

一、Baseline

我们选用来自阿里巴巴 AI Lab[1]中提出 BFENet 作为 Baseline。

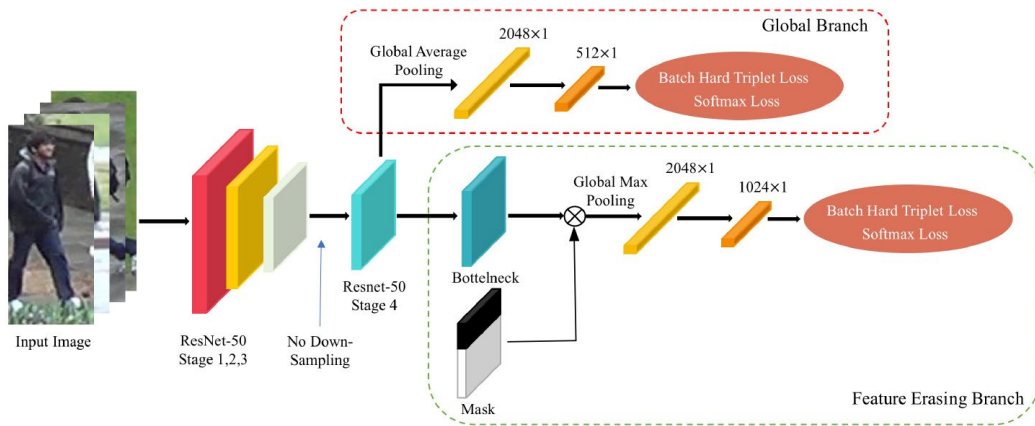


图 1

BFENet 具有 2 个分支，结合了全局和局部信息进行训练能够更好的学习整体特征和一些细粒度的特征以达到更好的效果。

二、改进的网络

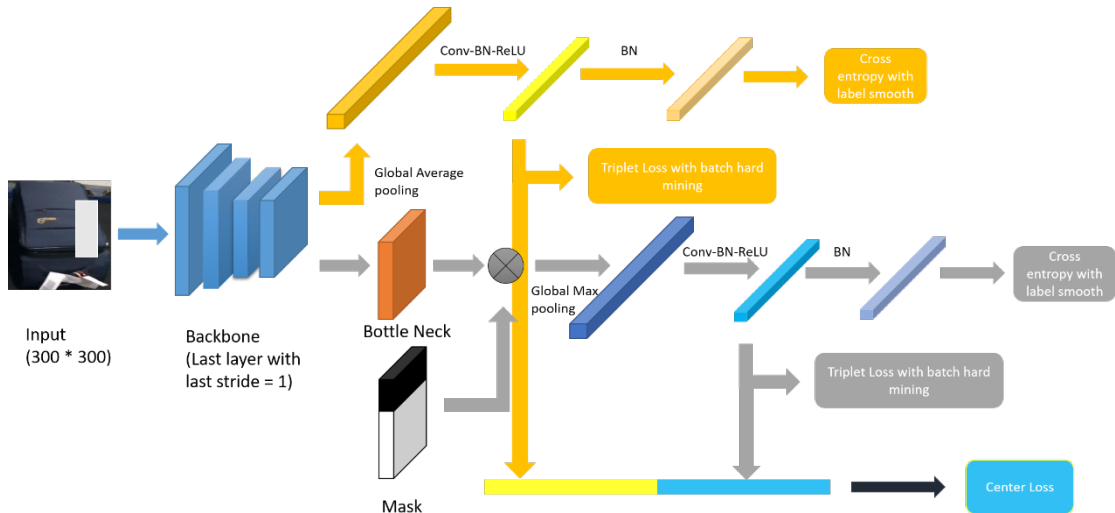


图 2

受 BFENET 启发，我们在此基础上提出了 Random Erasing Network，简称为 RENet。RENet 根据[2]对输入的箱包图像进行随机的遮挡，实现 Image-level 的遮挡，并 Resize 为 300 * 300。Backbone 我们使用 ResNet-50 及其变体的网络，并对最后一层的 stride 改为 1 以增大特征图的大小以便减少信息损失。接着会有两个分支的处理，其中上面的分支对整个特征图做一个全局平均池化处理来得到一个全局的信息，而下面的分支在经过一个 Bottle Neck 之后通过一个随机生成的 mask 对特征图做一个部分遮挡，接着通过一个全局最大池化之后能让这个分支更加关注到一个局部的信息。在训练的时候，我们会对这两个分支的特征进行一

个基于难样本采样的 Triplet loss 的度量学习，同时，我们还会对两个特征拼接后的特征计算 center loss 使得相同样本的距离更近；此外，我们将两个分支的特征通过一个 BN 层之后计算标签平滑的 Cross Entropy Loss，进行一个分类学习。最终，我们通过结合三个 loss 的学习使我们的模型对目标能有较强的表示能力。在进行推断时，我们使用两个分支在 BN 层之前的特征拼接作为该箱包的特征表示。

在特征提取结束后，进行检索时，我们采用了[3]提出的 ReRanking 策略对第一次检索后的结果进行重排列，ReRanking 本质上是利用多图信息来进行无监督的排列的一种后处理的方法。

三、实验细节

根据第二部分所描述的基本结构，我们训练了有细微差别的模型。具体如下表 1

Backbone		IBN-a-Net-ResNet50	ResNet-50	SE-RESNEXT-50
Global Branch	Feature dim	2048	2048	1024
	Conv-Bn-ReLU	✓		✓
	BN	Gamma, Beta	Gamma	Gamma, Beta
	Last Linear Layer	Weight, Bias	Weight	Weight, Bias
Local Branch	Feature dim	2048	2048	1024
	Conv-Bn-ReLU	✓	✓	✓
	Bottleneck	✓	✓	✓
	BN	Gamma, Beta	Gamma	Gamma, Beta
	Last Linear Layer	Weight, Bias	Weight	Weight, Bias

根据第二部分我们介绍的损失函数，我们采用了[4]基于批量的难样本采样的 Triplet Loss、来自[5]提出的 Cross Entropy Loss with label smoothing 以及[6]提出的 Center Loss。Triplet Loss 的 margin 的值设为 2，Center Loss 的特征维度为两个分支拼接后的维度，以及标签平滑系数为 0.1。在训练阶段，进行了随机水平翻折和随机竖直翻折的数据增强策略。Backbone 预训练于 ImageNet，以便加速收敛。此外，采用[7]Adam 优化器和 warm up 策略能得到更优的结果。在推断阶段，我们采用了多推断策略，仿造训练时候的数据增强方式，我们这里对图像采用水平翻折、竖直翻折生成共 3 张箱包图片和 1 张原图，分别通过网络进行特征提取，4 个特征按照 Element-wise 进行 max 操作，以得到最后用于检索时的特征。对

于 Reranking 的参数，此处我们设 $K1=5$, $K2=5$, $\lambda=0.3$ 。

我们在提交的验证集结果上达到了领先的水平，其中第一次和第二次是采用 ResNet-50 作为骨干网络的单模型，第三次提交采用的是三模型特征融合策略。耗时方面，我们在 Windows10、CPU 酷睿 I7-5700HQ 以及 GPU 为 GTX 960M 的笔记本上进行三模型融合的验证集的推断以及 csv 文件写入总耗时 1393s，平均 1.3s/张，符合挑战赛 2s 内要求。

四、总结

得益于 BFENet，我们提出了 RENET。RENET 利用随机擦除进行 Image-level 的随机遮挡操作，并根据 BFENet 对 backbone 提出的特征图进行随机的擦除进行 Feature-level 的随机遮挡操作，以便学习细粒度的特征。此外，我们还引入了 Batch Norm 层将特征归一化使得 ID Loss 更好的监督。对于 Center Loss，我们在训练中发现倘若分开优化特征的效果较差，几乎对网络的训练没有帮助甚至有时候比 Triplet Loss + ID Loss 还差，为此，我们思考在推断时候的需求，将网络的两个特征输出拼接起来用于 Center Loss 的训练。在推断阶段，提出了翻转增强策略并采用 Element-wise 的 max 操作来获得更多箱包的更多细节，以便提高检索能力

	Rank-1	Rank -3	Rank -5	Rank -10
第一次提交	0.8432	0.942	0.9667	0.9876
第二次提交	0.8498	0.9544	0.9724	0.9886
第三次提交	0.8869	0.962	0.9838	0.9905

参考文献

- [1] Z. Dai, M. Chen, S. Zhu, and P. Tan. Batch feature erasing for person re-identification and beyond. arXiv preprint arXiv:1811.07130, 2018. 7
- [2] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang. Random erasing data augmentation. arXiv preprint arXiv:1708.04896, 2017. 3
- [3] Z. Zhong, L. Zheng, D. Cao, and S. Li. Re-ranking person re-identification with k-reciprocal encoding. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 1318–1327, 2017. 7
- [4] A. Hermans, L. Beyer, and B. Leibe. In defense of the triplet loss for person re-identification.

arXiv preprint arXiv:1703.07737, 2017. 2, 7

- [5] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 2818–2826, 2016. 4
- [6] Y. Wen, K. Zhang, Z. Li, and Y. Qiao. A discriminative feature learning approach for deep face recognition. In European conference on computer vision, pages 499–515. Springer, 2016. 5
- [7] D. P. Kingma and J. Ba. Adam: A method for stochastic aoptimization. In ICLR, 2015. 5