0	96.1	89.4	94.8	86.7	93.8
Divide [27]	R^{128}	87.7	92.9	85.7	90.4	82.9	90.2
MIC [25]	R^{128}	86.9	93.4	-	-	82.0	91.0
FastAP [2]	R^{512}	91.9	96.8	90.6	95.9	87.5	95.1
Cont. w/ M	G^{512}	94.0	96.3	93.2	95.4	92.5	95.5
Cont. w/ M	B^{512}	94.6	96.9	93.4	96.0	93.0	96.1
Cont. w/ M	R^{128}	94.7	96.8	93.7	95.8	93.0	95.8

Table 5. Recall@ K (%) performance on VehicleID. @王珣

跨时空难例挖掘

□ 林火数据集实验结果

Method	Acc/50%训练数据
TRN	97.36
TRN+xbm	98.04