2020 DIGIX

全球校园AI算法精英大赛路演

赛道B：数码设备图像检索

xxx队

1.1 赛题理解

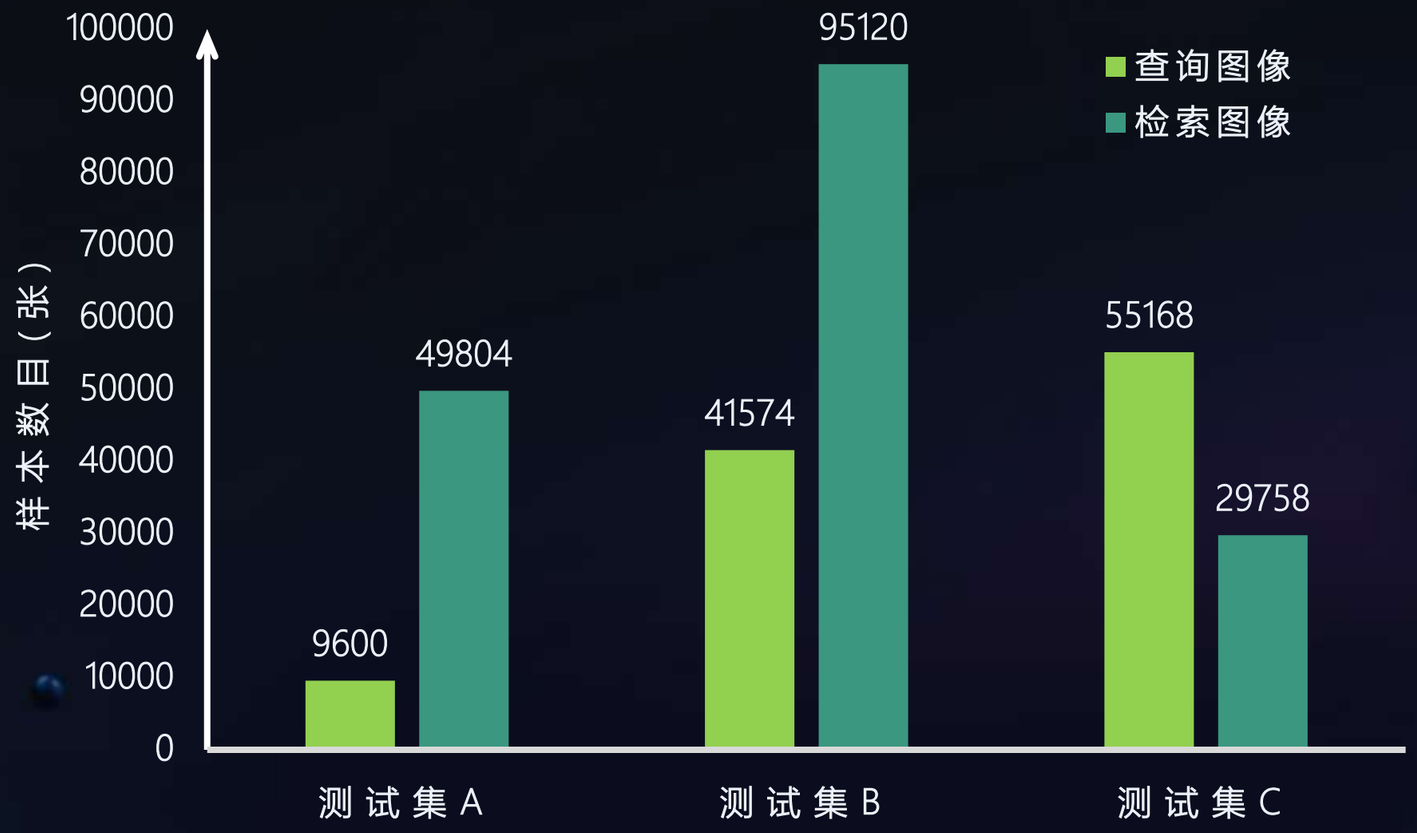
- **问题描述**：给定一张查询图像，在数据库中查找同类图像
- **模型限制**：大小之和不超过500MB
- **评分标准**：50% x Top-1 + 50% x mAP@10

查询图像

检索结果

1.2 数据分析

- 训练集（train_data）
  - 68810 张图片
  - 3097类
- 测试集 A（test_data_A）
  - 查询图像: 9600
  - 检索图像: 49804
- 测试集 B（test_data_B）
  - 查询图像: 41574
  - 检索图像: 95120
- 测试集 C（testdata_1019）
  - 查询图像: 55168
  - 检索图像: 29758

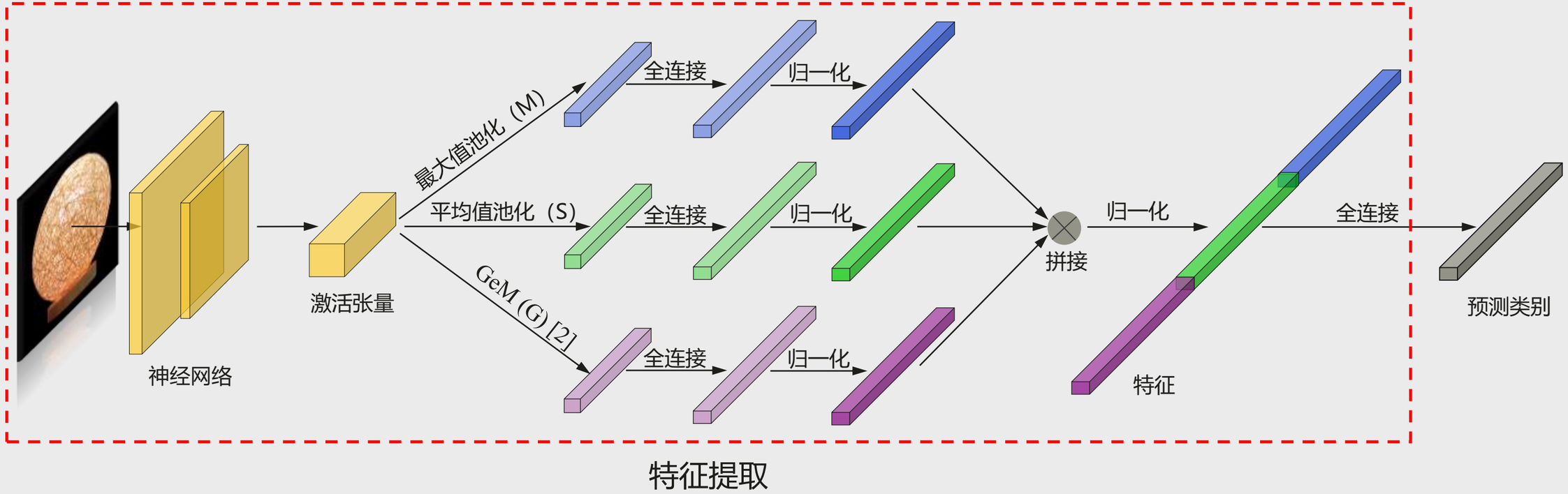


1.3 基本框架

- **核心思路：**
  - ✓ 特征提取
  - ✓ 后处理

2.1 特征提取

- CGD [1] 方法
  - 多个全局描述子 + 端到端训练



特征提取

[1]: Jun, HeeJae, et al. "Combination of multiple global descriptors for image retrieval." arXiv preprint arXiv:1903.10663 (2019).
[2]: Radenović, Filip, et al. "Fine-tuning CNN image retrieval with no human annotation." In TPAMI. 2018

2.1 特征提取

➢ 训练过程

- 从训练集中分出1/4的验证集
- 数据增广：随机旋转，颜色抖动，仿射变换
- 损失函数：CosFace [1] 损失

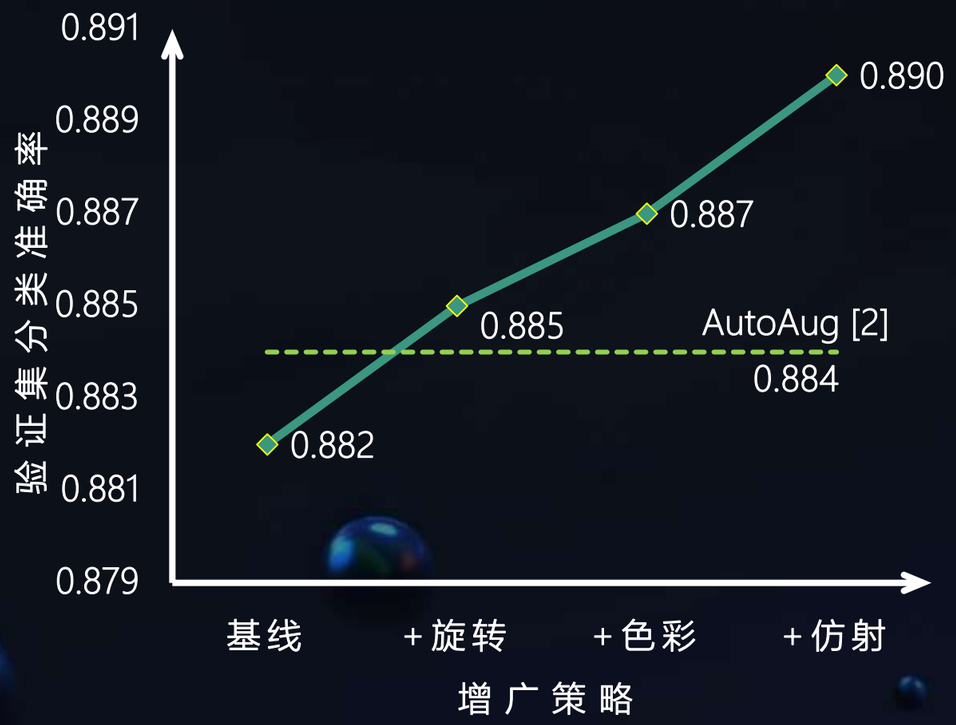
原始图像　　随机旋转　　颜色抖动　　仿射变换

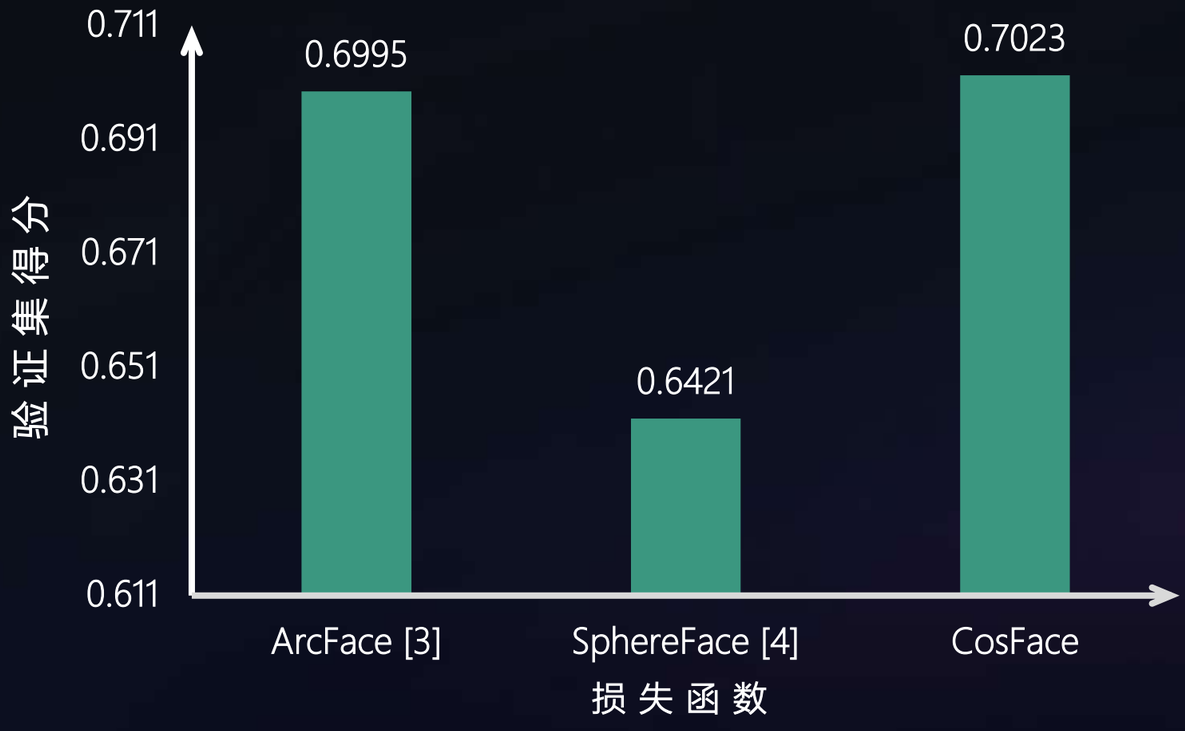
[1]: Wang, Hao, et al. "CosFace: Large margin cosine loss for deep face recognition." In CVPR. 2018

2.1 特征提取

➢ 数据增广与损失函数

- ResNet50 [1] 在验证集上效果





[1]: He, Kaiming, et al. “Deep residual learning for image recognition.” In CVPR. 2016.
[2]: Cubuk, Ekin D., et al. "AutoAugment: Learning augmentation policies from data.” In CVPR. 2019.
[3]: Deng, Jiankang, et al. "ArcFace: Additive angular margin loss for deep face recognition.” In CVPR. 2019.
[4]: Liu, Weiyang, et al. "SphereFace: Deep hypersphere embedding for face recognition.” In CVPR. 2017.

2.1 特征提取

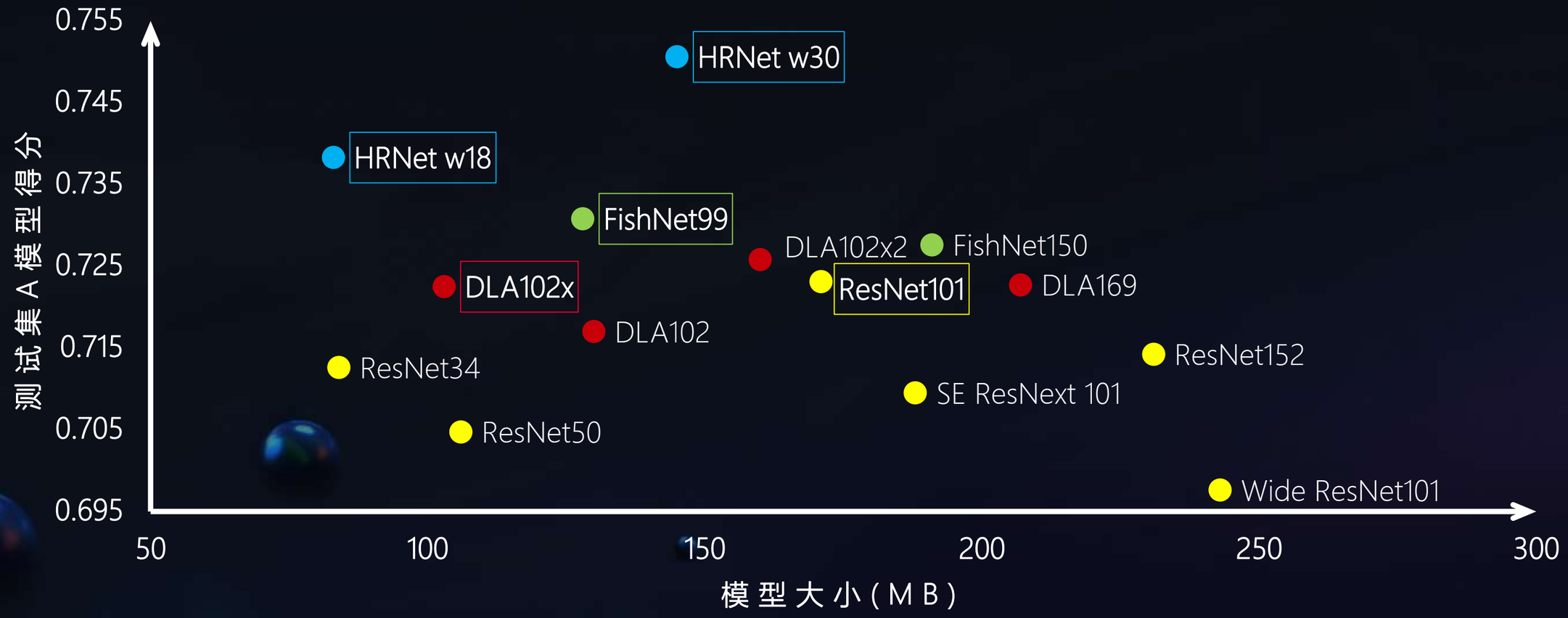
➢ CGD池化层设置与输出维度

- ResNet50在验证集上效果

2.1 特征提取

➢ 最终选择模型

- HRNet w30 & w18 [1], FishNet99 [2], DLA102x [3], ResNet101

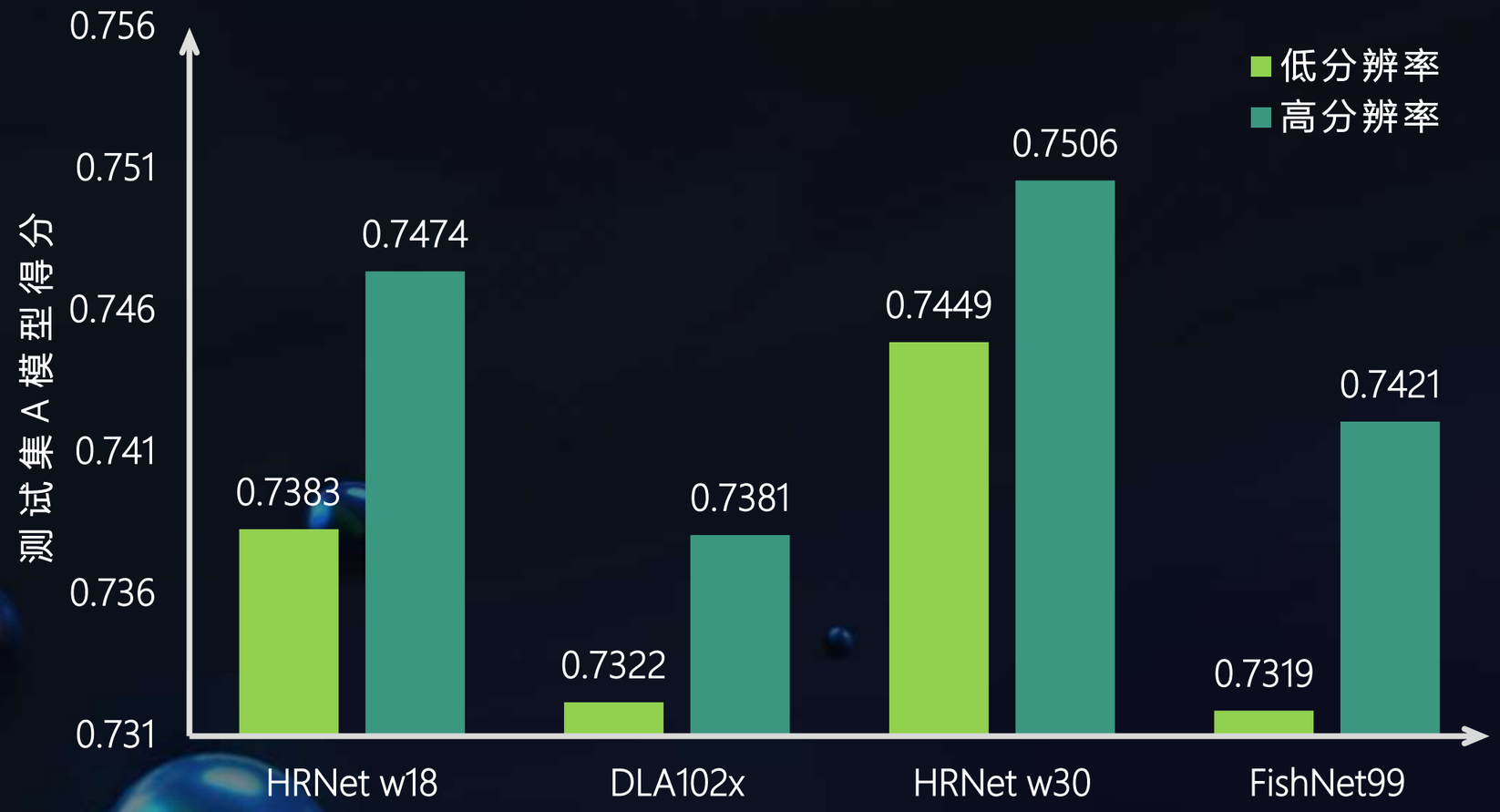


[1]: Wang, Jingdong, et al. "Deep high-resolution representation learning for visual recognition." In TPAMI. 2020.
[2]: Sun, Shuyang, et al. "FishNet: A versatile backbone for image, region, and pixel level prediction." In NeurIPS. 2018.
[3]: Yu, Fisher, et al. "Deep layer aggregation." In CVPR. 2018.

2.1 特征提取

➢ 分辨率

- 测试集A

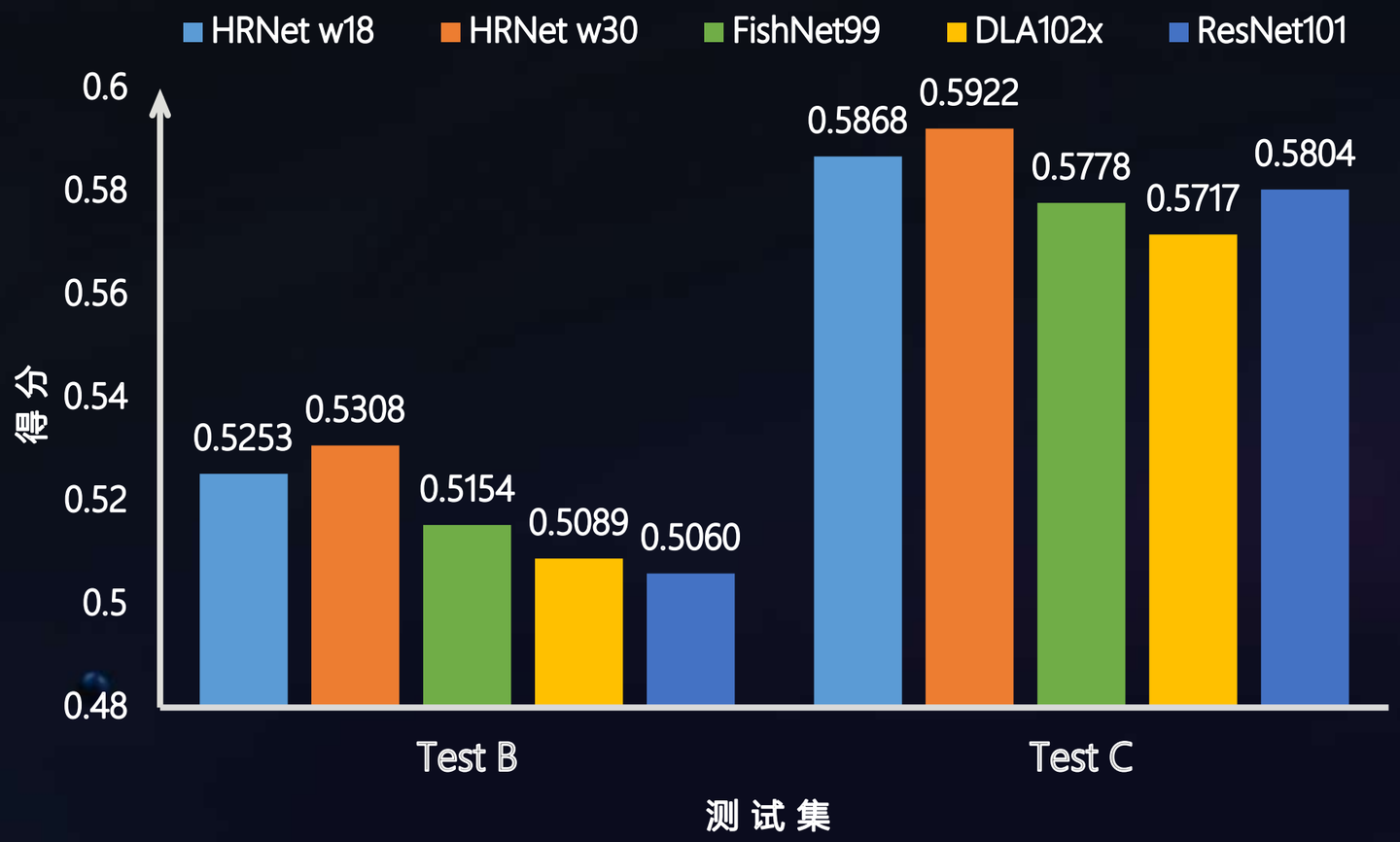


2.1 特征提取

➢ 最终表现

- 使用FP16保存模型

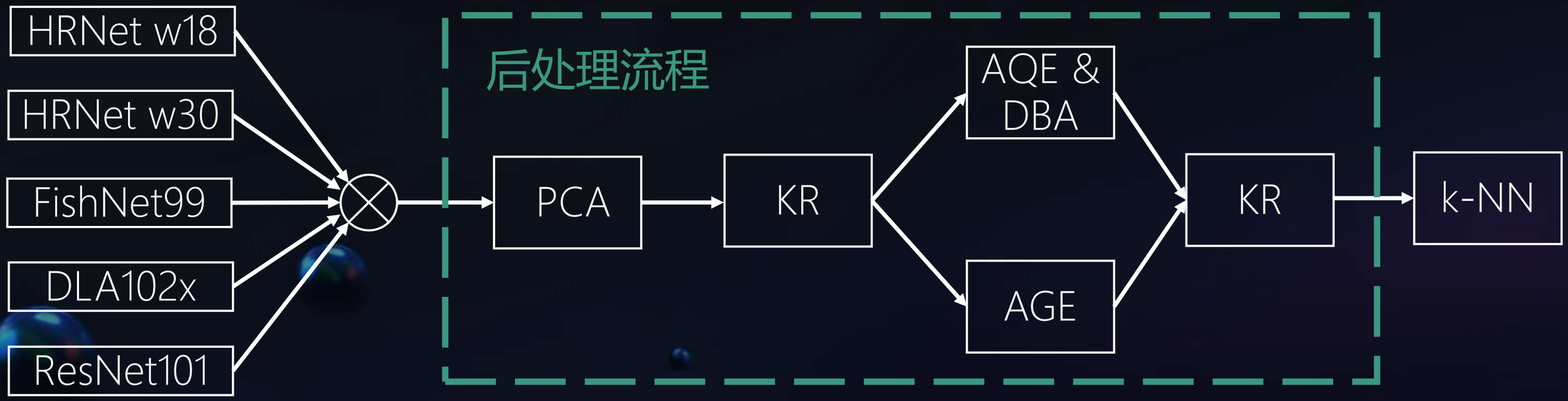
模型	大小（MB）
HRNet w18	62
HRNet w30	93
FishNet99	43
DLA102x	74
ResNet101	106
总和	378



2.2 检索排序

➢ 后处理算法

- KR [1]：原始图像和数据库中的同类图像可以互相检索
- AQE [2] & DBA [3]：使用同类图像特征增广原始图像特征

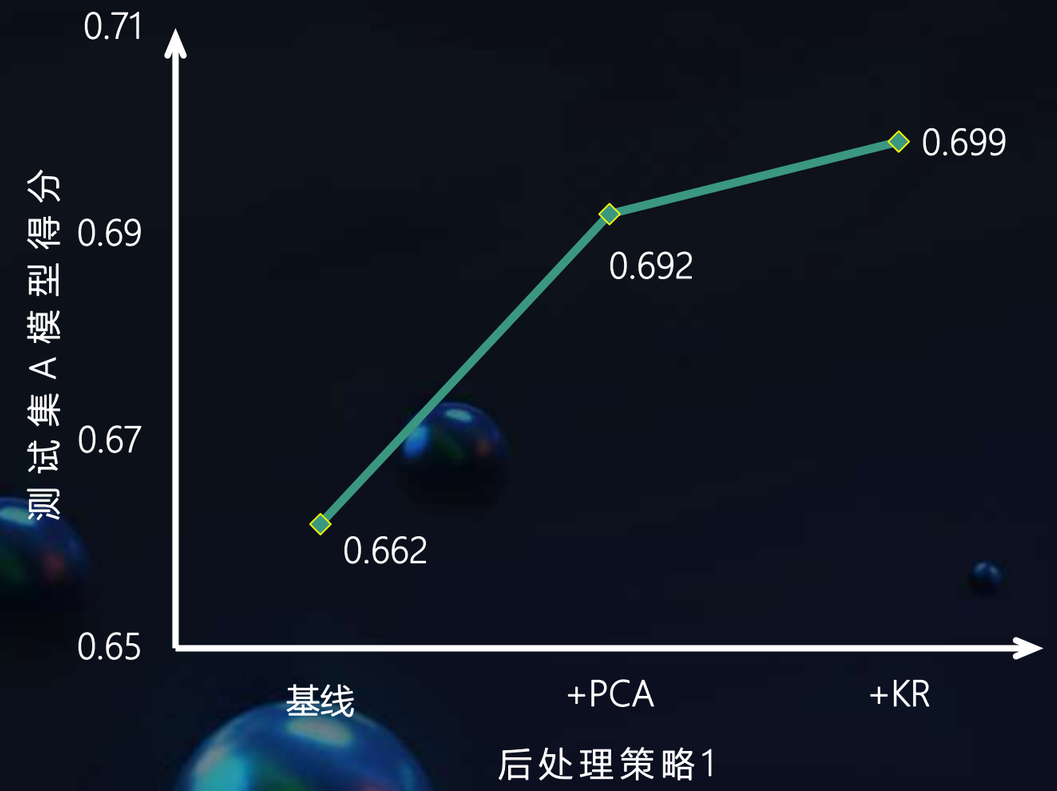


[1]: Zhong, Zhun, et al. “Re-ranking person re-identification with k-reciprocal encoding.” In CVPR. 2017.
[2]: Chum, Ondrej, et al. "Total recall: Automatic query expansion with a generative feature model for object retrieval." In ICCV. 2007.
[3]: Arandjelović, Relja, and Andrew Zisserman. "Three things everyone should know to improve object retrieval." In CVPR. 2012

2.2 检索排序

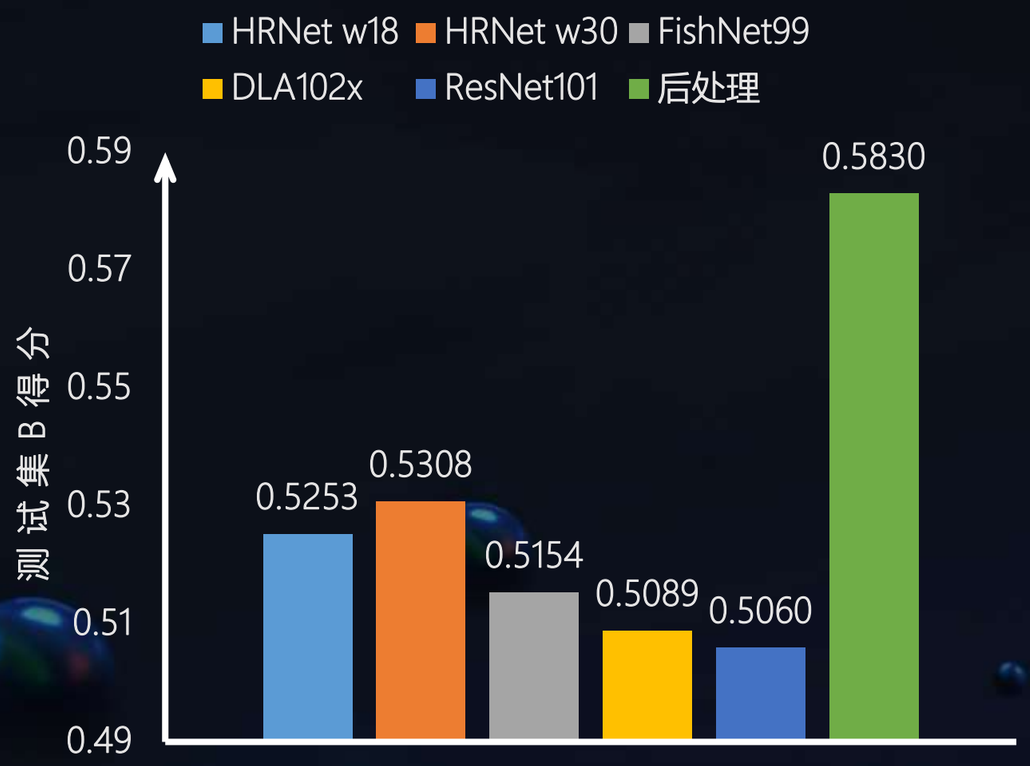
➢ 后处理流程

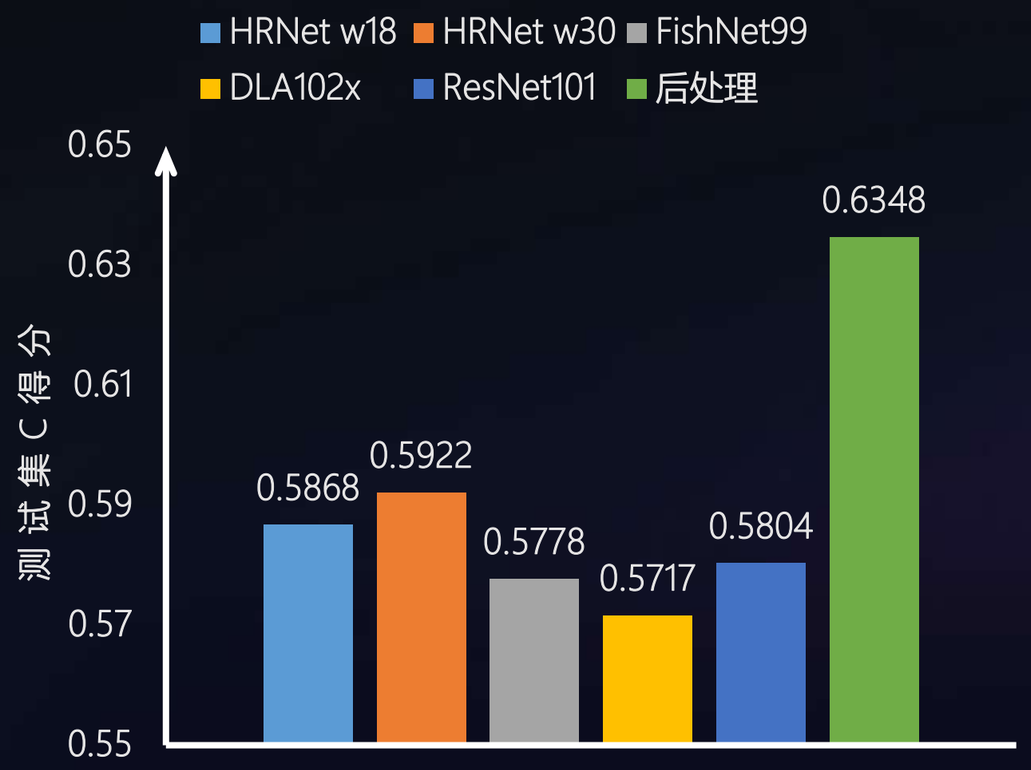
- 左图：策略1，ResNet50
- 右图：策略2，DLA102x + ResNet101 + FishNet99



2.2 检索排序

➢ 最终结果


HRNet w18
HRNet w30
FishNet99
DLA102x
ResNet101
后处理
测试集B得分
0.59
0.57
0.55
0.53
0.51
0.49
0.5253
0.5308
0.5154
0.5089
0.5060
0.5830


HRNet w18
HRNet w30
FishNet99
DLA102x
ResNet101
后处理
测试集C得分
0.65
0.63
0.61
0.59
0.57
0.55
0.5868
0.5922
0.5778
0.5717
0.5804
0.6348

2.3 代码结构

➢ 精简图示

根目录	目录或文件	代码说明
src/	dataset/	数据集相关
	dataset/datasets.py	数据增广代码
	model/	模型定义
	model/CGD_margin_loss.py	CGD架构
	tests/	提取特征
	train_retrieval/	训练模型
	losses.py	定义损失函数
post_process/	rank.py	后处理流程
	util.py	KR & AQE & DBA

3.1 方案优势

✓CGD方法

✓数据增广策略

✓后处理流程

3.2 相关讨论

- ✓有高底层语义混合的模型架构表现更好
- ✓用于人脸识别的损失函数优于基于样本对的损失函数
- ✓后处理参数选择对结果影响较大
- ✓后处理流程设计至关重要

3.3 赛后致谢

✓感谢华为主办方的付出

✓感谢队友之间的通力合作

谢谢
Thank you.