

苏宁图像智能分析平台及实践

二〇一六年六月



内容

1

图像智能分析系统组成及平台介绍

2

图像智能分析应用

3

技术案例: 图像自动标注方法

图像与大数据的关联

1. 图像是一种重要的数据载体。
2. 图像数据的爆炸，“图像大数据”成为现实。
3. 图像分析技术的进步，挖掘图像大数据成为可能。

图像大数据对苏宁的价值



线上电商



线下店商



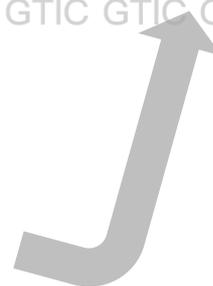
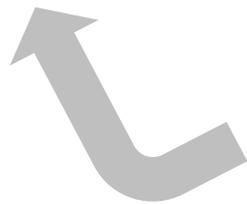
图像智能分析系统 组成及平台介绍

图像智能分析系统组成:

智能监控系统

线上应用系统

离线算法平台



线上应用系统:

应用

拍照购

价签购

人脸分析

Logo购

...

数据

图像数据解析

结果整合

数据库管理

接口

图像算法接口

推荐平台接口

搜索平台接口

算法

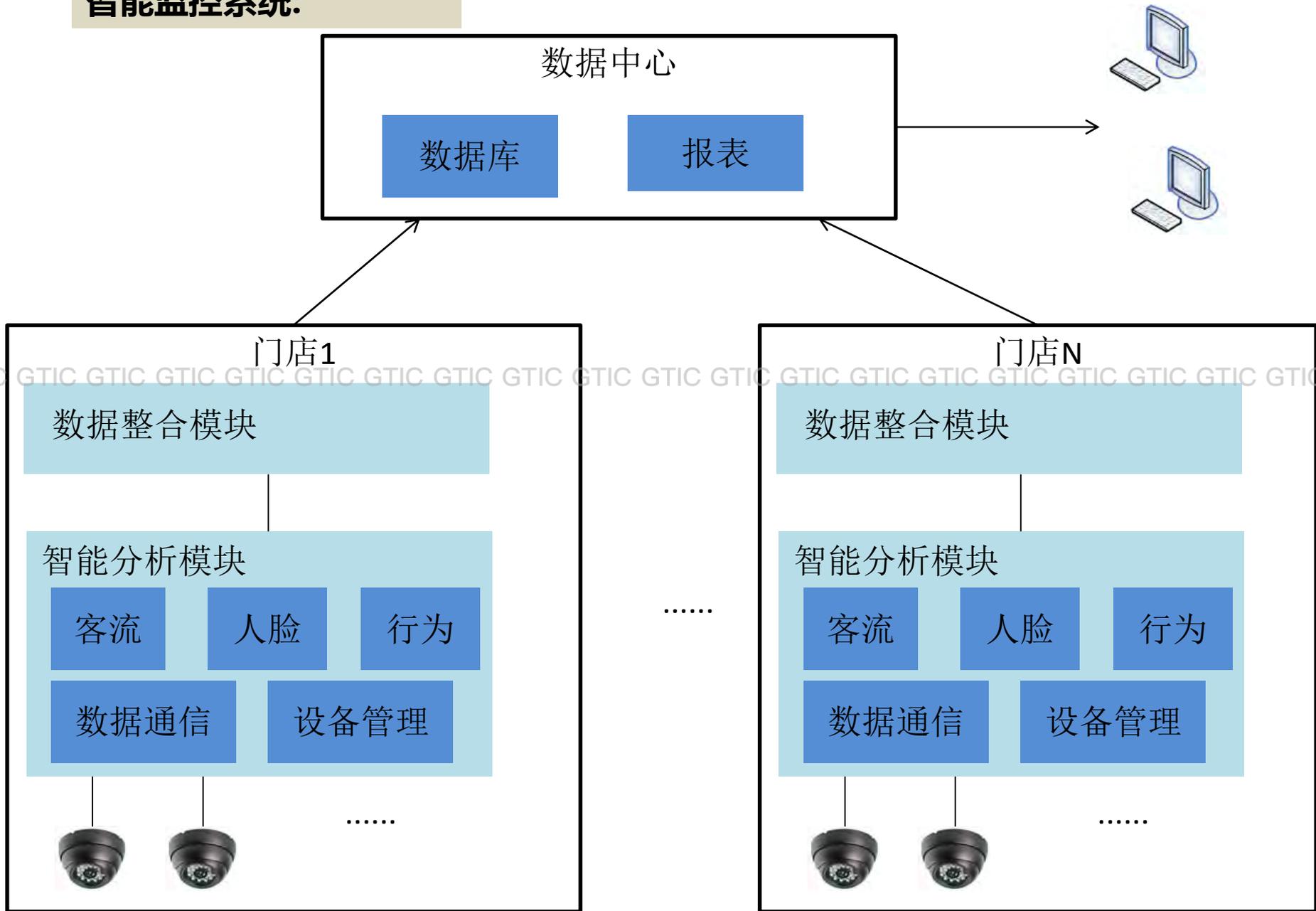
商品识别

文字识别

人脸分析

目标检测

智能监控系统:



离线算法平台:

效果层

商品识别

人脸算法

OCR

检测算法

...

模型层

深度学习模型训练

模型测试验证

检索模型

传统模型训练

特征提取及融合

数据层

图像获取

图像预处理

图像标注

传统特征提取

图像智能分析平台

目的: 为不同的应用需求提供一致性的算法接口, 满足各种图像智能分析需求, 为内部和外部用户提供智能分析服务。

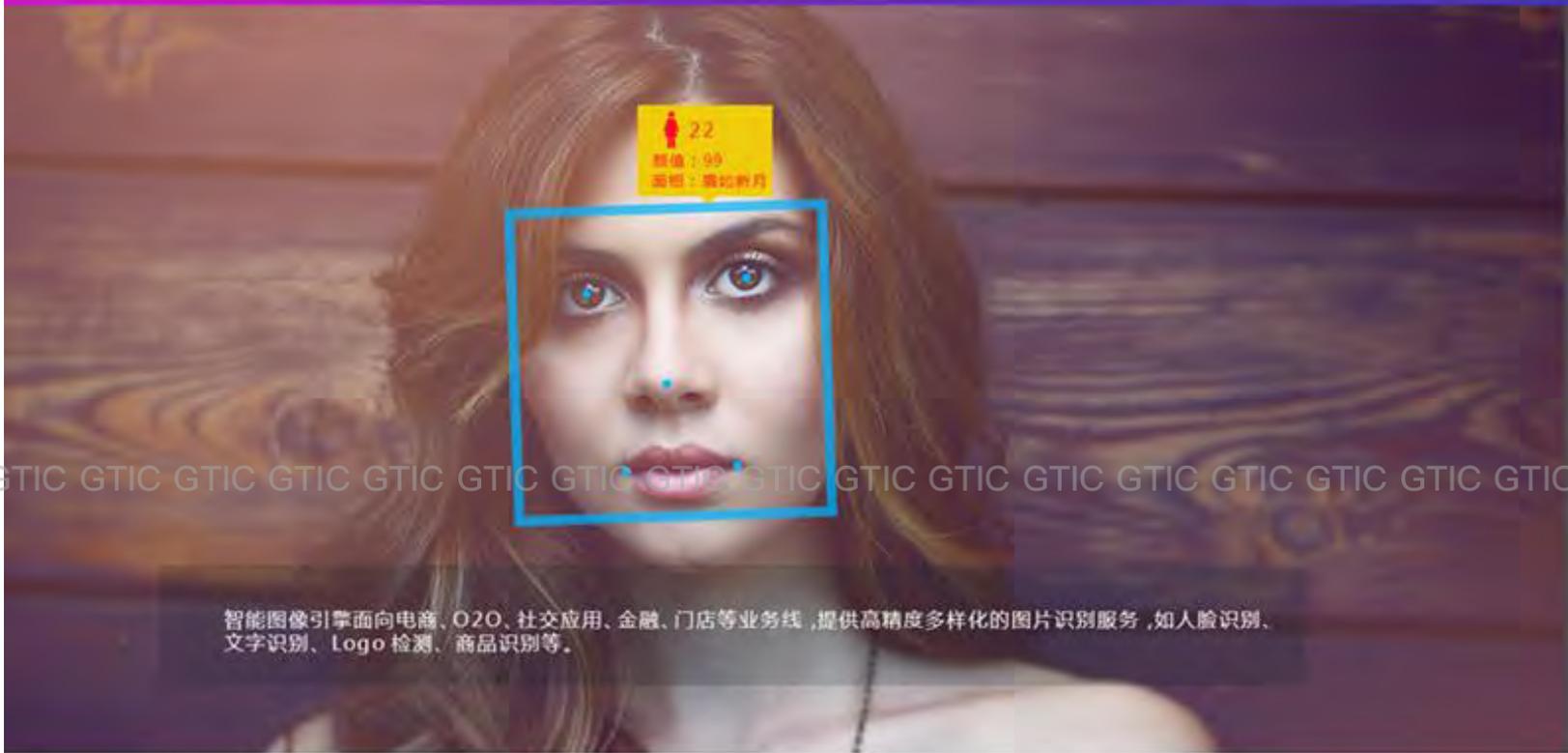
功能: 商品识别

Logo检测

人脸分析

OCR应用

.....



智能图像引擎面向电商、O2O、社交应用、金融、门店等业务线,提供高精度多样化的图片识别服务,如人脸识别、文字识别、Logo检测、商品识别等。

应用场景



电商 O2O

识别图像中商品,实现自动化精准化的商品检索和推荐,增强用户购买体验;识别图像中的 Logo,实现品牌的精准化推荐。



金融

使用文字识别 OCR 对身份证、银行卡等开户证件进行识别,减少用户输入,增强用户体验;对关键交易使用人脸识别进行刷脸支付,控制风险。



社交

对社交应用中上传的图片进行审核,高效识别色情、涉政等违规图片;对图片中人脸特征进行分析,并实现换脸、美颜、PK 大咖等娱乐体验。



门店

使用文字识别 OCR 对门店价签进行自动识别和搜索;检测和统计顾客基本信息,用于客流统计、用户分析、商品推荐等商业应用。



图像智能分析应用

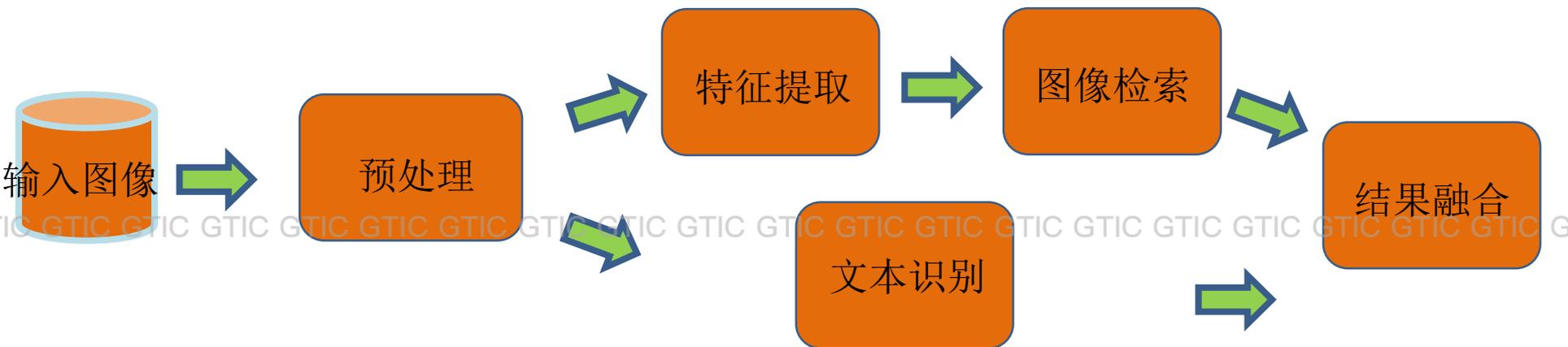
拍照购（商品识别）：介绍

以商品识别算法和图像检索系统为基础，利用深度神经网络强大的表达能力，抽取图像中商品的有效特征，并结合文本识别的结果，获取高相关度商品的推荐结果。

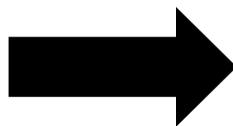
特点:

1. 支持广泛品类的商品识别(今年内基本做到全覆盖)。
2. 与文本识别相结合。
3. 较快的识别与检索速度，相关性较高的返回结果。

拍照购（商品识别）：算法系统组成



拍照购（商品识别）：使用示例



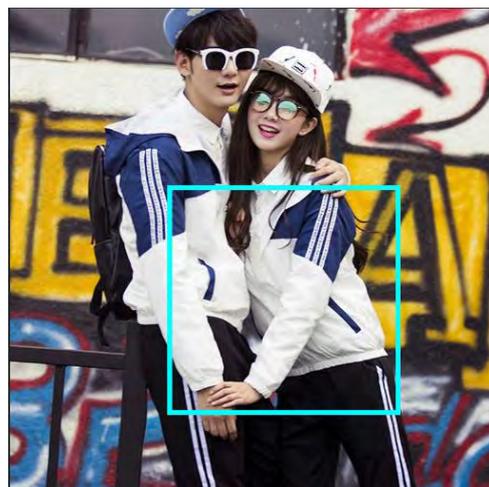
拍照购（商品识别）：优化策略

有用策略:

1. 补充数据。增加有效的真实训练样本，提升模型的表达能力，并提高检索精度。
2. 主体检测。有效检测出图像中商品主体，减少背景的干扰，提升识别精度。



拍照购（商品识别）：优化策略



智能门店监控：介绍

目的: 实现门店监控视频的智能分析，以获取准确的客流人数、客流成分组成、商品热区、顾客滞留分析等丰富的商业数据信息，进而为商业智能系统提供数据保障。

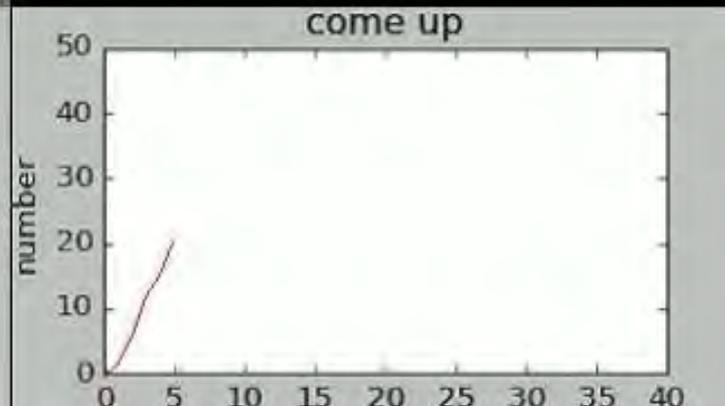
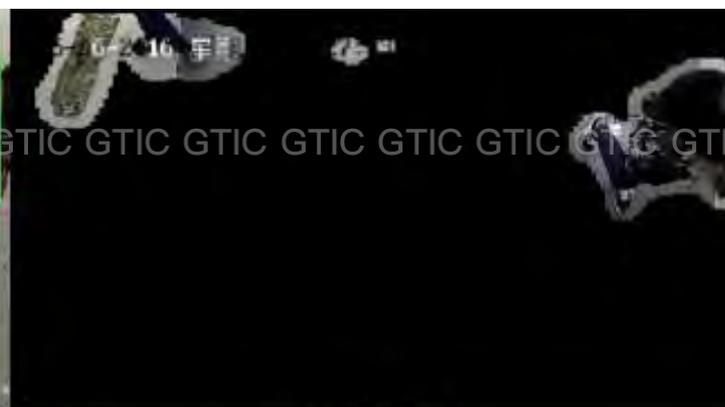
特点:

1. 能够统计不同环境、不同状态下的客流人数。
2. 能够检测不同环境下人脸，并分析其年龄、性别等信息。
3. 能够得到不同排布商家的商品热区信息。
4. 能够分析顾客的滞留，进而对顾客行为进行判断。

智能门店监控：客流统计

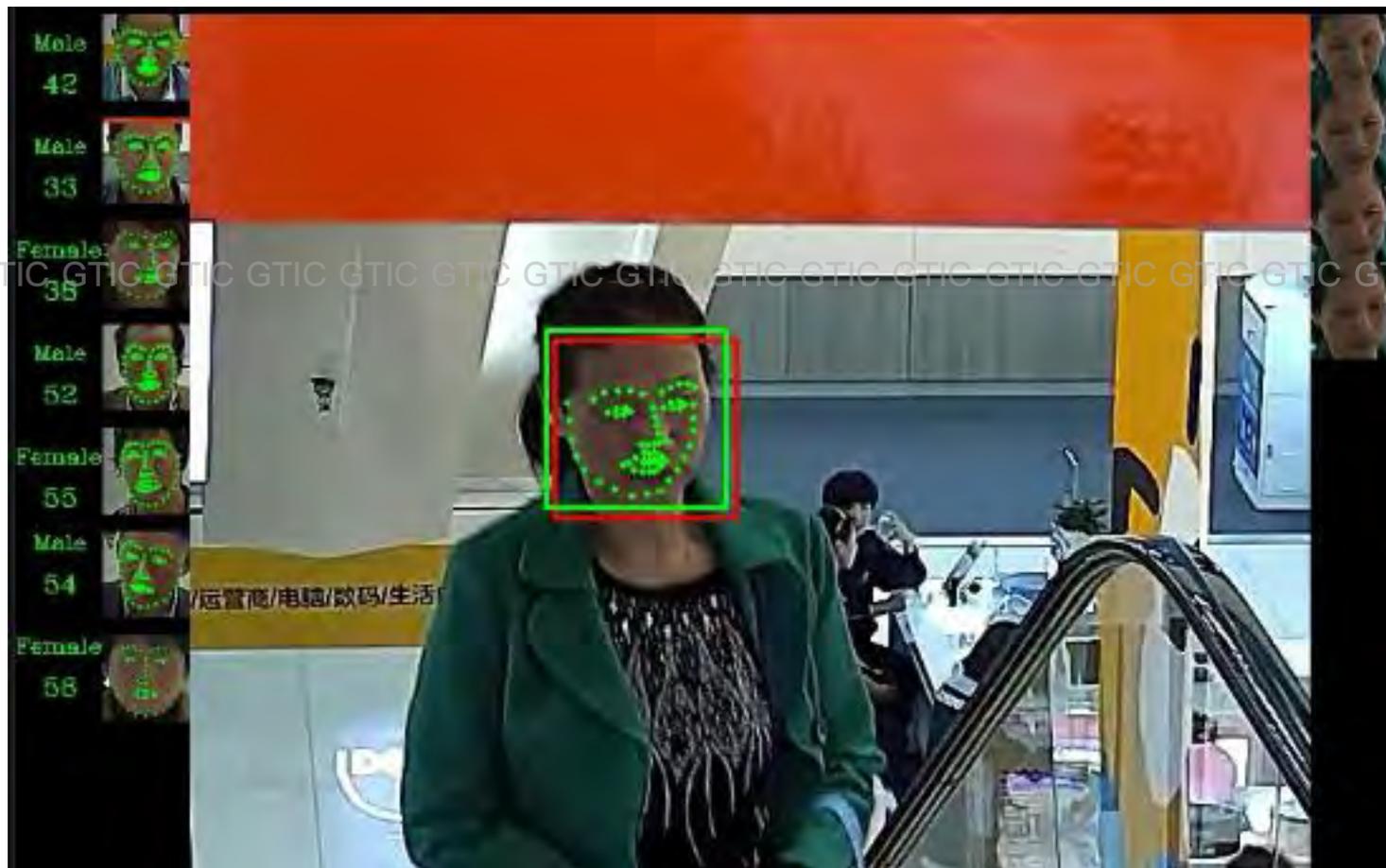
方案：

1. 采用置顶数字摄像头
2. 检测进入区域的人体
3. 跟踪并确认有效客流



智能门店监控：人脸分析

特点：利用特征点跟踪有效估计人脸姿态，利用正面人脸进行去重分析。



文本图像识别：介绍

应用:

1. 对价签等丰富文本载体进行分析，得到文本内容。
2. 对自然场景中文本进行识别，辅助物品识别。

技术方案

1. 检测图像中可能的文本区域。
2. 对文本区域进行验证和校正。
3. 迭代寻找最优的切分和识别结果。
4. 利用语言模型进行校正。

文本图像识别：示例



做工细致，符合人体工学，握持舒适，充分体现品牌精湛的制作工艺

做工细致，符合人体工学，握持舒适，充分体现品牌精湛的制作工艺

别致的手提设计

别致的手提设计

文本图像识别：价签购示例



背景介绍

— 来源

- 拍照购
- 图像识别引擎
- 深度学习

— 需求

- 海量的已标注图像数据



背景介绍

图像来源

1. 拍摄: 成本太高
2. 商品主图: ps过, 与真实场景不符
3. 网络图像: 晒单等。 噪声太严重

需求: 标注出商品图像和噪声图像

难度: 人工标注效率太低



技术现状

1. 学术圈关注度低
2. 很常见的现实问题

可能方案

1. 标注少量数据进行冷启动

标注数据影响结果

2. 半监督学习?

数据太杂乱，不符合一般半监督学习方法的假设

3. 主动学习

Good idea.

技术思路

希望

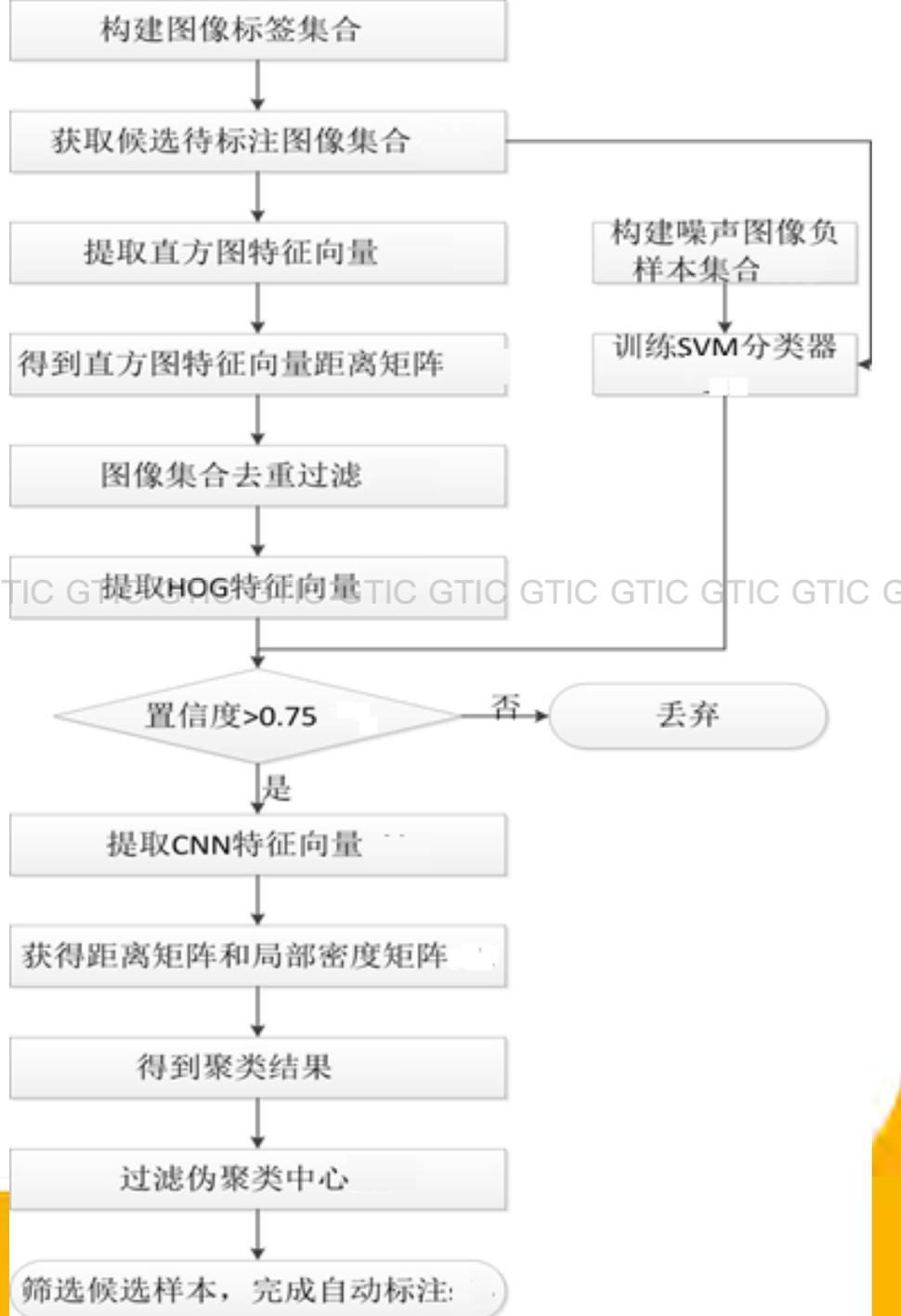
尽量少人工干预，人工干预对最终结果的影响很小。

效率尽量高，且结果无需迭代式优化。

1. 针对噪声图像多且杂，采用多重过滤
2. 选择合适的无监督聚类方法，自动获得好的聚类中心
3. 利用聚类的结果进行标注，数据量大且分布较为紧凑的聚类极有可能就是真实样本图像

技术方案

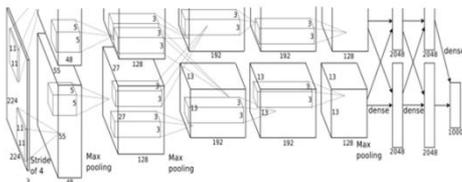
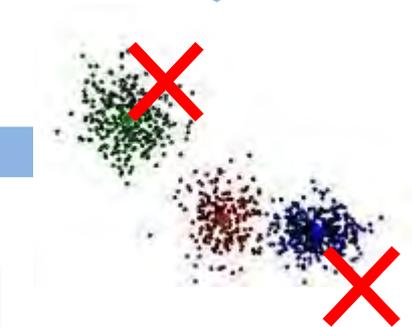
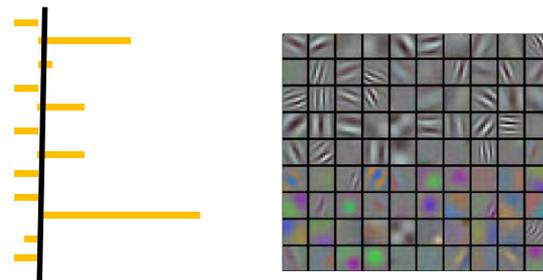
1. 图像采集
2. 重复图片过滤
3. 非商品噪声图片过滤
4. 图像聚类自动标注



End-to-end 数据过滤

搜索

舒肤佳沐浴露



progressive CNN



无监督聚类方法

常见方法:

1. 基于划分: k均值, k众数, k中心点, Canopy
2. 基于层次: BIRCH, Chameleon
3. 基于密度: DBSCAN, DENCLUE
4. 基于网格: STING
5. 基于图论: Spectral Clustering

常见问题:

1. 适用范围窄
2. 参数不易设置, 对结果影响大
3. 聚类中心选择准则过于简单或过于复杂, 适应性差
4. 高维情况下表现不好

无监督聚类方法

Clustering by fast search and find of density peaks, Alex Rodriguez and Alessandro Laio.

步骤:

1. 聚类中心选择
2. 根据聚类中心进行聚类

特点

1. 基于密度的方法，能够发掘不同形状的聚类，也能很好描述离群点
2. 提出一种寻找最佳聚类中心的新颖方法
3. 形式简单，效率较高
4. 唯一参数需要设定，具有弹性，使用于不同情况

无监督聚类方法: 聚类中心选择

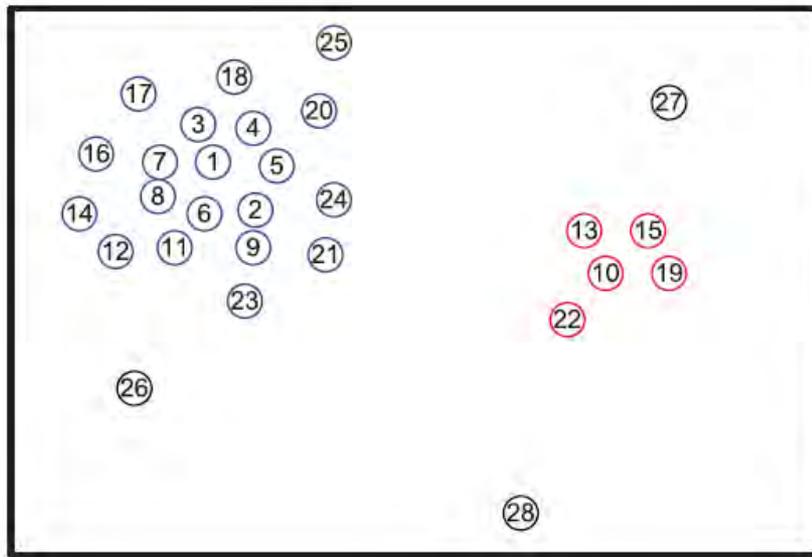
基本思想: 寻找合适的聚类中心, 使其满足

自身的密度大, 密度大于它的所有邻居

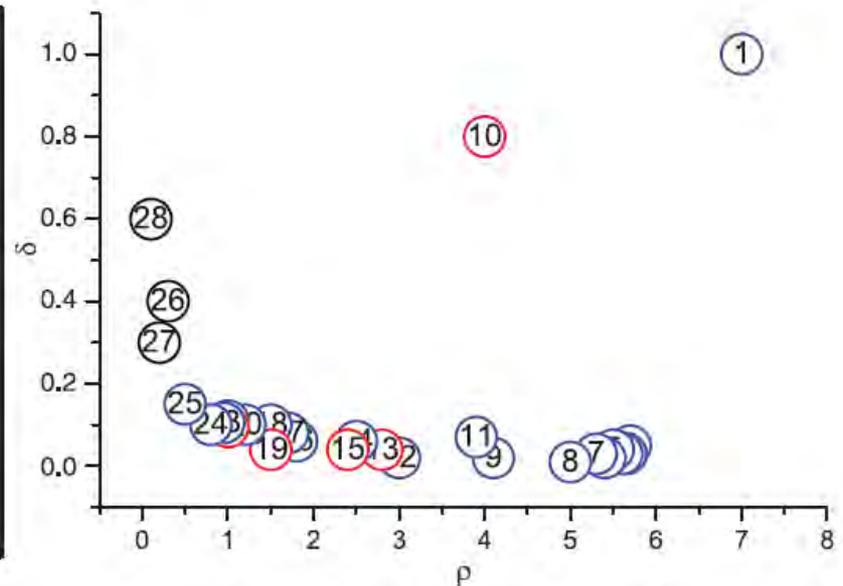
与其他密度更大的数据点的距离相对更大

定义局部密度 ρ_i 和距离 δ_i

A



B



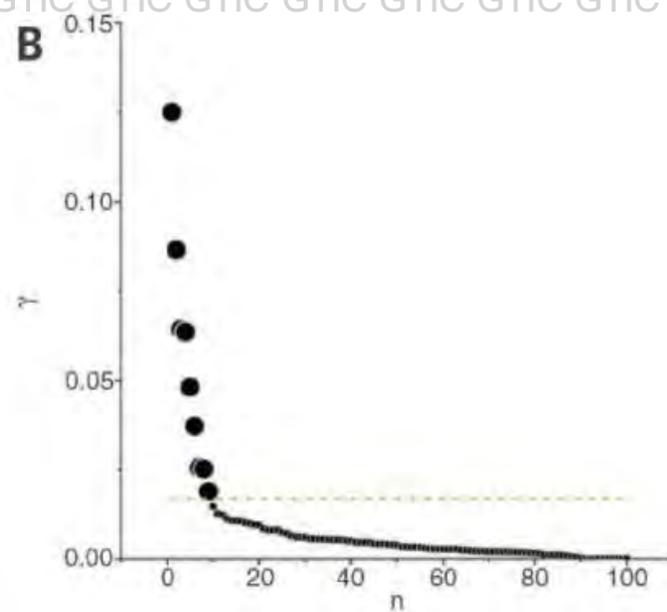
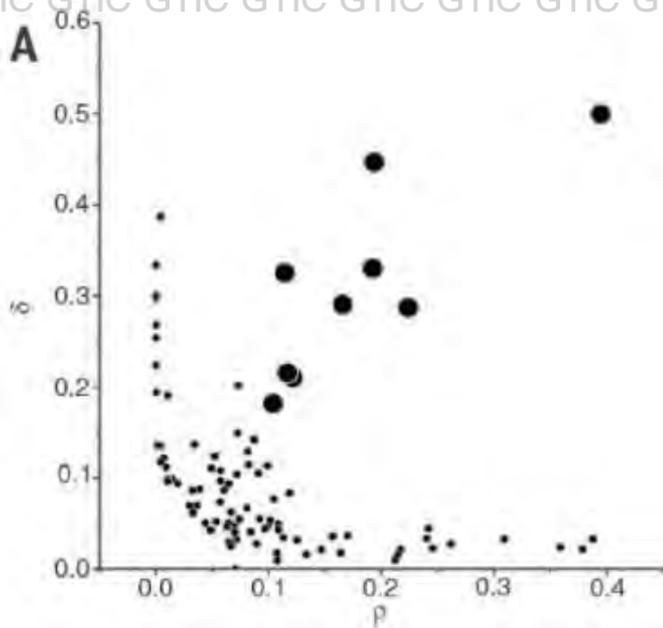
无监督聚类方法: 聚类中心选择

当聚类个数不能轻易获得时

定义

$$\gamma_i = \rho_i \delta_i, \quad i \in I_S$$

当 γ_i 越大时, 越可能是聚类中心。



无监督聚类方法: 聚类策略

步骤:

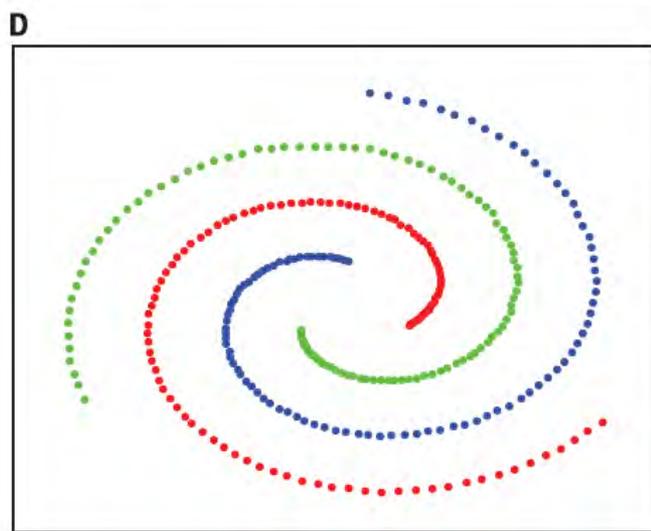
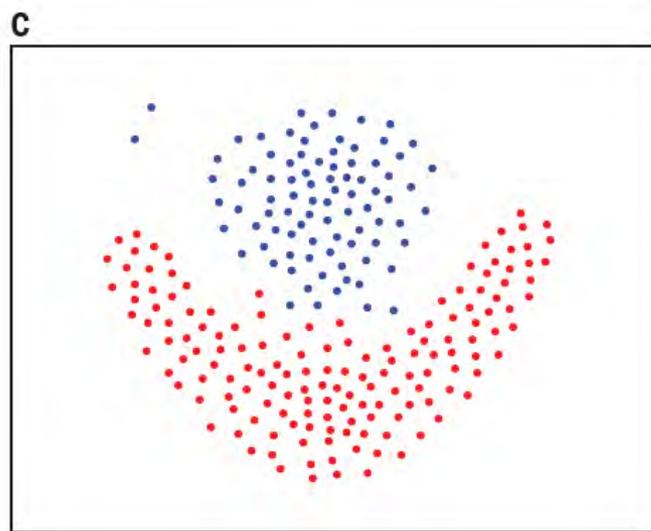
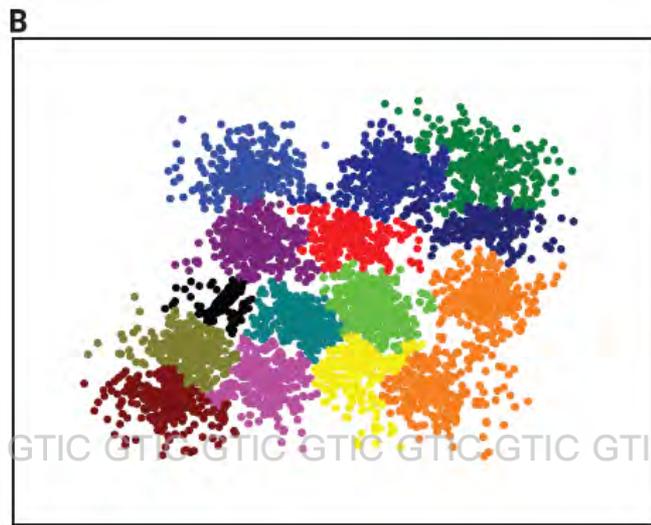
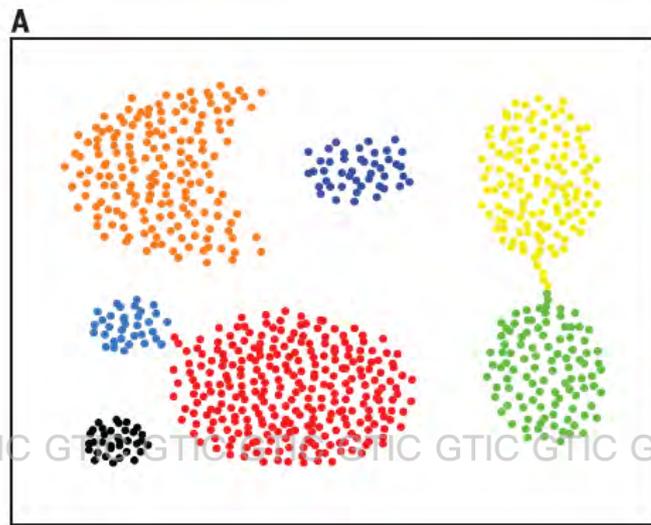
1. 初始化各个聚类, 以聚类中心为初始点
2. 根据距离判断其他非中心数据点的类别, 将其划分到某一聚类
3. 将每个聚类中的点, 进一步划分为核心数据点和边缘数据点

截断距离参数 d_c 的选取:

$$\rho_i = \sum_j \chi(d_{ij} - d_c) \quad \chi(x) = \begin{cases} 1, & x < 0; \\ 0, & x \geq 0, \end{cases}$$

1. 太大的话 ρ_i 太大, 区分度太低; 反之, 会得到细碎的聚类
2. 建议值: 使得每个数据点的平均邻居个数约为数据点总数的1%~2%

无监督聚类方法: 聚类效果



实际效果



big_55d2d0dcN
8a737c80.jpg



big_55d9bf5bN
1d78a281.jpg



big_55d9c4daN
69d3054a.jpg



big_55d9d509N
96f42cee.jpg



1-1410311F12S
27.jpg



2f5ca0f1c9b6d0
2ea87df74fcc_5
60.jpg



15594R919-7.jp
g



7411759_16415
7418126_2.jpg



9374377_13170
7272000_2.jpg



big_55d94bc2N
3f6dfb7d.jpg



big_55d930d6N
49ee3c95.jpg



big_55d958f7N
303847b0.jpg



big_55d999e2N
b5712413.jpg



1414299718611
88177.jpg



1416487171841
88746.jpg



1417484818148
58625.jpg



1418465167940
49557.jpg



1418610731881
77264.jpg



big_55dab3ceN
b290f35d.jpg



big_55daed80N
973163dc.jpg



big_55db38beN
d9e33b5f.jpg



big_55dc665bN
7c267574.jpg



1425130752445
50125.jpg



1425524877772
73455.jpg



1425549667299
30645.jpg



1427112394333
05852.jpg



1429626334774
19781.jpg



big_55dd073eN



big_55dd96d8N



big_55dd612aN



big_55dd5395N



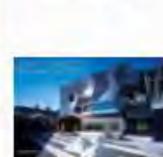
1446183292952



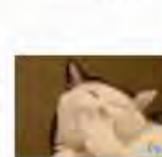
Antony_Leung_1



Arnold_Palmer_



astrid-lind_17.jp



bca3622fjw1dzjt

总结

目标: 为图像自动打标签, 为图像识别引擎准备训练数据

特点:

全过程无需人工干预。

效率很高, 和人工标注相比可提高100倍。

效果很好, 图像相关度可达95%以上

