

# 中国计算机视觉行业研究报告

2017年



# 摘要



深度学习主要提升的是计算机视觉领域分类任务的准确率；开源环境仅降低计算机视觉领域的入门技术门槛，前沿算法的技术壁垒依然存在；计算机视觉比赛成绩、论文成果不直接代表技术团队解决实际业务问题的能力。



2017年中国计算机视觉规模预期为40亿，凭借安防领域的爆发性增长，预期2020年将增长至725亿。



前端嵌入式智能系统的渗透率将逐步提升，与后端协同智能计算，加速产业智能升级。



算法迭代将不断提升限定场景识别准确率，加速渗透为各行业应用赋能。



前沿算法之外，计算机视觉公司的商业壁垒有赖于产品、服务、市场等综合建设。

计算机视觉技术概述

1

计算机视觉行业概况

2

计算机视觉的应用场景

3

计算机视觉典型公司案例

4

计算机视觉行业发展趋势

5

# 计算机视觉的概念与行业主要应用

## 计算机视觉横跨感知与认知智能，现阶段应用以感知为主

视觉使人类得以感知和理解周边的世界，人的大脑皮层大约有70%的活动在处理视觉相关信息。计算机视觉即通过电子化的方式来感知和理解影像，以达到甚至超越人类视觉智能的效果。从1966年学科建立 ( MIT : The Summer Vision Project ) 至今，尽管计算机视觉在感知与认知智能方向仍有大量难以解决、尚待探索的问题，但得益于深度学习算法的成熟应用 ( 2012年，采用深度学习架构的AlexNet模型，以超越第二名10个百分点的成绩在ImageNet竞赛中夺冠 )，侧重于感知智能的图像分类技术在工业界逐步实现商用价值，助力金融、安防、互联网、手机、医疗、工业等领域智能升级。2017年下半年，数家计算机视觉公司单笔融资上亿美元，再次将计算机视觉推向人工智能领域最受关注的方向之一。本报告聚焦于计算机视觉技术现阶段在工业界的应用与研发，将在对相关技术热点及产业整体做概述性介绍的基础上，对典型应用的具体使用场景分领域论述，希望理清现状，写明征途，供产业界、投资界、政策制定者及关注人工智能领域的社会各界以参考。因报告研究对象以技术层创业公司为主，偏颇遗漏之处，敬请指正。

### 视觉感知小有所成，视觉认知道阻且长



### 计算机视觉的典型应用



# 计算机视觉的学术研究任务

## 现阶段有较好商业化进展的主要为语义感知中的分类任务

与人类实时选择性处理视觉信息不同（如人在驾驶时不需在意公路边草地的纹理或形状，也不用知道每辆车的确切形状），计算机仍难以从实际需求出发自主选择性输入并计算影像信息，通常需要人类对具体任务进行分解并使用与之匹配的计算方法，建立完整理想的智能视觉系统仍有很大挑战。另外，与可结合常识做猜想和推理进而辅助识别的人类智能系统相比，现阶段的视觉技术往往仅能利用影像表层信息，缺乏常识以及对事物功能、因果、动机等深层信息的认知把握。

### 计算机视觉的学术研究任务（部分）

语义感（认）知 建立影像信息与语言描述的映射	 <p>图像分类      物体检测      语义分割</p> <p>分类：图像分类、物体检测、语义分割等，分别为图像整体、局部以及像素级的分类任务</p> <p>对象关系识别：空间关系、动作关系、介词方位关系等</p> <p>行为识别及预测、追踪</p> <p>看图写话：场景理解，语句、段落等生成</p>	去模糊、超分辨率等技术 可有效增强影像信息
		3D目标分类及对象关系、姿势行为分析、相关场景理解
		3D重建：从影像信息中重建这个世界的几何模型，一般从二维影像出发得到场景的三维表示，可利用运动视差、双目立体视觉、多视图、纹理、明暗、轮廓、物体及场景几何结构等线索
	定位	基于影像信息（往往也结合惯性运动单元 IMU等其他传感器获取多维度信息） 机器人（包括自动驾驶汽车）在环境中对自身进行定位，进而导航
	控制	指导机器人操纵物体，拾起、抓住、转动等，自动驾驶领域则涉及横向、纵向及障碍物避让等控制 是一个需要不停观察、规划的连续决策问题

来源：参考资料 Stuart Russell.人工智能：一种现代的方法[M]等。

# 核心算法之深度学习

## 深度学习的概念及成就

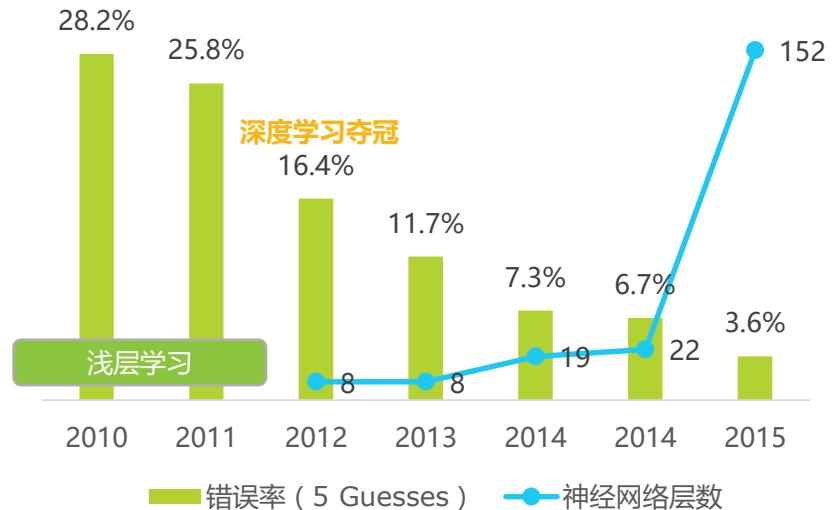
深度学习算法由多重非线性变换构成多个处理层，辅以复杂结构设计和各种梯度技术，通过对大量样本的输入与对应输出数据的抽象计算，拟合出一个可处理新输入信息的函数模型，解决其分类或预测问题。尽管神经网络（可理解为深度学习算法的前身或别称）的研究历史其实比计算机视觉的研究历史还要长，但在之前的研究过程中，深度神经网络的方法并未得到有效验证。直到2011年，语音识别领域凭借深度学习取得重大突破；2012年，AlexNet，一个8层的神经网络，以超越第二名10个百分点的成绩在ImageNet竞赛中夺冠（图像分类的Top 5错误率为16.4%），深度学习终迎来包括学术探索与工业应用中的热潮。不断提升的层数逐步提升计算机分类视觉的准确率，2015年夺冠的ResNet深达152层，以3.57%的错误率超越人类视觉的5.1%。

### 深度学习与传统方法的区别

传统方法，针对不同类别的物体，一般首先由研究员充分发挥聪明才智，手工定义不同的特征，然后利用不同的机器学习算法（分类器学习），这时的算法一般仅有两到三次非线性变换，学到的参数较少（浅层学习）；深度学习则通过机器学习自身来产生特征，因此特征和分类器学习不再有区分。如今深度学习的算法已可达上千层。



### ImageNet Large Scale Visual Recognition 2010-2015



来源：参考资料 周志华.机器学习[M]等。

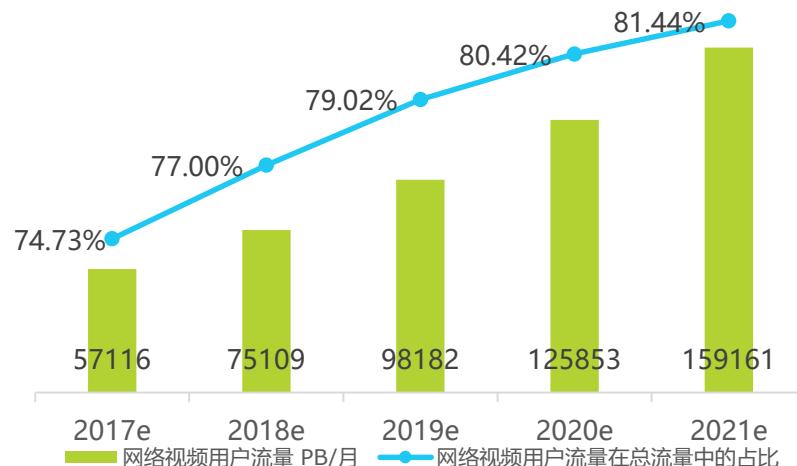
来源：参考资料 CS231n.Convolutional Neural Networks for Visual Recognition等。

# 核心算法之深度学习

## 数据与算力是深度学习的重要支撑

日益丰富的影像内容为深度学习算法提供了大量的数据支撑。据思科公司评估，2021年单月上传至全球网络的视频总时长将超过500万年，每秒将诞生1百万分钟的网络视频内容，网络视频流量将占据全球所有网络用户流量的81.44%。需要说明的是，现在的学习多为有监督学习（需要对数据进行充分标注），而且并非所有类型的影像数据都易得易标注（比如医疗影像数据需由专业医师标注病灶），业界领先的视觉公司一般会有数百人的标注团队（多为外包，但需专业培训，实时指导）。另一方面，深度学习学习过程中的“训练”与应用部署后的“推断”均涉及大量并行计算，传统CPU算力不足，而GPU、FPGA（现场可编程门阵列）、ASIC（TPU、NPU等AI专属架构芯片）等具有良好并行计算能力的芯片可提供数十倍乃至上百倍于CPU的性能，与云服务一起，大幅缩短计算过程（在过去，往往数周甚至数月才能跑出一次结果，然后调整模型架构，效率极低），易于短期调整多种模型架构，显著提升分类模型的进步速度。2010年以后，CPU内部晶体管数量的增长明显放缓，传统摩尔定律失效，而GPU类处理器依然保持着快速增长的势头（2016年GPU的计算力为10个TFLOP/S，2017年达到了120个TFLOP/S，TPU则实现了惊人的180个TFLOP/S），验证着AI时代的摩尔定律。

2017-2021年全球网络视频用户流量



并行计算能力有效提升模型计算效率



GPU、FPGA、ASIC等并行运算处理器  
在云端服务器、用户终端等  
助推模型训练与推断

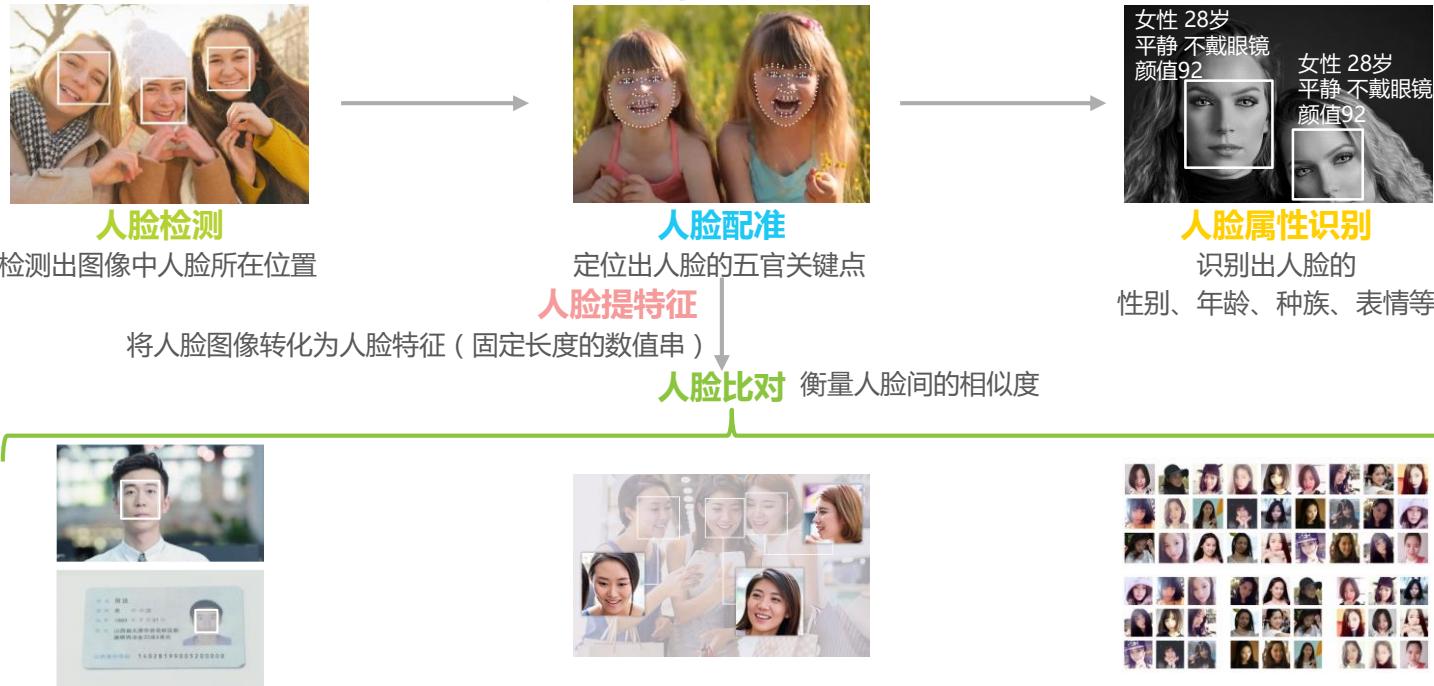
来源：艾瑞咨询研究院自主研究绘制。

# 热门技术之人脸识别

## 人脸识别是当下视觉领域热门应用的重要技术支撑

人脸识别可看做语义感知任务中针对人脸影像的分类问题，也是当下视觉领域热门应用的重要技术支撑，各个环节都因深度学习算法的推进实现了更优的计算结果。泛金融领域的远程身份认证、手机领域的刷脸解锁一般属于人脸验证（技术相对成熟）；安防影像分析一般为人脸识别，刑侦破案对亿级甚至十亿级比对有刚性需求，目前技术仍有很大进步空间，更多新功能、新场景的解锁依赖于最先进的算法团队和相关业务领域开拓者的共同努力。

### 人脸识别的技术流程



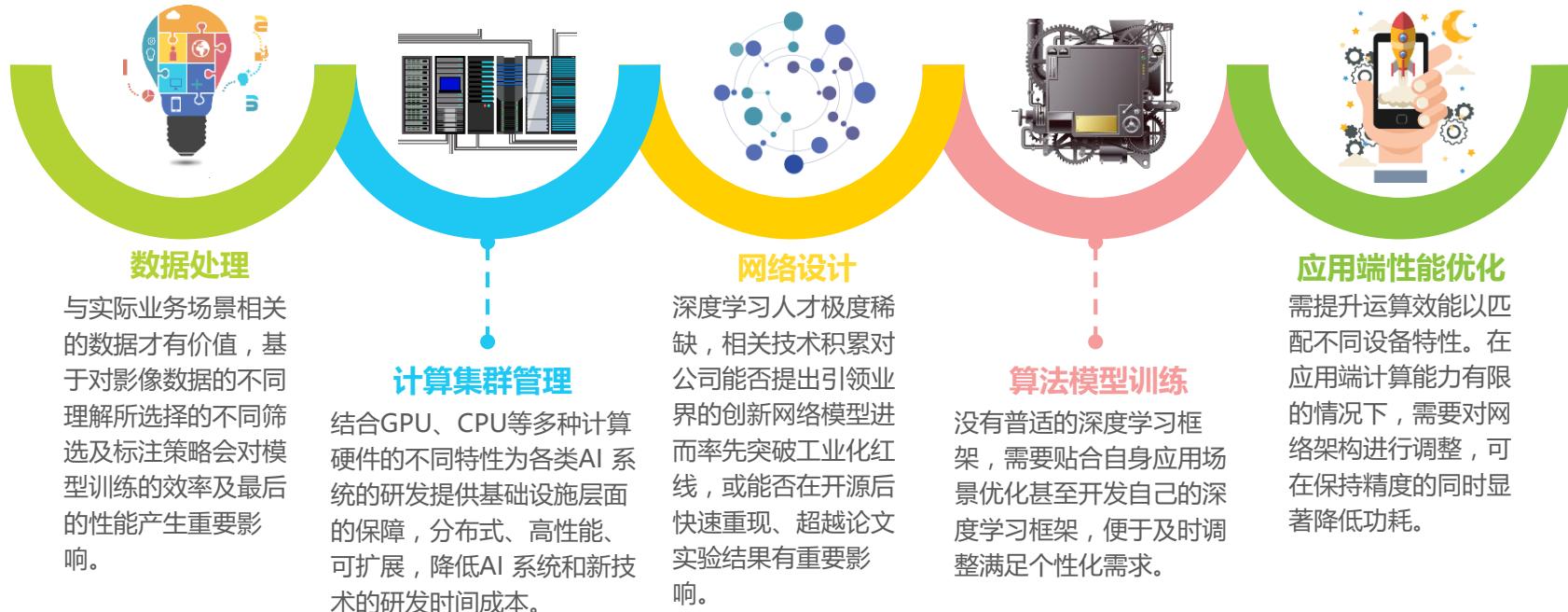
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 开源环境与技术壁垒

## 开源环境仅降低计算机视觉领域的入门技术门槛

工业界和学术界先后推出了用于深度学习模型训练的开源工具和框架，包括Caffe、Theano、Torch、MXNet、TensorFlow、PaddlePaddle、CNTK等等，极大降低了人工智能技术在工业实践中的入门门槛。尽管不同框架各有所长，但它们并不能真正满足企业在处理实际复杂业务时所面对的挑战，性能、显存支持、生态系统完善性、使用效率等不同层面的不足要求企业需要针对性的调整框架以适合自身业务所需。而在数据处理、计算集群管理、网络设计、应用端性能优化等若干重要环节都存在各种各样非开源技术或已成熟方案所能解决，极度依赖相关技术专家去探索求解的重要问题。对于前沿算法的突破创新以及算法在不同使用环境中的优化升级，不同公司的技术高低差异依然很大。

### 深度学习算法研发中的重要环节



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 计算机视觉比赛的是与非

## 比赛的意义在于推动算法思想的进步

2007年由李飞飞教授发起的ImageNet计划将人工智能领域的影像数据推向了前所未有的规模，至今已有1400万张经过人工清洗标注的图片，含有2万个分类，为计算机视觉领域做出巨大贡献。自2010年每年一度的ImageNet物体识别竞赛（对1000类接近50万张图片的单标签识别），更是成为了计算机视觉领域最受关注的比赛。2017年，ImageNet举行了最后一届图像分类竞赛，Top 5的错误率降至2.25%（大幅领先人眼），该竞赛完成了历史使命，而更多的关于图像语义分割（像素级的分类问题）、1:N人脸识别、图像及视频理解（看图写话）方向的数据集与比赛将逐步登上舞台，推动更多领域更贴近真实世界场景的算法革新。

### 计算机视觉的数据竞赛成果（部分）

<b>ImageNet</b> <b>2.25%</b> 图像分类（共1000类） Top 5错误率	<b>LFW</b> <b>99.82%</b> 1:1 人脸验证 准确率 13000多张面部图像
<b>WebVision</b> <b>94.78%</b> 图像分类（共1000类）准确率 训练数据非人工标注	<b>FRVT</b> <b>95.5%</b> 1:1 人脸验证 准确率 百万量级数据盲测 来自真实边检、刑侦场景
<b>COCO</b> <b>0.463</b> 图像语义分割 平均精确度	<b>MegaFace</b> <b>91.7%</b> 1 : N 人脸识别 准确率 百万量级测试数据

### 计算机视觉比赛的价值



视觉比赛的算法设计一般不太考虑对计算资源的消耗和实际场景的可用性，所以相关算法不能直接用于实战。比赛成绩或许可用来判断对方技术是否位于世界前十，但并不代表相关公司服务客户、解决实际问题的能力。比赛的意义在于推动算法思想的进步，能否提出创新的思想更重要。

来源：艾瑞根据专家访谈、公开资料等研究绘制。

计算机视觉技术概述

1

计算机视觉行业概况

2

计算机视觉的应用场景

3

计算机视觉典型公司案例

4

计算机视觉行业发展趋势

5

# 2017年中国计算机视觉行业图谱

iResearch  
艾瑞咨询

## 计算机视觉行业图谱

应用层  
技术使用者

公共安全 金融 互联网 手机 交通 医疗 零售 工业 智能硬件

技术层  
算法、产品及  
解决方案提供者

### 安防影像分析



### 泛金融身份认证



### 手机及互联网娱乐



### 批发零售 商品识别



### 嵌入式 智能系统



### 工业制造



### 广告营销



### 医疗影像分析



### 自动驾驶



### 服务器 芯片 传感器



基础层  
数据及计算资源

### 计算平台



数据

说明：本报告的调研对象以2012年深度学习爆发以后的创业公司为主，未能对起步于传统方法做视觉相关业务的公司做充分研究，因此本图谱未包含OCR、传统门禁等相关领域。

# 计算机视觉公司的主要商业模式

## 主要服务方式



### API

提供云端比对识别服务  
弹性、灵活、高效  
互联网金融公司多采用此种形式



### SDK

向客户提供核心算法模块  
在用户端或客户的服务器端完成视觉计算  
基于对数据的保护或实时性要求  
手机及互联网娱乐公司多采用此种形式



### 解决方案

向客户提供较为全面的软硬件集成解决方案  
可能涉及软件系统、嵌入式解决方案  
前端硬件设备、专有服务器部署等  
公安、银行多采用此种形式

## 主要收费方式



### 按调用量或包时收费

API服务多采用此种收费方式



### 结合授权设备量及授权周期定价

SDK服务多采用此种收费方式  
解决方案也可能采用



### 结合具体项目收费

后续每年可有升级维保收入  
解决方案多采用此种收费方式  
API和SDK也可能采用

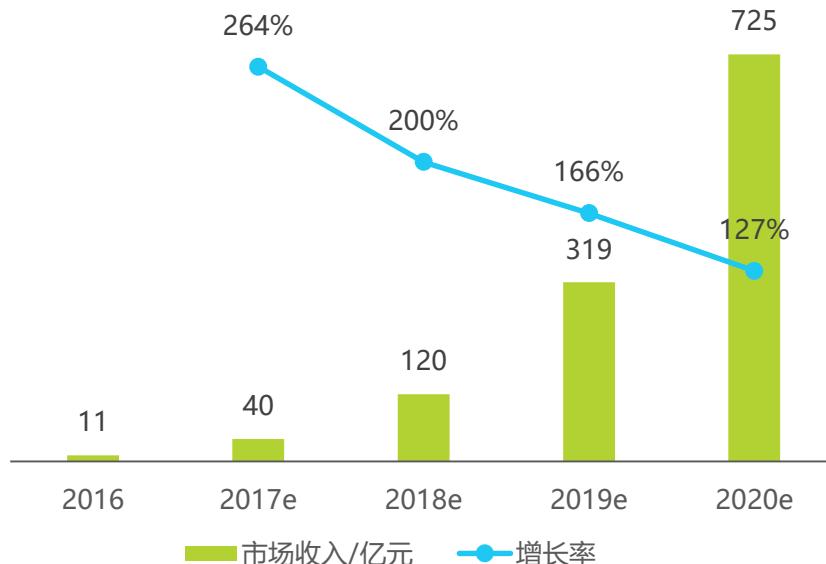
来源：主要为技术层公司的商业模式，艾瑞根据专家访谈、公开资料等研究绘制。

# 计算机视觉行业市场规模

## 中国计算机视觉行业市场规模

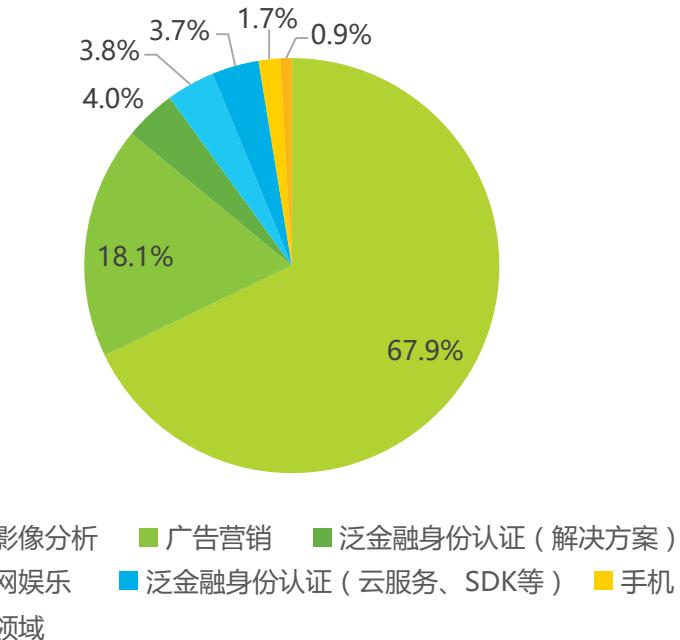
2016年下半年，1:N人脸识别、视频结构化等计算机视觉相关技术在安防领域的实战场景中突破工业化红线，敲响了计算机视觉行业市场大规模爆发的前奏。伴随人脸识别、物体识别等分类、分割算法不算提升精度，在2017年占比较高的安防、视频广告、泛金融、手机及互联网娱乐领域之外，医疗影像、工业制造、批发零售等现阶段的创新领域也将逐步解锁，成为行业整体快速发展的重要支撑。

2016-2020年中国计算机视觉行业市场规模



来源：艾瑞咨询研究院自主研发绘制。市场规模的统计推算以产业图谱中的技术层企业收入为主，侧重于以视觉算法为主要构成的云服务、SDK及软件产品的相关收入，也包含一定集成视觉算法及其他丰富功能的软件系统，还有服务器、传感器、整体硬件设备系统的收入，其中泛金融解决方案以后两种收入为主。鉴于自动驾驶技术相对综合且处于早期研发阶段，与驾驶相关的视觉技术（包括司机疲劳检测、ADAS、自动驾驶等）收入并未包含在内。

2017年中国计算机视觉行业市场构成



来源：艾瑞咨询研究院自主研发绘制。2017年中国计算机视觉行业市场构成为预期构成。

# 嵌入式感知系统加速产业智能升级

## 前端智能的渗透率将逐步提升，与后端协同智能计算

伴随AI专用芯片及嵌入式感知系统的成熟研发，前端设备即可完成相对基础的视觉感知，并将识别、分类的结果实时应用，实现前端智能。对于需要大量计算、存储资源，利用多维度数据关联分析的诸多场景，后端服务器计算依然是当前主流。前端智能的渗透率将逐步提升，与后端智能协同加速产业升级，不仅可满足特定场景对实时性、隐私性的要求，还可在前端成像（千万像素、千帧/秒）提供越来越多数据信息的背景下，让前端智能选择预处理过的结构化的高质量数据及分析结果传输至后端，减少因带宽压力过大导致的丢包、压缩所引起的信息丢失或误差，提升智能分析的准确性。

前端智能系统的功耗、稳定性、存储空间、数据及系统更新的网络传输等诸多重要问题仍有待解决，现阶段主要通过压缩算法模型、挖掘硬件潜力以及在压缩模型的同时针对现有芯片进行优化的方式来解决功耗有限、运算能力低的问题。

### 前端智能 VS 后端智能



来源：艾瑞根据专家访谈、公开资料等研究绘制。

计算机视觉技术概述

1

计算机视觉行业概况

2

计算机视觉的应用场景

3

计算机视觉典型公司案例

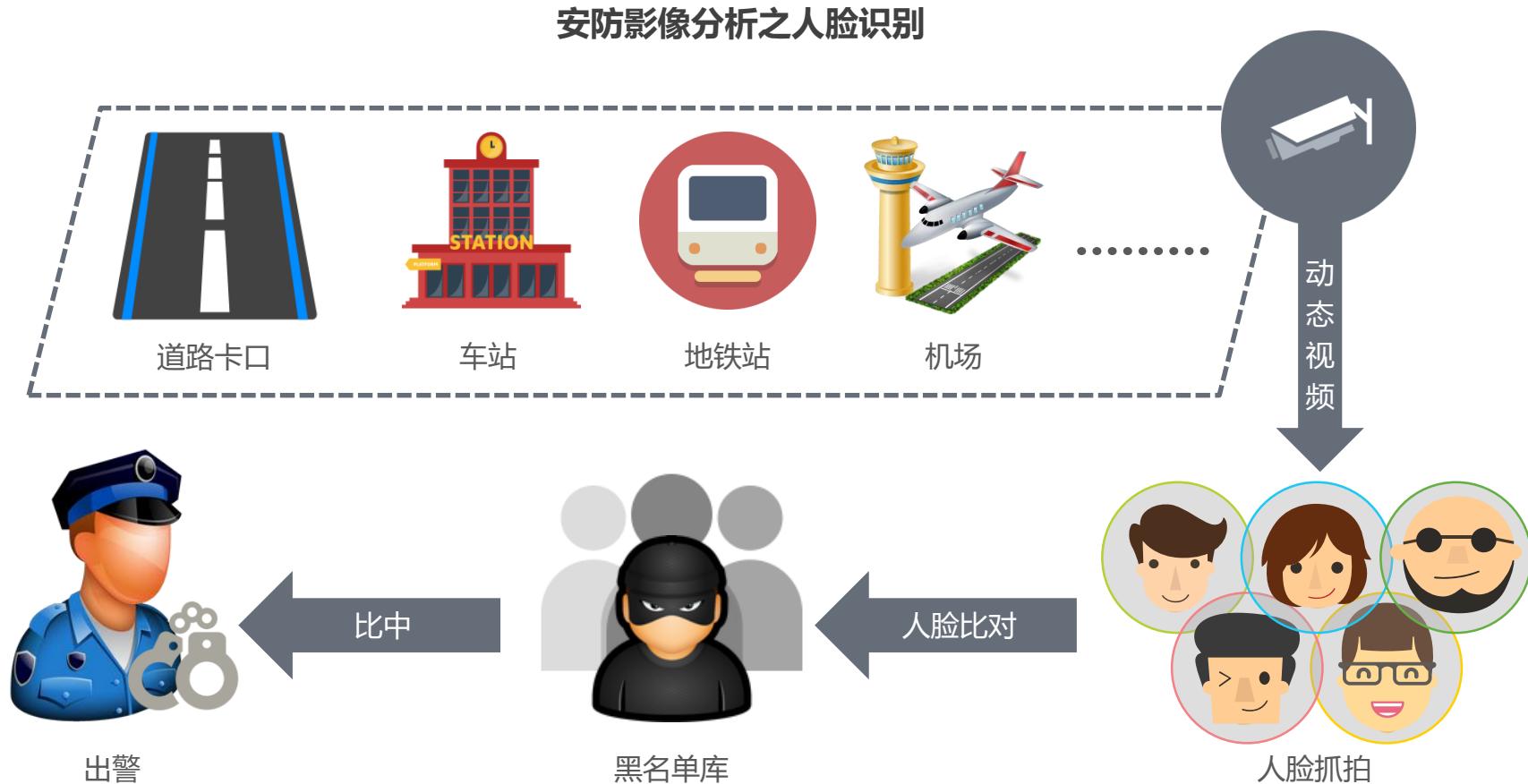
4

计算机视觉行业发展趋势

5

# 安防影像分析领域 主要应用场景之人脸识别

对道路卡口、车站、地铁站、机场等地方的监控视频进行智能分析，检测出动态视频中的人脸与黑名单库中的影像记录做实时比对，比对成功则立即报警推送给警务人员处置。



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 安防影像分析领域

## 主要应用场景之视频结构化

2012年南京“1.6”案件发生后，南京警方从全市1万多个摄像头共提取了近2000T的视频数据，调动1500多名公安干警查阅搜索视频线索，共耗时一个多月。海量视频的有效利用存在巨大挑战，完全依靠人工费时费力，而安防影像智能分析则可有效缓解这一问题。视频结构化可针对已经生成的海量视频内容进行自动化处理，提供行人、机动车、非机动车等关键目标的监测、跟踪、属性分析，辅以以图搜图等检索功能，让案件侦办和治安布控更加智能便捷。

### 安防影像分析之视频结构化



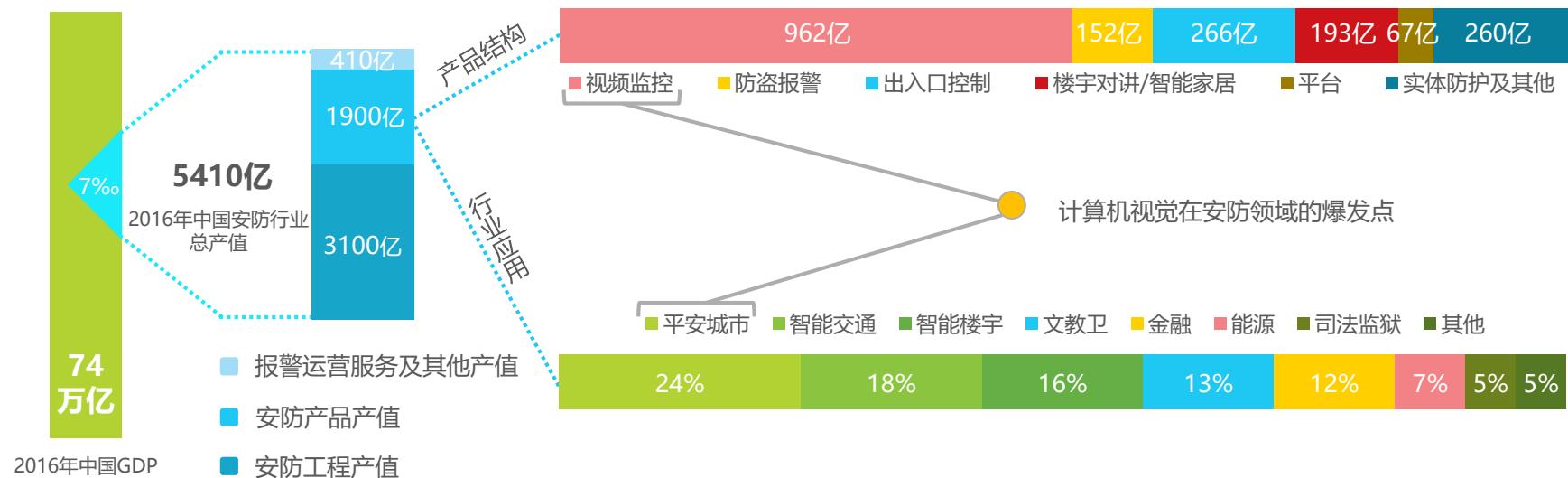
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 安防影像分析领域

## 安防行业的千亿市场为视觉智能改造提供充分空间

2016年中国安防行业总产值为5410亿，占据2016年中国GDP的7%，其中1900亿属于安防产品产值。从产品结构角度，安防产品可划分为视频监控、防盗报警、出入口控制等，安防影像的智能分析属于视频监控产品的升级改造，是各地区平安城市建设中的重要组成。2011年至2016年中国安防市场连续5年保持2位数增长，结合国家政府对建设更高水平的平安中国、进一步提升人民安全性的要求，未来中国的安防市场依然会保持稳定增长，对人工智能、计算机视觉技术的升级改造也提出了更高的期待。

### 2016年中国安防行业市场结构



来源：艾瑞根据专家访谈、中安网等公开资料研究绘制。

# 安防影像分析领域

## 智能安防相关政策分析

伴随从平安城市、智慧城市到雪亮工程等公共安全相关政策，以及人工智能相关国家战略政策的逐步深化，智能安防也将与高清联网一样，由重点区域、有条件的地区起步，开始新一轮的从局部到整体的全国性拓展。

### 公共安全与人工智能相关政策



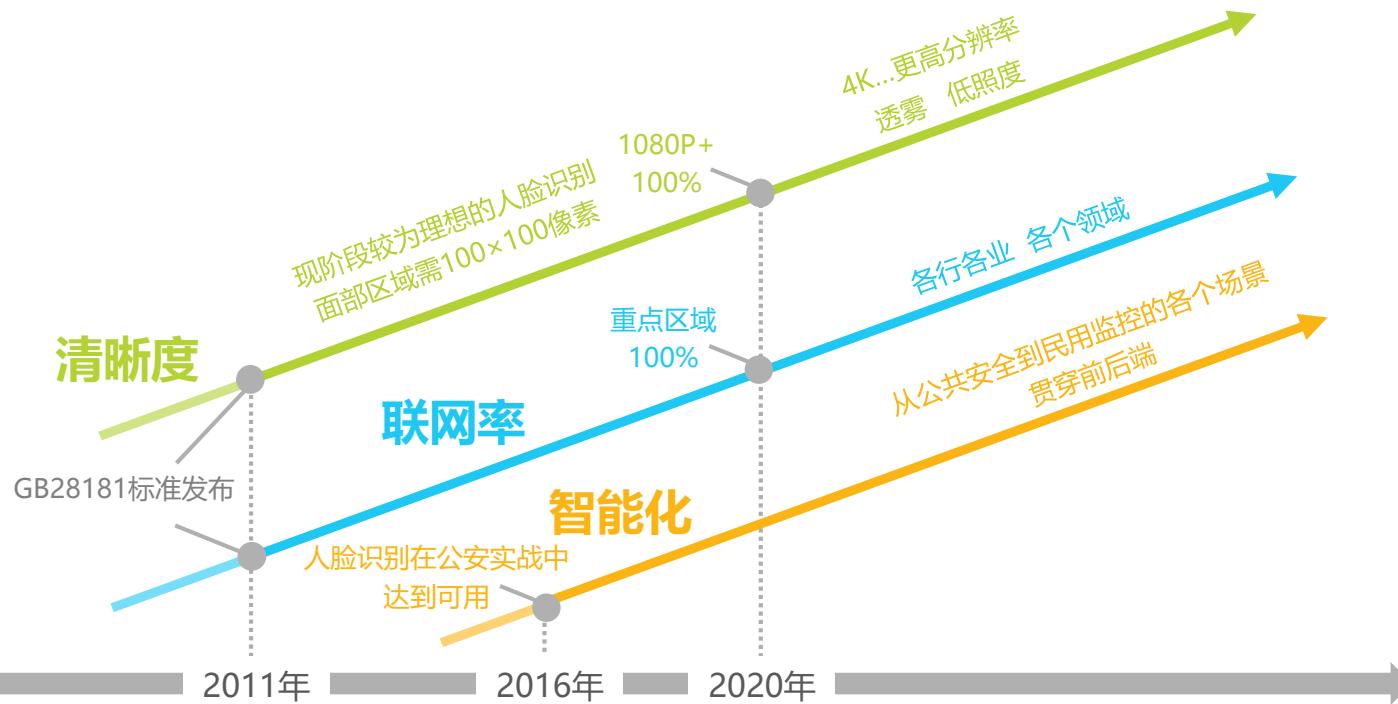
来源：艾瑞根据公开资料研究绘制。

# 安防影像分析领域

## 视频监控步入智能时代

2015年，国家发展改革委、中央综治办、公安部等九部委下发《关于加强公共安全视频监控建设联网应用工作的若干意见》，提出到2020年基本实现“全域覆盖、全网共享、全时可用、全程可控”的公共安全视频监控建设联网应用，视频监控的清晰度和联网率得到快速发展，为智能升级奠定了较好的基础条件。需要说明的是，相比高清联网的“看到、看清”，智能升级的“看懂”对前端设备的部署位置、高度及角度等有着更苛刻的要求，现公安联网的数千万摄像头中的80%~90%仍然需要针对性的调整改造。

清晰度与联网率为视频监控的智能升级提供基础保障



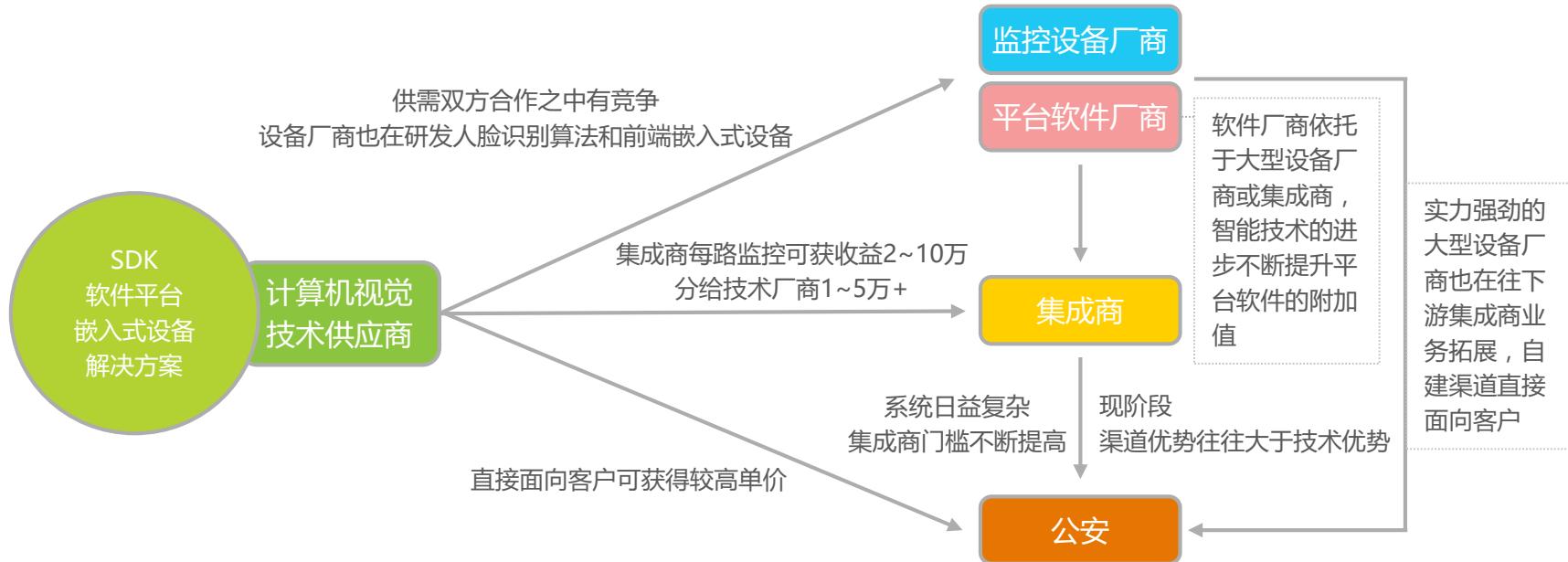
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 安防影像分析领域

## 智能安防相关产业链条分析

计算机视觉技术供应商在提供智能安防影像分析的时候，可能面对监控设备及平台软件厂商、集成商、公安等三大类客户。在最终的使用者公安以外，其他三类参与者关系复杂，各自的业务定位都在发展变化，合作间有竞争。

### 智能安防相关产业链条分析



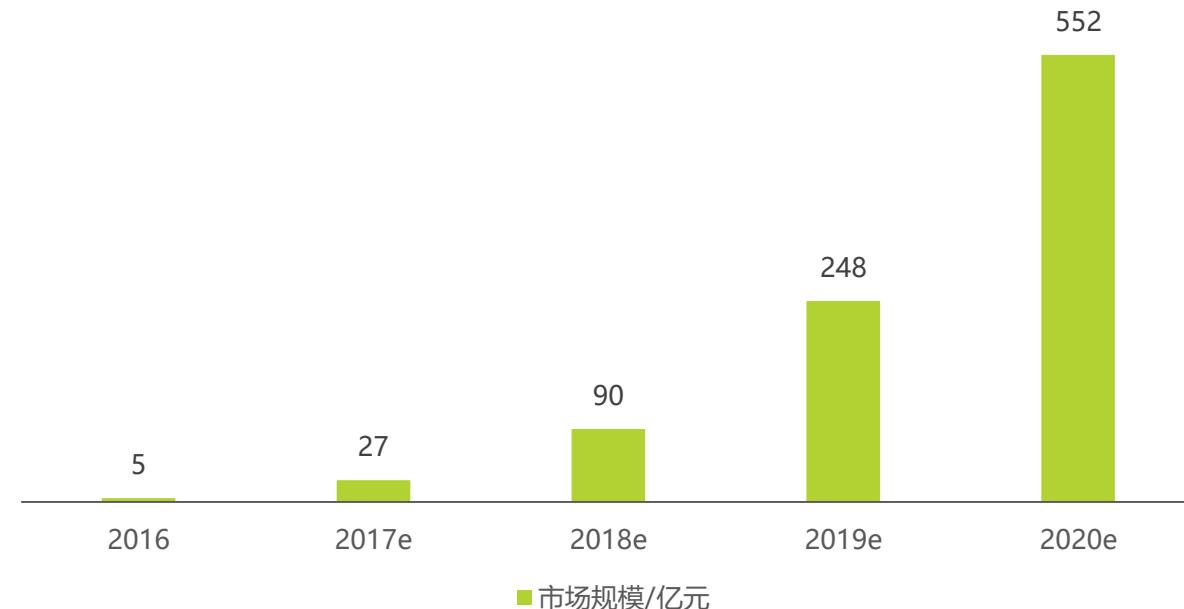
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 安防影像分析领域

## 智能安防影像分析的市场规模

《中国安防行业“十三五”（2016-2020年）发展规划》中提出到2020年，安防行业总收入达到8000亿元左右。结合对计算机视觉技术进步在视频监控前端、软硬件、公安刑侦各场景迅速渗透的预期，我们认为未来三年智能安防影像分析的相关收入在安防产业中的占比将以1~2倍速度增长，由此得出如下市场规模以供参考。

2016-2020年中国智能安防影像分析的市场规模



来源：艾瑞根据专家访谈、公开资料等研究绘制。可能高估的风险来自平安城市推进及人工智能、计算机视觉相关技术的发展、渗透不及预期，或成熟技术大规模落地后成单价降低；可能低估的风险来自相关技术的发展、渗透速度高于预期。

# 安防影像分析领域

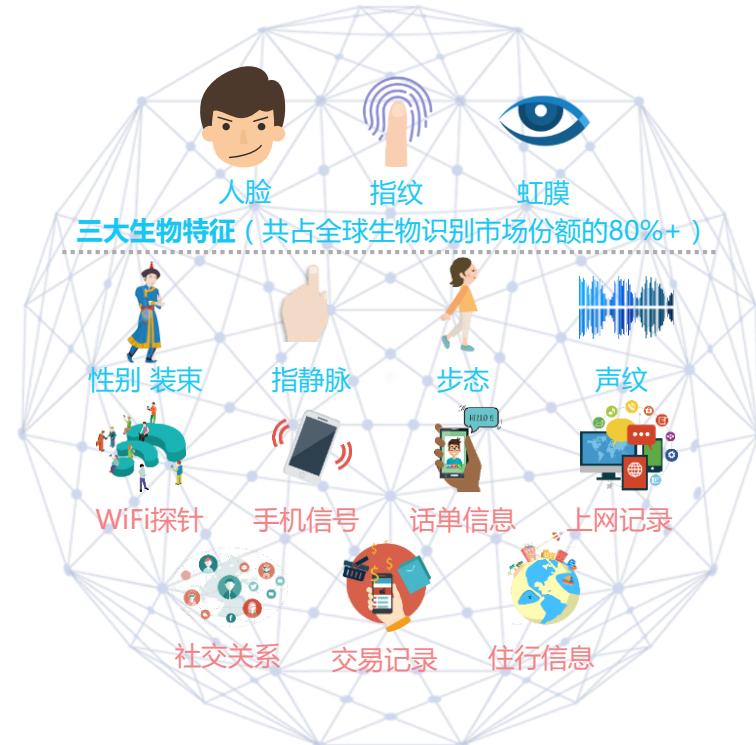
## 多模态融合、大数据多维分析共同增强智能安防分析能力

不足够理想的人脸影像捕捉及人脸本身的相似性导致识别精度有限，因此融合多种生物特征的多模态识别加多维数据的大数据分析等智能技术的交叉融合可大幅提升身份识别精度，结合公安战法有效辅助断案效率，从而进一步提升人工智能在安防领域的落地速度及渗透率。

### 若干因素限制人脸识别精度



### 多模态生物特征+大数据多维信息分析



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 泛金融身份认证领域

## 主要应用场景及相关影像采集设备

与安防影像分析中人脸的“1:N”识别不同，目前泛金融领域以人脸“1:1”身份认证为主，部分场景涉及“1:N”识别，如银行网点中对VIP客户的智能识别。

### 泛金融身份认证的主要应用场景及相关影像采集设备

金融应用场景	影像采集设备	互联网应用场景	影像采集设备
 远程开户  无卡取款  转账/交易/申贷确认	 网点VTM/ATM  个人终端	 注册、登录等身份认证	 个人终端
 协助网点柜员确认客户身份  VIP迎宾 智能分析网点客户	 摄像头	 手机号实名认证	 网点自助终端  个人终端
 金库等区域出入管理  押运员身份确认	 人脸、虹膜、指纹等智能识别设备	 刷脸支付	 网点自助终端  个人终端

来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 泛金融身份认证领域

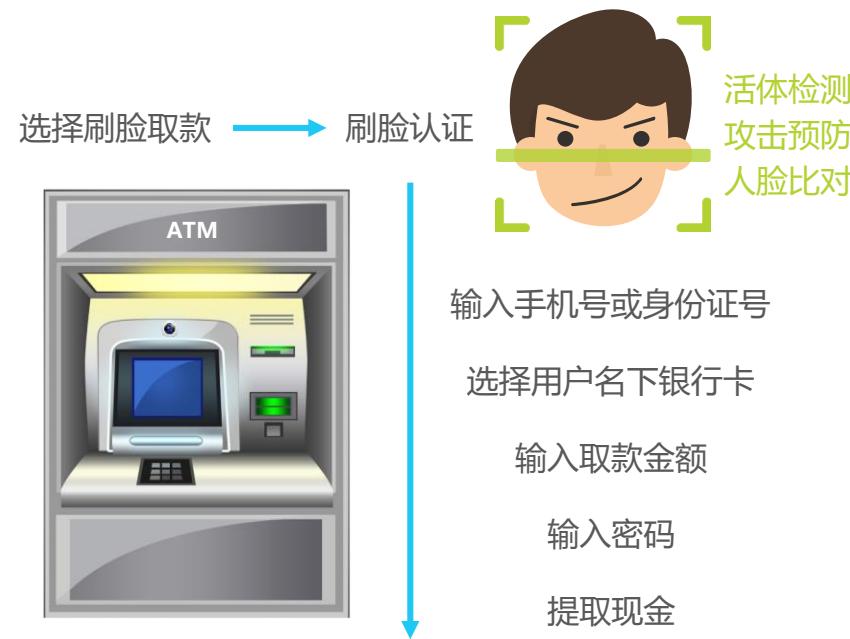
## 刷脸认证与其他认证方式的关系

在实际应用场景中，人脸识别通常需与其他认证方式进行组合，共同保障交易安全。刷脸认证可在部分场景替代具体实物（如银行卡、优盾等）或操作较为繁琐的身份认证方式（如短信验证码），甚至成为唯一的身份验证方式（如小额交易场景），提升用户体验，较低交易成本。

### 人脸识别与多种方式组合认证



### 无卡刷脸取款示意图



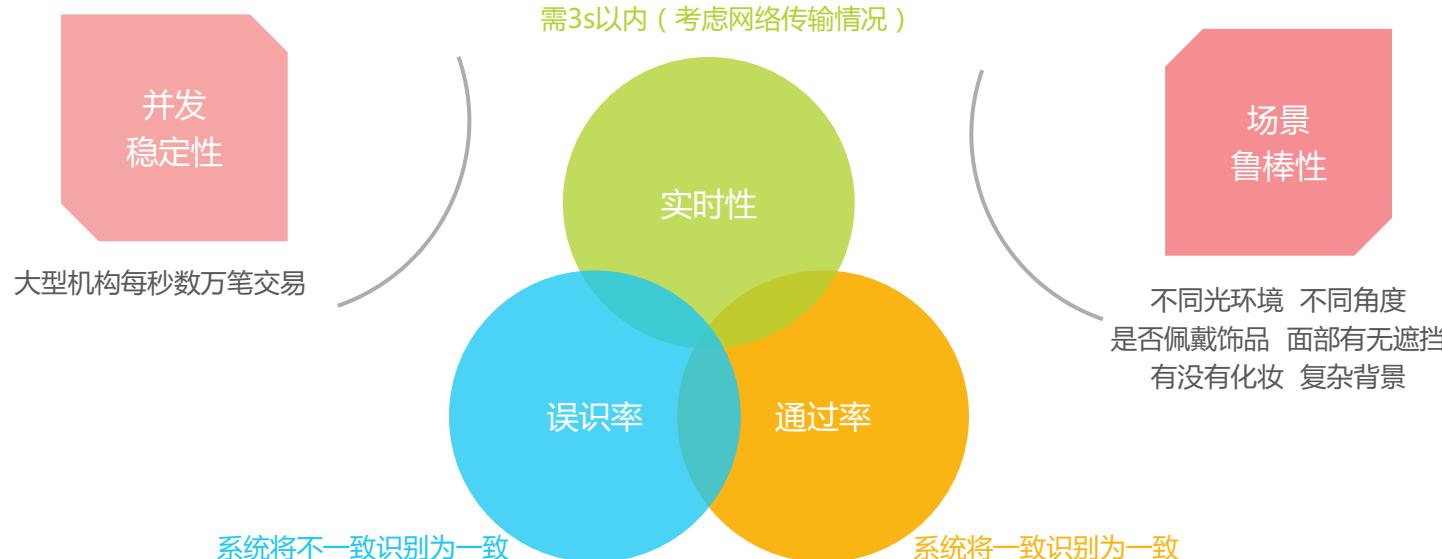
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 泛金融身份认证领域

## 刷脸认证，各种限制因素下的取舍平衡

泛金融领域中，银行对人脸识别的各项技术指标的要求相对较高，需要在实时性、误识率、通过率等各种限制因素下取舍平衡。少数银行和第三方支付已开始小范围尝试刷脸转账、刷脸消费及刷脸购物，个别鼓励技术创新的银行已开始在全国全面试点刷脸取款，刷脸认证在银行领域的商业化刚刚开始。

### 金融级人脸识别的三大技术指标与工程实践要求



误识率越低则安全性越好，但通过率随之越低，合规用户的体验将越差

2017年，适用于金融行业大范围推广的标准如下

误识率为0.001%时，典型场景通过率为99%以上

误识率为0.0001%时，典型场景通过率在95%以上

来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 泛金融身份认证领域

## 刷脸认证的优势与功效

刷脸认证容易落地，并可帮助企业有效防范风险，降低成本

### 容易获取

- 银行无需过多改造
- 人脸天然性
- 公安部已有人脸数据
- 手机等影像采集设备较为普及

### 人脸认证

- 防范冒名开户  
互联网冒名开户比率从万分之五几乎降为零
- 防范黑产交易  
核实客户身份及投资、申贷意愿

### 降低人力成本

- 人工审核成本

### 降低交易成本

- 部分场景可替代短信验证码  
某行可节省3600万/年的短信费用

容易落地

防范风险

降低成本

### 容易使用

- 免签约
- 免学习
- 非接触

与指纹、指静脉等不同

刷脸不需要肢体接触传感器

### 人脸存档

- 辅助交易举证
- 辅助内部管控  
防止银行内外勾结诈骗资金

### 降低风险损失

- 欺诈交易带来的损失  
某行信用卡欺诈案件预防成功概率已达95%  
预计减少损失800万/年

来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 泛金融身份认证领域

## 泛金融身份认证的主要服务方式

与计算机视觉公司的主要商业模式一致，泛金融身份认证也有API、SDK及解决方案三种服务方式。相比API及SDK的服务方式，一体化解决方案拥有更高的客单价和市场空间，但因包含诸多软硬件集成服务，难以将核心视觉技术相关收入单独剥离计算。

### 泛金融身份认证领域的三种服务方式

1 API

为客户提供  
人脸识别云服务  
一般按次计费

3 一体化解决方案

为客户提供  
服务器部署、生物识别系统集成、硬件改造集成等  
各类软硬件定制产品及服务  
2017年已有单一客户千万元项目

2 SDK

为客户提供  
人脸识别算法模块  
客户可将其部署于移动端或服务器端  
可结合授权设备量及授权周期定价



自助机具改造

数千元  $\times$  100万  
客单价 全国机具存量



智慧网点改造

数百万  $\times$  20万+  
客单价 全国网点量

来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 手机及互联网娱乐领域

## 为消费级产品带来全新智能体验

计算机视觉技术的成熟进步为同质化的手机产品及互联网娱乐应用带来了新的活力。2017年诸多国内外手机厂商推出了具有刷脸解锁的旗舰机型，而手机与影像相关的拍照优化、相册分类、编辑处理等也于近几年得以智能升级。人脸识别、特征点定位以及场景识别、物体识别等技术也丰富了直播、短视频等互联网应用的娱乐性，同时为影像内容的智能审核及分类做出贡献。

### 手机及互联网娱乐领域的主要应用



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 商品识别领域

## 拓宽信息边界，连接人与商品

键入关键词，搜索引擎可连接人与信息，大幅提升人类获取信息、搜集知识的效率，为世界创造巨大价值。计算机视觉则将信息的边界再度拓宽，缩短设计、原料采购、生产制造、线上与线下零售等各个环节的人与商品的距离，为商品供应链带来效能提升。

### 计算机视觉赋能商品供应链

#### 设计



为设计师提供配色参考  
可对时装周走秀图片做智能分析  
实时呈现系统化色彩逻辑

#### 原料采购



结合设计趋势推荐相关原料  
以图搜图提升采购效率

#### 生产制造



拍摄样品给出专业制作流程  
智能质检保证生产品质

#### 电商



以图搜图提升购物效率  
智能推荐提升成交转化率

#### 线下零售



智能识别所购商品 提升结账体验  
智能监督 降低店面货品损耗  
优化商品摆放运营

来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 工业制造领域

## 产品质检及3D分拣

伴随年轻人从事重复性体力劳动的意愿降低，相关领域的劳动力成本极速上升，工厂对智能自动化设备的需求日益凸显，为计算机视觉技术在工业制造领域的研发落地提供了市场基础。产品（尤其3C产品）质检是现有视觉技术有望在未来3-5年在工业制造领域大规模落地的应用场景，目前产品质检依然依靠大量人力做肉眼判断，效率低、成本高、漏检误检严重，而深度学习算法可支持多种缺陷类型，增量学习也能不断提升产品适用性。另外，工业场景中存在大量的冲压件、组合件等不规则物体，不规则物体的分拣（无序抓取并有序放置，涉及物体识别、姿态估计、尺寸测量、运动规划等）需借助3D视觉技术，即使技术相对领先的国外厂商相关产品也不够成熟，出现问题难以即时响应（往往邮件沟通，售后服务难以保障），给予目前尚处于产品研发测试阶段的国内新兴厂商反超机会。

### 智能质检



中国集中了全球70%的3C产能  
单个智能检测系统的成本大致相当于2.5个质检人员2年薪资  
3C智能质检在中国的市场空间预计在千亿元级别

### 3D分拣



需求相对杂乱  
技术不够成熟 依赖售后服务  
国内厂商拥有地域及价格优势

来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 广告营销领域

## 智能挖掘影像内容广告位，构建新型营销模式

计算机视觉技术可在长视频、短视频等点播平台、直播平台以及利用手机摄像头的AR应用中，为广告主提供多种形式的互动化、与内容强相关的场景广告。

### 点播平台



时间、空间、人物、事件、衣物等多维属性的视频内容结构化，还可通过三维定位与重建，实现视频场景中加入虚拟物品的逼真效果，完成广告内容的自动精准投放以及电商闭环。

### 直播平台



在直播等视频内容中主播可通过比心、挥手等交互动作触发广告，广告可实时贴合人物动作，也可智能替换原有背景。

### AR应用

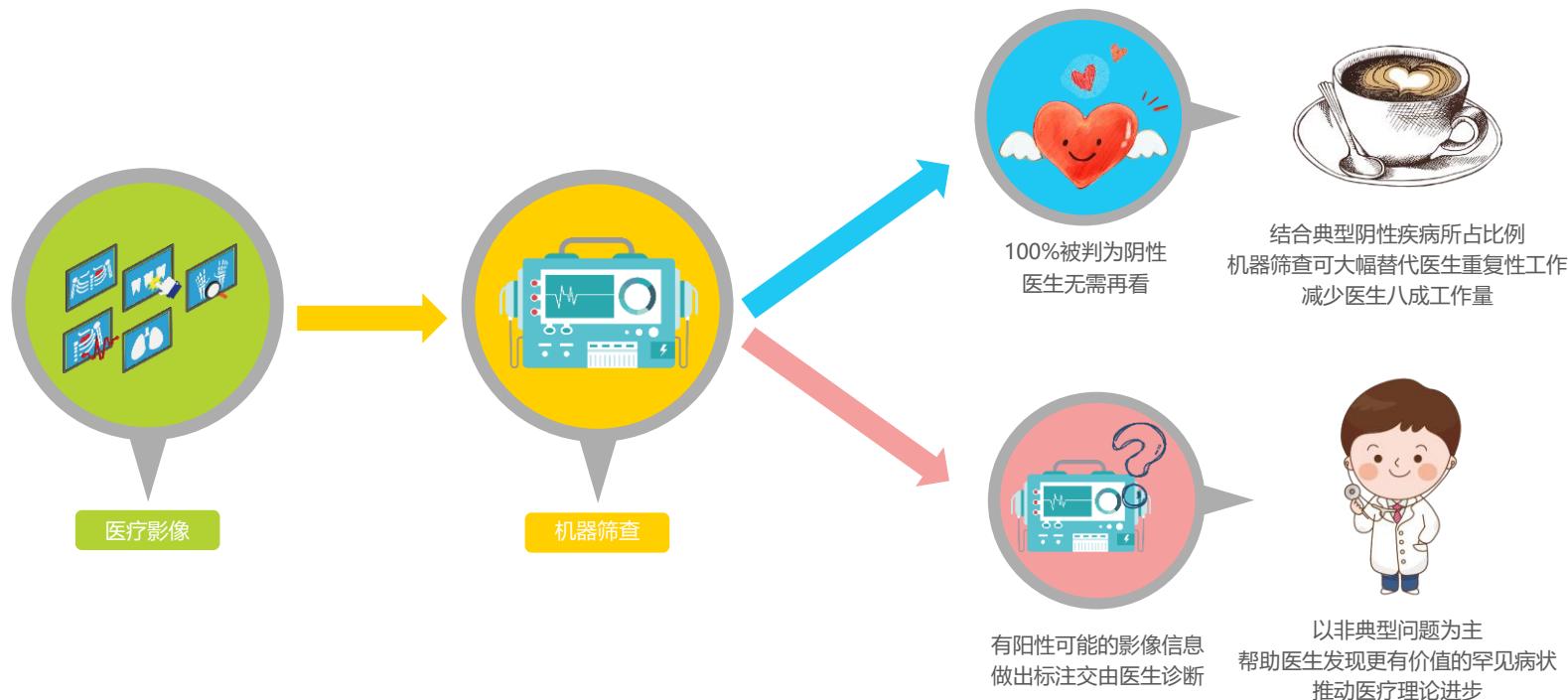


通过手机的“相机”或“扫一扫”，扫描广告主设置的物料实现触发，也可直接将APP中的任一广告位进行AR赋能，实现用户与广告的趣味智能交互。

来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 医疗影像分析领域

## 医疗影像智能分析的可能形态与潜在价值



说明：另一种可能形态为机器仅做标注，所有影像均需交由医生诊断。艾瑞根据专家访谈、公开资料等研究绘制。

# 医疗影像分析领域

## 智能医疗影像分析对病种的要求及重要指标

相比计算机视觉在其他领域的数据标注工作，医疗影像的标注门槛较高，需由专业医师标注，而且非典型病例的标注意见较难统一。标注工作之外，医疗影像分析对数字化程度、数据量、临床路径、对应检测量都有着苛刻的要求，不同病种的情况不同，难以一概而论。另外，对具体系统分析能力的考核，不能仅依据简单的准确率，特异性与敏感性是最基本的两个指标。

### 敏感性与特异性



### 智能医疗影像分析对病种的要求



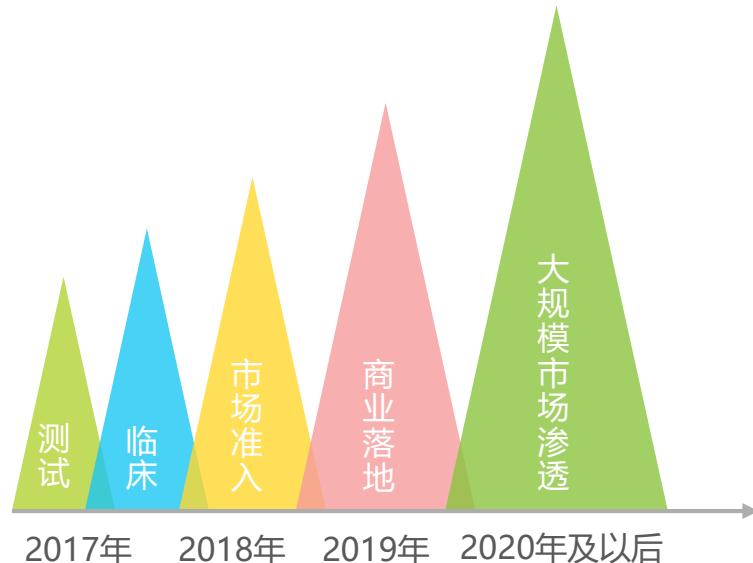
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 医疗影像分析领域

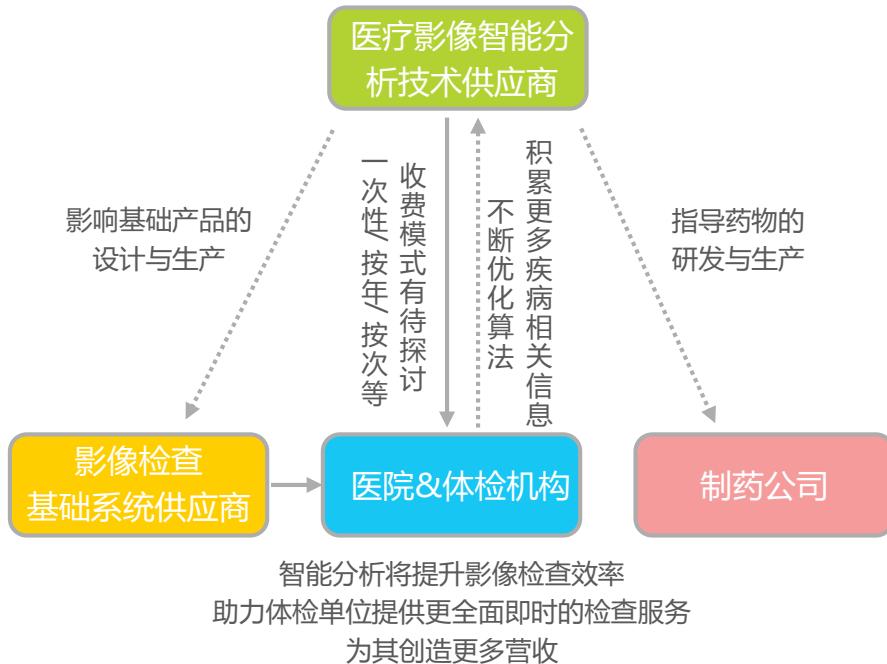
## 具体商业模式有待探索

虽然个别智能医疗影像分析系统已进入医院临床应用，但与获取市场准入进而产生相关收入依然有段距离，产品仍需打磨完善。另外，智能医疗影像分析产品如何融入医疗产业链，如何收费等问题也有待进一步的探索。

### 医疗影像智能分析预计在2019年落地



### 智能分析技术供应商在产业链中的节点预期



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 自动驾驶领域

## 自动驾驶技术剖析

自动驾驶系统主要涉及传感器融合、感知、高精地图、定位、规划及控制等若干技术环节，以解决“我在哪儿，周围有什么，环境将发生什么变化以及我该怎么做”等四个问题，计算机视觉则在环境感知（周围有什么）与地图绘制（我在哪儿）中发挥重要作用。

自动驾驶系统示意图



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 自动驾驶领域

## 从起步到落地，自动驾驶仍需经历长期测试

自动驾驶汽车主要由车辆本身、内部硬件（传感器、计算机等）以及用于做出驾驶决策的自动驾驶软件等三个子系统组成。车辆本身需由OEM认证；内部硬件也需在各种极端条件下充分测试其稳定性，达到车规级要求；自动驾驶软件方面，相关系统需经过百亿甚至千亿公里以上的测试来充分验证其安全性（人类司机平均每1亿公里发生致命事故1~3起，自动驾驶技术要想大规模落地应用，必须优于人类司机的安全性）。与此同时，大规模路测也是收集相关场景数据以便改进感知、决策等智能技术的必要手段。然而，100万辆10万公里/年的车辆行驶1年才能达到千亿公里的数据量级（Waymo在过去8年积累350万英里以上的自动驾驶数据），仿真环境下的虚拟路测（如今每天有多达25000辆虚拟的Waymo无人车在模拟器中驾驶高达八百万英里的里程）与不涉及实际控制的影子模式可作为常规测试的补充，有效降低路测成本。

### 自动驾驶软件测试



仿真环境  
虚拟路测

在结合真实世界和挑战性场景构建的仿真环境中进行虚拟路测，通过后可部署于实体车辆上



封闭环境  
结构化路测

构建专有试车跑道，可变换地形、道路类型、速度范围、气候、时间点以及法律法规等，在各种特定情形或极端条件下进行小规模路测



公路测试  
涉及实际控制

开放环境下公共道路测试，在小部分车辆验证安全性后将系统更新至整个车队



公路影子模式  
不涉及实际控制

在人类司机控制车辆的前提下，通过对比系统人类司机的输出结果的差异来测试系统，可省去造车、改装控制系统等环节，众包路测

来源：艾瑞根据专家访谈、公开资料等研究绘制。

## 自动驾驶，起步于限定场景

数据驱动的感知及决策算法难以应对开放、动态的环境的千变万化，自动驾驶车辆如何理解人类意图、如何与人工驾驶车辆的司机沟通交互也面对巨大挑战。在清晰简单的限定场景中，规则易总结，数据易收集，相关算法就越容易达到安全性要求。因此，相比开放环境下大众乘车出行的一般场景，自动驾驶技术将先在高速货运、低速摆渡、特定生产等场景落地应用。大众出行领域，自动驾驶也会逐步在特定速度限制下（时速60公里内的L3级自动驾驶汽车已有量产），停车场，乃至高速或环线等相对简单的封闭道路中替代人类驾驶。另外在自动驾驶实现以前，视觉监控系统也可对车内驾驶员进行疲劳检测、注意力检测和手势识别等，为L3级自动驾驶人机控制权的交接提供支撑，兼顾安全、驾驶辅助和车内互动娱乐。

### 三年内有望量产落地的限定场景



中国拥有数千万货车司机  
中国每年道路运输费用在3万亿以上  
卡车车身较大，易部署、方便进行后期改装



机器具备高精度感知和执行能力  
可有效降低事故率  
还能有效降低人力成本  
延长工作时间



在算法难以完美的时候  
降低时速可有效保障安全性  
时速低于20公里时若发生撞击  
人的生还率约为90%



满足车主高频刚需  
自动找到停车场入口  
在停车场内自动驾驶停入车位

来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 自动驾驶领域

## 智能出行公司为大众带来无人驾驶出行体验

美国交通部和美国高速公路安全管理局（NHTSA）在今年9月份发布《自动驾驶制度方针2.0》，预计将在2025之后实现全面自动的安全功能以及高速公路的自动驾驶。国内科技公司相对乐观，纷纷发声将在2021年前后实现仅在特殊情况需人类介入的L4级自动驾驶乘用车的量产。考虑到现有算法技术的能力边界，2021年其实难以实现通用场景的L4级自动驾驶，小概率的意外缺陷都有可能引发致命事故，但2021年成为创业者标杆的时候，它将促进“预言的自我实现”，技术的突破性进展及不断拓宽的行驶场景依然值得期待。

自动驾驶的到来需要在汽车中装配大量的软硬件设备，共享出行可减轻自动驾驶在推进消费市场时的阻碍，由出行服务商评测系统安全性，承担并消化成本。智能化与共享化是汽车产业生态的重要发展趋势，车企、科技公司、出行服务商间（三方均已布局自动驾驶研发）的合作结盟也会愈发频繁，智能出行公司随之诞生。

### 共享出行摊低落地成本，自动驾驶引发产业融合



来源：艾瑞根据专家访谈、公开资料等研究绘制。

计算机视觉技术概述

1

计算机视觉行业概况

2

计算机视觉的应用场景

3

计算机视觉典型公司案例

4

计算机视觉行业发展趋势

5

## 坚持原创，让AI引领人类进步

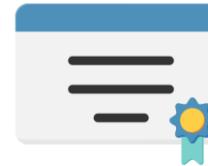
商汤科技创立于2014年，公司专注于计算机视觉与深度学习底层算法的研究与开发，并作为人工智能算法供应商赋能多个行业。商汤科技核心团队具有20余年研发经验，在2011年到2013年间，于世界顶级计算机视觉会议CVPR和ICCV上发表了14篇关于深度学习的论文，占据全球共29篇的近一半。从2015-2017年，公司在顶级期刊CVPR、ICCV、ECCV，发表论文总数为119篇，在学界、产业界排名世界第三。商汤与香港中文大学共建的香港中大-商汤科技联合实验室于2016年被评为世界十大顶尖AI实验室。2014年，商汤团队的系列人脸识别算法以打败Facebook的成绩成为业绩首个超越人眼识别率的算法；2015年，在ImageNet竞赛中取得两项世界第一，并于2016年蝉联并拿下三项冠军。

技术的突破与领先让资本市场看到商汤创造差异化产品及蓝海市场的巨大空间，2017年7月，商汤科技宣布完成4.1亿美元B轮融资，创下当时全球人工智能领域单轮融资最高纪录。



### 人才

拥有18名教授、150余名博士  
来自MIT、CMU、香港中文、清华、北大等一流  
学府及微软、谷歌、百度、联想等知名企



### 研究成果

在顶尖学术杂志和会议上发表研究报告400余份  
人脸识别技术超越Facebook  
图像识别技术比肩Google



### 深度学习平台

深度学习超算中心  
深度学习软件框架 Parrots  
为算法研发提供基础设施



### AI独角兽

4.1亿美元B轮融资  
顶级VC、PE及产业战略投资者入股

来源：艾瑞根据专家访谈、公开资料等研究绘制。

## 坚持原创，让AI引领人类进步

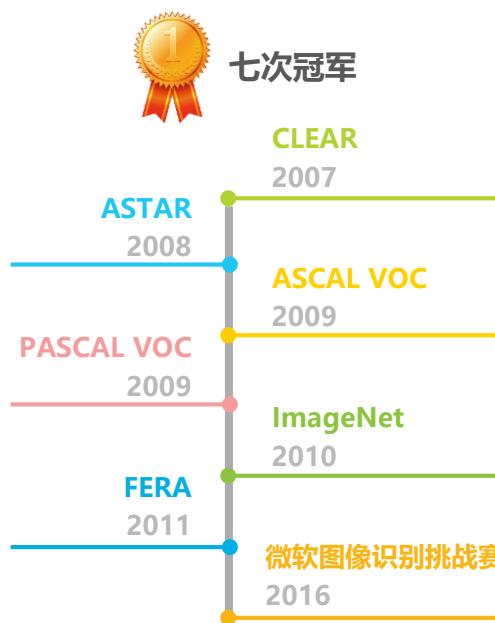
商汤科技凭借坚持原创、技术产品化、产品规模化的发展模式，广泛服务于安全监控、金融服务、手机、移动应用等领域。公司目前已服务超过400家客户，包括Qualcomm、英伟达、中国移动、银联、万达、苏宁、海航、京东、华为、小米、OPPO、vivo、微博、科大讯飞等知名企业和政府机构，提供包括智能视频、增强现实、手机图像处理、身份验证、智慧商业、遥感影像等十个应用领域。商汤科技将继续深耕人脸识别、图像识别、视频分析、无人驾驶（已与国际知名车企签署合作协议，共同开发自动驾驶技术）、医疗影像识别等各类AI应用技术，打造中国的深度学习引擎和人工智能生态。



来源：艾瑞根据专家访谈、公开资料等研究绘制，数据截止2017年8月。

## 专注视觉智能，深耕安防、银行、机场等重点领域

云从科技成立于2015年4月，是一家孵化于中国科学院重庆研究院的计算机视觉企业。云从核心技术源于四院院士、美国UIUC图像实验室主任黄煦涛教授，研发团队曾于2007年至2016年7次斩获智能识别世界冠军，包括ImageNet、PASCAL、FERA等顶级比赛。公司目前已建立由UIUC和硅谷的两个前沿实验室，中科院和上海交大的两个国内学术联合实验室，以及上海、广州、重庆、成都四个研发中心所组成的三级研发机构。2017年3月，国家发改委确定云从科技承担国家发改委重大工程——“人工智能基础资源公共服务平台”建设任务，为未来社会运行提供AI基础设施与服务。11月，云从科技正式完成B轮融资由顺为资本、元禾原点、普华资本联合领投的5亿元人民币融资，加上此前广州市政府对云从科技的20亿政府资金支持，此次总计获得25亿元发展资金。



来源：艾瑞根据专家访谈、公开资料等研究绘制。

## 专注视觉智能，深耕安防、银行、机场等重点领域

目前，云从科技主要深耕安防、银行、机场等重点领域，在前沿算法的三级研发架构以外，为了更好的服务行业客户，云从科技也先后与公安部、四大银行、民航总局等产业界成立联合实验室，共同打磨更符合落地场景需求的工程化产品。云从在公安领域推动中科院与公安部全面合作，通过公安部重大课题研发火眼大数据平台等智能化系统，相关产品已在23个省上线实战；银行领域，云从从客户实际需求出发，打造包含远程人脸识别、活体检测、自助终端、网点VIP迎宾、客户分析、自助刷脸支付购物机、金库区域布控等覆盖银行各项服务的全链产品及完善的解决方案，已服务农行、建行、中行、交行等100多家金融机构；在民航领域，云从与中科院重庆院合作已覆盖80%的枢纽机场。



来源：艾瑞根据专家访谈、公开资料等研究绘制。

## 拓展人工智能新疆界

依图科技由朱珑博士（从事计算机视觉的统计建模和人工智能的研究，师从霍金弟子艾伦·尤尔教授）和林晨曦（前阿里云计算资深专家，组织搭建了飞天分布式云计算操作系统，曾于微软亚洲研究院从事机器学习、计算机视觉、信息检索以及分布式系统方向的研究工作）创立于2012年，公司致力于将先进的人工智能技术与行业应用相结合。在2017年11月的人脸识别挑战赛FRPC中，依图获得了全球人脸识别算法检索准确率冠军。此前，依图科技曾在NIST（美国国家标准与技术研究院）组织的人脸识别技术测试FRVT上，在四项测试中获得第一，也是第一个获得此项测试冠军的中国团队。冠军算法与强大工程能力相结合，依图将解锁安防、医疗健康、金融、城市交通等更多场景落地应用，持续服务城市安全和人类生活。2017年5月依图完成由高瓴资本领投的3.8亿人民币C轮融资，11月完成招商银行领投的C+轮融资。

### 安防



在通过率为90%的情况下，依图的智能安防平台误报率从2015年3月的千分之一，提升至2017年6月份的十亿分之一，并将不断攀升。目前已服务了中国20多个省公安厅、百来个地市公安局。

### 医疗健康



依图胸部CT智能辅助诊断系统已在全国多家三甲医院进入临床工作流，儿童骨龄智能辅助诊断系统也已在浙江省医学院附属儿童医院正式临床应用，并在进行基于自然语言理解、知识图谱构建的医疗AI技术探索。

### 金融



依图双目活检系统能在1秒内完成活体检测和人脸识别，为招行、农行等更多客户带来刷脸取款的便捷体验。在刷脸取款以外，依图还有智能网点、远程核身等解决方案。



### FRPC 人脸识别算法检索挑战赛

依图的夺冠成绩远超俄罗斯公司Vocord、法国知名安防大厂Morpho等公司。项目要求1:N 动态检索匹配，在海量人脸数据库中快速找出目标人脸或判断其不在库中。比赛查询照片都来自监控视频的截帧，包含识别对象不在摄像头前特定位置的情况，比如曝光过度、逆光、侧脸、远距离等。衡量指标在于搜索的精确度，以误报率在七亿分之一水平下的漏报率为评估标准。不论是测试集或训练集都不会公开给参赛者，由开发者发送预先设置好的软件库给组织方，算法在隔离的照片库上运行。

来源：艾瑞根据专家访谈、公开资料等研究绘制。

## 赋能机器之眼，构建城市大脑

旷视科技成立于 2011 年 10 月，三位创始人印奇、唐文斌、杨沐均出自清华“姚班”。“姚班”创办人、世界著名计算机科学家、首位图灵奖华人得主姚期智院士也于 2017 年 11 月受邀担任旷视科技的首席顾问。旷视科技首席科学家兼研究院院长孙剑博士曾于 2015 年带领团队开发出 ResNet，赢得 ImageNet 和 MS COCO 等五项世界竞赛冠军，揽获 CVPR 2016 最佳论文奖，成为计算机视觉领域最流行的架构之一。2017 年 10 月 31 日，公司完成 C 轮 4.6 亿美金融资，本轮由中国国有资本风险投资基金领投。旷视科技将进一步依托独创深度学习引擎 MegBrain，加强金融安全、城市大脑、手机智能三大领域的人工智能技术工程化、产品化和产业化。截至目前，旷视科技 Face++ 已经与十余个省会级城市实现了智慧城市战略合作，并与国内多家一线手机厂商在智能化方面达成合作。

### 金融安全

人脸识别云平台 FaceID，为中国上百家互联网金融公司提供线上刷脸身份认证服务；利用 Face++ 技术实现刷脸支付商业落地，探索新零售、无人店的赋能升级。



### 城市大脑

打造安防“云和端”的结合，把安防、金融、零售、物流、制造等城市各场景数据连通，形成智能化平台化的计算范式，成为大数据的未来的构建者和运营者。



### 手机智能

为 vivo v7+、小米 Note 3、坚果 Pro2 等国内手机厂商的新机提供了完整的人脸识别解锁方案，助力国产厂商将智能手机升级为可以交互感知的智慧手机。

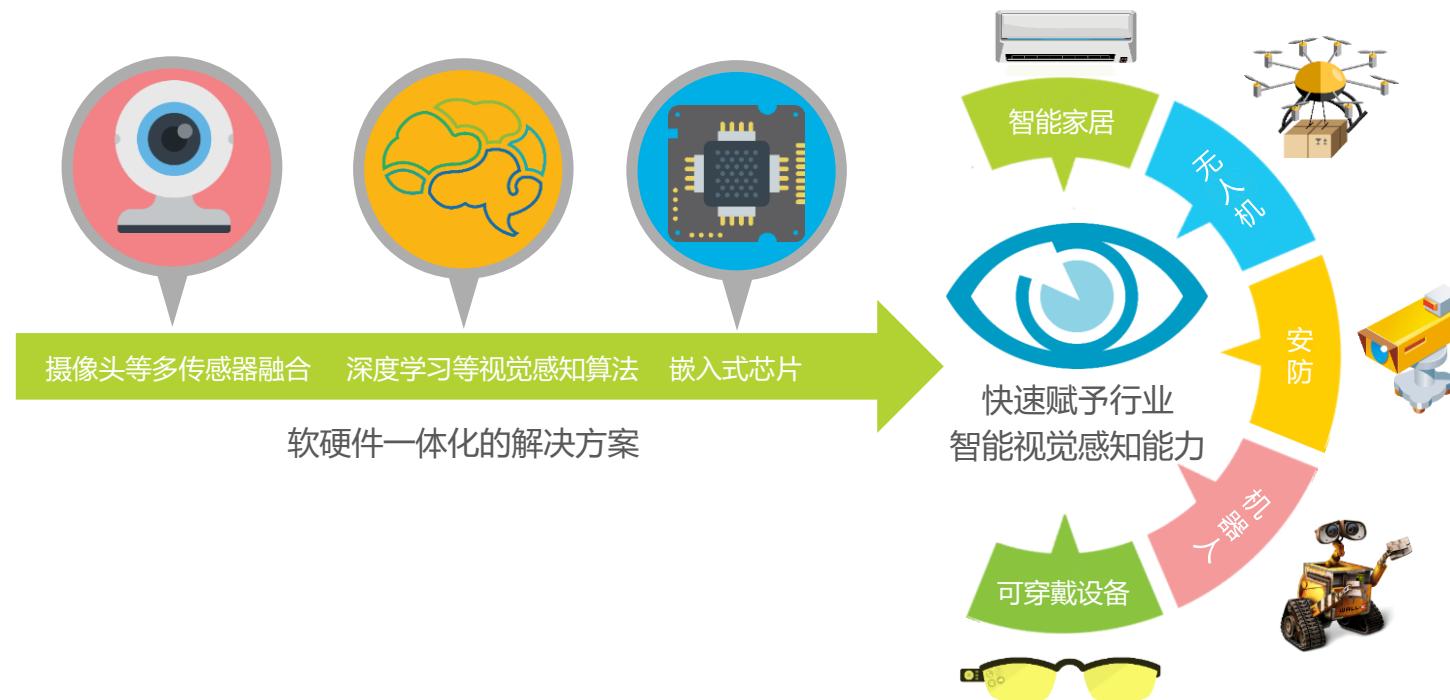


来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 触景无限

## 用前端视觉感知赋能万物

触景无限由国家“千人计划”专家、北京市“海聚工程”专家陆凡博士与前UIUC高等媒体研究中心图像组项目负责人、前IBM资深信息架构师肖洪波先生联合创办，公司致力于构建嵌入式人工智能行业解决方案，将视觉等感知智能赋予机器万物。2017年3月，触景无限获得5000万人民币A轮融资，并于6月份联手Intel推出点睛计划（为行业领军企业提供定制化解决方案，解决其感知算法和嵌入式芯片两大难题），8月份携手中科院深圳先进技术研究院打造“嵌入式人工智能及机器视觉”联合实验室，产学研协同创新，助推人工智能的大规模产业化与新生态建设。



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 触景无限

## 用前端视觉感知赋能万物

触景无限的嵌入式人工智能感知平台基于先进的视觉感知技术，能够在复杂背景中和高动态光照条件下完成对目标图像的提取、识别和分类，实时感知摄像头、目标物体、周边环境的空间关系，分析和理解实际的场景，以完成智能设备特定的功能。视觉卡模块产品具有低功耗、体积小、处理能力强等特点，满足智能设备对广域、移动、便携等方面的需求。智能引擎“瞬视”系统全嵌入式搭载于摄像机，帮助其实现前端智能化，看清与看懂，是全球首款基于英特尔Movidius芯片开发的人脸抓拍、人脸识别系统；智能引擎“盾悟”系统，既可以独立构成一个人脸识别系统，也可以仅向后台或云端传输有意义的识别结果，从而降低系统成本、提升响应速度并减缓带宽压力。

基于英伟达Tegra K1芯片开发  
开发自由，可轻易移植  
适用于安防摄像头

### 视觉卡一代



全嵌入式实时识别  
复杂环境成像更清晰  
精确识别空间人物与环境

### 前端智能引擎-瞬视系统

基于Intel-Movidius芯片开发  
功耗控制佳  
适用于无人机等高续航要求设备

### 视觉卡二代



前端服务器模式（外置、内嵌皆可）  
支持非联网移动交互  
可多路摄像机同时接入  
低成本实现高效人脸识别

### 前端智能引擎-盾悟系统



来源：艾瑞根据专家访谈、公开资料等研究绘制。

## 探索商品识别领域的技术创新与商业落地

码隆科技成立于2014年，公司专注于深度学习与计算机视觉技术创新，希望为行业提供先进的智能商品识别服务。码隆核心技术人员毕业于牛津大学、清华大学、中科院等知名院校，50%成员曾就职于微软、谷歌、百度、腾讯等世界500强企业。团队成员于2017年7月份首届由谷歌研究院、苏黎世联邦理工学院共同举办的WebVision全球图像识别挑战赛上，基于最新的“弱监督学习”研究成果（充分利用互联网和实际生产环境中的各类含噪音图像数据进行算法模型训练，这一核心算法创新节省了过去所需的千万级人工标注数据的成本），对图像识别正确率达到94.78%，领先第二名2.5%夺冠。同年，码隆科技代表中国人工智能企业参加APEC峰会，并就新技术带来的机遇与挑战进行了主题演讲。2016年码隆完成A轮6200万融资，发布ProductAI人工智能商品识别平台，并与清华大学成立人工智能联合实验室，获得当地政府超过1500万的经费支持。2017年11月码隆完成由软银中国领投的2.2亿元的B轮融资，成为软银中国在华投资的第一家人工智能公司。



团队成员毕业于  
牛津大学、清华大学  
北京大学、中科院  
等知名院校



团队成员曾就职于  
微软、谷歌  
百度、腾讯  
等知名企业

IMAGENET

ImageNet  
计算机视觉领域最受关注的比赛  
团队成员获得优胜奖



A轮  
6200万



与清华大学  
成立联合实验室



发布视觉应用平台  
ProductAI

webvision



WebVision  
94.78%夺冠

B轮  
2.2亿

2014年 公司成立

2015年

2016年

2017年

来源：艾瑞根据专家访谈、公开资料等研究绘制。

## 探索商品识别领域的技术创新与商业落地

纺织与时尚是码隆科技最早主打的应用行业，目前识别柔性物体准确率可以达到94%以上，高于人眼识别率。目前，码隆科技的商业场景已拓展至家居家具、电商、快消等方向，拥有家图网、聚合网、京东商城、唯品会、可口可乐、蒙牛等知名客户，为其提供商品属性识别、海量商品图像检索等服务，将视觉技术应用到设计、生产、批发、零售等整个产业链。同时，码隆科技正在与国际零售龙头企业联合开发AI新零售解决方案，将商品识别技术落地到无人支付、无人商超、线下零售升级等场景，预期明年初落地进而规模化。码隆科技现阶段合作主要基于ProductAI平台，通过平台化自助式服务，大幅降低垂直行业客户接入智能商品识别服务的门槛。此外，码隆科技通过违禁商品检测布局安防行业，项目已落地；同时，通过和政府、医疗机构合作进行肺癌早筛、孕妇超声检测等项目，未来将进一步部署医疗行业。B轮融资之后，码隆科技也将积极推進海外布局，已与美国、日本、拉丁美洲等诸多客户达成合作意向。



### 商品属性智能识别

自动识别商品的各种属性，如衣服的颜色、元素，家具的风格，材质，陶瓷的产地、工艺，零售商品的品牌、型号等。



### 海量商品图像检索

给客户的商品库建立以图搜图引擎通过图片和视频实时检索出商品库中的相同商品或相关商品并可实现跨品类检索。



### 无人零售与智能导购

提供纯视觉的新零售解决方案，以AI替代收银员和导购员，有效控制成本，规模化拓展落地。

来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 360人工智能研究院

## 探索AI前沿技术，为360提供AI技术支持

360人工智能研究院成立于2015年9月，由计算机视觉与机器学习专家颜水成博士担任院长，由国际计算机视觉和图形学专家谭平博士及360资深工程师和系统架构专家韩玉刚担任副院长。研究院立足于深度学习及3D技术研发能力，着眼于大数据和云端计算的契机，向360智能硬件、短视频等业务提供计算机视觉、语音语义及大数据的技术支持，并完成人工智能相关方向的原始技术积累和前沿探索。

计算机视觉、大数据分析、语音分析、语义理解等四大研究方向



为360智能硬件及内容产品等提供AI技术支持



**360人工智能研究院**  
360 Artificial Intelligence Institute



来源：艾瑞根据专家访谈、公开资料等研究绘制。

## 用计算机视觉做医疗图像识别和早期筛查

羽医甘蓝 (DeepCare)成立于2016年，是一家将人工智能应用于医疗图像的识别和早期筛查的医疗科技公司。公司目前专注于研发病理图像的检测、识别和分析技术，通过融合深度学习、计算机视觉、以及大数据挖掘技术，致力于为广大医疗机构和医疗器械厂商提供优质的产品和技术服务。羽医甘蓝团队汇集了来自达特茅斯、卡内基·梅隆、凯斯西储、清华、北大、浙大等众多知名学府的精英，并在人工智能、医学图像与商业领域深耕多年，拥有丰富的商业运营经验及强大的技术背景。2016年6月羽医甘蓝获得峰瑞资本的600万天使轮投资，2017年7月，羽医甘蓝获得中关村发展集团的数千万战略投资。公司目前主打宫颈癌前病变辅助筛查，已开发出相关智能辅助筛查系统，同时在肺部肺小结节检测、乳腺癌组织病理切片识别分类、肺癌组织病理切片识别分类、糖尿病眼底病变分级，肠镜息肉检测等亦有模型储备。

### 智能辅助筛查系统



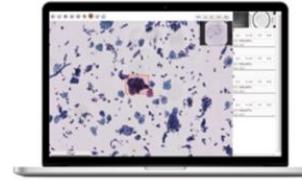
#### 数据管理与查询

- 涂片管理
- 用户管理
- 报告管理
- 多维信息查询



#### 标记病灶区域

- 多倍率查看
- 标注病灶区域
- 长度和面积测量
- 生成截图
- RGB、 $\gamma$  等参数调整



#### 智能辅助诊断

- 自动标注可疑病灶区域范围
- 标注可能亚型及其概率
- 智能分析结果



#### 自动化生成报告

- 可嵌入HIS/PACS
- 病变区域自动化截图
- 智能分析结果

来源：艾瑞根据专家访谈、公开资料等研究绘制。

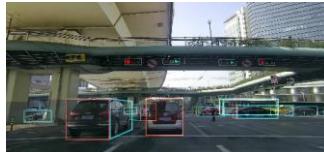
## 让无人驾驶成为可能

Momenta 致力于打造自动驾驶大脑，核心技术为基于深度学习的环境感知、高精度地图、驾驶决策算法，产品包括不同级别的自动驾驶方案，以及衍生出的大数据服务。公司团队成员来源于清华大学、麻省理工学院、微软亚洲研究院等，技术积累雄厚，包括图像识别领域经典框架Faster R-CNN和ResNet的作者任少卿，曾获得ImageNet 2015、ImageNet 2017、MS COCO Challenge 2015等多项世界级比赛的冠军。2017年10月，Momenta 获得由凯辉基金领投的B+轮融资，资金将用于AI人才组建、产品落地以及L4级技术研发在高频刚需场景下的进一步巩固与扩展，此前7月份完成由蔚来资本领投，戴姆勒集团、顺为资本等跟投的4600万美元B轮融资。



### 高精度语义地图

通过提取从多辆汽车拍摄的2D图像语义点  
融合来自GPS和IMU的数据  
创建更高精度地图



### 人工智能重现视觉感知

视觉感知车辆、行人和道路环境  
并进行高精度地图的实时定位



### 数据驱动的驾驶决策

众包路测获取高精度语义地图中的海量驾驶轨迹  
通过对海量驾驶轨迹的学习  
算法可以根据当前环境感知和高精度地图信息  
实现由数据驱动的驾驶决策规划

来源：艾瑞根据专家访谈、公开资料等研究绘制。

## 开创AI新纪元，基于深度学习推动安防智能升级

北京深醒科技是一家从事人工智能方向，集研发、生产和销售为一体的高科技创新公司，以基于深度学习的人脸识别技术为突破口，致力打造科技安防、智慧金融、智能交通、平安校园、平安小区等智能综合解决方案。中国科学院院士、人工智能泰斗张钹院士为深醒科技首席科学家，联合创始人袁培江则师从张钹院士和加拿大两院院士Patel教授，在计算机视觉、人脸识别与智能监控等领域有丰富的研发经验。作为重点发力领域，深醒为科技安防打造的大数据一体化防控平台，以人像识别信息为主线，形成人、证、照三位一体信息数据，可实现各类案件事前、事中、事后情报预警和智能分析，满足公安技战法业务需求。深醒科技于2016年底完成亿元级别A轮融资，由昆仲资本、经纬中国双领投，清科创投等机构跟投。

### 静态人脸识别比对系统

在亿级人脸数据库中以图找人，秒级返回比对结果，千万人脸数据库毫秒级返回比对结果，应用于公安、金融等领域，快速确认被查询人身份

### 人证合一身份认证产品

结合人脸识别技术与身份证件读取技术，实时采集人脸信息与身份证件信息进行人证合一比对，推动边防、酒店、展馆、机场等场所的身份认证智能化升级

### 动态人脸识别系统

深醒科技明星产品，通过前端摄像机主动捕捉移动的人脸，与系统黑名单库做快速比对，比对成功则立即报警并推送给警务人员，目前已在全国十余个省市投入使用，协助警方抓获各类犯罪嫌疑人达千余人

### 动态车辆识别管理系统

汇聚海量卡口数据与视频资源，动态、准确识别车辆特征，实时预警和情报推送，实现智能化交通管理

来源：艾瑞根据专家访谈、公开资料等研究绘制。

计算机视觉技术概述

1

计算机视觉行业概况

2

计算机视觉的应用场景

3

计算机视觉典型公司案例

4

计算机视觉行业发展趋势

5

# 计算机视觉技术展望

## 从表层感知到深层认知

数据驱动的深度学习算法通过一个函数来实现诸多场景与物体的准确分类，但泛化能力差，碰到新考题缺乏分析能力，问题的解决有赖于常识的建立与基于先验假设的逻辑推断，这将使计算机视觉与语言接轨，由感知智能上升至认知智能。再进一步，分类仅为视觉系统应有的基础功能之一，终极目的应为打造出可与世界交互的机器人智能视觉系统，由机器人所要解决的更加综合复杂的现实问题来驱动其选择要感知的事物与感知的精准度，支撑其圆满完成任务。

### 计算机视觉的学术研究愿景

计算机视觉的任务不只是分类问题  
而是建立使用视觉观察世界、与世界交互的机器人智能视觉系统  
解决现实问题所驱动的视觉认知计算

增强对功能属性、物理关系、因果逻辑、动机预测等的认知  
结合常识与先验假设进行逻辑推断

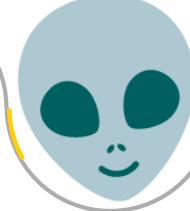
罕见角度？



罕见光线？



罕见事物？



遮挡？



现在的深度学习为大数据学习，鲁棒性差  
对于训练数据未能充分覆盖的物体、场景难以准确识别分类

机器人智能视觉系统



计算机视觉需跟语言接轨  
由感知智能升级为认知智能



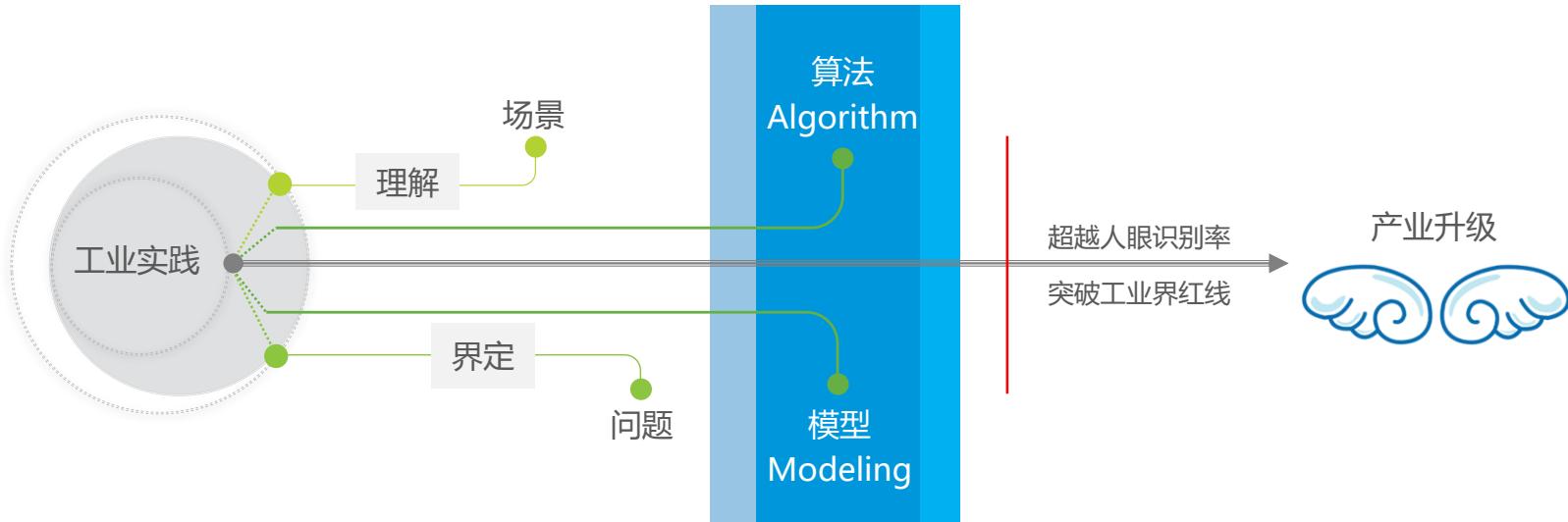
来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 算法迭代加速为各领域商业赋能

## 不断提升限定场景识别准确率，优化性能渗透更多行业应用

虽然终极愿景道阻且长，但分类任务的日益精准已解锁并将不断解锁更多场景应用。如同过去5年计算机视觉技术在人脸识别上的不断突破，误报率从2015年的千分之一提升至2017年的十亿分之一（在通过率为90%的情况下），商业服务、城市安全、大众娱乐等诸多场景均体会到不同层次的智能升级，商品、道路环境、医疗影像、遥感影像等更多对象的识别、分类问题也将会逐步突破工业化红线，从仅做辅助补充的非关键性应用拓展到切实提升核心业务效率的关键性应用。各行各业的创新型智能应用将纷至沓来，而人脸识别的性能亦将继续攀升，追求百亿、千亿规模上的可行性。

### 算法迭代驱动产业变革



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 技术供应商将继续完善商业服务链条

## 前沿算法之外，商业壁垒有赖于产品、服务、市场等综合建设

对于更为广泛的传统行业或线下使用场景的潜在客户，计算机视觉的技术落地往往涉及对具体业务场景的硬件设备改造、软件集成以及本地计算设施的部署，算法、技术的实际功效更需要建立在对客户真实业务场景的深层理解之上的针对性开发。不断增长的市场需求要求更加全面、及时的售前、售中、售后服务，而在对视觉技术能够达到的有效帮助缺乏足够认知或部分暂时缺乏科学完善的评测标准或技术相对同化的业务场景，市场销售的重要性尤为凸显。既要注重前沿算法研发，又要注意现阶段商业落地与市场拓展，这些都为以高新技术人才为主的计算机视觉公司提出了更为综合的挑战。

### 计算机视觉技术大规模应用的重要因素

#### 学术前沿算法

可通过会议及期刊论文发表情况  
计算机视觉比赛成绩等体现



#### 软硬件产品

易用性 流畅度 稳定性  
各方面体验满足实际业务所需



#### 市场营销

资源渠道建设 销售体系扩充  
为技术落地垂直业务领域加速



#### 工程落地技术

架构搭建 系统配合  
流程控制 质量监督  
需贴合客户业务场景做算法优化



#### 客户服务

售前 售中 售后  
需求响应是否及时 解决方案是否全面  
问题出现能否及时修复



来源：艾瑞根据专家访谈、公开资料等研究绘制。

# 公司介绍/法律声明

## 公司介绍

艾瑞咨询成立于2002年，以生活梦想、科技承载为理念，通过提供产业研究，助推中国互联网新经济的发展。在数据和产业洞察的基础上，艾瑞咨询的研究业务拓展至大数据研究、企业咨询、投资研究、新零售研究等方向，并致力于通过研究咨询的手段帮助企业认知市场，智能决策。

艾瑞咨询累计发布数千份新兴行业研究报告，研究领域涵盖互联网、电子商务、网络营销、金融服务、教育医疗、泛娱乐等新兴领域。艾瑞咨询已经为上千家企业提供定制化的研究咨询服务，成为中国互联网企业IPO首选的第三方研究机构。

## 版权声明

本报告为艾瑞咨询制作，报告中所有的文字、图片、表格均受有关商标和著作权的法律保护，部分文字和数据采集于公开信息，所有权为原著者所有。没有经过本公司书面许可，任何组织和个人不得以任何形式复制或传递。任何未经授权使用本报告的相关商业行为都将违反《中华人民共和国著作权法》和其他法律法规以及有关国际公约的规定。

## 免责条款

本报告中行业数据及相关市场预测主要为公司研究员采用桌面研究、行业访谈、市场调查及其他研究方法，并且结合艾瑞监测产品数据，通过艾瑞统计预测模型估算获得；企业数据主要为访谈获得，仅供参考。本报告中发布的调研数据采用样本调研方法，其数据结果受到样本的影响。由于调研方法及样本的限制，调查资料收集范围的限制，该数据仅代表调研时间和人群的基本状况，仅服务于当前的调研目的，为市场和客户提供基本参考。受研究方法和数据获取资源的限制，本报告只提供给用户作为市场参考资料，本公司对该报告的数据和观点不承担法律责任。

## 联系我们

咨询热线 400 026 2099

联系邮箱 [ask@iresearch.com.cn](mailto:ask@iresearch.com.cn)

集团网站 <http://www.iresearch.com.cn>



艾瑞咨询官方微信

生活梦想 科技承载  
TECH DRIVES BIGGER DREAMS



艾瑞咨询