

移动电商中的图像算法 应用

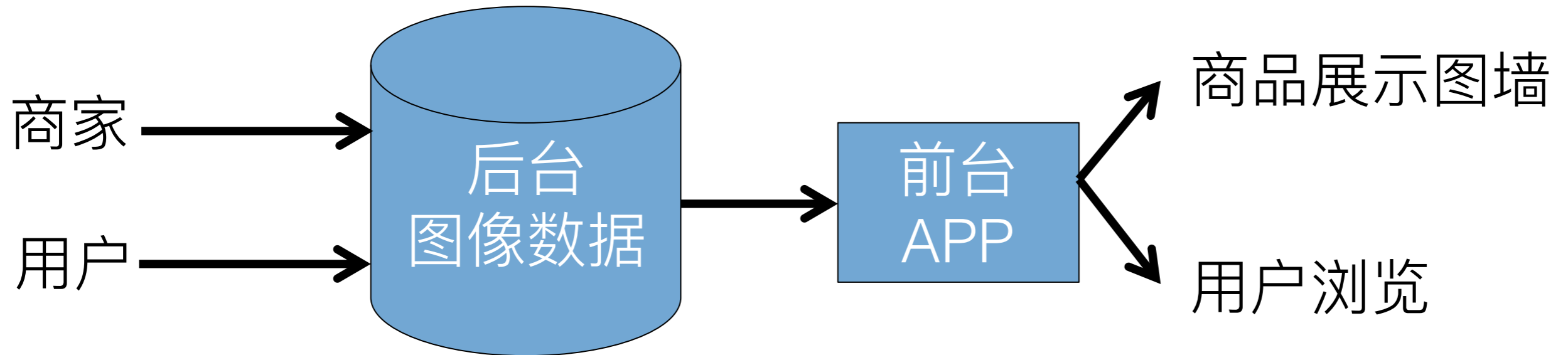
张洪明(民达) / 美丽联合集团 技术专家

- 美丽联合集团：女性时尚媒体和时尚消费平台，通过整合现在已有的资源，包括电商、社区、红人、内容等，来服务于不同的女性用户
- 旗下包括：蘑菇街、美丽说、uni引力等产品与服务，集团拥有超2亿注册用户
- 蘑菇街：定位于年轻女性用户的时尚媒体与时尚消费类App，核心用户人群为 18-23 岁年轻女性用户

提纲

- 1. 电商中为什么需要图像处理技术？
- 2. 图像算法的运行模式：云端和移动端
- 3. 应用实例：图像标签 / 人脸特征点跟踪
- 4. 未来的方向

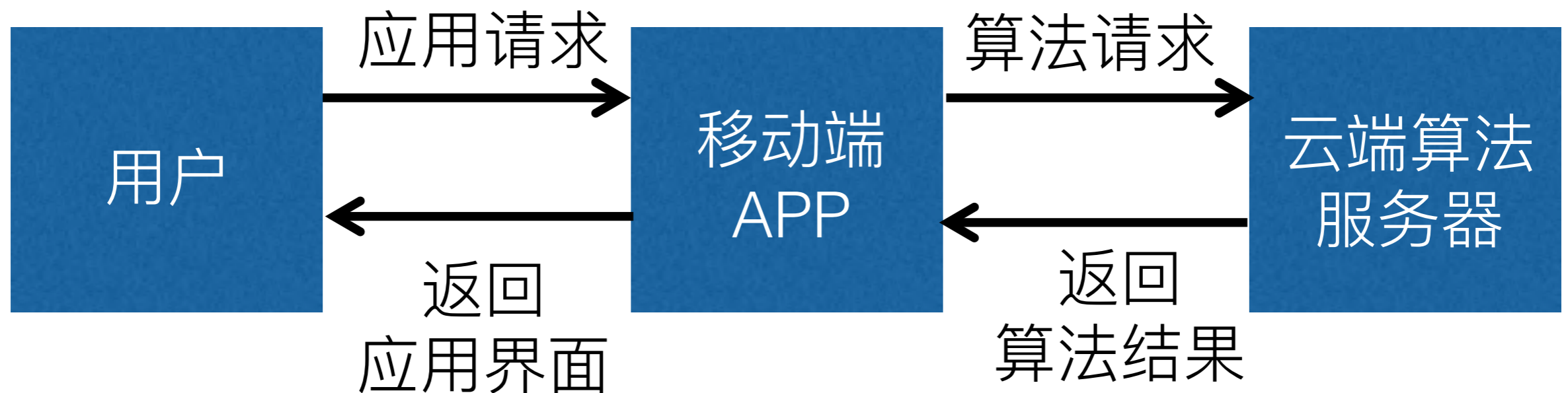
电商中的图像数据



- 海量的图像数据
- 商品信息结构化：类目 + 属性
- 丰富的业务场景

第1种运行模式

- 在云端执行算法



- 例子：图像自动标签，运用于相机功能

图像标签

通过图像算法，自动识别图片内容，如场景、风格、主体名称/颜色/图案等



背带裤 蓝色

护肤品

旅行箱

标签定义(1)

- 三个维度：图像类目、颜色、元素
 - 类目：72类，涉及服装、化妆品、生活类
 - 颜色：10类(白色/黑色/粉色/红色/紫色/绿色/黄色/蓝色/灰色/棕色)
 - 元素：9类(条纹/格子/字母/卡通/碎花/波点/迷彩/蕾丝/破洞)
- 更细维度
 - 化妆品品牌：33类

标签定义(2)

- 类目：72类标签

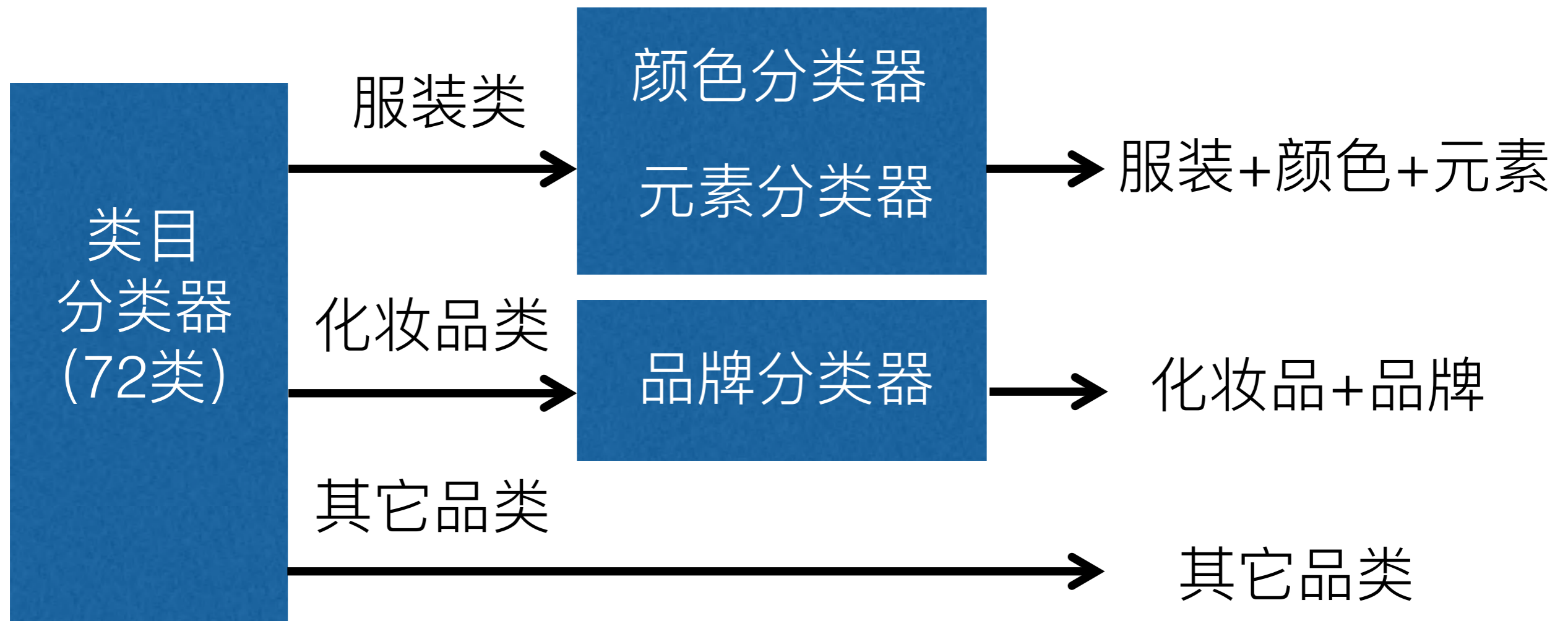
0-9: 衬衫, T恤, 背心, 蕾丝衫, 卫衣, 毛衣, 风衣, 棒球外套, 牛仔外套, 黑色小脚裤
10-19: 牛仔裤, 休闲裤, 背带裤, 短裤, 连衣裙, 半身裙, 泳装, 口红, 粉饼, 睫毛膏
20-29: 香水, 眼妆, 唇妆, 美甲, 新发型, 自拍杆, 街拍, 风景, 野炊工具, 美食
30-39: 喵星人, 汪星人, 男神, 女神, 合影, 家居, 平跟鞋, 腮红, 帽子, 手表
40-49: 墨镜, 配饰, 高跟鞋, 西装, 面膜, 情侣装, 多肉植物, 花束, 花瓶, 摆件
50-59: 杯子, 零食, 蛋糕, 餐具, 厨具, 家具, 灯具, 收纳, 笔记本, 手机
60-69: 相机, 单肩包, 钱包, 手拿包, 手提包, 双肩包, 指甲油, 旅行箱, 玩偶, 积木模型
70-71: 书, 护肤品

- 化妆品品牌：33类标签

Kiehls、It's-skin、Fresh、Sulwhasoo、Estee-Lauder、Innisfree、Clarins、Anessa、Jayjun、Fancl、Dior、3CE、Laneige、Armani、BanilaCO、Shiseido、SK-II、VDL、Lancome、Lamer、YSL、Clio、Missha、Decorte、NARS、L'Occitane、LV、CPB、Veet、Jurlique、Benefit、Clinie、Amore

解决思路

- 基于深度学习(CNN)的分类器
- 采用级联结构



算法训练

- 训练样本：近60万标注样本
- 网络结构：ResNet（残差网络）
- 实验结果

	类目 (72类)	颜色 (10类)	元素 (9类)	化妆品牌 (33类)
准确率	92.46%	93.92%	91.10%	96.68%

算法在云端的部署

- 评估网络结构：准确率、召回率、性能、内存、模型大小

上衣数据集（14类）分类算法比较

算法	SPPNet	ResNet-50	ResNet-18
准确率	96.20%	97.50%	96.91%
召回率	88.09%	96.87%	97.25%
性能	9.6ms	22.9ms	14.4ms
内存	563M	501M	361M
模型大小	301M	94.4M	90.8M

- 框架的选择

- 标签服务由将 caffe 框架切换到 mxnet 框架，mAP 相当
- 内存使用：caffe 3167M，mxnet 1041M（优化 67%）

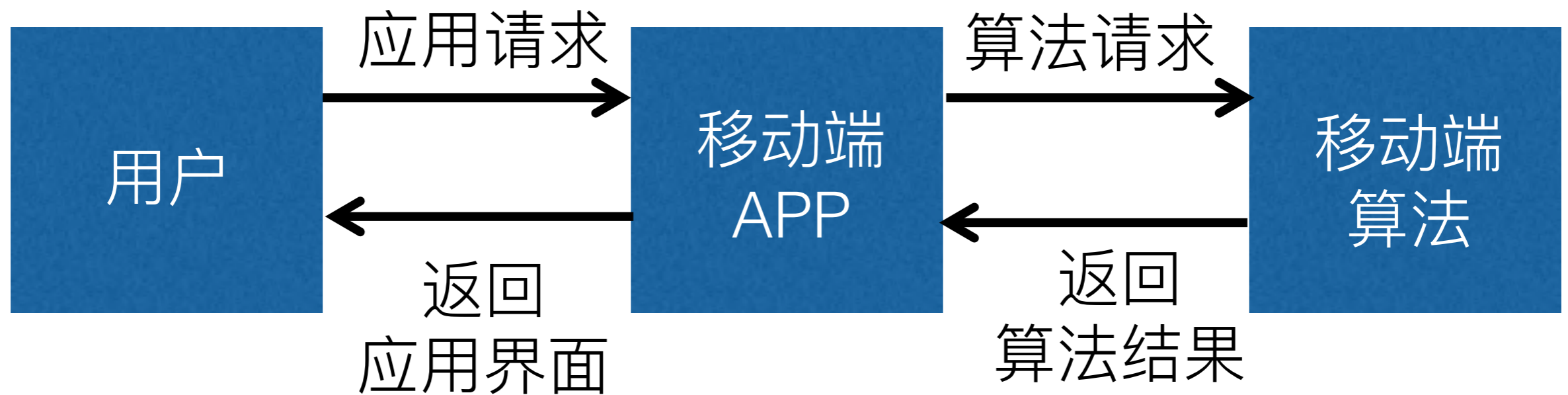
演示：图像标签

- 手机app上的演示



第2种运行模式

- 在移动端执行算法



- 例子：人脸特征点跟踪，运用于直播

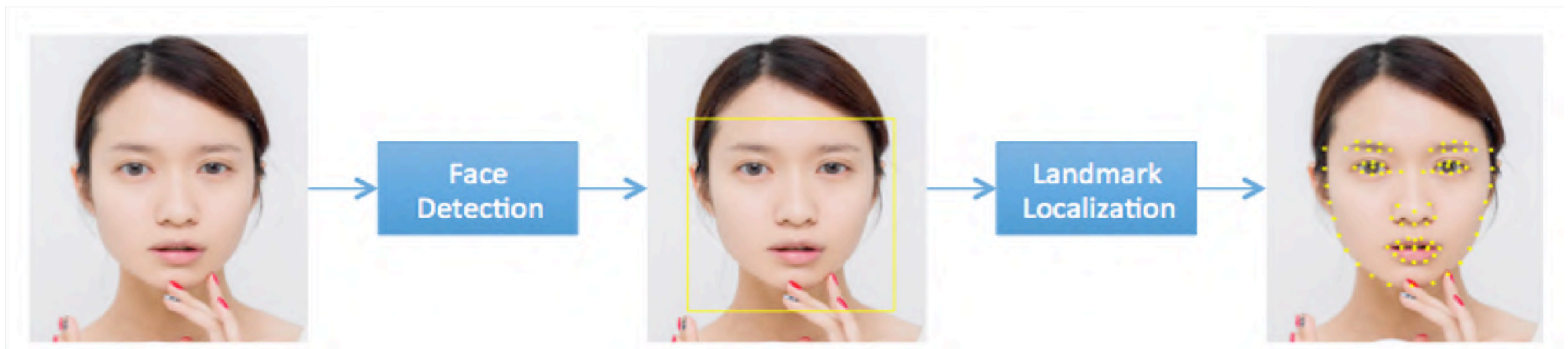
人脸相关技术概述

- 近几年来，基于深度学习的人脸技术取得显著成绩
 - 基于Faster-RCNN的人脸检测算法在FDDB上名列第一
 - Google的FaceNet在LFW上取得了99.63%的人脸识别成绩
 - Face++基于CNN的人脸特征点定位算法，在公开数据集300W上取得第1名

挑战：深度学习对大量计算资源的需求，以及复杂的模型(几十MB到几百MB)，使得上述技术难以应用在移动端

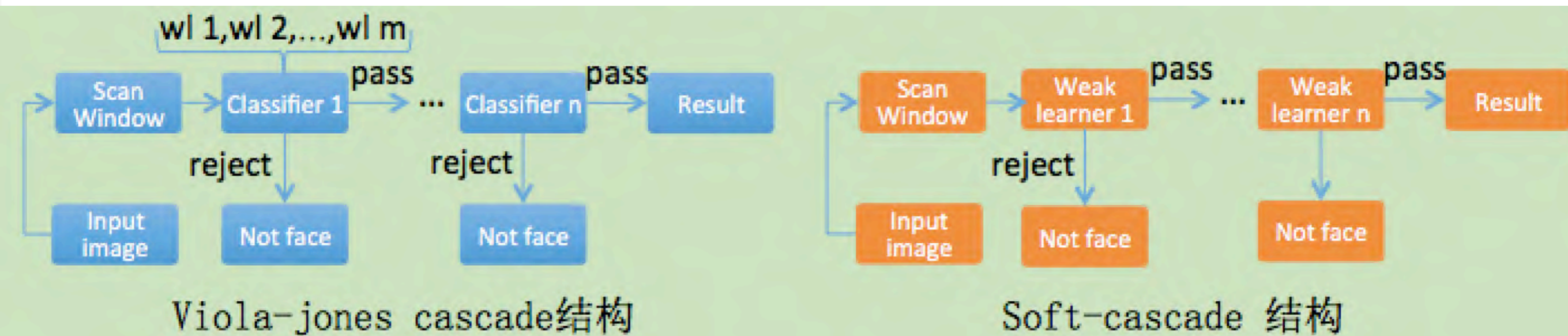
人脸相关技术概述

- 传统算法
 - 效果上不及深度学习算法
 - 计算资源消耗少、模型简单，更适合应用于目前的移动设备
- 随着处理器的发展和针对深度学习的优化技术逐渐成熟，基于深度学习的人脸技术势必在未来更加普及



移动端的人脸检测技术

- 难点：保证检测准确率的同时，需要有非常高的检测速度，以便留给后续模块足够的处理时间
- 基础算法：PICO--Object detection with pixel intensity comparisons organized in decision trees[1]



[1] Markuš N, Frljak M, Pandžić I S, et al. Object detection with pixel intensity comparisons organized in decision trees[J]. arXiv preprint arXiv:1305.4537, 2013.

人脸检测技术

- 图像特征：Normalized Pixel Difference (NPD) feature[2]

$$f(x, y) = \frac{x - y}{x + y}$$

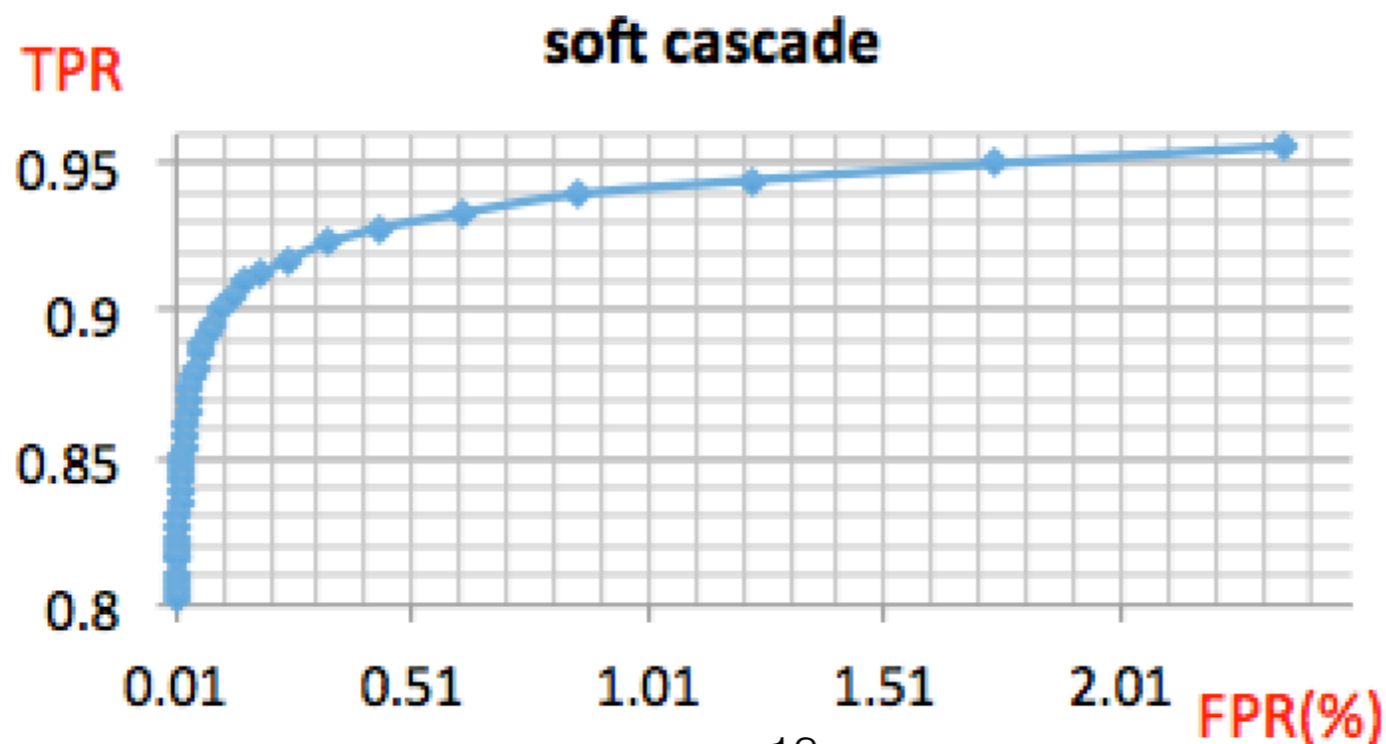
提高算法稳定性：光照鲁棒

[2] Liao S, Jain A K, Li S Z. A fast and accurate unconstrained face detector[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 38(2): 211-223.

- 训练数据集：添加公司自有人脸数据（约10万量级）
- 优化方法
 - 针对移动端CPU的特点进行了定点化
 - LUT(Look Up Table)

人脸检测SDK

- 应用场景：设备计算能力较弱，实时性要求高的场景。
- SDK容量：约300kB
- 检测速度：15ms/frame（直播，分辨率940*560）
- 适用范围：平面内 ± 30 度，平面外 ± 45 度（直播）



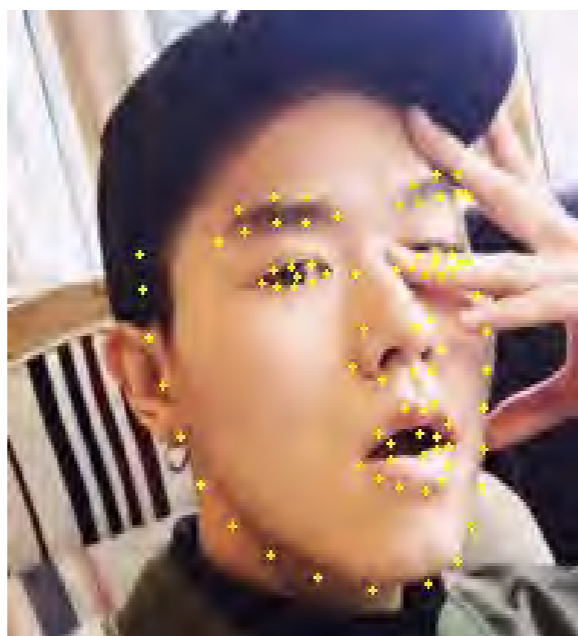
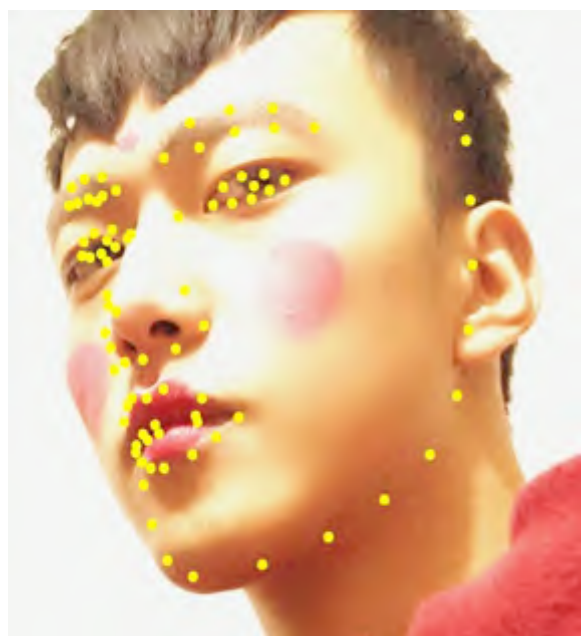
人脸特征点定位：算法开发

- 采用ERT[3]算法，做了一系列优化
 - 使用OpenMP并行技术对训练阶段进行了优化
 - 将PD特征替换为NPD特征，提高了算法对光照的鲁棒性
 - 针对实时直播，利用视频中前后两帧关系修正特征点预测结果
 - 通过参数定点化、huffman编码等方法对模型进行压缩，压缩后的模型容量约为原始模型的35%

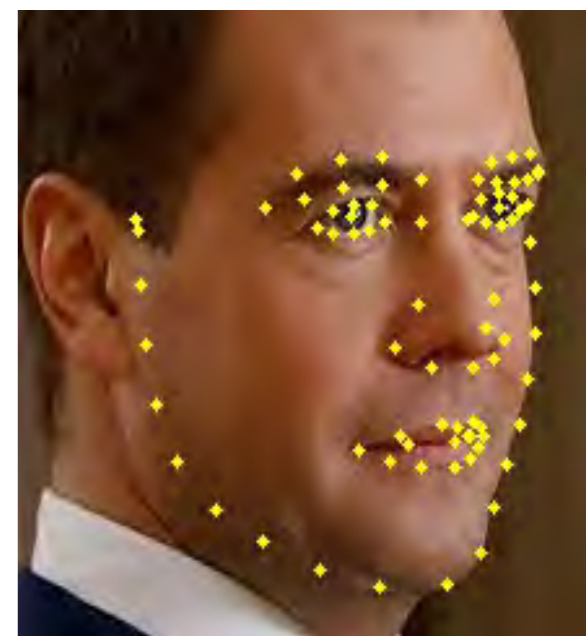
[3] Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees[C]Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 1867-1874.

人脸特征点定位SDK

- SDK容量：约4MB
- 特征点数量：83点
- 检测速度：约2ms / frame（直播，分辨率940*560）
- 适用范围：平面内 ± 30 度，平面外 ± 45 度

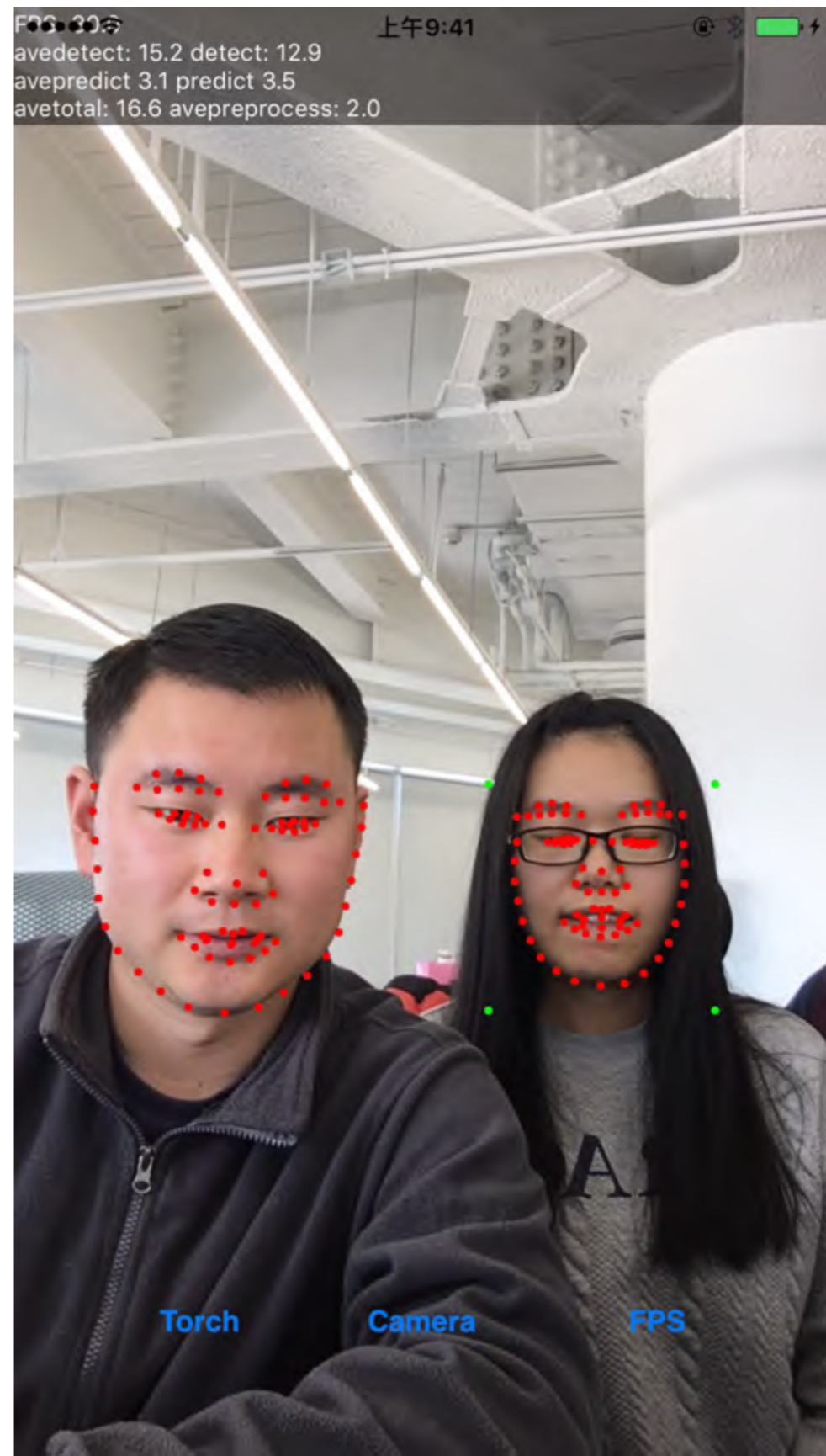


20



演示：人脸特征点跟踪

- Video



总结

- 算法部署的两种模式：云端和移动端
- 应用驱动的图像算法开发
- 根据场景中的数据，进行算法和模型的迭代
- 未来的方向
 - 深度学习方法的优化，适用于移动端
 - AR、VR：商品展示、用户体验

THANKS!





蘑菇街技术博客

敬请关注
“蘑菇街技术博客”
公众号



敬请关注
“美丽联合数据技术”
公众号