

中国AI Unicorn4 现状报告 -- 商汤, 旷视, 依图, 云从

首先来稍微对比下这几家公司的特点:

| | 商汤 | 旷视 | 依图 | 云从 |
|------|---|------------------------------------|------------------|-----------|
| 业务领域 | 金融, 安防, 手机智能, 智能商业, 自动驾驶/车载业务, 芯片, VR/AR | 金融, 安防, 手机智能, 智能硬件(物流机器人), 智能地产 | 安防, 医疗, 智慧城市, 金融 | 金融, 安防 |
| 核心业务 | 自动驾驶, 芯片, 手机智能 | 安防, 手机智能 | 公安安防, 医疗, 芯片 | 金融 |
| 特点 | 研究院, 技术积累多 | 工程化强, 开放平台 | 产品化强, 垂直领域 | 国家队 |

以下将分别从一下5个方向进行分析:

1. 发展经历
2. 核心技术
3. 产品服务
4. 技术团队
5. 行业布局(?)

1. 商汤科技(Sensetime)

1) . 发展经历



数据来源:鲸准: <https://rong.36kr.com/project/24439>

2014年成立,到目前完成7轮融资,最后一次为C轮融资,总融资金额超过**12亿美元**,C轮之后的估值超过**45亿美元**,目前员工超过1500人,专利84.

公司高层:

| 人员图片 | 姓名 | 职位 | 简介 |
|---|-----|-----------|--|
|  | 马天军 | 联合创始人 | 马天军，商汤科技联合创始人。 |
|  | 汤晓鸥 | 联合创始人 | 汤晓鸥，商汤科技联合创始人。 |
|  | 杨帆 | 联合创始人&副总裁 | 杨帆，商汤科技联合创始人&副总裁。 |
|  | 徐立 | 联合创始人&CEO | 徐立，商汤科技联合创始人&CEO。本科、博士分别毕业于上海交通大学和香港中文大学。拥有十余年计算机视觉、人工智能基础研究和产品开发经验。在视觉领域顶级会议、期刊上发表40余篇专著，论文引用率上千，获得... 详情 |
|  | 沈徽 | 副总裁 | 沈徽，商汤科技副总裁&工程院院长、首席架构师。在2001年就加入微软，从Windows网络高级开发主管一路升为微软（亚洲）互联网工程院副院长，主要负责Bing平台和AI基础架构。 |

(信息来源: <https://www.tianyancha.com/brand/b8f8b11152>)

- 商汤集团创始人: [汤晓鸥](#):香港中文大学教授,多媒体实验室创始人.
- 联合创始人: 徐冰(董事长,法人):[汤晓鸥](#)学生(2012-2016 phd)



Bing Xu Working at SenseTime

PhD., The Chinese University of Hong Kong, 2016. B.Eng. The Chinese University of Hong Kong, 2012. From 2012 to 2016, **Bing** Xu was an PhD student in the Multimedia Lab. His research interests include computer vision and machine learning.

- 联合创始人: [徐立](#)(CEO):港中文[贾佳亚](#)(目前加入腾讯优图([信息来源](#)))老师的学生(2007-2010phd,2010-2013 postdoc)
- 联合创始人: 马天军
- 联合创始人: 杨帆(副总裁)

在2006-2014年间任职微软亚洲研究院创新工程组，负责视觉算法的产品转化。

信息来源: <http://www.xtecher.com/Mobile/Article/view?aid=3228>

- 研发总监: [林倞](#)(CTO):中山大学教授,研究方向主要包括人机智能交互,视频分析,行为分析等(<http://www.linliang.net/index.php/home/research/>).
- 副总裁: 沈徽(工程院院长.首席架构师):

沈徽早在2001年就加入微软，从Windows网络高级开发主管一路升为微软（亚洲）互联网工程院副院长，主要负责Bing平台和AI基础架构。资料显示，沈徽在知识图谱、机器人、信息检索、机器学习、分布式存储和计算引擎、无线网络等方面有深入的理解和第一手的专业知识。

信息来源: <http://36kr.com/p/5123584.html>

- 商务副总裁: 柳钢

2). 核心技术

来源: 商汤官网: <https://www.sensetime.com/core#0>

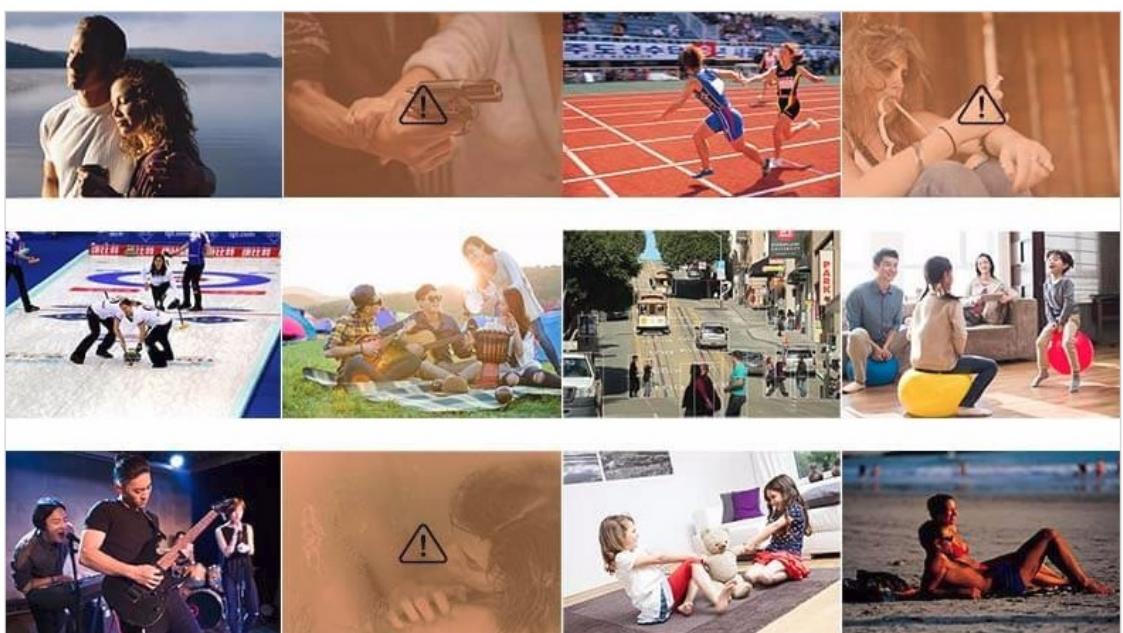
1. **人脸技术:** 检测跟踪, 关键点定位(21, 106, 240), 人脸识别, 人脸属性, 人脸活体(用户配合/不配合两套方案) ==> 衍生人脸聚类(人脸embedding), 身份验证(人脸识别+人脸活体), 人像美妆/美颜(关键点)

应用场景: 手机(相册, 自拍), 身份认证(金融, 安防), 用户画像(广告)

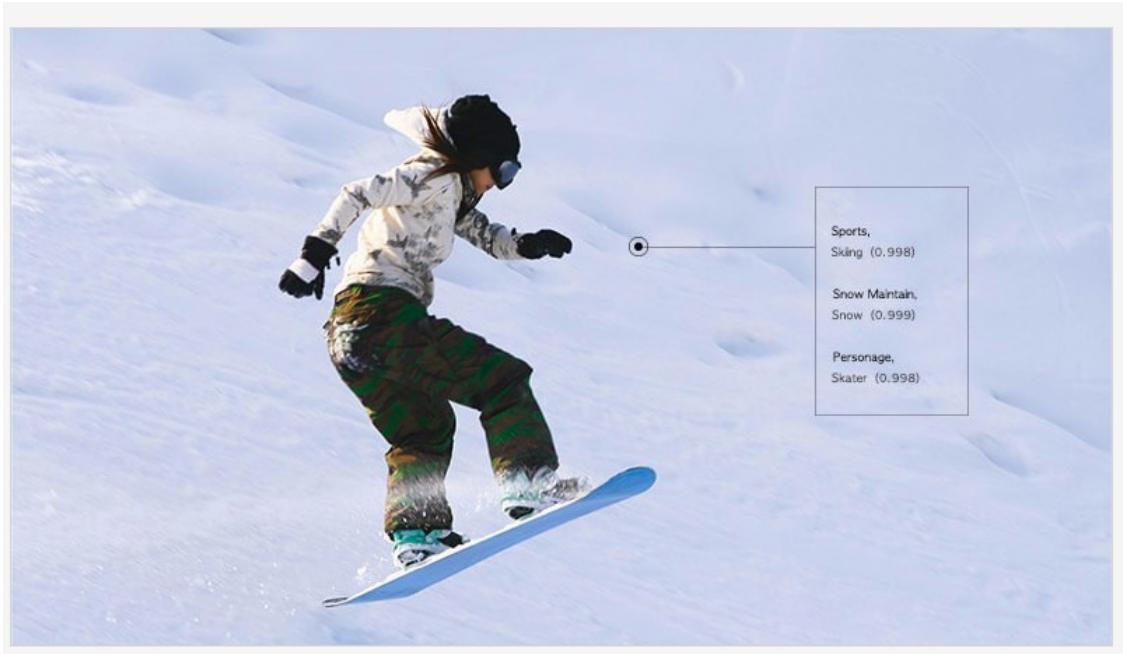
2. **通用图像识别:** 服装属性, 物体检测, 场景识别, 车型识别, 文字识别(票据, 卡证, 车牌)

3. **视频挖掘:**

- 视频内容审查(视频分类+人脸识别+视频指纹+场景识别);



- 短视频标签(视频指纹+场景识别+物体识别+行为识别+...);



- 视频内容结构化分析(名人识别,时装检测识别,商品识别,场景识别,行为识别,视频指纹)





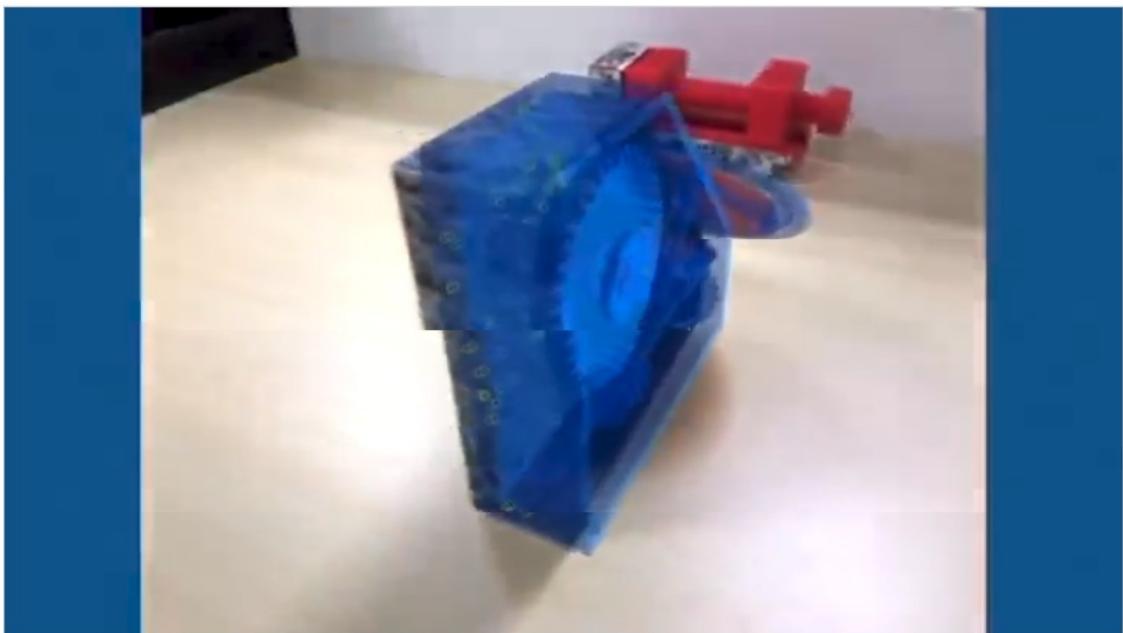
应用场景: 广告插入,视频审查过滤,推荐

4. **图像视频增强:** 去雾,超分,降噪(移动端),单图HDR,失焦修复,后聚焦(双摄),滤镜(包含风格迁移),光照估计(虚拟物体渲染,VR)

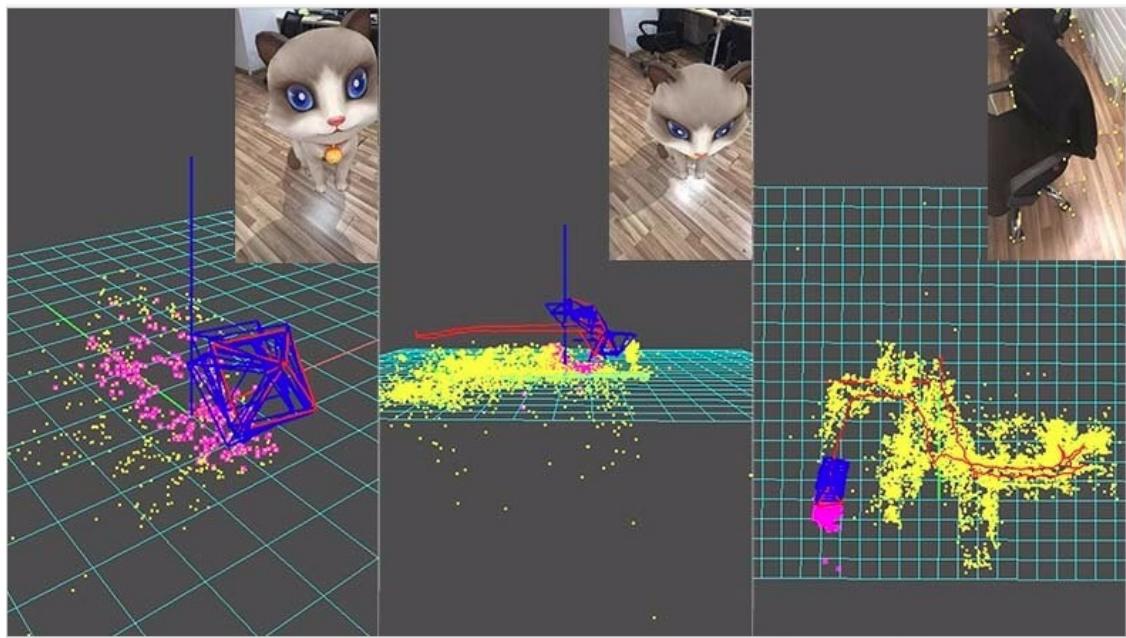
应用场景: 摄影/视频后处理;手机摄像头成像优化

5. SLAM和3D视觉:

- 平面/三维物体识别与追踪;



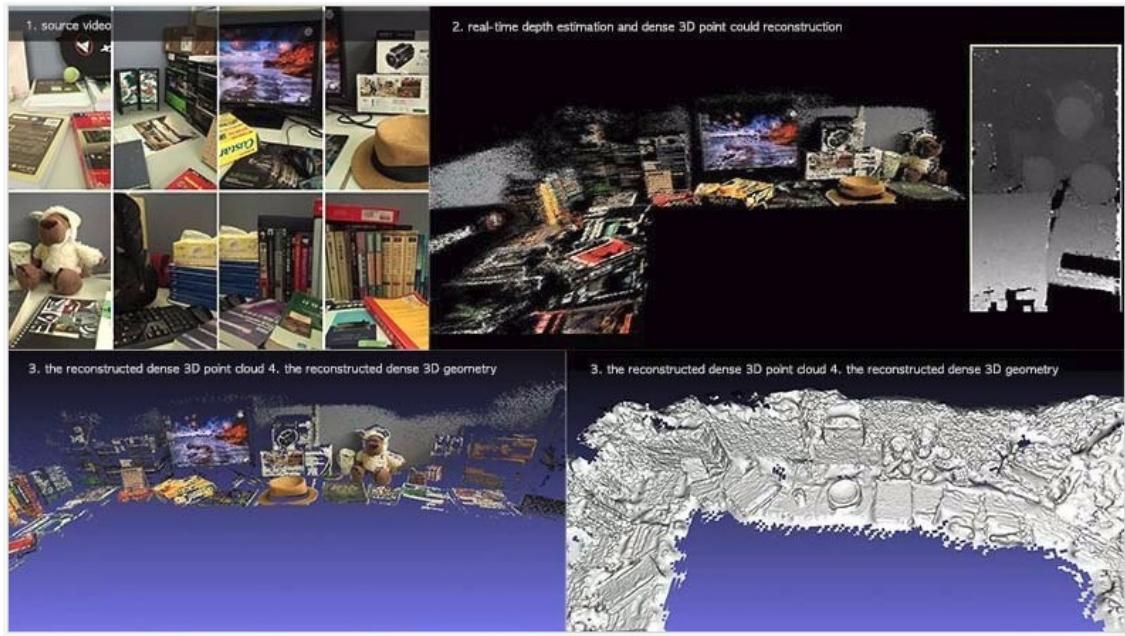
- 同时定位与地图构建(SLAM)



- 运动恢复结构(SfM):从连续2D图像重构物体的3D结构([具体说明](#))

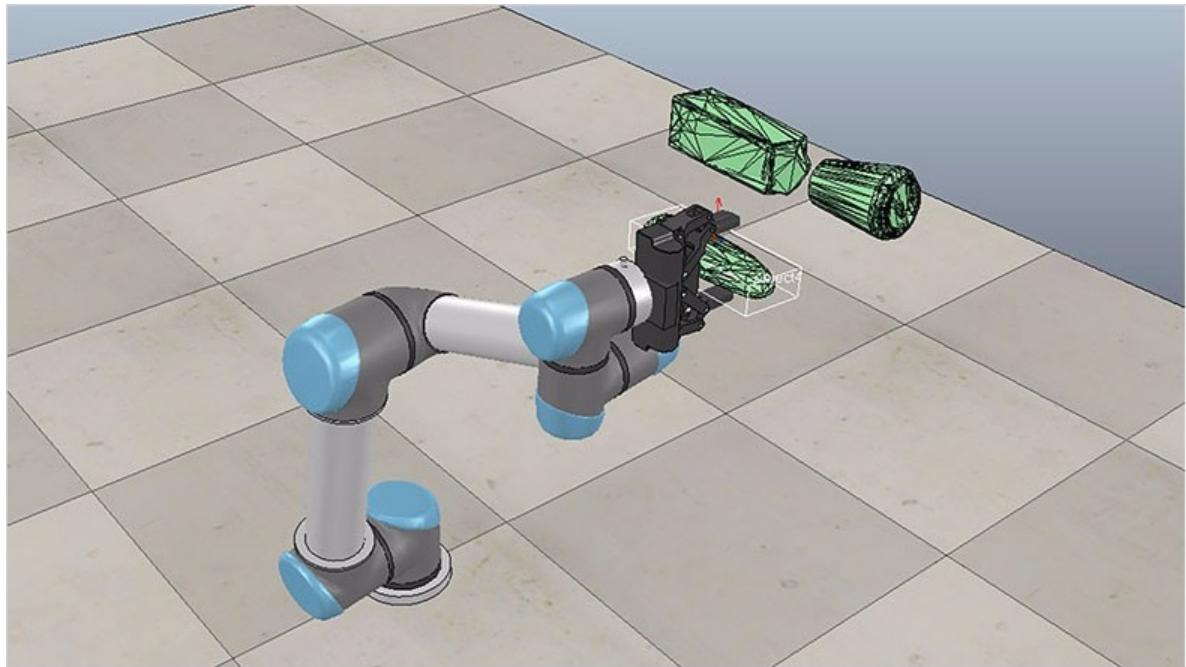


- 实时深度估计和三维重建:在线恢复视频每帧的深度图以及场景的三维模型(RGB->RGBD), 在**移动设备**上支持实时的景深虚化、双目立体化、真实感渲染、阴影合成、虚实物体的遮挡和碰撞等交互效果



应用场景: 跨平台轻量级AR/VR,无人驾驶

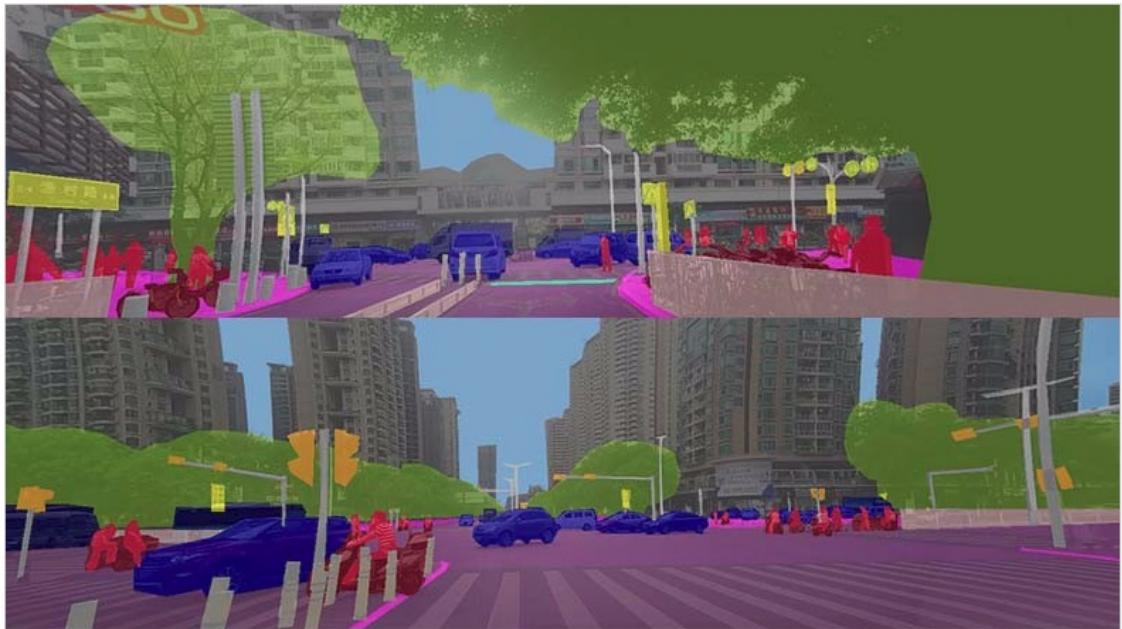
6. 机器人传感与控制: 机械臂自主抓取(视觉),RGBD传感器自动标定,机器人仿真平台(快速实验)



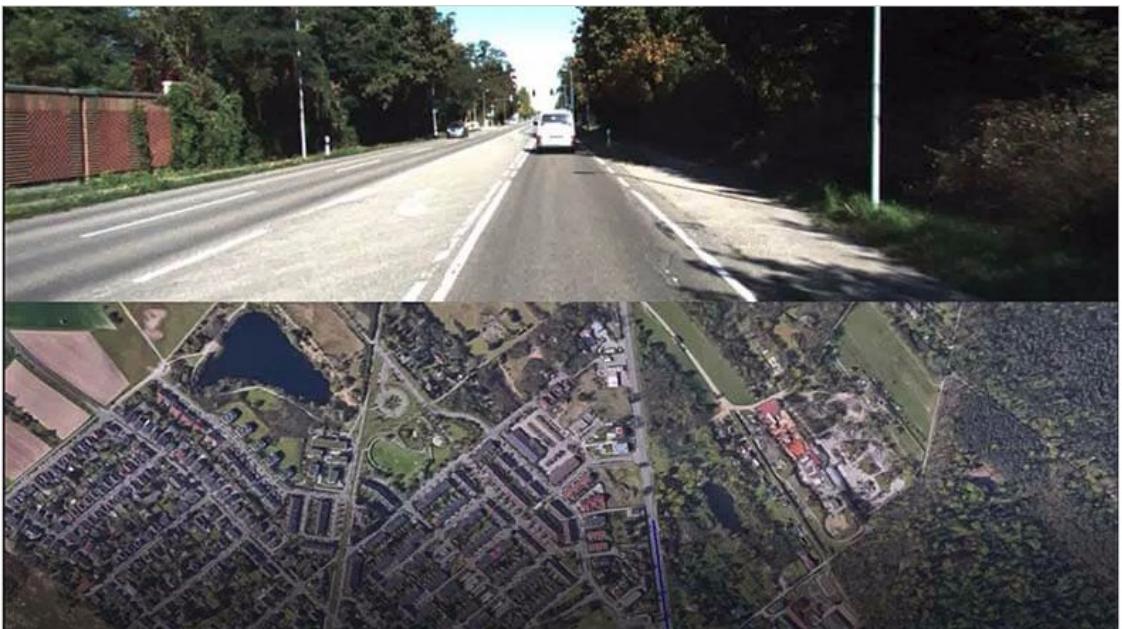
其他: 人体关键点检测, 遥感图像分割

7. 无人驾驶: 行人车辆检测, 车辆属性, 行人运动分析(运动方向), 车道线检测; 交通指示牌识别

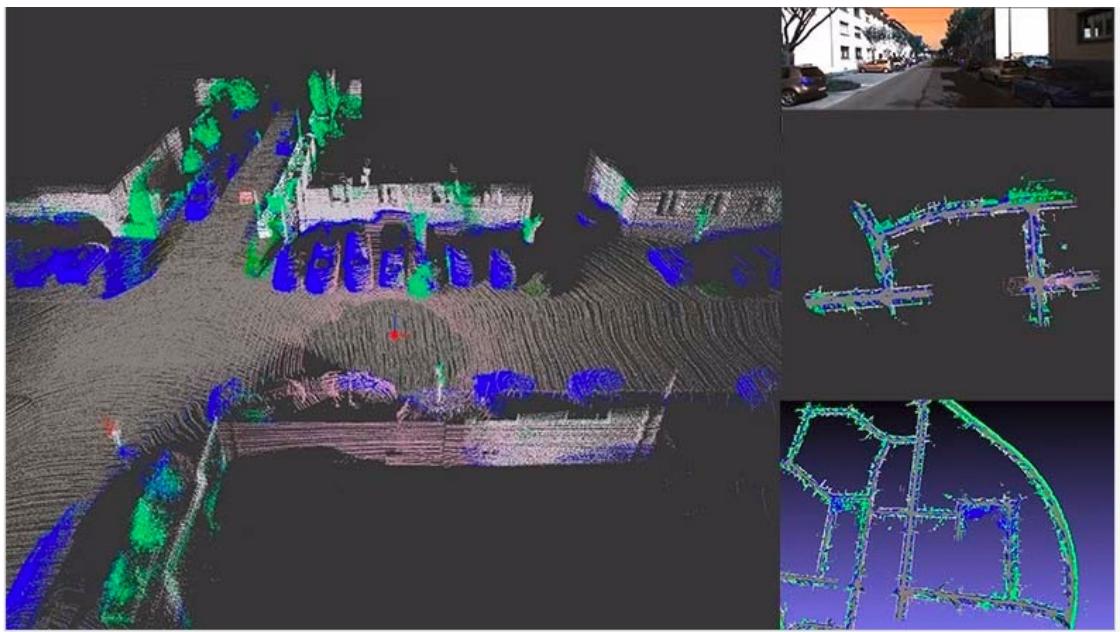
- 场景分割(细粒度检测,结合SLAM进行3D建模);



- 基于视觉的大规模城市场景定位



- 基于多目视频,雷达,卫星定位和惯性导航系统的大规模场景的高精度三维重构



8. 深度学习平台: GPU云(最高320个GPU)

3). 产品服务

来源: 商汤官网: <https://www.sensetime.com/authentication/87>

|  智能视频 |  身份验证 |  移动互联网 |  其它 |
|--|--|---|--|
| SensePortrait-S 人脸静态比对服务器 | SenseID 身份验证服务 | SenseAR 增强现实感绘制平台 | SenseGo 智慧商业解决方案 |
| SensePortrait-D 人脸动态比对服务器 | SenseKeeper 人脸识别机 (闸机版) | SensePhoto 手机图像处理解决方案 | SenseRemote 遥感影像智能解译解决方案 |
| SenseFace 人脸布控实战平台 | | | SenseDrive 驾驶员监控系统 |
| SenseTotem 视图情报研判系统 | | | SenseDrive 车道偏离预警 |
| SenseVideo-A 视频结构化解析服务器 | | | SenseDrive 前车碰撞预警和行人检测预警 |
| | | | SenseU 广告机 |

1. 人脸比对服务:

- 解决方案:
 - 静态比对服务器(基于单张图片): **千万级** 人脸库 **0.1s** 检索; **单GPU** 图片入库 **200万/h.**
==> 适用于人证核验(互联网金融)
 - 动态比对服务器(基于视频流): **秒级** 人脸比对, **单GPU** 图片入库 **3万/min;单GPU 12路** 高清实时视频. ==> 适用于公安刑侦,访客(会员)识别
 - 人脸布控平台(城市级大规模人脸): **分布式GPU集群 1000路+** 实时监控视频流人脸识别

别;10亿 级人脸抓拍**历史轨迹检索** ==> 适用于交通枢纽,边检等.

- 其他: 视图情报研判系统:人脸模糊检索;视频结构化解析服务器:行人车辆多属性分析.

- **产品:**

- **身份验证一体机: 云+端** 方式, 技术为卡证识别+活体识别(双目/结构光)+人脸识别 ==> 银行/政府办事大厅



- **人脸识别闸机:** 支持身份证+IC卡+二维码+人脸, 技术为活体识别(双目+结构光)+离线人脸识别(1s) ==> 所有使用闸机/门禁的场所



2. 移动端图像增强服务

- **实时视频视觉增强:** 表情涂鸦,背景替换,2D/3D渲染(AR/VR), 技术为人脸关键点,人脸手势检测,手势识别,背景分割==> 短视频,直播应用



人体姿势估计替换:

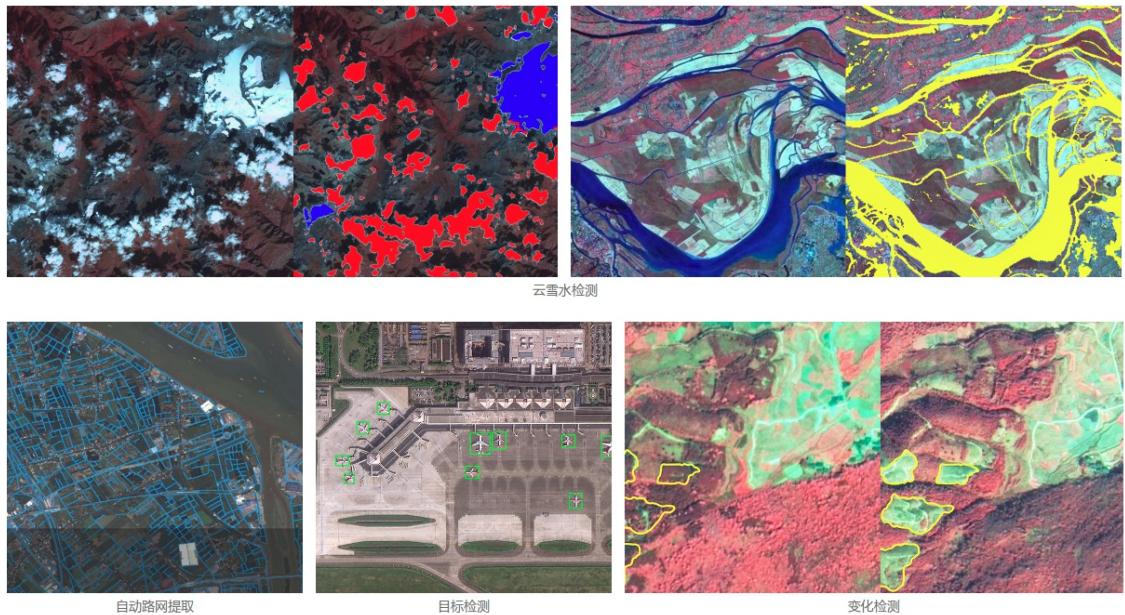


(*信息来源: <https://www.csdn.net/article/a/2017-05-19/15927974>*)

- **手机摄像头成像增强**
- **智能相册**: 人脸聚类, 照片分类, 场景分类
- **人脸解锁**

3. 无人驾驶

- **驾驶员监控**: RGB+红外成像方案, **离线嵌入/云+端** ==> 出租车登录, 疲劳检测/注意力检测; L3自动驾驶人机控制权交接(注意力检测); 数字化座舱功能控制和娱乐接口 (手势识别)
- **车道偏离预警**: RGB图像+**FPGA** 加速, 技术为车道线检测和识别(**厘米级** 精度) ==> 车辆道路定位+车辆状态预警(无人驾驶基础), 未打转向灯转向预警
输出车辆与车道线距离, 车辆
- **前车碰撞和行人检测预警**: 单目图像(RGB+红外?)+FPGA加速, **实时**(雨/雾/强光/黑夜)输出前方(100m)识别物体(机动车+非机动车+行人)的相对距离、相对速度, 以及相对碰撞时间和绝对碰撞时间。当报警时间达到临界值时, 输出报警信号
- **遥感影像解析**: 路网提取(图像分割), 其他用途包括: 气象, 国土资源, 环境监控, 国防(国家政府项目)



4. 智能零售

- **智慧商业解决方案(客流大屏):** 人脸识别+人脸识别+视频挖掘, **无感知**,场景不限, **云端** ==>顾客画像(消费行为分析), 客流统计分析,重点顾客分析(精细化运营分析决策)



- **广告机:** 人脸识别+人脸属性+AI互动游戏+行为分析



4). 技术团队

AI公司到最后还是需要拼技术实力的.目前通过公开尚不能比较完整了解商汤内部技术团队的构成,因此从以下三方面信息推测:1). 创始人**汤晓鸥**所领导的港中文多媒体技术实验室的研究和人员; 2).商汤及其联合发表的**学术文献**; 3). **技术岗招聘** 信息

1. 汤晓鸥实验室(信息来源 <http://mmlab.ie.cuhk.edu.hk/projects.html>)

- **总体概况:**

根据**GTC 2016**上的说法, 该实验室被列为全球top10的AI Lab,是由汤晓鸥于2001年创立的.汤晓鸥本人就不用多说了,大牛!



Xiaou Tang

The Chinese University of Hong Kong
Verified email at ie.cuhk.edu.hk - [Homepage](#)

Computer Vision Pattern Recognition

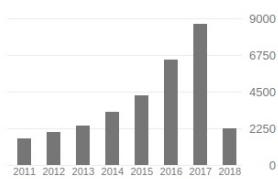
[FOLLOW](#)

[GET MY OWN PROFILE](#)

Cited by [VIEW ALL](#)

All Since 2013

| | | |
|-----------|-------|-------|
| Citations | 36196 | 27527 |
| h-index | 93 | 79 |
| i10-index | 324 | 259 |



Co-authors [VIEW ALL](#)

| | |
|--|---|
| | Xiaogang Wang Associate Professor of Electronic... |
| | Chen Change Loy The Chinese University of Hong ... |

TITLE

CITED BY

YEAR

Single image haze removal using dark channel prior
K He, J Sun, X Tang
IEEE transactions on pattern analysis and machine intelligence 33 (12), 2341 ...

2547

2011

Learning to detect a salient object
T Liu, Z Yuan, J Sun, J Wang, N Zheng, X Tang, HY Shum
IEEE Transactions on Pattern analysis and machine intelligence 33 (2), 353-367

1958

2011

Guided image filtering
K He, J Sun, X Tang
IEEE transactions on pattern analysis and machine intelligence 35 (6), 1397-1409

1631

2013

Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval
D Tao, X Tang, X Li, X Wu
IEEE transactions on pattern analysis and machine intelligence 28 (7), 1088-1099

801

2006

Image super-resolution using deep convolutional networks
C Dong, CC Loy, K He, X Tang
IEEE transactions on pattern analysis and machine intelligence 38 (2), 295-307

751

2016

Learning a deep convolutional network for image super-resolution
C Dong, CC Loy, K He, X Tang
European Conference on Computer Vision, 184-199

743

2014

(*来源: google scholar: <https://scholar.google.com/citations?user=qpBtpGsAAAAJ&hl=en>*)

汤晓鸥香港实验室目前Staff有**7人**,除汤以外还包括**2名**副教授,**1名** 助理教授,**3名** 研究助理教授,**3名** 博后,**1名** 副研究员,**19名** phd;联合实验室的中科院深圳技术研究中心有**2名**教授**1名**副教授,下属phd数量不详,大于**5人**.

而根据目前该实验室的**毕业生**去向(Kaiming He也是这个实验室的毕业生),8人去了商汤(总共68人),而计算14年商汤成立以后开始计算的话这个比例高达8/18.综上,汤晓鸥团队应该占了商汤所宣称的100人+phd技术团队的**40%以上**.业内优秀的有号召力的PI对人才吸引力可见一斑!

◦ 主要研究方向:

主要大方向如下:主要分3大块:**人脸分析,视频监控分析,图像视频检索**

Deep Learning



Deep Learning

Face Analysis



Face Recognition



Face Alignment/Tracking



Face Detection



Face Sketch Synthesis and Recognition

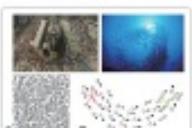


Face Hallucination

Visual Surveillance



Human Detection



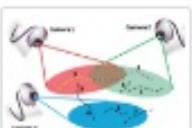
Crowd Analysis



Person Re-identification



Tracking



Camera Network

Image and Video Search



Re-ranking



Photo Quality

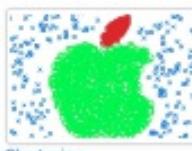


Action Recognition



Music

Machine Learning



Clustering

3D Line Drawing



3D Line Drawing

Image & Video Editing



Image and Video Editing

从实验室开放的公共数据集来看:**人脸**相关的有5个(celeba等),**行人**相关的有2个,**场景**相关的有6个,**服装**相关的有4个(deepfashion),**车辆**相关的有1个,**图像检索**相关1个.总结来看与主要研究方向比较吻合人脸分析和视频监控分析为主.

- i. **人脸分析:** 人脸相关的研究仍然是该实验室目前主要的研究方向,到目前的发表**人脸相关的文献为115篇**,其中2012年以后发表通过深度学习的为**40/86篇**.主要内容包括**人脸识别**(23篇文献),**人脸对齐**(7篇)/**跟踪**(2篇).**人脸检测**(5篇),**人脸素描合成和识别**(刑侦),**人脸超**

分辨率.

- ii. **视频监控分析**: 主要技术方向包括**行人检测(8篇),人群检测(16篇),行人识别(>6篇),跟踪(6篇),摄像头组网(7篇)**(主要用于安防和无人驾驶场景)
- iii. **图像和视频检索**: 主要包括**网络图像检索和重排序(5篇),图像质量和超分辨率(11篇),动作识别(10篇),音乐视频匹配(7篇)**
- iv. **深度学习和机器学习**:主要包括聚类(包括图像聚类)(**12篇**)
- v. **3D-空间/物体重构**: **12篇**
- vi. **图像/视频编辑**: 主要包括图像去雾,过滤,消光.(**9篇**)

PS:以上文献统计来源来自[实验室官网项目](#)和[google scholar](#)

◦ **联合发表的文献**: 主要集中在计算机视觉领域三大顶会之二:CVPR(CVPR2017 23篇), ICCV(ICCV2017 43篇), ECCV. 此外在AAAI、ICML、NIPS、TPAMI、IJCV、TIP也有不少.根据[google scholar](#)的信息,与商汤有关的论文大概在115篇左右. 整体方向与之前没有大致相同.不过值得注意的是不少文献是由商汤实习生参与的,说明其联合的实验室有不少学生直接参与商汤的日常科研项目.

◦ 招聘信息:

- 1). **超算平台相关: 高性能异构计算平台,大规模并行计算,异构优化;** 2). **芯片设计相关: ASIC设计,FPGA开发, SOC构架;** 3).**无人驾驶相关: ADAS开发,自动驾驶控制,** 4). **理论研究:深度学习核心构架,计算机视觉研究**,这4个方向是商汤在[拉钩](#)上社招的主要算法相关技术岗位,而商汤强项的人脸和视频监控相关的人员招聘不多,说明这方面的技术已经很强;而商汤官网上的招聘信息主要是校招和实习生,(猜测可能用于降低博士的研发压力以及培养人才).
- 杭州3D视觉研究中心: 浙大鲍虎军任主任, [张国锋](#)任副主任:主要招聘为**AR/VR,3D重建,SLAM,SFM**方向人员.>(*信息来源:http://www.sohu.com/a/163257609_715754*)
- 商汤日本研发中心: 主要与Honda联合开发自动驾驶方向为主:

■ Recruitment occupation

- ① Image processing technology researcher / engineer (deep learning experience is not required)
- ② Automatic driving technology researcher / engineer
- ③ Robot control technology researcher / engineer
- ④ Embedded software engineer
- ⑤ FPGA Engineer
- ⑥ QA engineer

信息来源:<https://www.sensetime.jp/job>

总结,从前面的招聘信息来看,商汤确实目前确实有自己开发的**超算平台(DeepLink)**和**底层计算库(plugin_sample)**.此外,2018.2,商汤与MIT宣布达成人工智能联盟,成为MIT Intelligence Quest (IQ)计划的一部分 ([消息来源](#)),算是搭上了全球AI研究第一梯队的顺风车

关键技术创新 - 平台架构

运算集群 + 深度学习操作平台

SENSE TIME 商汤集团

DeepLink

- 大规模运算集群
- 2000 GPUs, 4000CPUs
(目前亚洲最大, AlphaGo 280 GPUs)

搭建超过1000层的网络

自主开发的深度学习平台
高效、可拓展、灵活
性能优于目前深度学习主流平台, 如Caffe、Torch、TensorFlow等

Parrots

采集

关键技术创新 - 平台架构

PPL

SENSE TIME 商汤集团

PPL (Parrots Primitive Library)

- 经过深度优化的高性能计算库, PPL在不同处理器架构 (X86, ARM, CUDA等) 下有不同版本

以ARM处理器为例子: **PPL ARM vs. Caffe ARM**

- 支持iOS、Android、Linux和Windows
- 加速运算

| 任务 | PPL (ARM) | Caffe |
|--------------------|-----------|-------|
| deep_bound | ~1.2 | ~1.5 |
| classify_cnn | ~1.2 | ~1.5 |
| detect_craft | ~5.5 | ~1.0 |
| detect_hunter | ~1.5 | ~1.0 |
| dewatermark_rand | ~1.2 | ~1.0 |
| ent_helen_fig | ~1.8 | ~1.5 |
| feature_extractor | ~1.5 | ~1.5 |
| vehicle_analyze | ~1.2 | ~1.0 |
| vehicle_detect | ~2.5 | ~1.0 |
| bankcard_ocr_still | ~1.2 | ~1.5 |

faceU应用需要实时处理视频, 在性能方面要求严格, 以PPL做为后台, 减缓了研发压力

在商汤团队人脸技术的加持下, faceU在今年三月获得了apple store颁发的年度最佳应用奖

(*信息来源:<https://www.leiphone.com/news/201612/SgrlzOqKb4T8qdfH.html>*)

因此目前商汤最大的技术优势包括以下几点:1. 深厚的**人脸分析**,视频**监控/挖掘**技术背景,软件层面来说在安防领域的所需的算法都有比较强的积累; 2. 有自己的超算平台和庞大的博士团队,不光自己研发实验的效率能提升,转化测试学术界新算法的能力也很强; 3. 目前的算法研究重点向**自动驾驶**所需相关技术转移(SLAM,3维重建,视频结构化分割/检测)

5. 行业布局

首先来看下商汤官方宣称的行业布局:



(*信息来源:<https://www.sensetime.com/cases#0>*)

基于技术创新，商汤服务于金融、平安城市、智慧家居、虚拟增强现实、互联网娱乐等多个行业领域和场景

安防-智慧城市

- 人脸检测跟踪**
 - 对背景复杂低质量的图片或百人人群监控视频，实现毫秒级别的人脸检测
- 行人车辆检测**
 - 在视频监控、辅助驾驶等场景中，准确、迅速地找出场景中所有的行人车辆位置

金融/互联网身份认证

- 人脸身份认证**
 - 毫秒级检索大规模人脸数据库或监控视频，给出身份认证。在认证出96%的人脸时，误检率低于十万分之一
- 真人检测**
 - 检测摄像头前用户是否为真人操作，能有效分辨高清照片、PS、三维模型、换脸等仿冒欺诈

移动智能/娱乐互联网

- 人脸关键点定位**
 - 微秒级别眼、口、鼻轮廓等人脸106个关键点定位
- 人脸聚类**
 - 数十万人的人脸快速聚类，可用于基于人脸的智能相册以及基于合影的社交网络分析。让照片整理更直观，让社交关系更清晰

智慧商业

- 人群分析**
 - 估计人群数量和密度，实现重大活动、重要区域的人流统计。可适应各种光照变化、人群遮挡等复杂环境
- 服装属性识别**
 - 自动检测和识别图片、视频中的服饰，准确识别服饰品类、花纹、袖型、领型等特征，显著抵抗光照、动态变化等干扰因素的影响

(*信息来源:<https://www.leiphone.com/news/201612/SgrlzOqKb4T8qdfH.html>*)

官方 显示的合作伙伴如下:

战略合作伙伴



[展开更多](#)

移动互联网



[展开更多](#)

金融



[展开更多](#)

智慧城市



手机



[展开更多](#)

物联网



[展开更多](#)

遥感



以上4个方向确实是主要的落地场景,根据其投资人IDG资本的说法,商汤的收入主要来自**安防,金融,手机移动互联网3块**.

• 1). **安防/智慧城市**

从其销售方面的**(17年10月为止的数据)**来看,**安防** 方向是最多的,说明这块业务仍在快速增长和拓展中. 目前安防已有的合作方是广州和云南的**公安系统**(**消息来源**,但是网上并无其他来源的信息,可靠性存疑)以及与深圳公安局的**城市规模化安防**应用(2018.3.20)(**消息来源**); 还与**华为**联合发布了超高密度人脸识别一体化方案(服务器单机即可实现224路高清视频流人脸识别的实时处理),作为**华为平安城市方向**的重要一步(**消息来源**).



此外,去年11月与**上海市政府**签订全方位战略合作协议(消息来源:[1](#);[2](#);[3](#)).将在沪设立全球研发、**智能汽车**、**智能芯片**、智能教育等相关总部,全面推动人工智能创新应用和核心技术落地.在3.11与提供智慧城市解决方案提供与运营商**上海仪电集团**达成战略合作;在3.14日与**临港集团**签订战略合作协议,内容包括:加快商汤人工智能**超算中心**、**智能驾驶**等重大产业项目落地临港科技城,积极探索和推进与临港松江科技城等园区的载体型、平台型项目的合作;共同打造人工智能研究及**人才培养**平台,促进人工智能领域的技术创新、产业孵化和专业人才培养([消息来源](#)). (感觉这个套路有点像受深圳市政府扶植多年的华大基因)

总的来说,相比于上海的智慧城市规划,个人认为更多的是上海需要引进AI产业的以及商汤需要其安防领域算法能力的落地标杆的共同利益所致,而以商汤的体量和目前的大数据处理能力(虽然有在招大数据方向的人),很难说会有很突破性的实际应用落地;而深圳公安局的合作则是更加实际有力的成功案例,可能能够真正帮其打开这一领域市场的关键.

- 2). **互联网金融服务**

主要场景是**身份验证**,目前的合作伙伴包括**借贷宝(融360)(超过1亿次人脸验证)**;**换呗(分众)**;**中国移动(3亿次验证)** ([信息来源](#))



感觉这部分业务比较尴尬,阿里腾讯这两大互联网金融巨头都有自家产品,所以对商汤而言这个天花板是看得到的,可能会考虑海外市场(东南亚之类的),但是从人员招聘来看貌似并没有准备花大力气拓展.

- 3). 手机/移动互联网应用

合作伙伴包括:**美图、小咖秀等相机APP, OPPO,vivo,小米,华为等手机厂商,微博(社交应用),bigo(短视频)(消息来源)**.



个人认为,这部分应用用到的技术多且杂,而且由于端的限制对算法模型本身的要求会比较高,对AI厂商而言是秀肌肉和技术输出,产品落地的好试验场.正如google用自家强大的算法改造android底层,不光对自家APP提供强大的算法支持实现智能化,也在不增加一般开发者开发成本(不需要专业的算法背景)的基础上就能极大的提高安卓应用的能力和用户体验,这可以视为一种AI平台化的输出能力.因此,这部分不光是商汤目前的盈利点,也是实践用AI推动移动互联网这一垂直领域的生态进化的重要场所(这个行业都无法通过AI赋能更不用谈其他行业),这是AI改造其他行业的起点.

- 4). 智能商业

目前已知的合作伙伴为苏宁,具体落地场景就是苏宁无人店.



能查到的具体落地信息很少,貌似这也是AI厂商改造传统垂直行业遇到的普遍问题:行业理解还不太够,对于商汤这种学院派出身的尤其如此.

- 5). **自动驾驶/车载业务**

这部分主要是通过劳世竑(具体可以看[这里](#))在2017.12与本田达成为期5年的L4无人驾驶研发合作协议([消息来源](#))之后开始启动的,而且从目前招聘来看势头很猛,尤其是商汤的日本研发中心(见上文)主要研发的就是**自动驾驶/车载服务**相关的技术,可见商汤对这块还是非常重视的.

自动/无人驾驶作为全球科技巨头的比拼AI算法能力的重点领域,日本三大汽车厂商之一的本田选择商汤(商汤之前基本没碰过自动驾驶),一方面说明行人车辆方面场景检测方面的积累很深,另一方面说明商汤的研发能力在超算平台加持下确实也很强.这个领域是目前最能密集输出AI能力的地方,商汤的入场也是意料之中的.

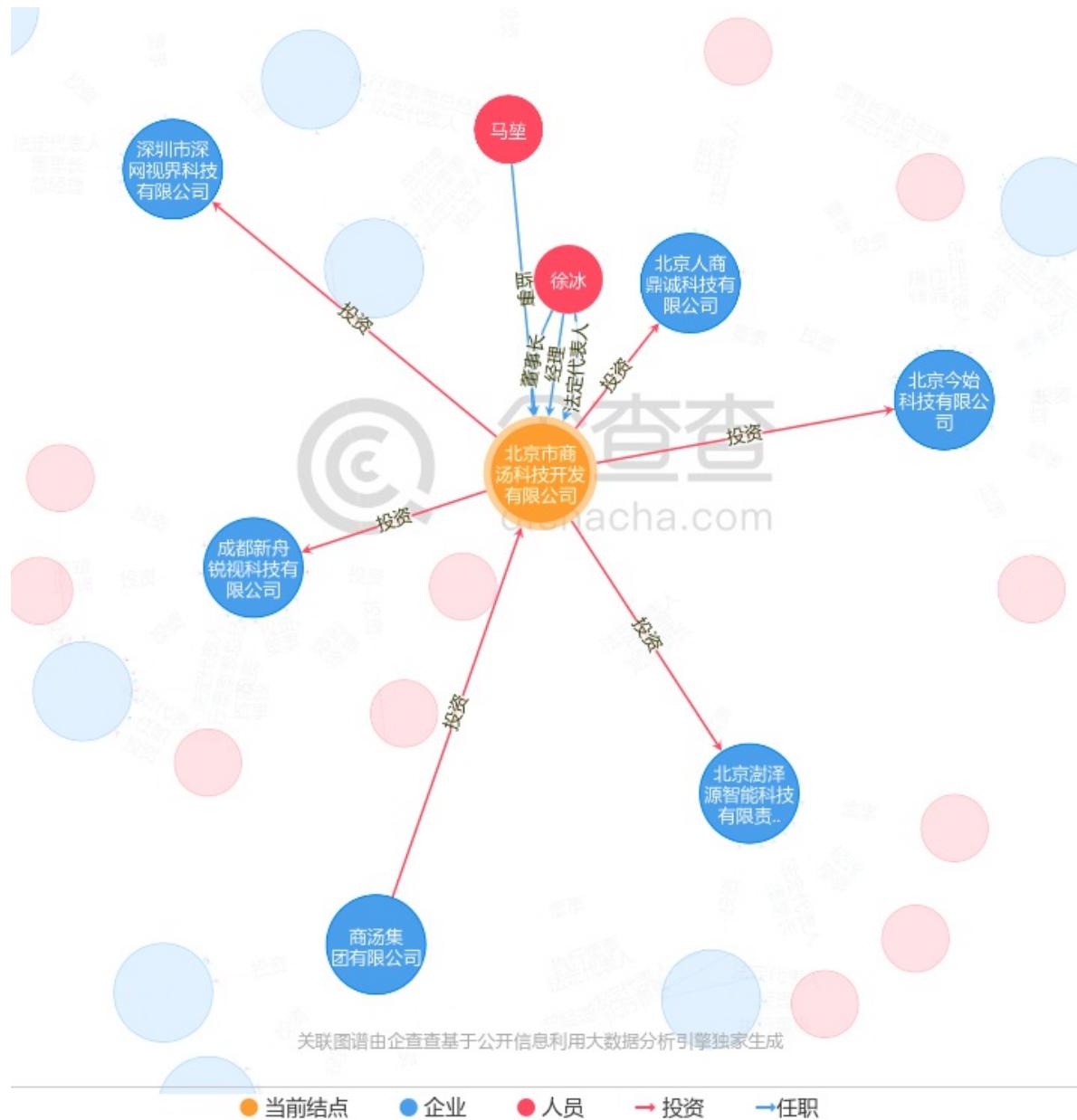
- 6). **芯片**

主要合作伙伴就是英伟达和高通:2017.10上GTC英伟达宣布与商汤合作([消息来源](#));2017.10高通战略投资商汤,并且宣布战略合作([消息来源:1;2](#)).目前招聘的情况来看来看在端上的芯片工程师确实比较多,这可能与商汤目前很多产品通过[云+端](#)的方式实现有关,根据商汤自己的定位来说还是算法公司,暂时没有打算单独进军AI芯片,以和专业公司合作为主([消息来源](#)).

端上算法模型的开发部署是目前工业界的一大热点,算法结合神经计算芯片(AI芯片)能够更有效的处理更多场景,所以商汤之后肯定会涉足其中的.

• 7). 行业生态(投资)

从工商局可以查到的商汤对外投资信息来看:



其投资的的成都**新舟锐视**主要开发智能摄像机(多用于安防领域);而**深网视界**也主要是基于监控视频的人脸检索和人群分析,所以目前商汤已投资的公司还是围绕在其深入**安防领域**市场的布局.

此外商汤对**51VR**投资2.1亿,其他计划投资的项目还有10多个,包括AI芯片、医疗、物联网、手机和互联网,还包括游戏([消息来源](#))

商汤确实在像小米一样构建自己AI产业链的生态...

• 8). 深度学习平台

这部分在之前的官网是找不到的,平台这种事目前只有google,facebook,amazon, apple,

microsoft, baidu之类的国际大厂在做,而且目前算法开源并且社区很活跃的只有google的Tensorflow,facebook的caffe+pytorch和amazon的mxnet,其他的比如微软的cntk和百度的paddlepaddle社区活跃度就很低了,基本只有自己公司的人在用,而这部分能力最适合的就是集成到自家的云计算平台上.个人认为即使今后开放了这个平台,客户定位依旧十分尴尬,而且从CEO的演讲也没有看到这方面的苗头.知乎这位同学的观点在下也十分赞同:

个人看法是平台型公司不一定是未来AI公司的出路, 把一个一个行业慢慢吃透, 用技术去推动行业生态的改变, 才是AI公司的出路。

最后,CEO徐立的关于学术和工业的思考([信息来源](#)):



总结:商汤由于其创始人汤晓鸥是学院派出身,之前与其说是公司更像是研究院,凭借港中文多媒体实验室多年在计算机视觉和图像领域的积累,其算法研发能力方面很强而且覆盖的领域也很广(文章数量多且领域分布广泛),不过目前的工程产品化方面关注较少(专利少).不过随着B轮C轮两轮融资和阿里的入局,今后其产品化应该会有比较大的动作(招聘信息也表明工程化人员在大量招募).

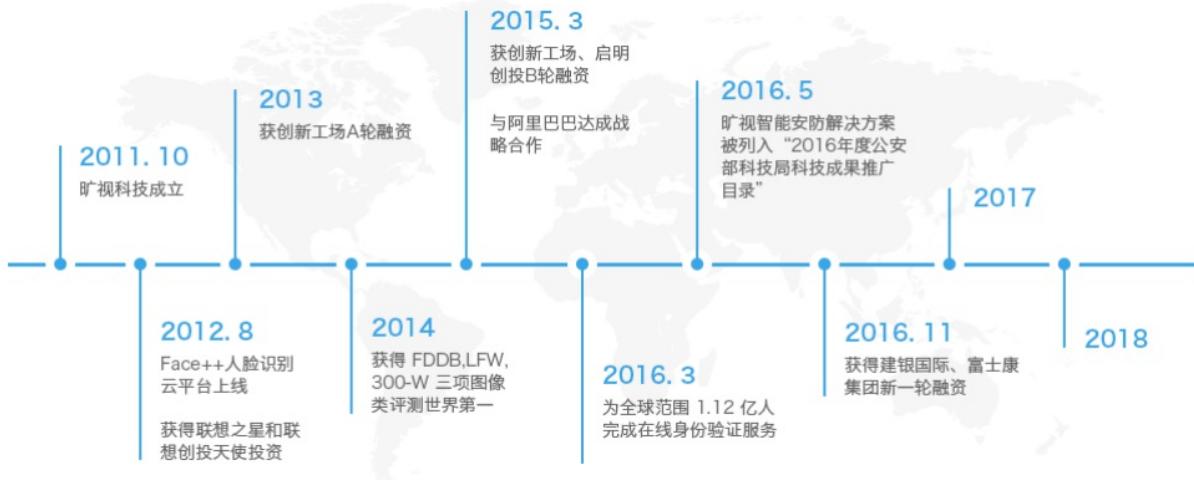
参考:

- 商汤官网: <https://www.sensetime.com/>
- 商汤招聘: <http://hr.sensetime.com/alljob/?p=1%5E2&PageIndex=3>
- [一文读懂30亿美元估值的商汤科技, 这些年做了哪些布局](#)
- [商汤徐冰：大牌投资人排队想投，阿里15亿投资完全是传言](#)

- 阿里要15亿投资不缺钱的商汤科技，打的什么算盘？
- 商汤、云从、依图、旷视和格灵深瞳五家公司的深度学习算法实现手段差别和公司产品优势有哪些？
- 融资6亿美金背后：商汤加速裂变
- 商汤联合华为发布“超高密人脸识别”解决方案
- 打造上海AI示范园 商汤科技与临港集团达成战略合作
- 商汤科技CEO徐立：让AI产业化落地，还要靠这双轮驱动？
- 商汤科技成为独角兽的秘诀，是人才“黑洞效应”
- 商汤研发总监林倞：20张PPT谈四大领域产业布局 | 2016 CAICC
- Nvidia's plan to turn data from 500 million cameras into AI gold
- MIT and SenseTime announce effort to advance artificial intelligence research
- SenseTime and Qualcomm to Collaborate to Drive On-Device Artificial Intelligence
- Honda and China-based SenseTime to Pursue Joint Research and Development in the Area of AI Technologies for Automated Driving

2. 旷视科技(Megvii/Face++)

1). 发展历程



(信息来源: 旷视官网: <https://www.megvii.com/about-us/>)

旷视科技成立于2011年年底,是4巨头中成立最早的,不过发展是从2014年印奇退博之后开始的,目前完成了6轮融资,最后一次为C轮,总融资金额为5亿美元,目前员工超过400人,专利267.

| 序号 | 日期 | 产品名称 | 级别 | 金额 | 投资方 |
|----|------------|------|-----|---------|--|
| 1 | 2017-10-31 | 旷视科技 | C+轮 | 3.6亿美元 | 国风投领投 蚂蚁金服 富士康跟投 中俄战略投资基金 阳光保险集团 SK集团 |
| 2 | 2016-12-06 | 旷视科技 | C轮 | 1亿美元 | 建银国际 富士康集团领投 |
| 3 | 2016-11-22 | 旷视科技 | B+轮 | 2000万美元 | 鸿海集团等 |
| 4 | 2014-11-06 | 旷视科技 | B轮 | 2200万美元 | 启明创投 创新工场 |
| 5 | 2013-07-18 | 旷视科技 | A轮 | 100万美元 | 创新工场 |
| 6 | 2012-01-01 | 旷视科技 | 天使轮 | 100万人民币 | 联想之星 乐基金 |

(信息来源:企查

查:http://www.qichacha.com/firm_52848015fa94a3e58cdaacd8c63eb737.html#run)

2). 核心技术

旷视的核心技术无疑是人脸技术,其成名之战就是2015德国Ce老马展示的Smile to Pay技术(信息来源).不过随着其业务线的扩展,其技术也扩展到文字识别,行人识别方向等.

| 人脸识别 | 人体识别 | 文字识别 | 图像识别 |
|-----------------------------|----------------------------|-----------------------------|--------------------------------|
| 人脸检测 | 人体检测 | 身份证识别 | 场景与物体识别 <small>beta</small> |
| 人脸比对 | 人体抠像 | 行驶证识别 | 人脸融合 <small>new</small> |
| 人脸搜索 | 人体属性 | 驾驶证识别 | |
| 人脸关键点 | 手势识别 <small>new</small> | 银行卡识别 <small>new</small> | |
| 人脸属性 <small>beta</small> | | 通用文字识别 | |
| 情绪识别 <small>beta</small> | | | |
| 颜值评分 <small>beta</small> | | | |
| 视线估计 <small>beta</small> | | | |
| 肤质评估 <small>beta</small> | | | |

1. **人脸技术: 非配合动态** 人脸检测,识别,关键点定位(106);人脸属性:年龄,性别,人种,情绪,颜值,头部朝向,视线估计,皮肤状态;在线/离线人脸**活体检测**(通过关键点实时定位和变化检测);超大

人像实时检索技术(对不同场景下的人脸进行优化)

The screenshot displays six main features of the Face++ service:

- 人脸检测**: Detects faces in images or video streams and returns bounding boxes. Includes a brief description and a "了解更多 >" link.
- 人脸关键点**: Locates up to 106 precise landmarks per face. Includes a brief description and a "了解更多 >" link.
- 人脸属性**: Extracts detailed facial attributes like age, gender, expression, and pose. Includes a brief description and a "了解更多 >" link.
- 人脸比对**: Compares two faces to determine if they belong to the same person, returning confidence scores and thresholds. Includes a brief description and a "了解更多 >" link.
- 人脸搜索**: Searches for similar faces within a pre-defined database. Includes a brief description and a "了解更多 >" link.
- 人脸贴纸**: A placeholder icon for the Face Swap feature.

应用场景: 各种场景(金融, 安防, 边检)的身份验证; 智能手机应用, 包括人脸编辑(美妆, 滤镜, 贴纸), 智能相册; 客户画像和推荐. 比如:

The section illustrates three practical applications of facial recognition technology:

- 应用场景**: A general heading for the examples.
- 人脸照片编辑**: Shows a woman taking a selfie with a coffee cup, demonstrating how the technology can enhance photos.
- 视频滤镜与贴纸**: Shows two women using a selfie stick on a beach, illustrating the use of filters and stickers in video.
- 3D人脸模型**: Shows a woman's face rendered as a 3D model, demonstrating the creation of 3D facial representations.

Below each example is a brief description:

- 人脸照片编辑**: "运用人脸关键点, 您的应用能够轻松编辑人脸照片, 并施以令人惊叹的精细效果, 如人脸妆容、贴纸、人脸五官或外貌变换等。"
- 视频滤镜与贴纸**: "Face++ Face SDK能够流畅处理视频流中的人脸, 您的应用将得以实时添加人脸滤镜或贴纸。"
- 3D人脸模型**: "利用高精度的人脸关键点, 您可以为视频或图片中的人脸打造纤毫毕现的3D模型。"

颜值评分:



美妆销售

利用 Face++ 颜值评分功能，您能获得顾客妆前、妆后的颜值分数，能够帮助您为顾客推荐、挑选最适合的美妆产品，为顾客提供完美的购物体验。



婚恋交友

Face++ 颜值评分能对照片进行颜值评分，让您的应用能根据用户的性别、要求偏好为用户推荐合适的人选，大大提升配对成功率。

肤质分析:



护肤品推荐

利用 Face++ 肤质评估能力，您能够准确获知用户的皮肤状态，从而有针对性的推荐适合用户的护肤品，有效提高推荐精准度和用户满意度。



在线医疗咨询

Face++肤质评估能力能够帮助您快速、准确的判断用户的皮肤状况，为您进一步分析、诊断提供可靠的信息，打造智能、可信的在线医疗咨询体验。

注意力捕获用户关注重点以及用户状态(疲态检测):



在线教育

利用 Face++ 的视线估计技术，教师能够了解学生的视线方向与停留时间，掌握学生的注意力状况和兴趣程度，从而显著改善教学体验。



广告精准投放

Face++ 能够帮助您实时分析受众群体的视线焦点，据此投放精准匹配的广告信息，并更好地评估广告效果。

2. **人体分析:** 人体检测,抠像(前后景分离);人体属性(性别,服装颜色);手势识别

人体属性

获取并返回图片中的人体属性信息，包括性别、上衣和下装颜色等，帮助开展基于人体属性的客户设计、人流统计等工作。

[了解更多 >](#)

人体检测

检测图片中的人体并返回人体矩形框坐标。广泛用于基于人体的人流数量统计，人体追踪抓拍等场景。

[了解更多 >](#)

人体抠像

识别图片中的人体轮廓并与背景区分，无需绿幕即可让您的应用轻松实现人景分离、虚拟背景等功能。

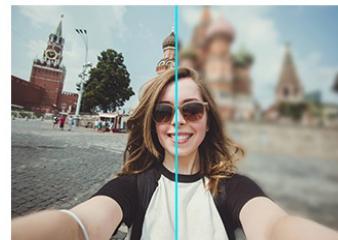
[了解更多 >](#)

应用场景: 人群检测, 人流数量统计(安防); 手机图像增强: 背景虚化, 合成, 替换(短视频, 直播); 客户画像推荐.



相片合成

自动将图像中人像抠出，选择一张新的图像做背景。轻松制作高质量的明星合影，景点纪念照，科幻照片等，实现你的每个精彩创意。



相片美化处理

先拍照，后对焦。自动识别图像中人体轮廓，并对背景进行虚化处理，轻松实现专业相机的成像效果。



娱乐应用特效

在娱乐类应用中，可以自动为图像中的人像添加预置的背景特效，得到前所未有的娱乐感染力。

3. 文字识别: 证件信息提取(身份证, 驾驶证, 银行卡); 通用文字识别(视频, 书籍, 广告牌, 单据等)

身份证识别

识别二代身份证上正反面的文字并返回结构化数据，可区分复印件、屏幕翻拍、合成图片等。

[了解更多 >](#)

驾照识别

识别机动车驾驶证主页正面的文字并返回结构化数据，用于各类驾驶证验证场景。

[了解更多 >](#)

行驶证识别

识别机动车行驶证主页正面的文字并返回结构化数据，用于各类行驶证验证场景。

[了解更多 >](#)

通用文字识别

识别图片中的文字信息并返回文字内容、文字框位置和文字类型，支持中英文。

[了解更多 >](#)

应用场景: 证件认证, 认证合一; 结构化信息提取, 内容审核管理

4. 场景物体识别: 场景检测, 物体检测. **应用场景:** 智能相册(相片聚类); 视频结构化处理(视频标

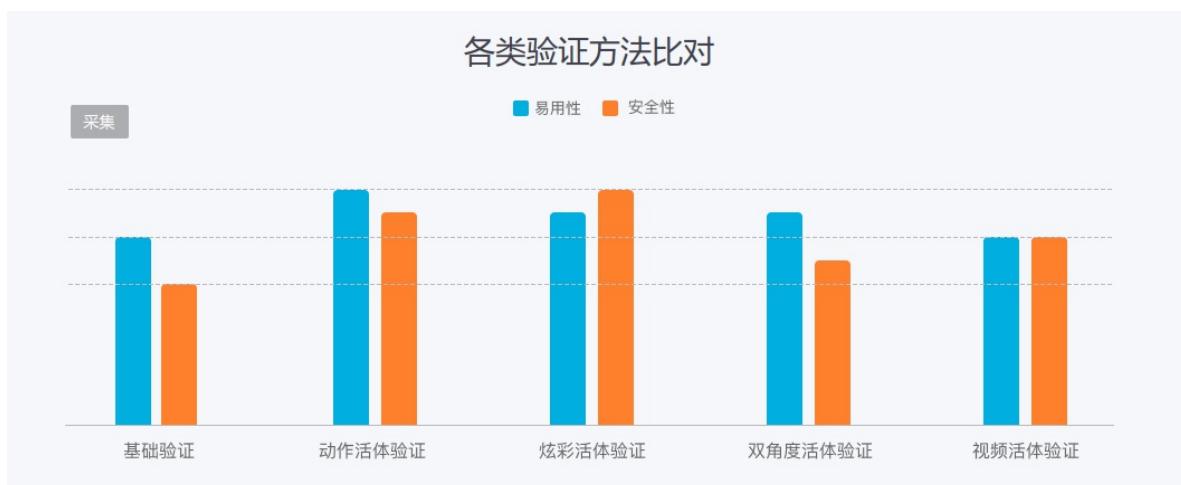
签);特定物体匹配追踪.

3). 产品服务

与商汤类似,旷视主要业务分也是以技术输出为主,而其在2012年就开始开放的Face++人工智能开放平台的API也为扩张和技术优化做出了不少的贡献.目前,旷视的主要产品集中在4个方向:^①FaceID服务,^②人脸分析服务^③智能安防^④智能地产.

1. FaceID服务

这个主要就是提供**金融行业**的在线身份验证服务,以**SDK,H5,小程序**等形式的接入,包含多种人脸验证方式:



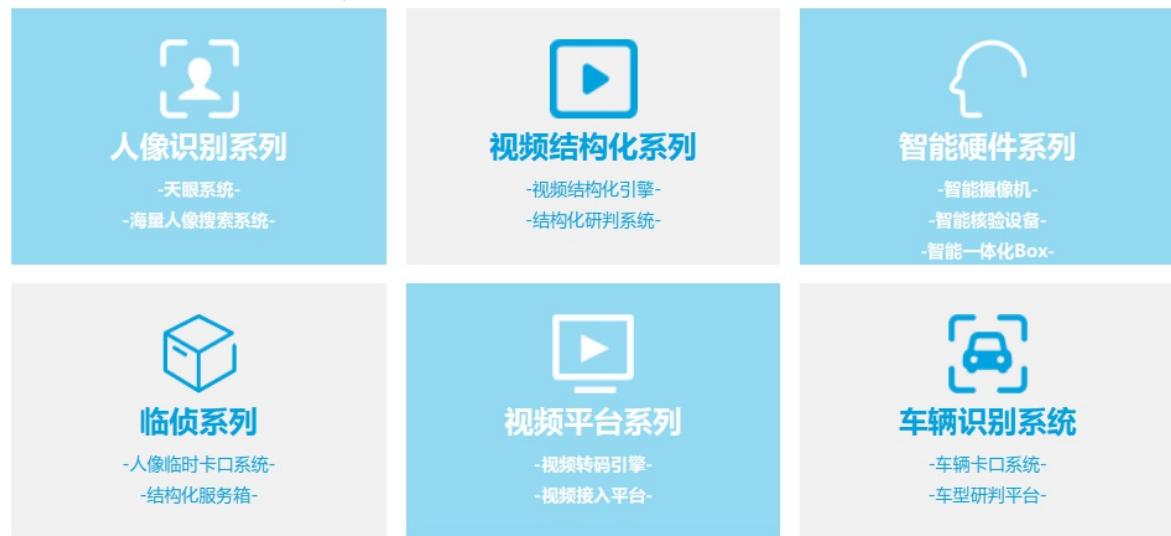
2. 人脸分析服务

- **产品: 移动端** 人脸关键点SDK,人脸比对SDK, 身份证质量SDK,视频美化SDK(面部AP贴纸,虚拟背景,美型美颜),并能达到**实时**的效果. ([信息来源](#))
- **解决方案: 定制化** 金融行业解决方案,手机行业解决方案,(虽然官网上还有在线营销解决方案,汽车行业解决方案,教育行业解决方案,但是没有具体的案例,可能是还没有具体案例或者还没有具体的落地方案)

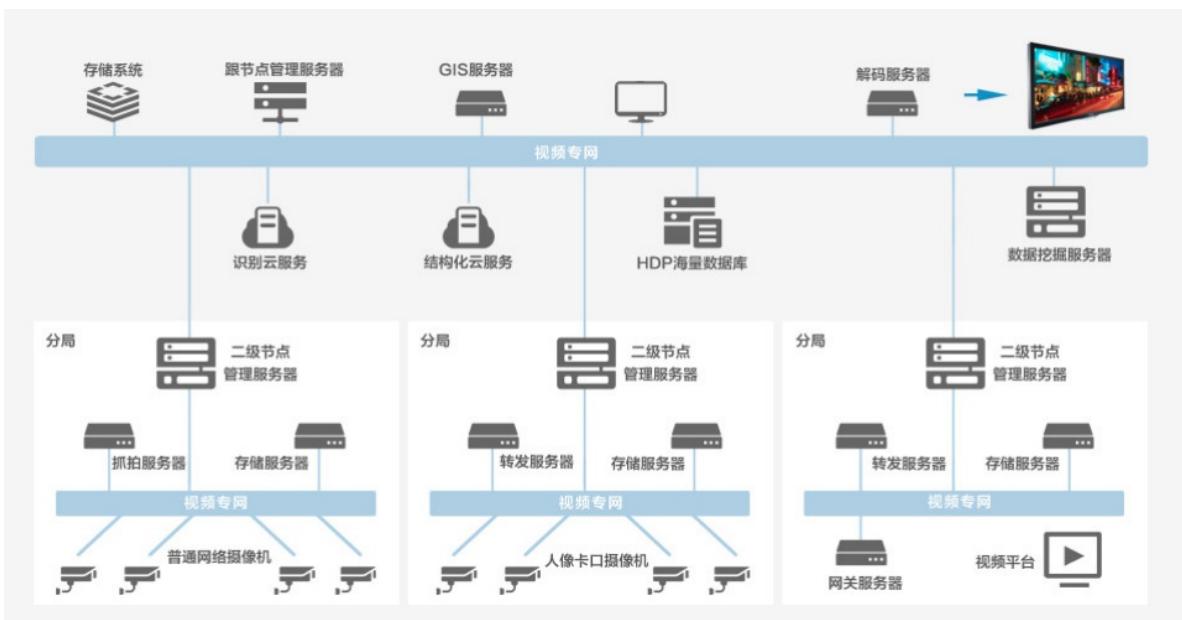
3. **智能地产解决方案:** 包括人员通行管理(门禁,访客管理,陌生人预警,会议室管理);车辆管理(停车缴费管理,车辆信息识别统计,寻车);视频监控(人员识别,行为分析,轨迹分析,场景图像分析);**智能零售**(会员识别,推送提醒,客流属性,消费行为,线上线下数据打通,精准营销).适合**企业楼宇园区,生活社区,校园管理,连锁商超**.



4. **智能安防解决方案:** 主要由以下资格子模块根据不同场景(包括重点场所监控,群体性维稳时间,景区防控等,具体看[这里](#))进行组合:



| 产品 场景 | 人像识别系列 | | 视频结构化系列 | | 智能硬件系列 | | | 临侦设备 | | 视频平台系列 | | 车辆识别系列 | |
|--------------|--------|--------|---------|---------|--------|--------|-------|--------|--------|--------|--------|--------|--------|
| | 动态卡口系统 | 海量搜索系统 | 视频结构化引擎 | 结构化研判系统 | 智能摄像机 | 智能核验设备 | 智能BOX | 人像临时卡口 | 结构化服务箱 | 视频转码引擎 | 视频接入平台 | 车辆卡口系统 | 车型研判平台 |
| 立体化治安防控 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 立体化综治 | ✓ | ✓ | | | ✓ | ✓ | ✓ | | | ✓ | ✓ | | |
| 智慧边检 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ |
| 反恐维稳 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 智慧监狱 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | ✓ |
| 智慧园区 | ✓ | ✓ | | | ✓ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ |
| 智慧小镇 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| 重大活动安保 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ |
| 智慧场站 机场车站 | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ |



产品特点为**系统库容量**为亿级,单机入库速度为1000人/s;支持高并发全帧率检测和识别,top10命中率99%;1500万底库检索时间小于1s;包含端上人脸识别算法和3D降噪技术的**智能摄像机**

5. **智能摄像头MegEye:** 暂时没有找到公开发售的聚到,可能只是作为其整体结局方案中的一环打包销售.([信息来源](#))

6. **人证一体机:** 使用红外活体模组实现活体检测([信息来源](#))

(信息来源: [1](#), [2](#), [3](#), [4](#))

总体来说,旷视的技术输出应该是工程化的最好的,其开放平台的调用磨炼了其核心的人脸技术大规模调用的可靠性.此前,旷视的主要产品就是人脸分析相关的SDK,但是目前官网上已经开始有成

型的整套解决方案,而且根据其构架图来看应该是有具体的落地场景的,因此似乎旷视正在从一家算法提供商慢慢转化成更具体的商业解决方案提供商.

4). 技术团队

与商汤不同的,旷视成立之初并没有像汤晓鸥这种在CV领域深耕多年的行业dalao(汤晓鸥在无论在学术界的号召力还是人脉资源都是要比旷视强不少的),但是创始人团队的3个人均是[清华交叉信息学院姚班](#)(这个班的水平可以参考[这里](#))出身的精英,其中唐,杨二人为IOI竞赛金牌出身,在国内OI界有比较强的号召力,因此吸引了一大帮OI界的天才.而随着2016.7 MSRA的首席科学家[孙剑](#)大牛的正式加入担任首席科学家(很早就开始有接触)以及同时加入的Adobe首席科学家[王珏](#),旷视除人脸方向的算法技术的研发实力也在慢慢补齐,而工程能力应该比商汤还强不少.

正如人才是每个AI公司的重要资产一样,除了需要暴露在闪光灯下的创始人团队和首席科学家之外,关于内部具体的牛人信息外露的比较少([知乎ZrpHercule的信息](#)),因此还是从以下三方面来研究下商汤的技术团队及其实力:**①创始人团队 ②孙剑 ③技术招聘信息.**

1. 创始人三剑客: **印奇,唐文斌,杨沐**

印奇:本科期间在MSRA实习,涉及**人脸识别**系统研发,哥伦比亚大学博士期间研究方向为**3D摄像**,博二放弃学业创立旷视(经历有点像盖茨).

唐文斌:清华OI总教练,IOI金牌.ACM/ICPC 2008 World Final 第6名。TopCoder前Target选手(ID: OpenGL)

杨沐:IOI金牌.

不得不说清华的圈子确实很强,旷视的重要人脉来源就是清华,其核心几乎清一色清华出身(配置豪华),清华本科Columbia毕业的张弛(内部称驰爷)于2015从Google离职加入Face++,负责行人车辆视频方向的研究([信息来源](#).而国内OI圈是一个小且精的圈子,前辈对后辈的影响非常大,经常抱团,所以这些OI界的传奇前辈会对后辈有不少的吸引力,像[范浩强](#)(清华人称范爷),周而进(IOI金牌),杨弋(NOI金牌)这些清华后辈加入旷视有很大原因在此,与汤晓鸥的商汤围绕其实验室和学生的人才体系很不一样.

2. 首席科学家**孙剑**

MSRA的首席科学家,是ResNet和Faster RCNN这两个目前深度学习非常常用的算法文章的通讯作者.其近些年的H指数非常高(与汤晓鸥差不多),证明其在深度学习领域的深厚功底.



Jian Sun

Chief Scientist | Managing Director of Research, Megvii (Face++)

Verified email at megvii.com - [Homepage](#)

Computer Vision Deep Learning Computer Graphics

[FOLLOW](#)

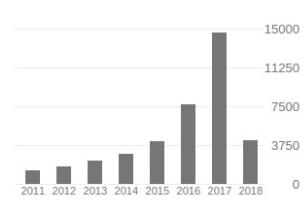
[GET MY OWN PROFILE](#)

| TITLE | CITED BY | YEAR |
|--|----------|------|
| Deep residual learning for image recognition K He, X Zhang, S Ren, J Sun Proceedings of the IEEE conference on computer vision and pattern ... | 7574 | 2016 |
| Faster r-cnn: Towards real-time object detection with region proposal networks S Ren, K He, R Girshick, J Sun Advances in neural information processing systems, 91-99 | 3081 | 2015 |
| Single image haze removal using dark channel prior K He, J Sun, X Tang IEEE transactions on pattern analysis and machine intelligence 33 (12), 2341 ... | 2548 | 2011 |
| Delving deep into rectifiers: Surpassing human-level performance on imagenet classification K He, X Zhang, S Ren, J Sun Proceedings of the IEEE international conference on computer vision, 1026-1034 | 2281 | 2015 |
| Learning to detect a salient object T Liu, Z Yuan, J Sun, J Wang, N Zheng, X Tang, HY Shum | 1959 | 2011 |

Cited by [VIEW ALL](#)

All Since 2013

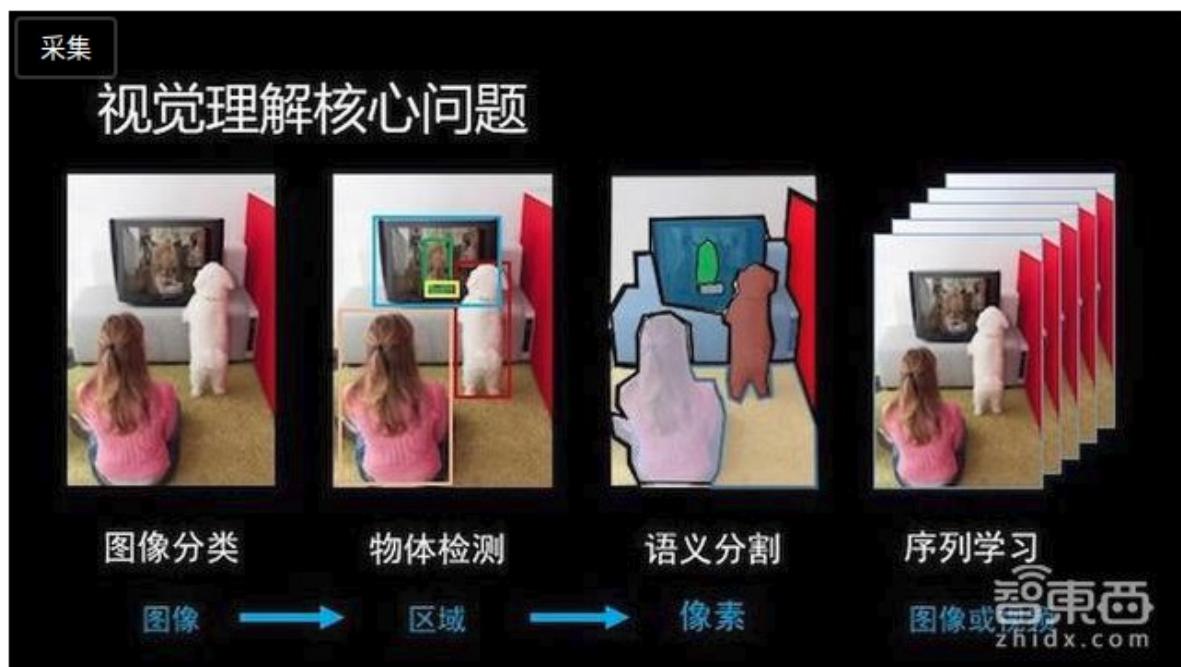
| | | |
|-----------|-------|-------|
| Citations | 43146 | 36226 |
| h-index | 68 | 64 |
| i10-index | 144 | 142 |



Co-authors [VIEW ALL](#)

(数据来源: google scholar: https://scholar.google.com/citations?user=7aQ_YLwAAAAJ&hl=en)

根据其自身的说法:视频理解的核心问题为:



根据其对Face++内部的描述(信息来源),face++对于算法模型的工程化部署非常重视:

我们的研发部门里有一个专门的小组负责研究如何训练最好的基础神经网络，并沿着以下三个子问题深入：

- 1) 针对不同计算复杂度下设计最优的神经网络；
- 2) 针对不同计算平台的实际要求，来设计最高效的网络；
- 3) 针对不同问题设计最合适的网络。

另外对神经网络模型的压缩和低比特化表示也是我们研究的重点之一。

其认为Face++与MSRA研发部门的人的特点很像：

- 1) 他们都同样有着既基于产品，又探索前沿技术的使命定位；
- 2) 他们都同样聚集着一群追求极致，有Geek精神，且高自我驱动的精英；
- 3) 他们都用同样的套路推进研究工作：确定问题->实现、研究和理解既有方法->进行持续改进或创新。

人才培养方向：

我们大致将人才向两个方向培养：研究科学家，和全栈人工智能工程师。研究科学家主要聚焦在算法上，寻求对问题的本质解，我们的培养目标是成为能独挡一面领域专家；全栈人工智能工程师是我们内部的叫法，目的是培养即能上九天揽月（算法设计和训练），又能下五洋捉鳖（算法的工程化，研究问题和方式系统化）的全能战士，他们既能做research,又懂system，能建系统、造轮子。针对目前AI发展的趋势，我们需要大量的全能人才来将AI“+”到不同的行业上，解决实际问题。这就对人才提出了更高的要求。我们相信即便没有AI背景的工程师，在这里工作1-2年后就能成为独当一面的人才。

做事方式：

有了一帮志同道合的小伙伴们，就要围绕这上面介绍的四个视觉理解核心问题开展研究、并将研究成果应用在具体的视觉识别场景中。Face++的研发团队扁平化，每个研究小组由2-4人组成，聚焦一个课题。课题可以是短期的，例如对已经应用的某个产品线上的算法的改进；课题也可以是长期的，例如持续提升识别的精度和性能。我们的每个研究员都可以在不同的课题之间自由切换，这样能最大程度发挥个人的长处和积极性，同时也让大家有机会短时间了解更多的问题，有更丰富的经历，能更快的成长。套用现在深度学习的精髓，每个人的学习也需要输入大数据。

研发环境：

做深度学习研究需要一个非常高效的训练引擎/平台和充沛的计算资源，Face++内部使用了近两年的“MegBrain”是一个全自主研发的训练引擎，它与目前流行的TensorFlow（Google一年前发布）设计相似，同属基于Computing Graph的新一代训练引擎。为什么非要自研系统呢？公司研究深度学习开展得非常早，当时还没有很好用的系统，并且MegBrain在同时满足灵活性及精简性的基础上，能最大限度提升工作效率。目前在AI创业公司中完全使用自研深度学习训练引擎的，可能只有Face++。

总的来说，孙剑的加入无疑使旷视有了更多接触和拉拢学界工业界巨头的机会，旷视美国研究院的成立和王珏的加入也说明了这一点：

“主要是三个方面的原因。一是和旷视有天然的亲近感。旷视是一帮有才华的清华年轻人创立的，在企业文化中有清华行胜于言，脚踏实地的价值观。对于作为清华校友的我来讲，对于这样的文化有天然的熟悉和亲近感。二是孙剑博士的诚意邀请。我和孙剑博士认识和同行多年，彼此都比较熟悉，他不远万里到西雅图来邀请我，我内心是无法拒绝的。第三是机遇。我个人觉得中国目前人工智能研究的环境，从数据规模到商业机会都比美国要好，但是在人才储备上仍然不占优势。能够帮助中国的人工智能公司，特别是初创公司更进一步，在美国吸引顶尖人才为我所用，对中国高科技发展是一件里程碑性质的事情。这样的机会在人生中并不多见，碰到了就应该抓住”。

(信息来源:http://www.sohu.com/a/164104646_418390)

此外,孙剑透露的旷视内部的研究状况的信息应该还是比较准确的,在大批OI金牌码农的强大码力下,旷视对于深度模型在不同硬件条件下的优化和部署很早就开始研究并且现在是由一定优势的.从研究方向来看,目前旷视发表的文献数量已经超过40篇([信息来源](#)),之前发表的文章以人脸分析为主,但包括不少对于深度网络结构和原理的研究;而孙剑加入以后2017年其领导发表的论文已经有9篇,主要涉及**行人检测,分割,模型压缩,检测网络框架**等,使旷视的技术和落地的范围其他领域扩展了不少.

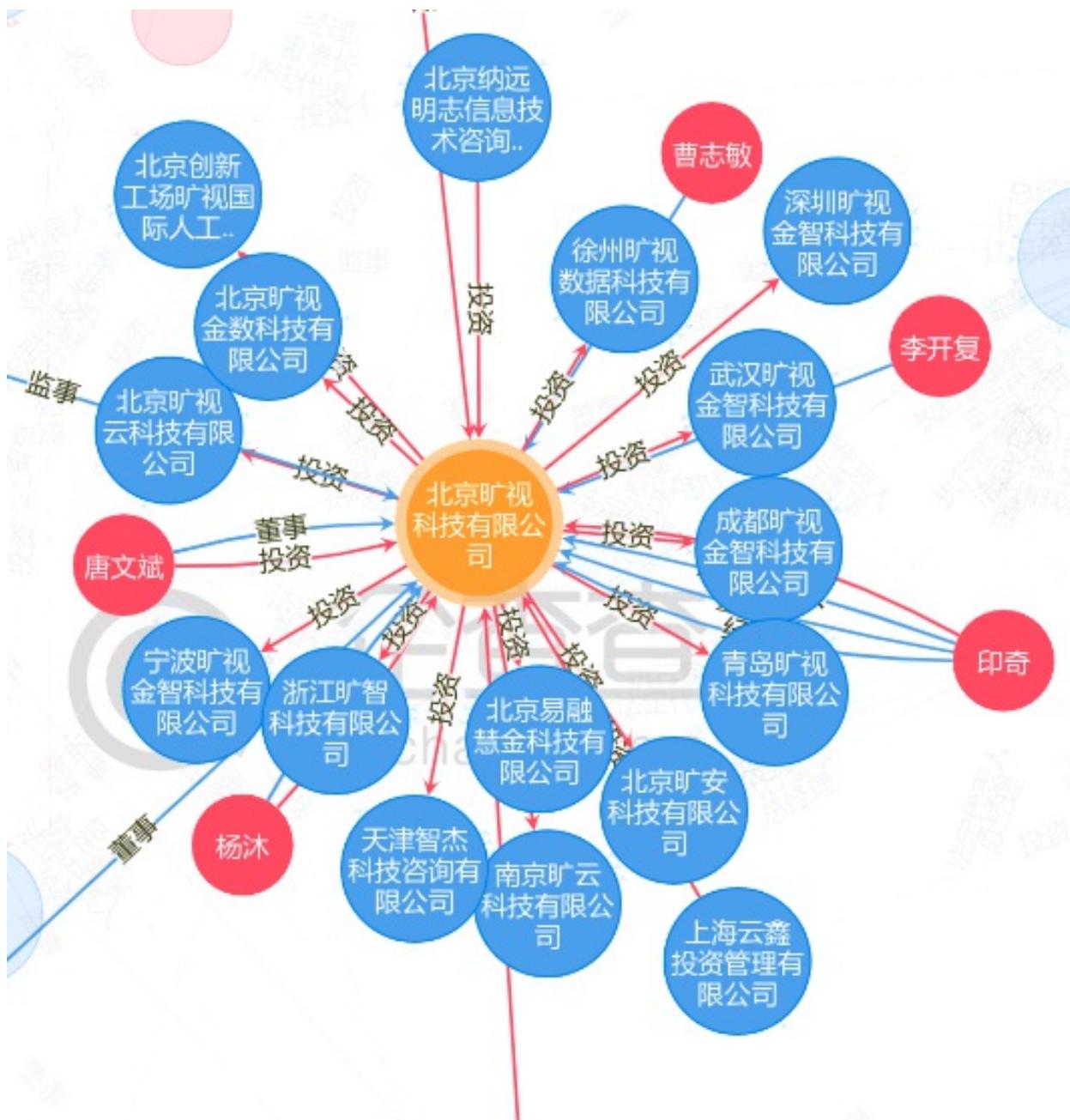
3. 招聘信息([来源:拉钩](#))

从招聘信息来看,目前旷视技术岗社会招聘大部分为**前后端开发工程师(包括java,安卓,ios,python)和测试人员**,虽然算法职位也有,但是以研究员的形式为主而且数量不多,所以与商汤类似,对于算法层面人才的培养是倾向于公司内部培养(可参见上面孙剑的描述);而大量前后端和测试的招聘说明目前旷视还是有大量技术准备落地,很可能是在他们的算法平台上发布的,所以未来一段时间内旷视主页上的核心技术部分的new和beta标签将会持续涌现.

总结,目前旷视最大的优势还是在于1.**人脸相关技术以及积累下来的人脸数据**(1.2亿人([信息来源](#)))(这一点目前远远超过商汤);2.完善的自主算法框架和深度学习平台(**MegBrain**),算法工程化能力很强.

5. 行业布局

根据旷视在官网上新上线的技术和解决方案来看,其目前的行业布局包括以下3个方面:**①安防②金融安全③手机智能**,其中,安防作为人脸识别技术的重要落地场景应该是旷视目前最重视的领域,而从旷视在各地成立分公司,很有可能就是安防业务的拓展和部署的需要(相比于其他两个,安防跟需要进行实地考察和调整).



根据36氪的[报道](#)(2017.10.31),旷视目前的合作伙伴包括:

根据公开信息,目前的合作伙伴包括**国家公安部、国家身份证中心,以及阿里巴巴、蚂蚁金服、华为、联想、富士康、中信银行、招商银行、小米、vivo、万科、凯德等近千家企事业单位**。旷视科技告诉36氪,在金融安全领域,旷视科技旗下的FaceID平台已为全球2.5亿人完成远程实名身份验证,是**支付宝、今日头条、滴滴出行**等知名公司的AI服务提供商;在城市安防领域,旷视科技推出了智能安防系列解决方案,已在全国26个省市上线,并被中国公安部列入重点技术推广目录,作为国家公安部重大活动智能安保方案主要提供商,旷视科技参与了**杭州G20峰会、厦门金砖峰会、海南博鳌论坛、上海劳力士大师赛**等国际级峰会、赛事的安保工作;在手机智能领域,旷视科技与**vivo、小米**等国内一线手机厂商实现合作,推出了人脸解锁手机;此外,在最近比较热门的商业物联领域,旷视科技先后与阿里巴巴、小麦铺,缤果盒子等新零售企业合作,为其提供视觉和感知技术。

- (1). 安防(城市大脑)

从旷视提供的解决方案的多样性来看,旷视应该做了不少安防类的实际项目,并且效果还不错(比如2017年的博鳌亚洲论坛的唯一智能视频技术应用提供商获得来自公安部的表扬([信息来源](#))),能查到的与公安的合作有湖北公安厅,合肥公安局,无锡公安局,铜仁公安局等.

2017年旷视挖来了负责过北京市奥运安保指挥系统、武警总体规划以及海外国国家信息化建设等项目的17年安防老兵[陈雪松](#),说明旷视在对安防行业落地的重视.根据陈的说法:

“AI初创公司纯做技术没有出路, 算法提供商一定要做端到端的整体解决方案, 如果不直接面对客户, 一定会被上下游碾压。”

“早在几年前, 我做奥运安保时就想过采用人脸识别, 那个时候基于模式识别的准确率在数字上达到了85%, 但经实测后发现根本行不通。85%的通过率与0本质上是一样的。而应用深度学习之后, 识别率大幅提升, AI在安防行业的价值才终于得以体现。”

“说到底, 就是要提升算法精度, 最少是小数点后的两个数量级以上。”

“安防行业未来的场景会愈来愈多元。对于试错率容忍度较高的安防“黑名单”追踪来说, 小数点后的几个数量级的增加的确作用不大；但对于红名单、白名单、灰名单而言, 精度提升的意义就非常明显, 99%和99.9%的可靠度差距并不是0.9%, 而是要反过来算, 差距是10倍。”

目前人脸识别落地安防的确是趋势所向, 但人脸识别只是AI落地安防行业的一小部分, 在很多实际场景中, 人脸识别作用并不大。

“‘人形追踪’技术 (或者说行人再识别 (ReID) 技术) 可以很好地解决这个问题, 这也是旷视接下来重点深耕的方向。”

以上说法一方面表明了[解AI的传统行业](#)资深专家在AI技术具体实际落地方案中制定和优化中起了重要作用;另一方面也透露出旷视对在安防领域的进行应用落地思考和努力,其提供解决方案更加贴近实际,相比于商汤强了不少. 此外,旷视开发智能摄像头(MegEye)也很有可能是其在安防领域项目之后的经验总结.根据旷视科技安防事业部解决方案总监安洋的以下说法,旷视对AI技术在安防领域的应用和其对安防产业的影响有自己的计划.

不拒绝做芯片级的算法来赋能其他传统安防厂商, 让后者有能力扩大整个AI安防市场的容量；

旷视不会仅仅充当嫁衣角色, 也会在自己擅长的领域做一些事情。



- (2). **金融**

与商汤一样,旷视在互联网金融领域的落地主要是在**身份认证**方面,这也是目前互联网金融对于CV技术主要且重要的需求.

- (3). **手机智能**

与商汤不同的是旷视在手机方向的合作还是主要以人脸解锁和基于人脸分析SDK的APP应用嵌入为主,针对手机摄像头成像优化,图像增强比较少,官网也确实没有类似商汤的图像增强的demo,说明旷视在这方面的积累比较少.而根据旷视云事业部副总裁吴文昊的判断--60%的摄像头应用在手机上([消息来源](#)),结合旷视强大的移动端算法优化能力,旷视不会放弃这个部分的市场(根据[官网的说法](#)vivo手机的影像优化有用face++的技术).

- (4). **智能零售**

最近(4.8)旷视宣布全资收购艾瑞思机器人(主要产品是仓储物流机器人)进入智能工业机器人领域(两者在2017年就达成合作)([消息来源](#)).而根据2018.4.12旷视科技副总裁敖翔的题为《人工智能公司从巨头公司转向商业化之旅》演讲(如下)中,其提到旷视已把零售作为人工智能的主战场.虽然这个说法可能有些夸张,但是也反映了旷视在以上三个领域取得一定成绩以后开始向其他领域渗透,而首当其冲的就是当下最火热的新零售.此前,与支付宝合作的KPro,缤果盒子也是旷视对于新零售的探索.

从去年8、9月份开始,我们看到零售已经是人工智能的主战场。讲零售就是讲人货场, 特别是我们做智能视觉公司, 一定还是从效率, 一定还是从精准性的角度出发, 做商业。从人的角度, 理解会员, 让商户感动会员。从货的角度, 比如仓储和选品。从场的角度, 我们可以尝试着彻底的数字化, 让所有现在大家已经非常擅长的线上数据化运营的能力和经验运用到线下店铺运营去。

理解一个人，我认识还不够，我要理解你。目前这方面我们在做一些工作，还没有完全商业化，如果说你在看一个广告，但是我们知道你在看广告，广告并没有在看你，传统的模式，我不知道在看我的，是什么样的人，胖与瘦，他有没有戴眼镜，这些信息在传统的体系里面是断裂的，这样一个千人千面的现实或技术对零售会有很大的促进作用。

关于货，两个切入点。一个点是仓储，我们觉得整个物流体系的高效完善一定能改善我们的体验，我们收购了一个公司艾瑞思，提供比较先进的视觉智能的仓储机器人，看起来是一个扫地机器人，但其实它可以搬起500公斤的东西，不需要辅助设备而可以完成搬运。我们尝试完全自主式的，没有预设条件的识别所有的商品，这个案例是一个货架，没有特殊标识，但是商品可以完全识别出来，所有的空档全部都可以识别出来，这样的技术成熟可以完善我们在无人便利店里面的消费体验。

关于场，对一个消费场所来说，我会觉得完全数字化要充分理解消费者、会员用户在一个消费场所里面与环境的互动，可能会铺设摄像头，像这个例子里面我们的案例摄像头。还有，我们可以做一些智能货架，我们已经成立了一个智能货架的产品，里面并不设摄像头，使用的是近距离感应的传感器，这个其实挺便宜的，有了这些之后，不需要真正的完成视觉，你也可以很清楚的感应出一个人的触动、走动，结合一般的视觉和这种相对低端但是近距离的传感技术，我们尝试把一个店面完整的信息抓取出来。设想一下，在一个线上店铺，比如淘宝、天猫或京东，消费者的每个行为都是得到京东捕捉的，也许我喜欢一个商品，我今天加到购物车，明天不想买了我会删掉，这些信息都会用上，这在线下全部是缺失的，我们很多线上的运营经验线下完全无法施展，有一句老话，线下是蒙着脸的，我们需要把所有的信息都搜集起来做数据分析，要了解人的运动轨迹，你有没有看到商品，我们的货品要包装，更令人亲睐，更能吸引人眼球。我们现在在做的事情甚至可以了解人口学的分布，甚至可以更清楚的得到一些更深入的数据，甚至进入到人，这样的数据如果跟店里面的销售数据打通的话，相信会带来巨大化学反应。

(信息来源: <http://news.10jqka.com.cn/20180412/c603884757.shtml>)

- (5). 其他: 从旷视官网上可以知道其提供的解决方案还包括在线营销,汽车行业和教育行业.但是目前并没有特别具体的案例,更多的只是当前技术的整合,尤其是汽车行业更是缺乏自动驾驶所需的关键技术.

最后,印奇的"4+2+x"的商业逻辑:

通过对脸、人、车和文字四个感知对象的识别对现实世界进行数字化的转化和抓取，再经由旷视领先的AI算法为决策和执行端的“手”和“脚”等载体赋能，最后投入众多行业去解决实际问题.

总结:总的来说,旷视的商业落地是做的比较好的,尤其是安防领域,挖了不少业内资深人士来优化解决方案的设计.旷视创世人清华姚班+OI金牌出身的背景使旷视吸引了一批及其优秀的学生和工程师,早在2012年就开始开放的Face++公开算法平台和背后的的Brain++算法引擎也表明其工程化和算法优化能力的优秀,甚至可能是4巨头里面最强的.公开算法平台帮助其磨炼技术的同

时也积累了大量的行业数据,也为其在后面的竞争和业务拓展铺下很好的基础,而在安防领域的解决方案便是其探索AI技术的落地的重要一步.随着孙剑王珏等学术大牛的加入,旷视在技术研发上会更加强大,从而往更多领域进行渗透,使在商业探索上会更加游刃有余.,

参考

- 旷视官网: <https://www.megvii.com/>
- Face++官网: <https://www.faceplusplus.com.cn/>
- 孙剑: 从微软研究院到Face++ 这半年的思考与观察
- 新智元专访 | 旷视科技美国研究院主任王珏 : 做能落地的AI研究
- 旷视科技28岁CEO印奇上榜福布斯 靠的是“刷脸”
- Face++团队, 目前发表的论文有哪些 ?
- 一文看懂旷视(Face++)全新智能人像抓拍机MegEye-C3S !
- 回一封来自公安部的感谢信
- 旷视科技如何推进人脸识别在安防领域的落地 ?
- 创收翻倍背后, 旷视安防的为与不为
- 17 年资历的安防老兵, 在旷视 Face++ 的所闻所感
- 36氪首发 | 旷视科技完成4.6亿美元C轮融资, 中国国有资本风险投资基金领投
- 旷视科技敖翔 : 零售是人工智能的主战场
- 巨擘思维 : 旷视科技占领智能机器人高地

3. 依图科技

1). 发展经历

| | | |
|-------|---|--|
| 2012年 | ○ | 依图科技成立。 |
| 2013年 | ○ | 获得真格基金天使轮融资。 |
| 2014年 | ○ | 获得红杉资本和高榕资本的A轮融资。 |
| 2015年 | ○ | <p>蜻蜓眼系统获得“公安部科技进步奖”。</p> <p>与阿里云合作搭建“贵州公安交警大数据实时作战云平台”。</p> <p>招商银行将“依图人脸识别技术”推广到全国1500家网点。</p> <p>浦发银行“人像识别平台及直销银行远程开户视频认证建设项目”投产，实现VTM和手机银行的人脸身份认证。</p> |
| 2016年 | ○ | <p>完成由云锋基金领投的数千万美元B轮融资。</p> <p>开始在医疗健康，城市数据大脑等领域的探索和实践。</p> <p>搭建全球最大的人像系统，覆盖超过18亿人像。</p> <p>为海关总署及中国边检提供人像比对系统。</p> |
| 2017年 | ○ | <p>完成由高瓴资本集团领投，云锋基金、红杉资本、高榕资本、真格基金跟投的3.8亿元C轮融资。</p> <p>在美国国家标准与技术研究院(NIST)主办的人脸识别测试(FRVT)中获得冠军。</p> <p>在美国国家情报高级研究计划局(IARPA)主办的全球人脸识别挑战赛(FRPC)中获得冠军。</p> |

(信息来源: 依图官网:<http://www.yitutech.com/intro/>)

| 序号 | 日期 | 产品名称 | 级别 | 金额 | 投资方 |
|----|------------|------|-----|---------|--|
| 1 | 2018-03-08 | 依图科技 | D轮 | 金额未知 | 高榕资本 红杉资本中国 高瓴资本 |
| 2 | 2017-05-15 | 依图科技 | C轮 | 3.8亿人民币 | 高瓴资本领投 云锋基金 红杉中国 高榕资本 真格基金跟投 |
| 3 | 2016-06-03 | 依图科技 | B轮 | 1000万美元 | 云锋基金 |
| 4 | 2014-12-01 | 依图科技 | A轮 | 100万美元 | 红杉中国 高榕资本 |
| 5 | 2013-08-14 | 依图科技 | 天使轮 | 100万美元 | 真格基金 |

(信息来源: 企查查: http://www.qichacha.com/company_muhou3?keyNo=cd87470c7d644d45da0ad8cbf177c1ac&name=%E4%B8%8A%E6%B5%B7%E4%BE%9D%E5%9B%BE%E7%BD%91%E7%BB%9C%E7%A7%91%E6%8A%80%E6%9C%89%E9%99%90%E5%85%AC%E5%8F%B8)

依图成立的时间也很早,发展的时间可能比旷视还要长一些,并于2017.5完成**3.8亿人民币C轮融资**,

目前正在D轮融资.其人数在B轮的时候只有**40人**,在C轮的时候差不多有一百人,被外界誉为小而美的团队(虽然周工作时间经常超过110h).专利**24项**.

2). 核心技术

由于依图并没有像商汤和旷视一样通过算法开放平台以API/SDK形式输出算法,更多的是以整个产品和解决方案的形式进行技术输出,因此对于其底层的技术只能从其提供的解决方案中来窥探一二(信息来源: [依图官网](#)).

- **1. 人脸技术:** 其人脸算法主要分为**动态人像**和**静态人像**两部分:其中静态人像包括**10亿级1:1,1:N,N:N人脸比对算法**,精度为**百万分之一**的误报,18亿级**秒级**返回,最小识别人像为**20像素眼间距**,支持**全人种,跨年龄段**;而动态人像包括**人脸跟踪捕获识别技术**,机器支持**万路横向扩展**,单台支持**200路实时**视频流.
- **2. 车辆技术:** 主要包括**车辆车型识别**(基于车头车尾车辆**品牌识别**),**车辆特征检索**(以图搜车),**车牌号识别**(假套牌识别),**图像分割?**(遮阳板,安全带)和**行为检测**(开车打手机)等技术
- **3. 医疗图像检测分割:** 肺结节检测(图像分割,检测,3维卷积),精度92%;骨龄检测,精度0.1岁;乳腺/甲状腺等器官的结节和肿块的秒级检测;手足口病、急性扁桃体炎、急性鼻咽炎、急性支气管炎检测;病历检索(文本分析);病历特征信息提取(NLP)
- **4. 智能硬件和AI芯片:** 目前只知道依图是通过投资Think Force进入智能芯片制造领域的,其他信息尚不明确,具体可以参考余凯博士的[地平线科技](#)

3). 产品服务

依图的核心产品是**智能安防**和**智能医疗**两大块,均有比较成熟的产品落地并大规模使用的.其他产品还包括智能金融,智慧城市,智能硬件.(信息来源: 依图官网:
<http://www.yitutech.com/corebusiness/3.html>)

- **1. 智能安防**

依图目前的核心产品就是在安防领域的**蜻蜓人像大平台**,核心算法就是方面的人脸识别比对,不过由于有公安的大量实际数据支撑,目前依图的人脸识别库已经达到**10亿**级别,应该是这4家里面人脸识别算法在实际应用中最强的.

蜻蜓眼™ 人像大平台

依图科技“蜻蜓眼”人像大平台，采用依图自主研发的人脸识别技术，可进行基于图片的静态人脸比对，和基于视频流、图片流的动态人像识别及比对。系统广泛支持x86、GPU架构；提供了完善，可定制的用户体系，授权机制，权限系统和监管能力；支持跨集群协同联动；针对不同应用场景，提供了一体机，便携式单兵工具等多种产品形态；目前已广泛应用于中国多地省市区公安系统，并取得了丰硕的实战战果。



• 2. 智能医疗

依图另一个核心产品就是医疗垂直领域用于辅助诊断的产品care.ai系列产品:包括care.ai胸部CT智能辅助诊断、care.ai儿童骨龄智能辅助诊断、care.ai儿科门诊智能辅助诊断、care.ai病历智能搜索引擎、care.ai临床智能科研平台,通过CV,NLP技术实现将多维度医疗数据进行整合,建立了一整套包含**数据ETL、数据标注、模型训练学习、临床工作流**集成的数据处理平台.

肺结节检测: 检测率超过92%,对病灶密度体积大小**形态特征**直接描述,根据医生手动修改直接生成报告,重构智能**数字肺**,精准定位

依图医疗胸部CT智能辅助诊断基于深度学习技术，自主迭代算法，不断提高检测敏感性与特异性。智能检测出胸部CT影像的病灶，并结合临床信息对病灶性质进行鉴别诊断最终将诊断结果形成结构化影像报告。



儿童骨龄检测: 精度小于0.1岁,适应范围从新生儿到青春期结算所有儿童.

依图医疗骨龄智能辅助诊断，将海量临床影像和权威医学知识相结合，通过深度学习技术，基于TW-3等主流标准对儿童骨龄进行智能检测。



肿块结节检测: 乳腺/甲状腺肿块秒级检测;在形态学,钙化程度,结构异常化,血管丰富程度等维度上进行精细的描述;降低小结节的漏检;输出BI-RADS/TI-RADS的分级报告

依图医疗超声智能辅助诊断系统，依托于深度学习技术及医生的临床诊断经验实现从多种超声影像中自动勾划多种器官中疑似病变的结节或肿块，并对其良恶性进行分级判断。



儿科常见病辅助诊断，支持手足口病,急性扁桃体炎,急性鼻咽炎,急性支气管炎检测(覆盖门诊总量90%),精度97%;对电子病例进行判读

依图医疗儿科门诊智能辅助诊断，是一款基于儿科专家经验、海量临床病历和先进深度学习技术的临床决策辅助产品。



病例智能检索引擎: 能够理解医疗概念和层级关系并提取搜索词条的医疗术语;理解临床场景下的医学概念和用户搜索意图;病历,病案,检验,影像**多类数据跨平台整合**;指定字段搜索,多词条搜索,检索词条指定逻辑组合 (类似谷歌学术的搜索功能)

依图医疗病历智能搜索引擎，是一个基于先进NLP技术和深度学习的智能病历搜索、查看分析平台。



智能科研平台：构建信息提取模型，自动从临床病历等数据中提取出数千个**特征**信息；深度理解医学术语、医学知识、临床科研路径，整合病历文本与医学影像检验**多类信息**；自动观测变量的相关因素，根据相关变量数据类型与分布自动选择建模算法，并自动完成建模分析、假设检验等步骤(**存疑**)；病例对照研究的数据制备及假设检验分析、典型病例的精确匹配及对照组建立、前瞻性科研自动化信息抓取及队列建立、横断面研究的数据筛选及危险因素自动挖掘(**存疑**)。

依图医疗临床智能科研平台，是一个覆盖临床科研数据制备、数据处理、数据分析全链路工作的科研辅助平台

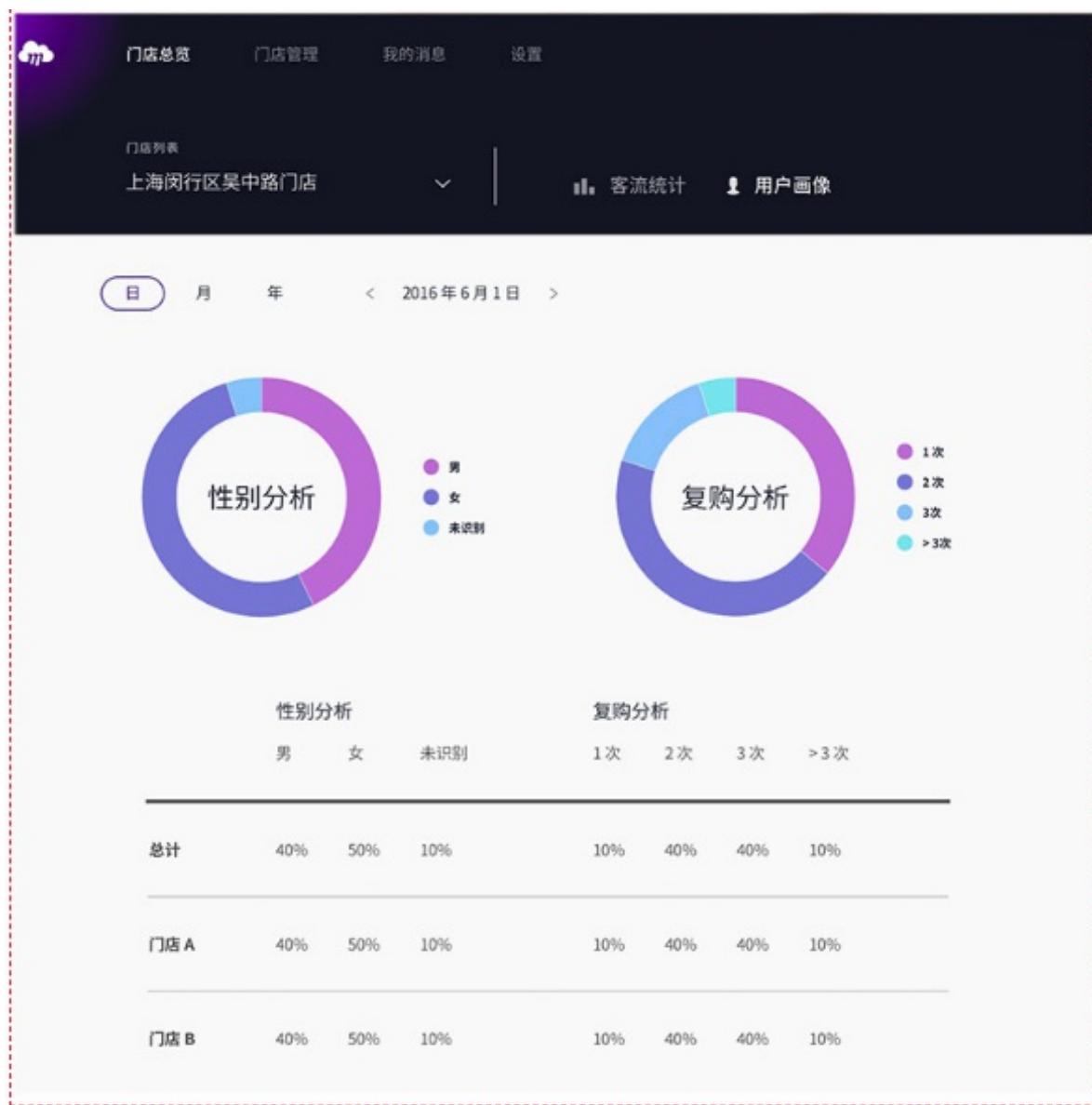


- 3. 智慧金融

这部分产品主要也是整合了**安防**和**人脸识别**技术进行搭建的,这个模式在4家公司中都是一样的,具体包括:刷脸取款解决方案;智能网点解决方案;远程核身解决方案

刷脸取款解决方案: **双目活体检测** 人脸识别系统(对于电子屏幕攻击、打印照片攻击、面具攻击等方式可以在真人通过率达到96%时, 异常情况拒绝率达到99%~99.99%),可识别**故意侧脸、有意遮挡脸部、戴墨镜、戴口罩**

智能网点解决方案: **身份识别(会员,黑名单);客流统计**



远程身份核验: 活体检测功能;人脸1v1比对功能;身份证OCR功能(人证合一)

- 4. 智慧城市: 路网状态建模,车辆行为建模,快速预测算法,信号灯周期相位优化 (杭州城市大脑)

4). 技术团队

朱珑: 目前来看,依图的技术团队担任首席科学家角色的还是创始人朱珑本人,其在传统计算机视觉领域和深度学习领域都有非常优秀的研究经历,个人在CVPR,ICCV,ECCV,NIPS,PAMI这类计算机视觉和算法领域顶会上均有文章,可见其功力非同一般.

美国加州大学洛杉矶分校 (UCLA) , 获统计学博士。师从霍金的弟子艾伦·尤尔 (Alan Yuille) 教授。从事计算机视觉的统计建模和人工智能的研究。麻省理工学院人工智能实验室担任博士后研究员, 深入研究大脑科学和计算摄影学; 纽约大学深度学习鼻祖Yann Lecun的Courant数学研究所担任研究员。

(信息来源: <http://www.yitutech.com/intro/>)

林晨曦: 另一位创始人则是有非常强的工程化能力,在阿里云的时候参与组建团队并搭建了飞天分布式操作系统;在MSRA的时候从事机器学习、计算机视觉、信息检索以及分布式系统方向的研究.2002年曾参加上海交通大学ACM/ICPC代表队并获得全球大学生程序设计竞赛世界冠军,其算法功底也是非常强悍的.

倪浩: 依图医疗总裁

吕昊(首席创新官): 前谷歌研究科学家,美国华盛顿大学计算机博士学位, 研究方向为智能人机交互, 研发和评估智能交互系统及其工具。([消息来源](#))

其他的成员包括MSRA和上海交大培养的本科生: MSRA: 刘缙,周健,吴岷(技术副总裁);上海交大: 朱一和,戎术,张至先,俞梦洁

依图对于人才和团队的理念如下:([信息来源](#))

- 注重任务完成的效率和质量
- 在AI 人才"市场"寻求志同道合者
- 利用大牛和理想吸引优秀的高材生

为了从高校获取优秀的人才, 他们邀请朱珑的博士导师、全球计算机视觉界奠基人之一 Alan Yuille 来华分享授课, 花时间与一流高校大学生交流计算机视觉产业的宏伟蓝图, 强调他们的才华应该用以解决真正的世界性难题。

我们偏向于扁平化, 因为越迟金字塔化, 公司越优秀, 效率就高。公司现在上海交大、清华、浙大刚毕业年轻人会稍微多一点, 创业上这样比较危险, 这也是我们团队比较特殊的地方, 我们的能力在于可以把年轻人培养起来。([消息来源](#): <http://biz.jrj.com.cn/2017/05/18140122498014.shtml>)

- 团队重视稳定性和协同能力, 重视个人能力, 结构相对扁平: 对人才的能力要求很强, 与其他人才连的边界越多, 比如既懂工程又懂算法, 既懂商务又懂产品, 能够跨越流程式的业务模式, 能

缩减管理成本，效率更高。

5). 行业布局

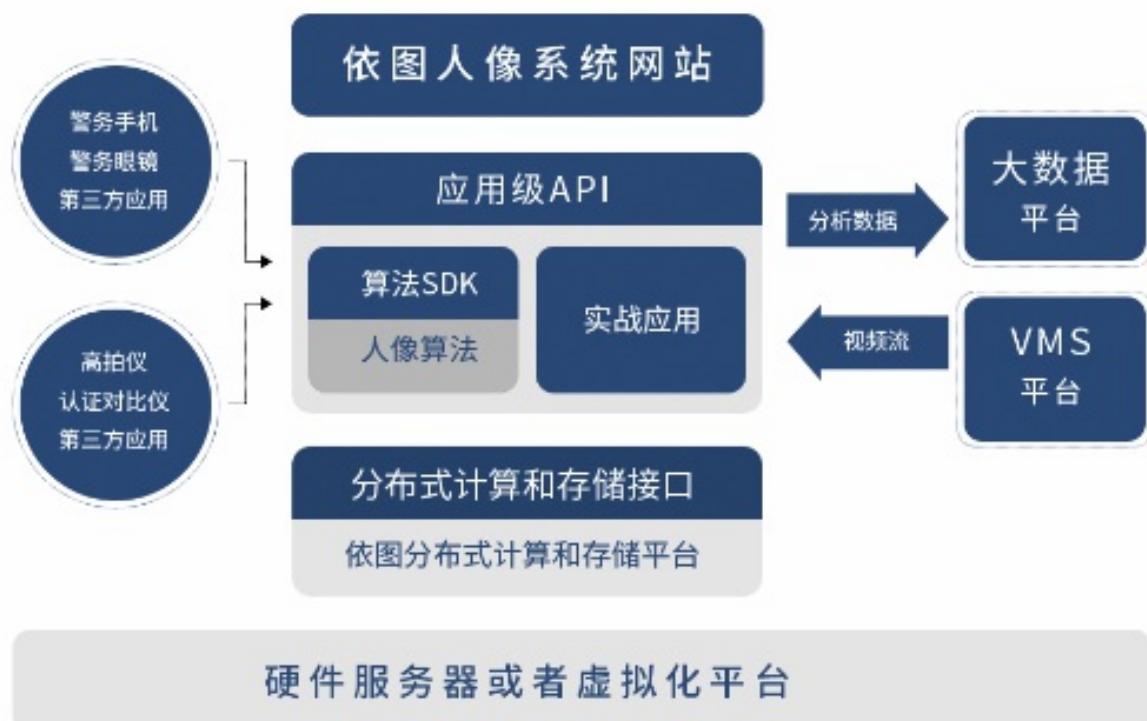
依图的行业布局相比于其他公司还是比较清晰明了的：目前主要就是其产品解决方案中涉及到的几个领域：**安全(安防和金融安全),医疗,城市交通,智能硬件.**

• 1. 安防

依图安防的产品“蜻蜓眼”应该是这几家公司在实际公安场景下功能最完善,最能符合公安部门需求,经过**实际场景**考验的产品,这部分是依图最早摸索和落地的,早在2012年就经过努力在苏州市公安局成功落地以后,借助全国公安系统内部的交流学习所示范和辐射的作用,依图的产品在全国范围内的公安进行扩展,一边在实际场景下打磨算法,另一边结合各地公安不同场景不同需求进一步扩充和丰富产品内容,是产品进一步符合更多用户的需求和认可. ([信息来源](#))

此外,借助公安的数据在搭建了国内乃至全球**最大的人脸比对库**, 依图将产品向海关边检进行扩展.虽然目前还没有像旷视一样在社区安防这类场景有明确的解决方案,但是借助其在公安场景下积累的声誉和产品化经验,个人认为其应该会进入其他安防场景并且竞争优势比较强. 2018.3.30的消息,华为携手依图发布全景式人脸识别解决方案面向的场景就是平安城市和智能社区,可以视为其在传统公安场景下对外拓展的一个尝试([消息来源](#)).

安防类场景的技术解决方案:具体场景包括公安(身份甄别、合成追逃、侦查破案、服务民生), 海关口岸(边检卡口),金融行业应用(实名认证),商业应用(客户人群&VIP识别)



合作伙伴:



◦ 2. 医疗

这部分业务现在应该算是依图区别于其他几家公司的一大特色:在融合大量专家知识的基础上用计算机视觉和自然语言处理技术对医疗垂直领域进行深入研究和开发.相比与之前很火热的移动医疗,依图对自己的定位是 服务于医院的人工智能 ,主要目标有两个: 1.通过计算机视觉技术对医疗影像数据进行分析,辅助医生进行智能诊断;2.提供给用户"医学知识的结构化输出", 建立医生和病人之间更友好的交互平台。因此,依图直接与顶级是三甲医院合作(主要顶级医院的专家心态更开放),并且直接全职招聘了医生来更好的获取专家知识.而依图对于AI在医疗行业理解的三句话可以扩展为AI用于某个垂直领域开发的经验:

公开数据集这件事算是我们的一个教训, 公开数据集可以让你起步, 但不能让你进入临床,因为数据集只局限于一个特定的场景.

第一句话：“仅谈算法是外行的表现，内行的人都谈标准”。算法目前已经趋同,但靠算法去保持先进性是不现实的;

第二句话：“医生大脑里的知识是非常宝贵的，但想挖掘很困难。”短时间无法将一个行业最精华的东西,最深的理解写出来，没有行业理解数据也不可能得到很好的训练

第三句话：“对数据理解的深度，决定了一个AI能够走多远。”对数据的理解，本质上是反映了你对场景的理解深度。

而其对于医疗人工智能的三个层次的理解也可以理解为依图在医疗领域每一步走的方向:

第一个层次是帮助医生去解决重复性问题;

第二个层次是帮助医生去聘请人工智能的专家顾问组，就像辅助诊断，让专家的思维逻辑摆在你的眼前;

第三个层次是帮助医学专家去拓展新的医学疆界，帮助不少医生去做课题，帮助他去发现

一些疾病原来可能没有办法去处理的，不同的特征之间的关联性到底有多大。

此外,相比于图像技术,医疗领域大量的病历,报告是以文本形式保存的,因此更需要使用自然语言处理进行文本挖掘.总的来说,依图对于医疗行业的野心还是比较大的,想成为AI行业深入改造医疗行业的行业制定者: 数据的标准, 临床路径的标准, 效果评估的标准

(信息来源: <http://www.yitutech.com/news/369.html>, 建议阅读 (原始视频可以看2017阿里云栖大会的[医疗专场](#))

合作伙伴:



• 3. 金融安全

目前依图对在金融场景的技术落地方案与其他几家没有区别:主要应用就是身份验证.产品开发并没有想医疗那么深入,个人猜测有两个可能:1. 金融行业对于AI技术的需求并没有那么强烈,改变其他场景的需求并不强烈;2. 依图尚未对金融这个垂直行业展开深入挖掘(人数限制)

合作伙伴:

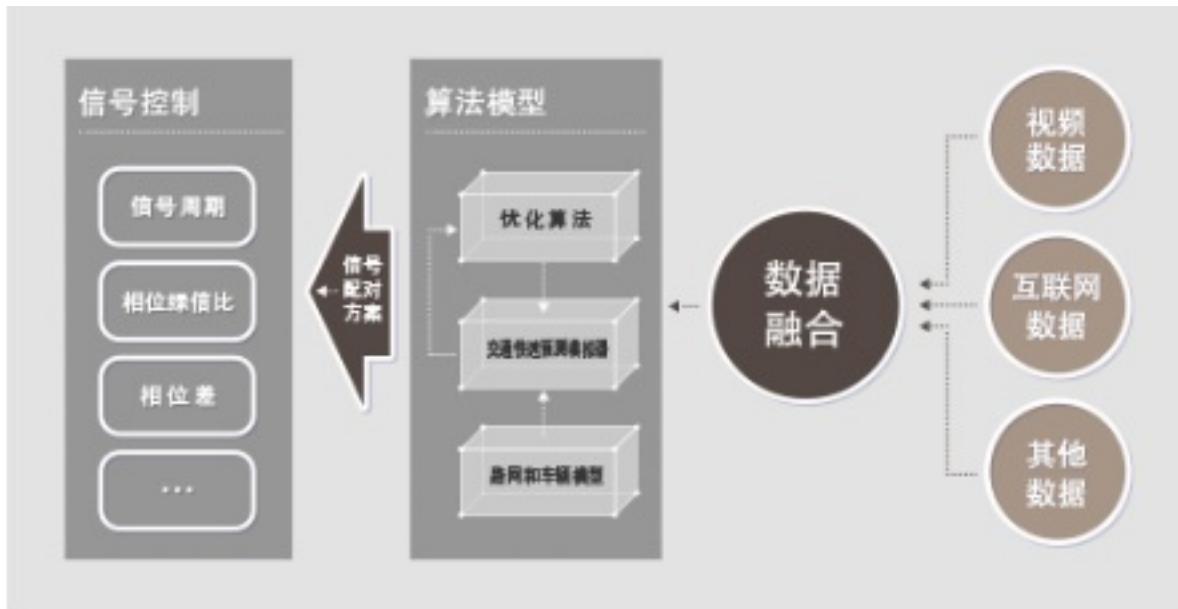


• 3. 智慧城市

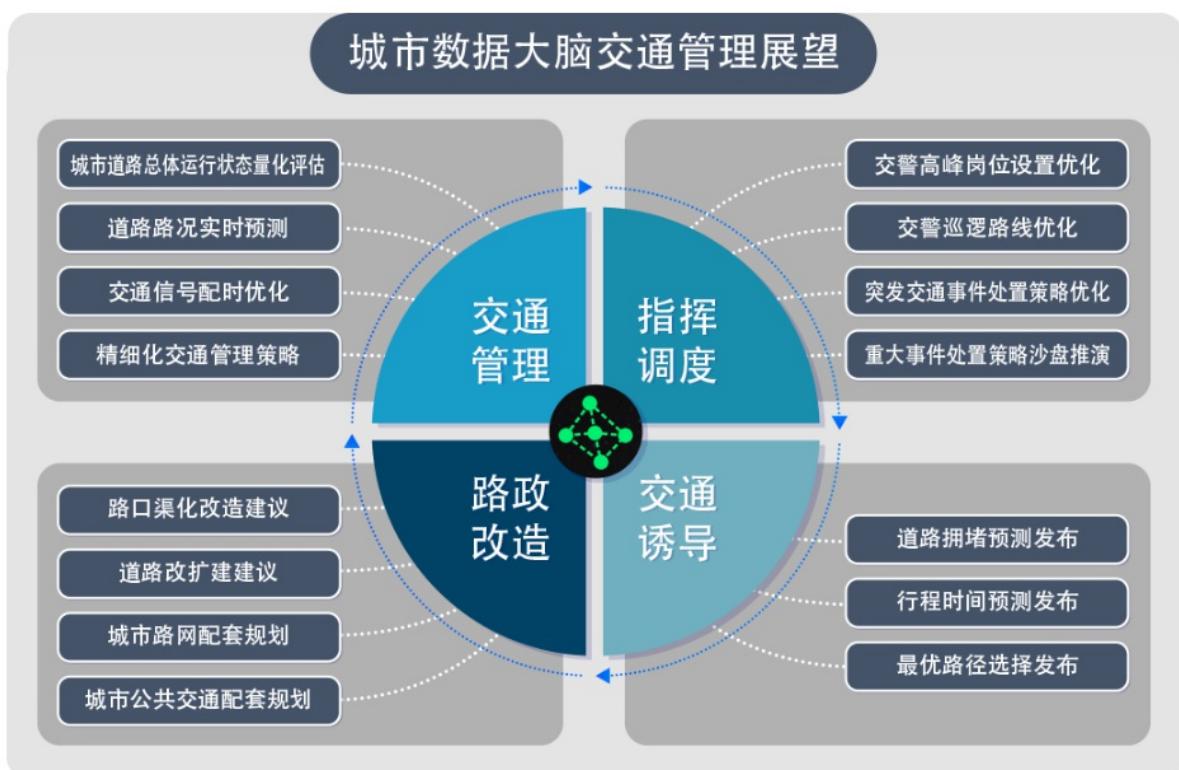
这个目前依图是参与阿里的城市大脑项目来进行落地的,目前最主要的作用就是治堵,那么依图在其中的主要贡献应该就是其车辆识别检测技术.这种项目与其说是AI技术项目,跟不如说是大数据和建模项目,因此个人并不觉得这方面会有很突破性的商业落地方案,更多是作为政

府项目来进行,所以依图会有多大的投入不好说(虽然依图拿政府项目拿的手软).(不过智慧城市这个理念不光在国内很火,在美国也很火,2017年的ICM大赛有一道题就是关于这方面的探索)

智慧城市技术构架:



技术目标:



- 4. 智能硬件

最早了解到依图想做智能硬件方向还是2017.11 AI芯片创业公司ThinkForce的A轮4.5亿融资中有依图的身影([消息来源](#)),虽然目前还没有硬件产品出现,更具官网上对智能硬件的描述,可以发现依图的是将AI芯片作为AI技术进一步渗透到安防金融民生等领域提高智慧体验的重要一步:

智能终端类型



应用场景



ThinkForce对于AI芯片的理念应该可以作为依图对此思考的参照:

AI时代是“算法即芯片”的时代。我们通过对各类AI算法的研究, 提取其共性的运算特征并将它们硬件加速, 辅以相对灵活的固件及驱动程序, 完成与上层应用软件的无缝衔接。依

图科技在这方面的经验和技术积累给我们带来了很大的帮助。同时，AI一定需要落到垂直行业的各个应用场景里面去才能产生价值，我们会和各垂直行业的企业深度合作，与其说用AI颠覆行业，我们更愿意用高效的算力协助业界伙伴创造更多价值。

对于AI创业公司来说,AI芯片是能够在广大算法趋同的AI初创公司中凸显差异和能力的一个重要领域,因为芯片需要比较高的技术要求.提高终端的性能一方面不仅能够提高"云+端"部署方式的灵活性,另一方面智能硬件所具备的**感知,互联,智能**三大要素在进行万物互联的物联网中能发挥更重要的功能,因此通过AI芯片似乎还能在IoT产业进行提前的探索和布局.

- **5. 招聘信息**

从[拉钩网](#)上的招聘信息来看,招募的都是大量的实施后售后人员,说明依图的业务落地已经是比较多的;而没有进行算法人员的招聘更多的能说明前面所说的依图选择从高材生阶层中自己培养需要的算法人才.

总的来说,依图在产品包装上还是很成熟的,其技术都是整体产品和解决方案进行输出的,在技术的使用上确保了专业性从而提升了产品整体使用精度和可靠性;另外,依图并没有像其他几家公司有公开算法平台,通过算法提供商的身份将算法开放集成到其他公司的产品中,依图更多专注于一个整体垂直领域的深入研究和开发保证了其算法对特定领域的针对性优化和先进性,结合该领域的专家知识对产品进行改善优化,能够显著提升算法并在该行业形成技术壁垒.

参考

- 依图官网: <http://www.yitutech.com/>
- 那个只肯接受徐小平100万美金的依图科技, 如今已经估值数亿
- C轮融资3.8亿 依图科技创始人说:我们很贵 别人买不起
- 依图医疗倪浩 :仅谈算法是外行的表现, 内行的人都谈标准
- 依图悄然进军AI芯片领域, 对标地平线, 机器视觉竞争进入下半场
- 资本聚焦AI芯片 :ThinkForce获得依图、云锋、红杉、高瓴4.5亿元投资

4. 云从科技

1). 发展经历



(信息来源:鲸准: <https://rong.36kr.com/project/27486>)

云从于2015.4成立,是4加公司里面成立最晚的,核心团队为从中科院重庆所(绿色智能技术研究院·智能多媒体技术研究中心)的团队,是中科院最大的人脸识别研究团队.目前完成B轮**5亿**融资+广州**20亿**政府资金支持;到目前超过400人(根据知乎小道[消息](#)),专利**16**项.

2). 核心技术

相比于其他平台型公司(旷视和商汤),云从的技术比较单一,目前主要集中在**人脸识别技术**(双层异构深度神经网络,2.4亿张人脸训练集([消息来源](#))),具体包括:人脸检测,跟踪,关键点检测,人脸对齐,活体检测(红外),人脸比对,人脸质量,属性分析,人脸聚类;其他还有OCR.

识别标准

由于目前人脸识别算法是基于人脸二维图片与视频进行的识别，算法系统对于图片和视频均有相应的基本要求。而视频又可以通过编解码技术转化为人脸图片帧进行处理，以下要求仅以图片提出要求：

| | |
|--|--|
|  支持格式 | 所有图片和视频帧格式 |
|  角度 | 旋转角<45°，俯仰角<20°，偏航角<30° |
|  光照 | 光照均匀，背光与曝光不严重均可做识别 |
|  模糊 | 人脸轮廓较清楚，五官位置可辨认 |
|  整体性 | 必须保证完整人脸，人脸轮廓全部在图片内；必须保证眼、鼻、嘴等重要部位不被遮挡 |
|  人脸像素 | 瞳间距15像素以上可检测；瞳间距30像素以上可识别 |
|  图片像素 | 系统支持30*30像素至5000*4000像素图片输入处理 |

算法性能

| | | | | |
|---|------|------|-------|--------|
|  FAR | 1% | 0.1% | 0.01% | 0.001% |
|  Correctness | >99% | >98% | >96% | >92% |

(信息来源:<http://www.cloudwalk.cn/sdkplus.html>)

3). 产品服务

目前主要输出**人脸服务API**和**SDK**(安卓,Linux,Windows平台)为主

(信息来源: <http://www.cloudwalk.cn/product.html>)

4). 技术团队

主要依赖中科院体系以及围绕"国家队"这一身份开展的联合实验室.

1. **周曨**: 中科院百人计划,phd赴UIUC交流阶段师从Tomas S.Huang,近些年发表的**文章**主要与人脸识别技术相关.单从学术水平上来说均不及前面3家公司的汤晓鸥,孙剑和朱珑.

2. **UIUC联合实验室**: 就是周曦老板Tomas S.Huang的实验室,此人被誉为计算机视觉之父,因此其实验室不光帮助云从建立一套核心技术体系,还给云从带来非常丰富的人脉资源.

3. Google联合实验室

4. 中科院联合实验室及上交联合实验室

和创始人[周曦]phd时期的UIUC的老板Tomas S.Huang的联合实验室(此人为被誉为计算机视觉之父)



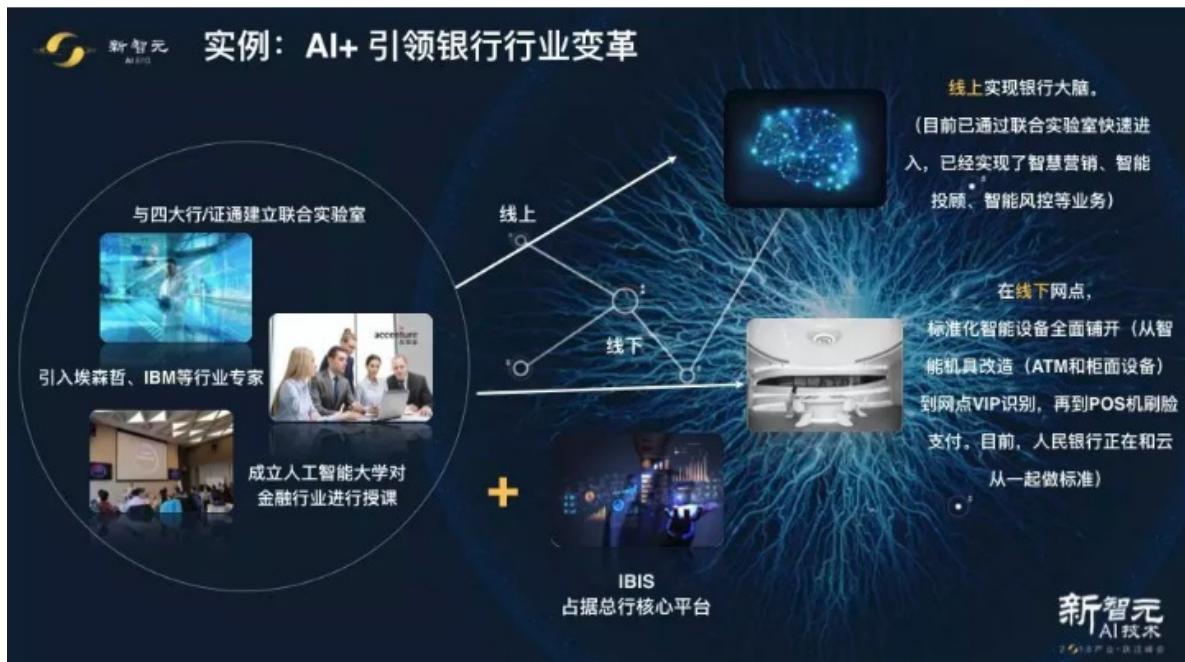
(信息来源: https://mp.weixin.qq.com/s?__biz=MzI3MTA0MTk1MA==&mid=2652016755&idx=2&sn=30b95b15cfad728e9a828c29052f1482&chksm=f121e482c6566d944d65ba04ae88885499cea7dd64a14b8ab8e8a0c12a2fc34e9738f94773b&scene=27)

5). 行业布局

由于云从国家队的背景,所以无论在人才,市场,资源方面都有很大优势:主要体现在1.**人才优势**:与国内外高校诸多的联合实验室;2.**平台优势**:和科大讯飞、百度、腾讯一起参加了国家发改委的项目,人工智能云平台。参与制定行业国家标准;3.**资源优势**:跟公安部, 四大行, 民航总局有联合实验室, 作为中科院的旗下企业, 中科院给我们支持还是比较大的, 建立联合实验室, 因此有大量的数据可以使用开发。

虽然说作为行业标准的制定者,云从可以参与的领域有很多,但目前主要集中在**安防**和**金融**,其中金融是目前的主要业务.

- **金融:** 主要业务与其他几家差不多,都是以**身份核验**为主,但是由于其国家队背景,轻松拿下了四大行(当当然官方宣称是第一),是国内银行方面的最大提供商,另一方面可以参与银行比较



- **安防:** 根据云从的认识: 安防市场比金融市场大几个数量级, 因为有一定规模的银行全国只有100余家, 而安防可以深入到县 ([消息来源](#)), 可以看出云从下一步的方向就是安防, 不过考虑到目前云从的技术只有人脸, 尚无**车辆识别技术**, 因此第一阶段应该还是主要会向银行安防入手, 慢慢向更多的安防场景扩展.

总结: 根据云从对友商不做评价来看([消息来源](#)), 云从核心的人脸技术与其他三家比应该并不占优势. 所以云从的崛起更多的是政策需要. 这一点云从自己也有比较清楚的认识:

三月份的两会上, 国务院总理李克强明确指出要加快培育壮大包括人工智能在内的新兴产业, “人工智能”也首次被写入全国政府工作报告。随后国家发改委公布了《2017年“互联网+”重大工程拟支持项目名单》中包含四个AI公司, 偏居西南的云从科技是其中唯一一家创业公司。

图像识别特别是人脸识别算是比较敏感的国家重点行业。云从代表创新、创业, 并且地处重庆, 全国一盘棋, 选择云从科技, 可以更好的服务于西部大开发, 符合国家整体战略。

信息来源: http://www.sohu.com/a/129575201_473283

云从目前更多是以技术提供商的角色提供sdk, 其官网上的行业解决方案的与其他几家相比还是比较简单的, 产品的包装和落地应该是远远不及其他几家的, 其技术进展与其说是云从推动行业进步到不如说是行业需求推动云从算法改进. 不过由于其国家队的背景, 应该会在国家需要领域占据主要的市场.

参考

- 云从官网: <http://www.cloudwalk.cn/index.html>
- 云从科技姚志强：AI落地需要合作伙伴支持，中国已处于国际领先地位
- 云从科技完成B轮融资 AI国家队继续巩固多行业龙头地位
- 云从孙庆凯：巨头芯片争霸，视觉识别技术是打穿上下游的关键
- 云从周曦：我为何拒掉千亿美元国际顶级资本 | 计算机视觉领军人物专访
- 云从科技创始人周曦：出租屋小公司成长为人脸识别“国家队”

5. 横向对比

| 技术方向 | 具体技术 | 商汤 | 旷视 | 依图 | 云从 |
|------|-----------------|------------|---------------------|-------|--------|
| 人脸分析 | 人脸检测 | √ | √ | √ | √ |
| | 人脸关键点 | √ (240) | √ (106) | √ (?) | √ (?) |
| | 人脸质量 | √ | √ | √ | √ |
| | 人脸活体 | √ (双目+结构光) | √ (视频,动作,炫彩,双角度) | √ (?) | √ (红外) |
| | 人脸特征 (比对,检索) | ** | *** | *** | ** |
| | 人脸聚类 | √ | √ | √ | √ |
| | 人脸属性 | ** | *** (颜值,肤质,情绪) | * (?) | * (?) |
| | 人像美妆 | √ | √ | ✗ | ✗ |
| | 视线估计 | ✗ | √ | ✗ | ✗ |
| | 人脸融合 | ✗ | √ | ✗ | ✗ |
| 人体分析 | 人体检测 | √ | √ | √ | ✗ |
| | 人体抠像 | ✗ (关键点) | √ | ✗ | ✗ |
| | 人体(服装)属性 | *** | ** | ✗ | ✗ |

| | | | | | |
|------|-----------------|----------------------|------------------------|------------|---|
| | 手势识别 | x | √ | x | x |
| OCR | 证件识别 | **(卡证+票据) | *** (各种证件, 通用文字) | * (身份证) | * |
| | 车牌识别 | √ | √ | √ | x |
| | | | | | |
| 车辆分析 | 车型识别 | √ | √ | √ | x |
| | 车辆属性 | √ | x | x | x |
| | 以图搜车 | x | x | √ | x |
| | 车辆轨迹 | √ | x | √ | x |
| 图像处理 | 视频图像增强 | *** (去雾,超分,滤镜...) | * (人像相关) | x | x |
| | | | | | |
| 视频挖掘 | 安防监控 | * | *** | *** | * |
| | 内容审核 | √ | x | x | x |
| | 视频标签 | √ | x | x | x |
| 3D视觉 | 视频内容结构化 | √ | x | x | x |
| | 3D物体识别跟踪 | √ | x | x | x |
| | SLAM | √ | x | x | x |
| | SFM | √ | x | x | x |
| | 实时深度估计 +3D重构 | √ | x | x | x |
| | 轻量化VR/AR引 | √ | x | x | x |

| | | | | | |
|------|-----------------|------|---|---|---|
| | 擎 | | | | |
| | 机械臂抓取算法 | √ | ✗ | ✗ | ✗ |
| 自动驾驶 | 行人车辆检测 +方向估计 | √ | ✗ | √ | ✗ |
| | 车道线检测 | √ | ✗ | ✗ | ✗ |
| | 场景理解 (物体分割) | √ | ✗ | ✗ | ✗ |
| | 交通表示识别 | √ | ✗ | ✗ | ✗ |
| | 大规模3D高精度重建 | √ | ✗ | ✗ | ✗ |
| 医疗分析 | 肿瘤结节检测定位 | ✗ | ✗ | √ | ✗ |
| | 骨龄检测 | ✗ | ✗ | √ | ✗ |
| | 儿科常见病识别 | ✗ | ✗ | √ | ✗ |
| | 病历检索 | ✗ | ✗ | √ | ✗ |
| | 医疗信息提取建模 | ✗ | ✗ | √ | ✗ |
| 芯片 | AI芯片 | √(?) | ✗ | √ | ✗ |
| | | | | | |
| 通用分析 | 通用物体检测 | √ | √ | ✗ | ✗ |
| | 通用场景识别 | √ | √ | ✗ | ✗ |
| 平台 | 开放算法平台 | √ | √ | ✗ | ✗ |