

机器之眼，看懂世界
----计算机视觉行业研究报告

36氪研究院
2016年9月

目录 Contents

一、计算机视觉行业概述

- 1.1 定义与研究意义
- 1.2 行业驱动—数据量、运算力、算法技术
- 1.3 政策法规
- 1.4 投资热度
 - 全球投资热度
 - 国内投资热度
- 1.5 市场规模

二、计算机视觉技术与应用场景

- 2.1 通用视觉识别技术流程
- 2.2 识别技术与应用场景简介
 - 生物特征识别
 - 光学字符识别
 - 物体与场景识别
 - 视频对象提取与分析

三、计算机视觉行业与创业公司梳理

- 3.1 产业链
- 3.2 全景图
- 3.3 商业模式
- 3.4 行业竞争
 - 市场格局
 - 创业公司分析
- 3.5 行业思考
 - 商业模式
 - 技术趋势

Chapter 1

计算机视觉行业概述

- 定义与研究意义
- 行业驱动— 数据量、运算力、算法技术
- 政策法规
- 投资热度

全球投资热度

国内投资热度

- 市场规模

1.1 定义与研究意义

1.2 行业驱动---

数据量

运算力

算法技术

1.3 政策法规

1.4 投资热度

全球投资热度

国内投资热度

1.5 市场规模

定义

让机器之眼代替人眼，看懂所在世界

"To know what is where by looking."

(通过看去了解所在之处。)

----亚里士多德

计算机视觉是指用计算机来模拟人的视觉系统，实现人的视觉功能，以适应、理解外界环境和控制自身的运动。

概括的说，视觉系统主要解决的是**物体识别**、**物体形状和方位确认**以及**物体运动判断**这三个问题。而计算机视觉的研究，则是专注于让机器代替人眼，解决这些问题。

(1) **物体的识别**：即理解物体是什么，对物体的识别主要体现在两方面，第一是将不同物体归类，第二是对同类型物体的区分与鉴别，如不同人脸的鉴别。物体识别要求既能抽象出物体的共同属性，又能分别出相似物体间的细微差别。



人体识别



文字识别



物体识别

(2) **物体形状和方位的确定**：判断物体的形状和方位是为了让物体在视觉的三维空间里得到记忆的重建，进而进行场景分析与判断。



(3) **物体运动的判断**：和物体形状方位的确定一样，对物体的运动的判断也是一种对于场景的重建和理解，用于进行视觉主体（人或机器）对场景的分析，乃至做出自身行动的决策，实现视觉主体和场景的交互。



1.1 定义与研究意义

1.2 行业驱动---

数据量

运算力

算法技术

1.3 政策法规

1.4 投资热度

全球投资热度

国内投资热度

1.5 市场规模

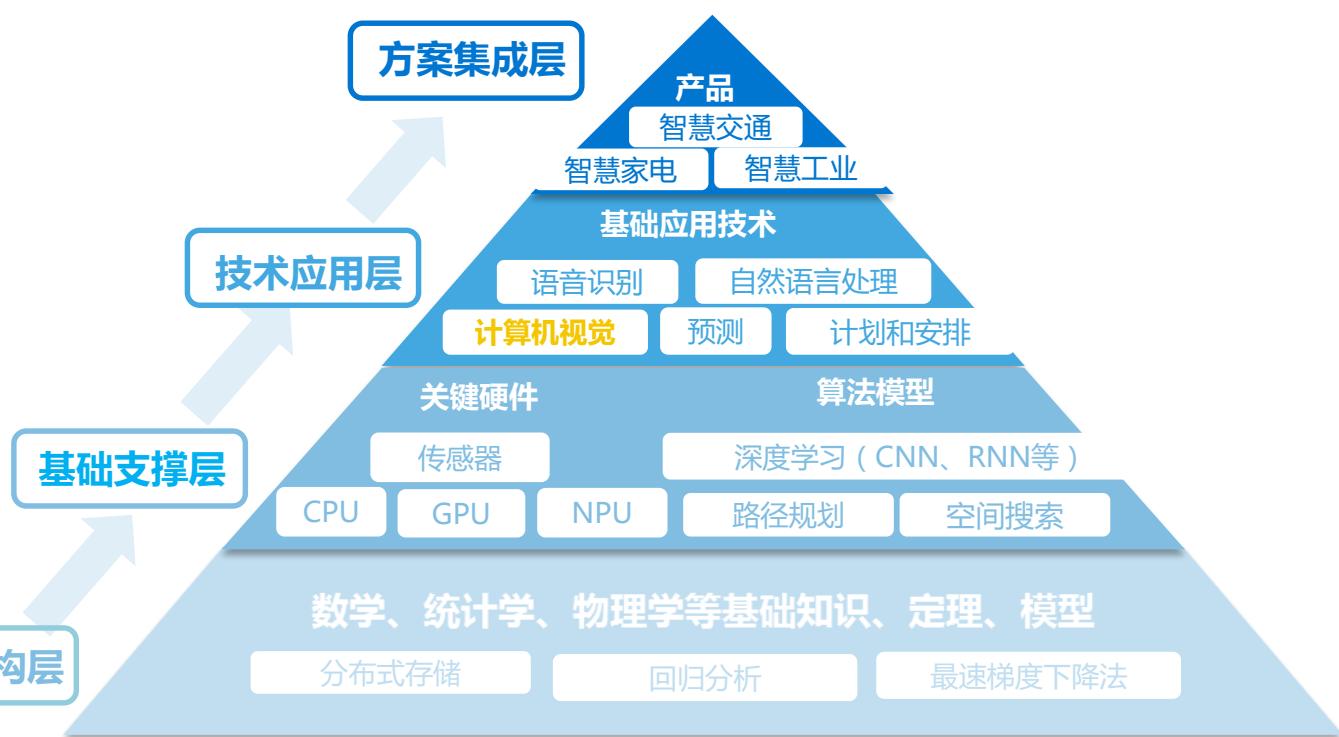
研究意义

计算机视觉赋予机器“看”和“认知”的功能，是人工智能的一类基础应用技术

计算机视觉作为计算机的眼睛，是机器认识世界、看懂世界的一种方式。而认识世界、看懂世界是人工智能产品或方案不可或缺的重要部分。唯有看见，才能够做出分析判断，进而代替人类完成更多的任务。它与语音识别、语言识别一并构成了人工智能的感知智能，让机器完成对外部世界的探测，进而做出判断，采取行动，让更复杂层面的智慧决策、自主行动成为可能。

基于深度学习算法模型和CPU、GPU等关键硬件的支撑，计算机视觉技术应用得以实现，并最终集成于多类产品和应用场景之中（如机器人、无人驾驶等等）。

人工智能基础架构



来源：36氪研究院

计算机视觉能够增强、改善人们的生活，代替人类完成更多的任务。

一方面，它为人类自身视觉提供强有力的辅助和增强，极大的改善人与世界交互的方式。比如，我们可以通过图片搜索引擎找到与之相关信息；另一方面，机器可以准确、客观而稳定的看见，突破人类视觉的局限，代替人类完成更多的任务。比如，机器可以24小时不间断、不疲倦的进行场景监控。

1.1 定义与研究意义

1.2 行业驱动---

数据量

运算力

算法技术

1.3 政策法规

1.4 投资热度

全球投资热度

国内投资热度

1.5 市场规模

行业驱动 · 数据量

海量数据为计算机视觉发展提供燃料

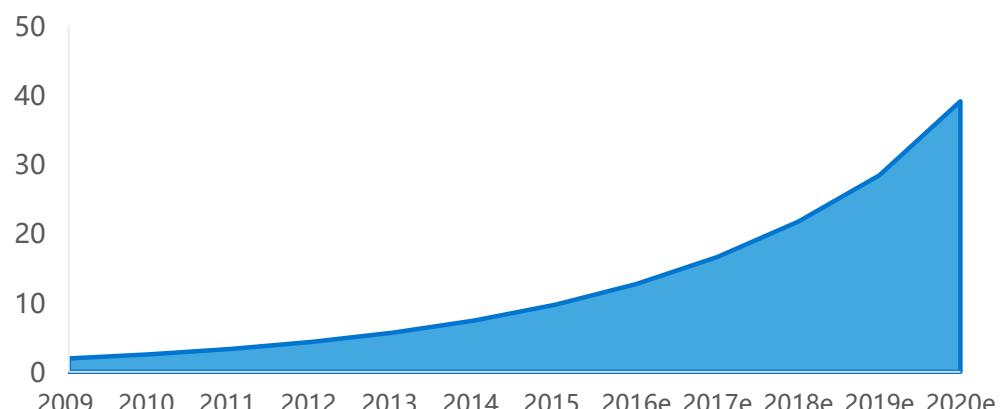
数据量、运算力和算法模型是影响计算机视觉行业发展的三大要素。2000年之后，数据量的上涨、运算力的提升和深度学习算法的出现极大的促进了计算机视觉行业的发展。

要理解数据量的重要性，得先从算法说起。数据量和算法可以分别比作人工智能的燃料和发动机。算法是计算机基于所训练的数据集归纳出的识别逻辑，好的算法模型可以实现精准的物体和场景识别。而数据集的丰富性和大规模性对算法的训练尤为重要。因此可以说，**实现机器精准视觉识别的第一步，就是获取海量而优质的应用场景数据。**以人脸识别为例，训练该算法模型的图片数据量至少应为百万级别。



2000年以来，得益于互联网、社交媒体、移动设备和廉价的传感器，这个世界产生并存储的数据量急剧增加，这为通过深度学习的方法来训练计算机视觉技术提供很好的土壤。IDC数据显示，从2011年起，全球所产生的数据量已达到ZB级别（1ZB约为10亿GB），此后仍将高速增长，预计2020年达到近40ZB。海量的数据将为计算机视觉算法模型提供远远不断的素材，助力机器视觉的精准度提升。

2009-2020年全球总体数据量（单位：ZB）



来源：IDC，36氪研究院

1.1 定义与研究意义

1.2 行业驱动---

数据量

运算力

算法技术

1.3 政策法规

1.4 投资热度

全球投资热度

国内投资热度

1.5 市场规模

行业驱动 · 运算力**运算力的大幅提升推进计算机视觉发展**

在计算机视觉领域，图像、视频作为数据密集型问题，需要大量的矩阵计算操作，因此特别适合使用并行运算芯片进行数据处理。而传统的CPU一次只能同时做一两个加减法运算，无法满足并行运算的需求。

在GPU出现之前，算法运行的速度是很慢的，即使是一个简单的神经网络数据的培训，也得花费几天、甚至几周的时间。这对算法的优化调整带来巨大瓶颈。而由于机器学习的理论刚刚起步，需要不断的试验、迭代，这时运算能力就变得尤为关键。

1999年，Nvidia公司在推销自己的Geforce 256芯片时，提出了GPU (Graphics Processing Unit, 图像处理器)这个概念。GPU是专为执行复杂的数学和集合计算而设计的数据处理芯片。它的出现让并行计算成为可能，对数据处理规模、数据运算速度带来了指数级的增长与改善，极大的促进计算机视觉的发展。

世界上第一款GPU-GeForce 256

GPU与传统CPU相比，在处理海量数据方面有压倒性的优势。据Rajat Raina与吴恩达的合作论文“用GPU进行大规模无监督深度学习”（“Large-scale Deep Unsupervised Learning using Graphic Processors”）显示，在运行大规模无监督深度学习模型时，使用GPU和使用传统双核CPU在运算速度上的差距最大会达到近七十倍。在一个四层，一亿个参数的深度学习网络上，使用GPU将程序运行时间从几周降低到一天。

在今天，数据处理速度不再成为制约计算机视觉发展的瓶颈。想要发挥专用芯片的计算效率优势，需要芯片结构和软件算法两者相匹配。目前的趋势是，随着对人工智能各类应用需求的不断增强，专门用于加速人工智能应用的AI-PU或将成为计算机另一个标配组件。

1.1 定义与研究意义

1.2 行业驱动---

数据量

运算力

算法技术

1.3 政策法规

1.4 投资热度

全球投资热度

国内投资热度

1.5 市场规模

行业驱动 · 算法技术**深度学习算法极大提高视觉识别准确率 (1/2)**

深度学习可以做到传统视觉识别方法无法企及的精度，它的出现让视觉识别准确率大大提高。计算机视觉技术的发展大致经历两大阶段：

Stage1：计算机的视觉识别原理多为通过寻找合适的特征来让机器辨识物品状态，识别准确率低。人们对经验归纳提取，进而设定机器识别物体的逻辑，机器通过对逻辑的学习来识别物体。由于这个逻辑是人设定的，不能穷举各种复杂的情境，因而有很大局限性，识别准确率较低。

Stage 2：深度学习的出现先让识别逻辑由认为设定变为自学习状态，数据量和计算机硬件性能也有了大的发展飞跃，助力精准度提升。

计算机视觉算法发展历史

1.1 定义与研究意义

1.2 行业驱动---

数据量

运算力

算法技术

1.3 政策法规

1.4 投资热度

全球投资热度

国内投资热度

1.5 市场规模

行业驱动 · 算法技术

深度学习算法极大提高视觉识别准确率 (2/2)

深度学习是一种基于多层神经网络的，以海量数据为输入的规则自学习方法。它不是通过固定的公式或程序的描述来做决定，而是根据提供给它的大量的实际行为（训练数据集）来自我调整规则中的参数，进而调整规则。在和训练数据集类似的场景下，可以做出一些很准确的判断。

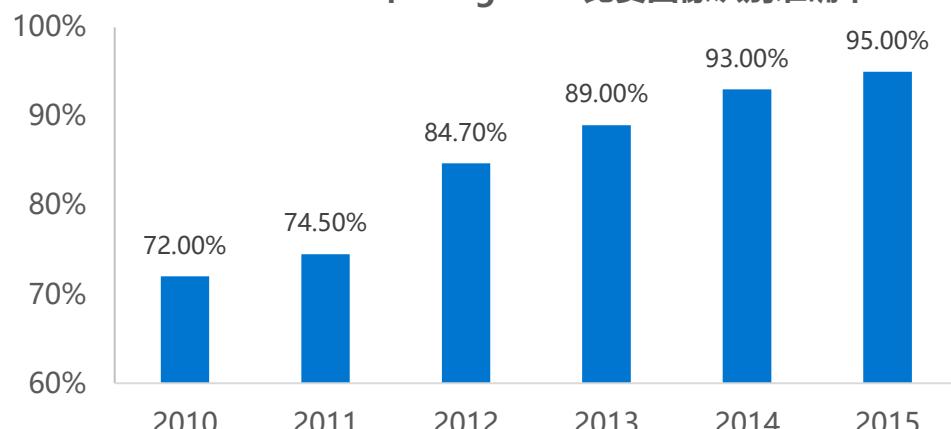
深度学习与传统神经网络的区别

对比项	传统神经网络	深度学习
数据	需人工标注分类特征	未标注
结构层数	一般最多2-3层	目前一般20层左右，最高已上百
参数初始化	随机给定	逐层训练获得
原理	监督学习+误差反向传播	非监督学习+监督学习+误差反向传播

来源：东方证券，36氪研究院

2006年，Geoffrey Hinton 和合作者发表论文，“A fast algorithm for deep belief nets”，此后“Deep Learning（深度学习）”的概念被提出。深度学习出现之前，基于寻找合适的特征来让机器辨识物体状态的方式几乎代表了计算机视觉的全部。尽管对多层神经网络的探索已经存在，然而实践效果并不好。深度学习出现之后，计算机视觉的主要识别方式发生重大转变，自学习状态成为视觉识别主流。即，机器从海量数据库里自行归纳物体特征，然后按照该特征规律识别物体。**图像识别的精准度也得到极大的提升，从70%+提升到95%。**

2010-2015年 ImageNet 比赛图像识别准确率



注释：ImageNet是计算机视觉系统识别项目。

来源：互联网，36氪研究院

1.1 定义与研究意义

1.2 行业驱动---

数据量

运算力

算法技术

1.3 政策法规

1.4 投资热度

全球投资热度

国内投资热度

1.5 市场规模

政策法规**国家政策支持人工智能发展，重点支持安防、交通领域**

人工智能的出现为很多曾依赖人类技巧和判断的领域带来大的改变，比较典型的领域有交通出行、安防、医疗等。各国政府也在积极出台政策促进人工智能对人类生活的改善和提升。以美国为例，2016年美国交通运输部提出议案并给出4000万美元奖励以鼓励中等规模城市布局智慧城市交通。

2015年以来，我国国务院相继发布文件支持人工智能行业的发展。计算机视觉作为基础技术应用，与人工智能各细分领域的发展都息息相关。在金融领域，实名制号角已吹响，这代表着人脸识别迎来巨大市场。在安防领域，国家出台文件加大公共安全视频监控，加强建设智慧城市，计算机视觉作为主要技术手段将发挥重要作用。

实施时间	颁布主体	法律法规	支持对象	相关内容
2015/5/6	发改委	《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》	安防领域视觉识别	要求逐步建立国家级、升级公共安全视频图像数据处理中心，深化预测预警、实时监控、轨迹追踪、快速检索等应用
2015/7/4	国务院	《国务院关于积极推进“互联网+”行动的指导意见》	人工智能行业	明确提出人工智能作为11个重点布局的领域之一，促进人工智能在智能家居、智能终端、智能汽车、机器人等领域的推广应用。
2015/7/9	中央办公厅、国务院	《关于加强社会治安防控体系建设的意见》	安防领域视觉识别	加大公共安全视频监控覆盖，将社会治安防控信息化纳入智慧城市建设总体规划，加深大数据、云计算和智能传感等新技术的应用。
2016/3/18	国务院	《国民经济和社会发展第十三个五年规划纲要（草案）》	人工智能行业	人工智能概念进入“十三五”重大工程。
2016/6/28	国家网信办	《移动互联网应用程序信息服务管理规定》	交通出行、金融等领域人脸识别	第七条 移动互联网应用程序提供者应当严格落实信息安全管理责任，依法履行以下义务：（一）按照“后台实名、前台自愿”的原则，对注册用户进行基于移动电话号码等真实身份信息认证。
2016/7/1	人民银行	《非银行支付机构网络支付业务管理办法》	金融领域人脸识别	要求支付机构对客户实行实名制管理，并在与客户业务关系存续期间采取持续的身份识别措施。

1.1 定义与研究意义

1.2 行业驱动---

数据量

运算力

算法技术

1.3 政策法规

1.4 投资热度

全球投资热度

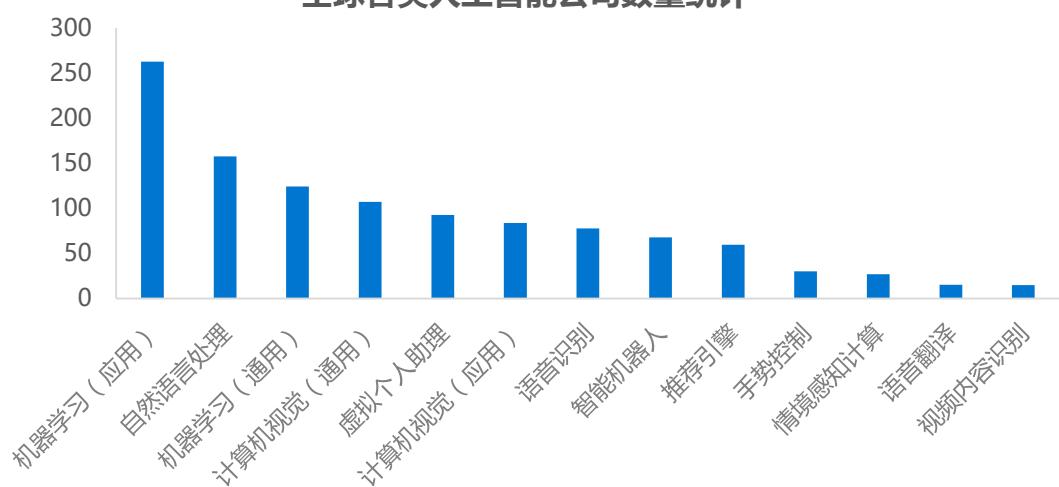
国内投资热度

1.5 市场规模

投资热度 · 全球**全球认知技术领域创业活跃，计算机视觉投资热度高**

Venture Scanner 追踪了全球957家人工智能公司，总融资额为48亿美金。其中，机器学习、自然语言处理、计算机视觉等认知技术占据当前市场主导地位。

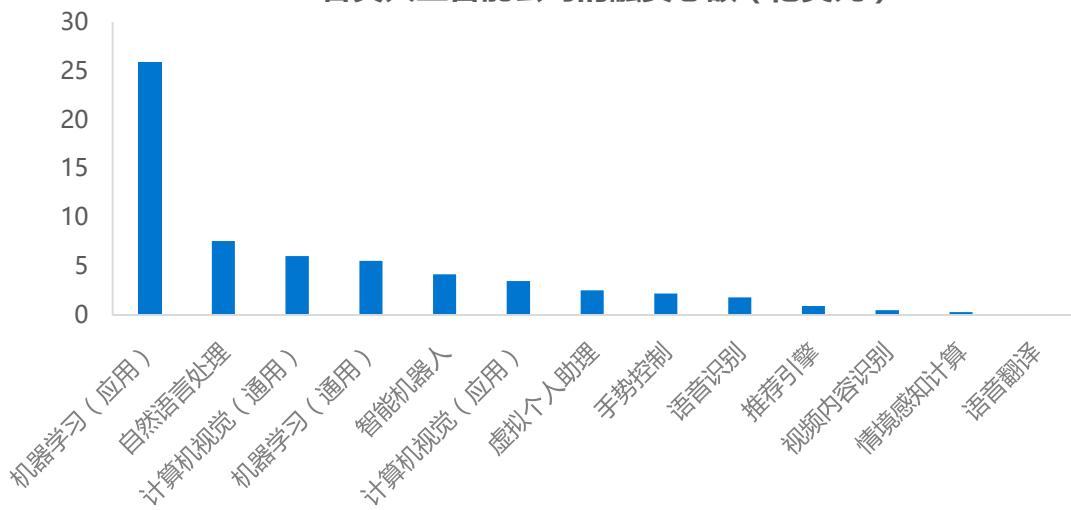
- 公司数量**：2016年Q1数据显示，从事计算机视觉的公司数量约在180+左右，约占人工智能创业公司的1/5.

全球各类人工智能公司数量统计

注释：数据统计截至2016年第一季度。

来源：Venture Scanner，36氪研究院

- 融资额度**：2016年Q1数据显示，计算机视觉融资额度在16亿美金左右，在细分领域中位居前列。

各类人工智能公司的融资总额（亿美元）

注释：数据统计截至2016年第一季度。

来源：Venture Scanner，36氪研究院

1.1 定义与研究意义

1.2 行业驱动---

数据量

运算力

算法技术

1.3 政策法规

1.4 投资热度

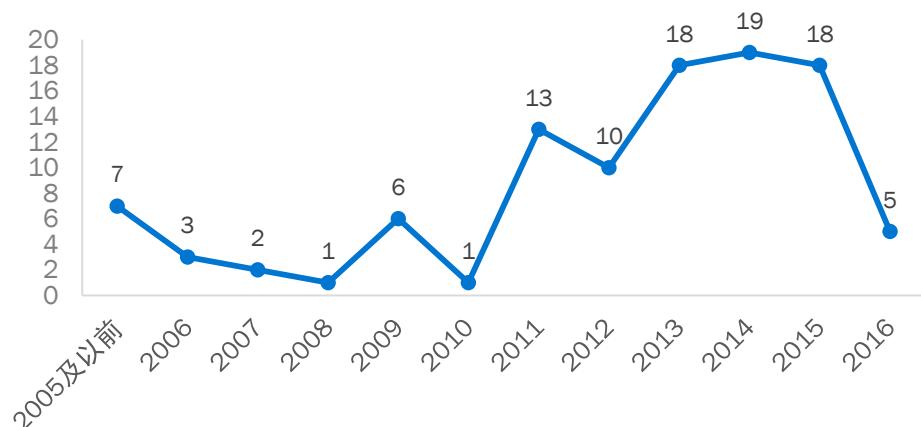
全球投资热度

国内投资热度

1.5 市场规模

投资热度·中国**国内公司2011年后出现创业热潮，融资阶段偏早期****• 创业公司数量：**

国内从事计算机视觉领域的公司在2011年后显著增加，至今热度不减。据36氪研究院不完全统计，目前国内从事计算机视觉技术的公司有104个。

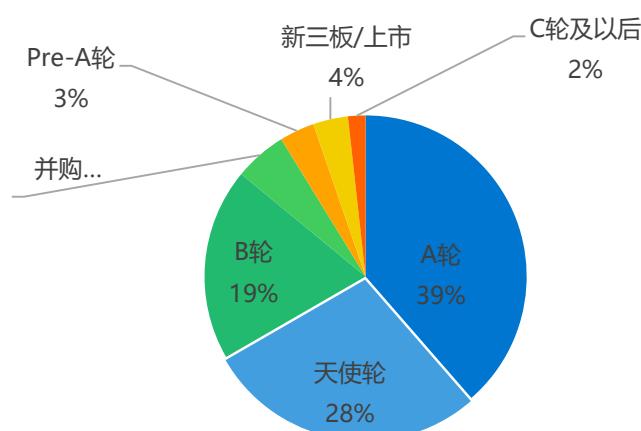
1997-2016年中国计算机视觉公司成立数量(个)

注释：数据统计截止至2016年8月5日。

来源：36氪研究院

• 融资阶段：

从事计算机视觉的创业企业目前发展偏早期，A轮和天使轮居多，占据创业公司总量的67%。

计算机视觉公司融资阶段分布(%)

注释：在104家从事计算机视觉的企业中，有融资数据的为57家，其融资分布如上。

来源：36氪研究院

1.1 定义与研究意义

1.2 行业驱动---

数据量

运算力

算法技术

1.3 政策法规

1.4 投资热度

全球投资热度

国内投资热度

1.5 市场规模**市场规模****人脸识别和视频监控领域仍有较大上升空间**

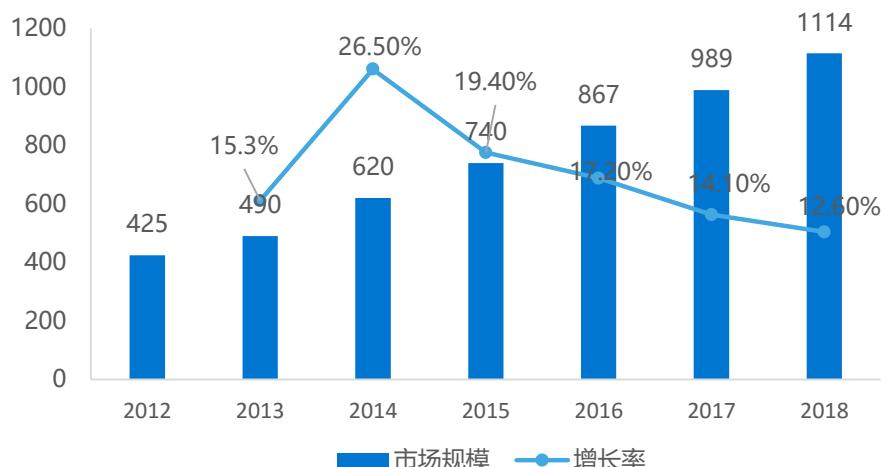
计算机视觉作为一种人工智能的基础技术应用，使用场景多样，市场潜力巨大。其中，**人脸识别和视频监控作为计算机视觉的重要应用领域，2015年市场容量已达十亿和百亿级别。**

- 2015年中国人脸识别市场为十亿级别：**以人脸识别为例，Capvision 数据显示，2015年全球人脸识别市场规模已达200亿，中国约为70亿。该数据的统计基于传统应用场景，如安防、门禁方面的人脸识别应用。

人脸识别市场规模 (亿元人民币)

来源：Capvision, 36氪研究院

- 2015年中国视频监控领域视觉识别市场容量约为200亿：**从应用场景来看，视频监控作为计算机视觉的一个重要应用场景，其市场规模在2015年达到740亿，该数据的估算依据来自视频监控领域企业的营收。由于目前国内视频监控领域投入大多在硬件领域，视频和数据分析占比较低。若30%的比例用于视频和数据的分析，则此部分的市场容量约为200亿。

2012-2018年中国视频监控市场规模

来源：中国安防, 36氪研究院

Chapter 2

计算机视觉技术与应用场景

- 通用视觉识别技术流程
- 识别技术与应用场景简介
 - 生物特征识别
 - 光学字符识别
 - 物体与场景识别
 - 视频对象提取与分析

2.1 通用视觉识别技术流程

2.2 识别技术与应用场景简介

2.2.1 生物特征识别

2.2.2 光学字符识别

2.2.3 物体与场景识别

2.2.4 视频对象提取与分析

技术流程

通用流程包含目标检测、目标识别、行为识别

从技术流程上看，计算机视觉识别通常需要三个过程：目标检测、目标识别、行为识别，分别解决了“去背景”、“是什么”、“干什么”的问题。



(1) 目标检测解决的是从背景中找出我们关心的物体的问题，祛除了背景中不相关的信息。首先，采取平滑去噪、调整对比度与尺寸空间的方式将图片进行预处理，使其可以被检测到并适合局部应用。然后，将每块文字/图像分检出来交给识别模块识别。再次，将图像分割成若干个特定、具有独特性质的区域，进而寻找出感兴趣的目标。

目标检测技术关键在于动态复杂场景中背景模型的建立、保持与更新。当背景发生动态变化或被遮挡时，检测难度会大大增加。

(2) 目标识别解决的是“是什么”的问题。在提取图像信息后，决定每个图像的点是否属于一个图形特征。将图像上的点分为不同的自己。通过对图像内容、特征、结构、关系、纹理及灰度等多维度的对应关系、相似性和一致性的分析，寻求相似影响目标，得出相似度比值。

目标识别的难点在于光照条件的改变和角度的不同，地点、光线、玻璃都会影响识别的准确性。

(3) 行为识别解决的是“干什么”的问题，是一种高技术难度的识别问题，需要对动态多帧图像数据进行理解，并构建相应动作行为模型进行比对。

行为识别是一种高层次的人工智能，因为它和预测、分析息息相关。在防恐防爆领域，行为识别可以帮助监控恐怖分子动向，并实现实时报警功能。

注释：部分内容来自兴业证券《视频大数据》报告，感谢整理与分享。
来源：36氪研究院

2.1 通用视觉识别技术流程

2.2 识别技术与应用场景简介

2.2.1 生物特征识别

- 技术简介
- 应用场景

2.2.2 光学字符识别

2.2.3 物体与场景识别

2.2.4 视频对象提取与分析

技术简介 · 生物特征识别

技术较为成熟，人脸识别将成主流

生物特征识别技术是一种通过对生物特征的识别和检测从而对人的身份实行鉴定的技术。从统计的意义上讲，人类的指纹、掌形、虹膜等生理特征存在着唯一性。因而这些特征都可以作为鉴别用户身份的依据。目前，**生物特征识别主要用于用户身份的鉴定，主要包括指纹识别、人脸识别、静脉识别、虹膜识别、语音识别等方式。**



指纹识别

- 造假成本低
- 普及率高
- 易用性强

人脸识别

- 造假成本高
- 正在推广普及
- 易用性强

虹膜识别

- 造假成本极高
- 普及率低
- 易用性适中

市场现状：指纹识别普及度最高，但市场份额呈下降趋势，人脸识别将逐渐成为主流。

由于身份识别的应用场景不断被开拓延伸，生物特征识别市场规模总体呈不断增长态势。目前，指纹识别的普及度是最高的，约占生物特征识别的60%。但由于指纹具有易磨损、伪造成本低的特点，长期稳定性和安全性较其他特征识别偏低，因而市场份额呈下降趋势。

尽管在虹膜识别是几类生物特征识别中安全性是最高的，但其识别成本和图像采集成本也较高。而**人脸识别因其非接触性**（用户无需和设备直接接触）、**非强制性**（可主动获取识别队形人脸图像）、**高并发性**（同一应用场景下可同时进行多个人脸分拣），**应用场景更为丰富**，与其他生物特征识别技术相比，前景广阔。

安防领域应用性



2.1 通用视觉识别技术流程

2.2 识别技术与应用场景简介

2.2.1 生物特征识别

- 技术简介
- 应用场景

2.2.2 光学字符识别

2.2.3 物体与场景识别

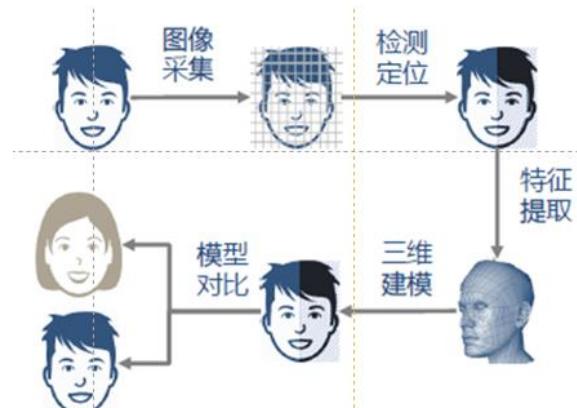
2.2.4 视频对象提取与分析

技术简介 · 人脸识别

1：1较为简单，1：N更具挑战

识别步骤：与其他生物特征识别步骤类似，人脸识别包括图像采集、检测定位、特征提取、模型对比等步骤。

人脸识别流程



来源：Capvision, 36氪研究院

识别难点：1：N识别更具挑战性

按照技术难度的不同，人脸识别可分为1：1 和1：N两种人脸识别应用。其中，1：N的识别难度更具挑战性。

1：1识别应用

1：N识别应用

主要特点

- 用于证明自己是自己的场景

- 用于动态和非配合状态下进行身份鉴定的场景

识别难度

- 通常会在识别对象配合的状态下进行图像的采集，识别难度低，精准率高

- 通常识别对象不用感知到摄像头的位置或进行配合；识别的图像为动态的视频流，且光线、遮挡等因素都会影响识别准确度

应用场景

- ✓ 机场安检
- ✓ 银行远程开户
- ✓ 上班考勤
- ✓ 公司门禁

- ✓ 疑犯追踪
- ✓ 商场用户画像分析
- ✓ 防暴预警
- ✓ 人流分析

2.1 通用视觉识别技术流程

2.2 识别技术与应用场景简介

2.2.1 生物特征识别

- 技术简介
- 应用场景

2.2.2 光学字符识别

2.2.3 物体与场景识别

2.2.4 视频对象提取与分析

应用场景 · 生物特征识别

适用各类身份认证相关场景，人脸识别应用尤为广泛

生物特征识别作为一种身份鉴定的识别技术，应用场景广泛。

- 指纹识别、静脉识别、虹膜识别因其图像收集的强制性（必须主动对着采集设备进行图像采集）更多的用于安防与考勤。
- 人脸识别因其非接触性和非强制性，在各类场景中广泛应用。

身份鉴定



金融安全：在金融领域，人们可以通过自拍照片进行身份验证，计算机自动比对是否为本人开户，大大提升了线上金融的便利性。另一方面，在网络安全越来越被重视的今天，人们可以刷脸支付，避免了手机被盗短信验证支付的安全隐患。这类技术已趋于成熟，不久前，阿里巴巴就和Face++携手推出“微笑支付”。



出行安全：网络约车的安全性一直广受诟病。不久前，Uber和滴滴相继上线“人像认证”功能以确认司机身份，在注册前、行车中均可通过“人脸识别”、“声纹识别”等生物识别技术对司机的身份信息进行确认，认证通过后司机才可以接单。



商业服务：人脸识别可以帮助企业自动识别VIP客户。当VIP客户被摄像头自动检测到，计算机就会将企业数据库中该客户的各类偏好信息提供给在场的工作人员，进而让VIP客户享受到主动的、及时的、个性化的商业服务。

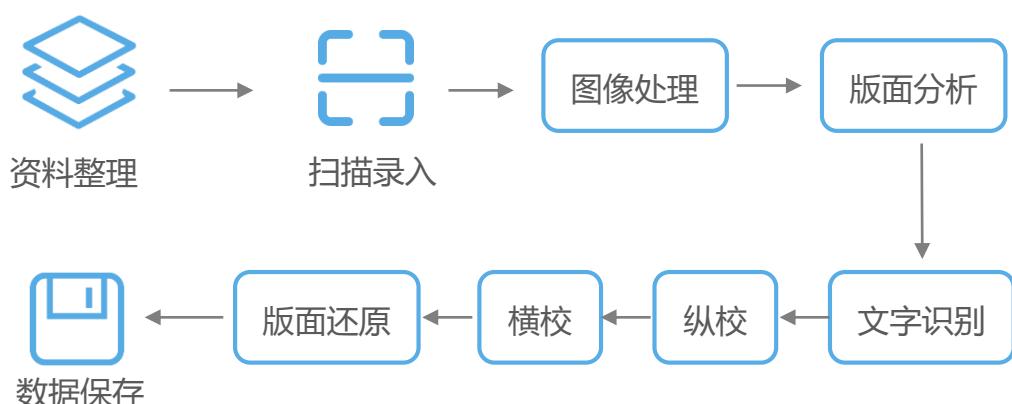
- 2.1 通用视觉识别技术流程
- 2.2 识别技术与应用场景简介
 - 2.2.1 生物特征识别
 - 2.2.2 光学字符识别**
 - 技术简介
 - 应用场景
- 2.2.3 物体与场景识别
- 2.2.4 视频对象提取与分析

技术简介 · 光学字符识别

技术较为成熟，版式、光照为识别难点

光学字符文本识别技术又称OCR技术 (Optical Character Recognition)。计算机通过光学设备检查纸上打印的字符，通过检测暗、亮的模式确定其形状，然后用字符识别方法将形状翻译成计算机文字的过程，完成计算机对文字的阅读。

识别步骤：OCR技术要经历扫描、版面分析、文字识别、版面还原等过程。



识别难点：照片中多字符文本识别

在受约束环境下（如文档处理）的字符识别技术已经被研究得很透彻。然而，对照片中的多字符文本识别却是个难题。由于照片中文字的版式不一（字体、颜色、样式、方向、排列），加之光照、阴影、镜像、遮蔽等环境因素影响，还有图像本身的分辨率、焦点模糊、抖动等问题，照片中的多字符文本识别具有挑战性。

OCR识别难点



- 过曝
- 光线不足
- 遮挡



- 焦点模糊
- 抖动
- 分辨率低



- 版式不一
- 方向排列不同
- 图文边界

- 2.1 通用视觉识别技术流程
- 2.2 识别技术与应用场景简介
 - 2.2.1 生物特征识别
 - 2.2.2 光学字符识别**
 - 技术简介
 - **应用场景**
 - 2.2.3 物体与场景识别
 - 2.2.4 视频对象提取与分析

应用场景 · 光学字符识别

适用各类信息录入场景，垂直领域大数据分析为创业公司商业化着力点

OCR技术具体到场景，就是通过扫描，将各类印刷体的文档证件（名片、身份证件、银行卡、财务单据、保险单据等）自动识别录入电脑或手机等终端，让高效的、大量的信息录入成为可能，对非结构化数据信息的采集具有重要意义。

字符识别应用场景示例



快速识别各类证件，导出证件信息，接入联网业务系统



识别文档图片文字，避免繁琐手工录入



扫描并录入表格、单据、发票，方便各类业务联网



减少客户等待，优化用户体验

主要应用：身份证件识别

通过手机拍摄证件，手机会快速扫描并读取图像中各栏信息，并存入证件信息数据库。虽然手动输入该行驶证信息并不繁琐，但当这类工作变为常规、大量重复的行为时，机器的自动识别会大大提升效率，节省时间。

主要应用：单据识别

通过拍摄病历、药品单、缴费单等单据，计算机可以将各类单据自动识别为文字。一方面电子化的录入大大提升了人工效率，另一方面，这让信息的流动成为可能。例如，通常医院和保险公司的数据是不流通的，欺诈理赔时有发生。保险公司通过使用OCR技术自动识别报销单据，并进行单据的验伪，有效防止骗保情况的发生。

在具体应用中，只要存在字符信息录入、印刷文件处理、表单验伪等需求的行业，均为OCR的潜在应用行业。据业内专家介绍，**目前OCR技术在金融、保险、物流等行业的应用已较为广泛成熟，将来会扩展到医疗和O 2 O。**

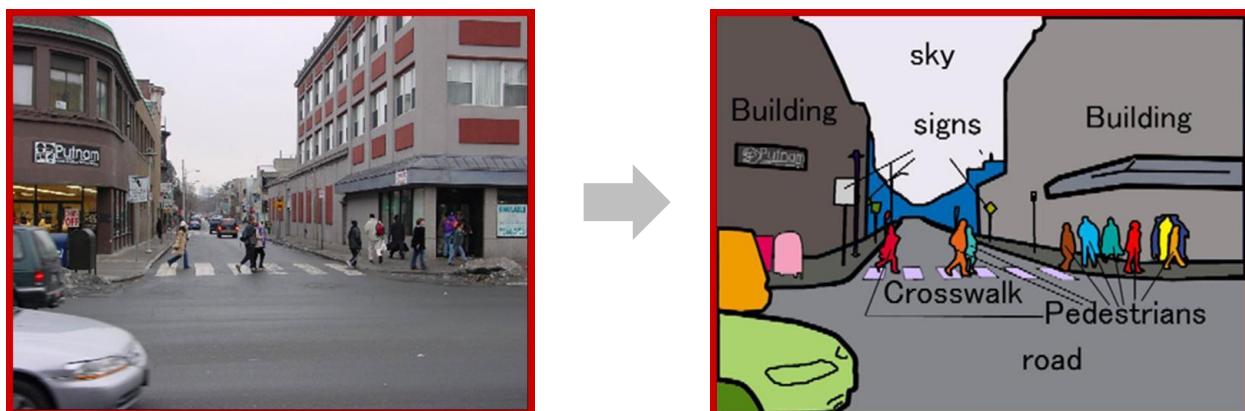
- 2.1 通用视觉识别技术流程
- 2.2 识别技术与应用场景简介
 - 2.2.1 生物特征识别
 - 2.2.2 光学字符识别
- 2.2.3 物体与场景识别**
 - 技术简介
 - 应用场景
- 2.2.4 视频对象提取与分析

技术简介 · 物体与场景识别

技术刚刚起步，现阶段多数公司着力数据标注

识别步骤：物体、场景识别是计算机将图片内容与数据库中的资料进行比对，进而识别图片或视频中的各种物体与场景的识别技术。
识别的过程与人脸识别类似，即通过检测定位、特征提取、三维建模、模型比对的过程进行物体的识别。下图即为一个典型的驾驶环境下的物体与场景识别。

物体与场景识别



识别难点：由于同类物体的形状、颜色不一，难以寻找同一规律，目前物体识别技术仍在研发的早期阶段。这个领域很多公司着力于给不同图片打标签，帮助机器学习同类物体的不同形态，用大的数据量来帮助训练物体识别算法模型的精准度。PK图像识别算法的最有影响力的比赛--ImageNet的基本任务就是训练一个1000类物种的图像识别器。然而，真实视觉的物种远远超过1000类，物体识别的研究早期且有限。基于此，**高效的自动或者半自动的数据集创建是实现物体识别的第一步**。不少视觉领域的公司雇佣外包团队进行数据标注。在这一点上，中国由于人力成本低，比美国有更多优势。



- 2.1 通用视觉识别技术流程
- 2.2 识别技术与应用场景简介
 - 2.2.1 生物特征识别
 - 2.2.2 光学字符识别
- 2.2.3 物体与场景识别**
 - 技术简介
 - 应用场景
- 2.2.4 视频对象提取与分析

应用场景 · 物体与场景识别

应用场景想象空间巨大，各场景探索刚刚开始

物体和场景是被是一种更高难度，更复杂层面的识别，几乎涵盖了所有计算机视觉的范畴。当机器代替人眼，应用场景有无限想象空间。目前，计算机视觉已广泛覆盖军事、医疗、工业、商业等领域。

物体与场景识别



应用对象	应用场景	应用方式
军事	定向武器投射	基于机器视觉，帮助武器自动进行目标定位。
医疗	医疗影像分析	通过机器学习和计算机视觉技术，实现机器对医学影像的分析判断，为医生的影像诊断提供辅助。
工业生产	成品无损检测	通过机器视觉进行产品质量检测，替代产品质检人员，避免接触损伤，提高瑕疵品检测率。
	高位作业监控	通过机器视觉，完成不适合人工的危险工作环境下的视觉监控。
消费生活	图片搜索引擎	通过图片本身来搜索相似图片。
	无人驾驶	利用机器视觉识别和发现周围场景及各类物体动向，进而辅助车辆自主进行驾驶决策。
	机器人	利用机器视觉，让机器人识别周围环境。
	无人机	帮助机器识别周围视觉基本属性和大致情况，避免高速碰撞。
	AR/VR	通过计算机视觉技术来模拟人类视觉。

来源：36氪研究院

- 2.1 通用视觉识别技术流程
- 2.2 识别技术与应用场景简介
 - 2.2.1 生物特征识别
 - 2.2.2 光学字符识别
- 2.2.3 物体与场景识别**
 - 技术简介
 - 应用场景
- 2.2.4 视频对象提取与分析

应用场景 · 物体与场景识别

家用机器人：处于产业化初期，整体智能程度低

受益于人口老龄化的趋势，家用机器人领域研发热潮不断。当机器人具有了“视觉”，便可以按照行为路径移动，躲避障碍物，进行家政清洁。一些视觉识别技术更成熟的家庭机器人，可以识别家庭成员身份，识别人类表情，进而与人类进行互动。

家用机器人Jibo



来源：互联网，36Kr研究院

典型家用机器人示例

机器人名称	Buddy	Jibo	Pepper
发售公司	Blue Frog Robotics	Jibo Inc.	软银&Aldebaran
发售时间	2016年12月	2016年10月	2015年2月
价格(美元)	699	749	1610
功能	<ul style="list-style-type: none"> · 附有3个轮子和1个8寸屏幕，可以自由移动 · 能够感知外在环境，进行物体与人脸检测，通过视频传输可让用户远程监控家中电器 	<ul style="list-style-type: none"> · 主打实体Siri,通过语言和摄像头实现语音通话、视频、拍照等 · 可识别家中成员身份 	<ul style="list-style-type: none"> · “全球首款情绪机器人”，能够随着人类情绪改变互动模式 · 结合IBM的“华生”系统，能通过机器学习不断进步

来源：Capvision，36氪研究院

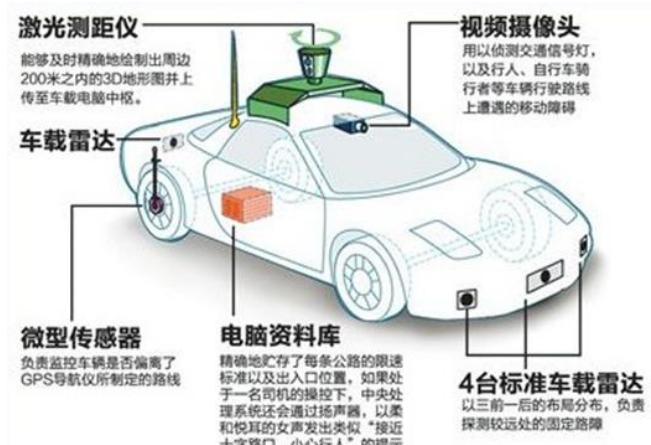
发展现状目前家用机器人仍在**产业化初期**，因其技术瓶颈大，市场上虽有大量服务机器人出台，但整体更偏向于**智能不足的玩具**。

- 2.1 通用视觉识别技术流程
- 2.2 识别技术与应用场景简介
- 2.2.1 生物特征识别
- 2.2.2 光学字符识别
- 2.2.3 物体与场景识别**
 - 技术简介
 - 应用场景
- 2.2.4 视频对象提取与分析

应用场景 · 物体与场景识别

自动驾驶：辅助驾驶成熟度更高，完全自动仍待时日

计算机视觉作为机器感知周围环境的基础技术，对驾驶的自动化起到重大作用。计算机视觉可以帮助汽车完成：(1) 环境数据和地标收集(2)车道的定位、坡道与弯度的检测(3)交通信号的识别(4)车辆、行人等运动目标检测。



来源：互联网，36Kr研究院

自动驾驶的发展需要产业链上诸多企业的关注与参与，人工智能技术的复杂度非常高，比如，自动驾驶的实现需要汽车零件、传感器、软件服务系统等不同公司基于产业链的通力协作才能实现。

全自动驾驶仍不成熟：从技术成熟度上讲，已基本实现从感知到控制的跨越，但产业化水平整体更偏辅助驾驶及其延伸，仍需人的参与，而非全自动无人驾驶。



来源：BCG，36Kr研究院

- 2.1 通用视觉识别技术流程
- 2.2 识别技术与应用场景简介
 - 2.2.1 生物特征识别
 - 2.2.2 光学字符识别
 - 2.2.3 物体与场景识别
- 2.2.4 视频对象提取与分析
 - 技术简介
 - 应用场景

技术简介 · 视频对象提取与分析

基于动态图像进行识别，光线、遮挡为识别难点

严格来讲，视频对象提取和物体与场景识别技术、人脸检测技术存在应用场景的重合，之所以将其另分为一类是为了强调两点：一是此类技术是基于动态图像的识别，二是强调视觉识别单项技术的应用即可成为此类应用场景的直接产出。

识别步骤：视频识别包括以下三个环节。在进行视频识别与分析时，需要前端摄像头设备来收集和传输数据，同时，需要通过大数据训练、具备云计算能力的深度学习图像分析系统来实时进行视频检测和数据分析。



识别难点：

- 一方面，**动态视频的识别中对抗光线、遮挡等干扰的难度更大**，对机器进行图像分拣、滤除干扰的提出了很高的要求；
- 另一方面，**动态视频识别对机器识别的速度有较高要求**。一些公司通过智能前端化的方式来提升分析速度，即在智能前端摄像头搭载强并行计算能力处理器的方式，以提供更高效、更实时、不依赖网络传输的智能服务。



来源：互联网，36Kr研究院

- 2.1 通用视觉识别技术流程
- 2.2 识别技术与应用场景简介
 - 2.2.1 生物特征识别
 - 2.2.2 光学字符识别
 - 2.2.3 物体与场景识别
 - 2.2.4 视频对象提取与分析**
 - 技术简介
 - 应用场景

应用场景· 视频对象提取与分析

监控系统、行为识别、电商营销为主要应用场景

监控系统、行为识别、电商营销是目前视频识别与分析应用商业化较为成熟的三个领域。

视频对象提取与分析



主要应用：监控系统

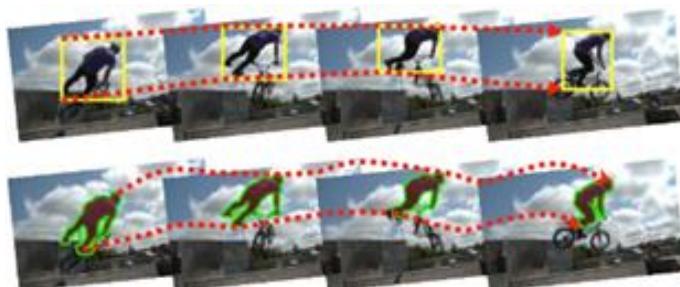
由于机器不会出现疲劳，且可以全面识别整帧图像信息，通过使用该项技术处理海量监控视频，可大大降低交管部门、公安部门人工监控的负担，提升监控效率。具体的应用场景有车牌识别、非法停车检测、嫌疑犯追踪等。国内的依图科技、商汤科技就已与交管、公安部门展开合作。



主要应用：恐怖分子检测

通过对一些典型恐怖分子丢包、使用武器等行为动作数据集的学习，计算机可以在视频中准确识别特定的“疑似恐暴动作”，发出预警。

动作识别示例



应用场景· 视频对象提取与分析

监控系统、行为识别、电商营销为主要应用场景



主要应用：电商营销

基于视频图像的分析技术可以广告与客户需求进行更精准匹配，在视频中精准植入广告，提升转化率。简单来说，就是通过视频识别，对识别对象、物品的建立判断标签，而后根据标签内容进行商品个性化推荐。比如，不少女生执着于韩剧女主同款，使用视频识别与分析技术，可以边看韩剧，边买同款。



来源：互联网，36氪研究院



主要应用：门店用户画像

通过视觉识别技术，计算机可对线下客群实时标注，包括年龄、身高、性别等个人特征信息和区域、通道和留滞时间等购买偏好、用户行为信息。目前，以旷视科技为代表的视觉识别公司已与国内零售、快餐等服务行业展开合作，为企业提供客流分析。



来源：互联网，36氪研究院

Chapter 3

计算机视觉行业梳理

- 产业链
- 全景图
- 商业模式
- 行业竞争

市场格局

创业公司分析

- 行业思考

3.1 产业链

3.2 全景图

3.3 商业模式

3.4 行业竞争

市场格局

创业公司分析

3.5 行业思考

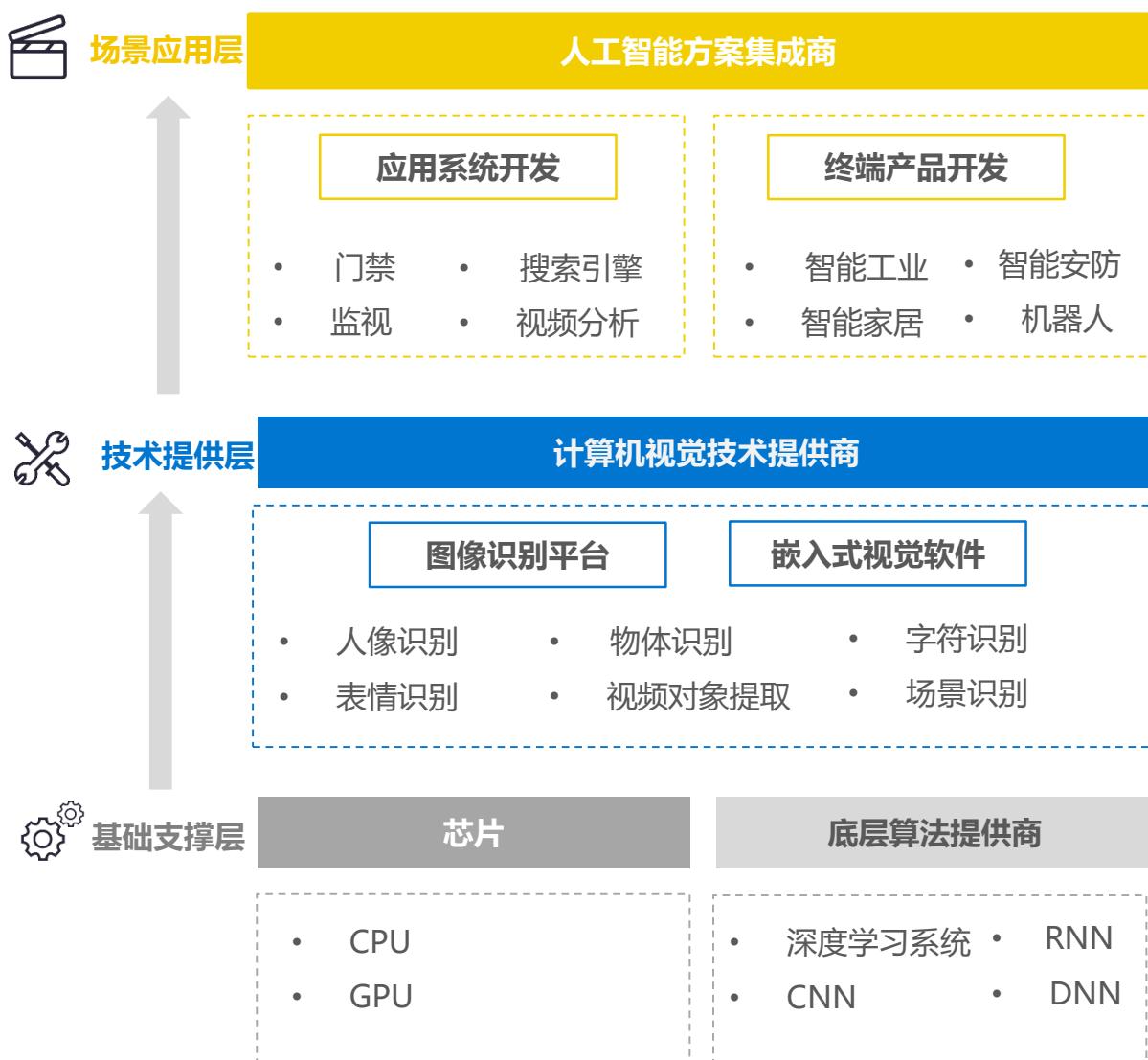
商业模式

技术趋势

产业链

大量创业公司集中于技术提供层

计算机视觉产业链全景图



基础支撑层：包含芯片和初级算法提供商两部分。初级算法提供商是指Google、Microsoft这类算法最强、最新的公司，他们会提供一些开源的系统和算法库供视觉识别的公司使用。芯片主要是为了解决运行算法、模型时的运算能力。

技术提供层：包含图像识别平台和嵌入式视觉软件两类。前者直接提供应用服务，后者需要和硬件进行系统集成后在终端产品中使用。

场景应用层：直接解决具体应用场景的需求，产品的形式可能是应用系统，也可能是软硬件一体的终端产品或服务。

- 3.1 产业链
- 3.2 全景图
- 3.3 商业模式
- 3.4 行业竞争
 - 市场格局
 - 创业公司分析
- 3.5 行业思考
 - 商业模式
 - 技术趋势

全景图

大量创业公司聚焦人脸识别领域

该全景图所包含企业覆盖上页产业链图谱中技术提供层和场景应用层的部分厂商。全景图行业划分未能涵盖全部，仅挑选部分代表领域。我们对市场上以下四类企业进行研究分析，发现人脸识别领域创业公司竞争激烈。

人脸识别



光学字符识别



物体与场景识别



视频对象提取与分析



注释：以上公司所属细分行业均按主要业务划分，仅选取行业部分企业，排名不分先后。

来源：36氪研究院

商业模式

软件服务：从事To C模式服务商较少，更多集中To B 模式

3.1 产业链

3.2 全景图

3.3 商业模式

3.4 行业竞争

市场格局

创业公司分析

3.5 行业思考

商业模式

技术趋势

从事计算机视觉技术的公司按照提供产品或服务的不同，可大致分为两类：软件服务商和软硬件一体解决方案服务商。

(1) 软件服务

此类企业通过售卖人脸检测软件、视频管理软件、视频分析软件等为市场提供视觉识别服务。按照数据处理方式及数据存储位置的不同，可分为在线API、离线SDK、私有云等服务方式。



- **在线API**：服务商提供接口，客户远程调用服务商提供的软件服务，数据的处理在服务商端进行。



- **离线SDK**：服务商向需要离线使用场景的客户授权离线SDK，数据的处理在客户端进行。

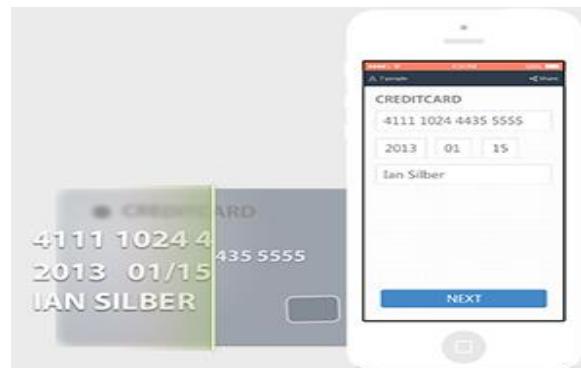


- **私有云**：服务商为企业定制个性化服务，调用私有云内的远程接口，数据的处理在客户端进行。

按照是否面对终端客户，视觉技术服务商又可分为：

- **To C 的软件服务**：这类公司直接面向用户，因而具有C端数据获取和数据反馈优势。合合信息旗下的名片全能王、扫描全能王就是通过软件服务的方式向客户提供服务。除了提供个人免费App版，还提供企业付费App版，客户可以通过APP进行名片和文件的识别。

使用名片全能王APP进行银行卡扫描



来源：合合信息，36氪研究院

- **To B to C的软件服务**：这些算法提供商不直接和最终客户接触，通过软件服务的方式向集成商提供视觉识别技术支持，团队以算法工程师为主，一般没有强大的市场团队，倾向于和渠道多的集成商进行合作，议价能力较弱。

商业模式

软硬件一体的技术解决方案：整体用户更偏企业，软硬一体更具技术壁垒

(2) 软硬件一体的技术解决方案

此类企业提供的是具体应用场景的技术解决方案。通过售卖硬件+软件服务，更加个性化的满足客户需求。

软硬件结合的优点是能够在**前端硬件设备上嵌入算法软件，实现更快速、更高精度地数据处理**。此外，能够让客户更直接的应用视觉识别技术，**避免对接复杂的软硬件的兼容集成**。

按照是否面对终端客户，视觉技术服务商又可分为：

- **To B的解决方案**：格灵深瞳的威目车辆特征识别系统就是通过软硬件一体的方式让客户直接应用车辆识别技术。

威目车辆特征识别系统



来源：格灵深瞳，36氪研究院

- **To B to C的解决方案**：此类公司不直接和最终客户接触，更多的是采用平台的做法向智慧家电、机器人等厂商提供视觉技术支持，即中间件服务。软硬件一体的解决方案技术壁垒强，与只提供软件相比，议价能力更高。此外，由于大的厂商具有获取客户的优势，与其合作成为越来越多创业公司的选择。例如，地平线机器人、阅面科技就致力于为智能家电、机器人提供软硬件一体的技术解决方案，实现低功耗、本地化的环境感知和人机交互。

行业竞争

创业公司可利用算法集成技术优势争取一席之地

3.1 产业链

3.2 全景图

3.3 商业模式

3.4 行业竞争

市场格局

创业公司分析

3.5 行业思考

商业模式

技术趋势

我们将从事计算机视觉的公司分为三类：工业巨头、互联网巨头和创业公司，并总结了市场竞争现状与发展预期。

- 工业巨头**：直接面向该领域客户，在该行业已有较长时间积累，市场份额大；
- 互联网巨头**：从事计算机视觉相关研究的互联网巨头，技术水平领先；
- 创业公司**：从事计算机视觉技术服务的创业公司。

企业类型	工业巨头	互联网巨头	创业公司
代表企业	<ul style="list-style-type: none"> 安防领域有： 海康威视、浙江大华等 智慧家电领域有： 美的集团、海尔集团等 	Google、Microsoft、Facebook、IBM、Intel、百度等	旷视科技、商汤科技、依图科技等
技术获取方式	<ul style="list-style-type: none"> 并购或合作 	<ul style="list-style-type: none"> 开设实验室或并购技术团队 	<ul style="list-style-type: none"> 创建技术团队
优势	<ul style="list-style-type: none"> 产业链布局全 渠道、获客能力强 上下游议价能力强 应用场景数据获取能力强 	<ul style="list-style-type: none"> 拥有最顶尖的技术团队，算法技术先进 盈利压力小 压倒性的数据获取优势 	<ul style="list-style-type: none"> 高端人才团队 基于应用场景的算法技术先进
劣势	<ul style="list-style-type: none"> 部门协作成本高，创新力较弱 	<ul style="list-style-type: none"> 由于不急商业化，技术发展缺乏用户数据反馈 	<ul style="list-style-type: none"> 需要较快商业化 渠道、获客能力弱 应用场景数据需合作获取
发展预期	<ul style="list-style-type: none"> 通过并购打通产业链 用多品类构建生态圈 	<ul style="list-style-type: none"> 技术引领 	<ul style="list-style-type: none"> 突破新应用场景算法技术，定制行业解决方案

行业竞争

人脸识别领域竞争激烈，机器人视觉或将成为下一热点，各应用场景探索仍待深入（1/2）

目前，**人脸识别是计算机视觉领域的竞争热点**，金融、安防领域的人脸识别为重点布局场景。而此类公司多有未来涉足机器人视觉、无人机视觉领域的计划。此外，智能家电、医疗领域、图像搜索引擎也有创业公司布局。我们认为，各应用场景的视觉识别技术探索刚刚起步。

国内部分创业公司信息梳理（待续）

公司名称	成立时间	融资信息	核心业务类别	合作企业	未来发展预期
旷视科技 (Face++)	2011	2015年获启明创投、创新工场等多家机构B轮4700万美元投资	人脸识别	蚂蚁金服、万科、美图秀秀、世纪佳缘、中信银行、小米金融、公安机关等	将人脸检测领域做深做透，可能会进军机器人视觉技术领域
商汤科技	2014	2014年获IDG1000万美元融资，2016年Star VC金额未公开	人脸识别	华为、Nvidia、京东、银联、华为、小米、微博、科大讯飞、宜信等	做强人脸识别技术，可能会进军机器人视觉技术领域
格灵深瞳	2013	2014年获红杉资本A轮数千万美元投资	人脸识别	北京新天地、中国农业银行等	做强人脸识别领域，未来可能进军无人车行业
依图科技	2012	2016年获云锋基金B轮数千万美元投资	人脸识别、车辆识别	招商银行、上海浦发银行、贵州省交通厅、江苏省公安厅、厦门市公安厅、福建省公安厅等	做强金融和安防领域的人像和车辆识别
云从科技	2007	目前融到A轮	人脸识别	御银股份、重庆大学、中国银行、新疆西北星公司、海通证券、重庆银行等	做强安防、金融等领域人脸识别，拓展教育、等领域人脸识别
飞搜科技	2015	2016年获天使轮1000万人民币投资	人脸识别、黄暴视频识别	移动互联网，金融，智能硬件集成商等	短期专注于面向政府和企业的人脸识别业务，长期将布局机器人视觉技术领域

注释：此处为国内部分创业公司信息梳理，排名不分先后
来源：公开信息，36氪研究院专家访谈

3.1 产业链
3.2 全景图
3.3 商业模式
3.4 行业竞争

市场格局

创业公司分析

3.5 行业思考
商业模式
技术趋势

行业竞争

人脸识别领域竞争激烈，机器人视觉或将成为下一热点，各应用场景探索仍待深入（2/2）

国内部分创业公司信息梳理（续前页）

公司名称	成立时间	融资信息	核心业务类别	合作企业	未来发展趋势
阅面科技	2015	2016年获A轮融资数千万人民币投资	智能家居、车载设备等领域的嵌入式视觉	智能机器制造商	依托人脸识别、手势识别、场景检测等技术，提供软硬件一体视觉解决方案
地平线机器人	2015	2016年获A+轮融资数千万美元投资	智能家电、智能驾驶领域的机器人大脑系统研发	美的集团、汽车零部件供应商	专注集成深度神经网络算法的OS和芯片研发，重点布局智能家电、机器人、智能驾驶领域
合合信息	2006	目前正在C轮融资	各类证件、表单、文件识别与大数据解析	中国银行、陆金所、太平洋保险、EBAY、顺丰、三星、华为、中兴、小米等	以OCR技术为数据收集切入点，拓宽数据场景，做强各个应用场景的大数据分析
Deepcare	2016	2016年获峰瑞资本天使轮600万人民币投资	医疗影像识别	医院、医疗机构	依托影像智能识别，协助基层医疗机构进行影像诊断
拍医拍	2015	2016年获A轮融资3000万人民币投资	医疗领域OCR识别与人工智能数据解析	华道、达康、百度、慈铭、科研宝等医疗机构	依托深度学习技术，挖掘与理解医疗垂直领域数据，提供多维度技术服务

注释：此处为国内部分创业公司信息梳理，排名不分先后
来源：公开信息，36氪研究院专家访谈

3.1 产业链
3.2 全景图
3.3 商业模式
3.4 行业竞争
市场格局
创业公司分析
3.5 行业思考
商业模式
技术趋势

行业思考 · 商业模式

智能前端化大势所趋，软硬件一体解决方案更具竞争力

• 智能前端化大势所趋：

从需求层面讲，一些应用场景对实时响应是有很高的要求的。然而，当前大多数互联网硬件秉承“轻前端、重云端”的思想，让实时响应难以实现。例如，进行疑犯追踪时，假如在数据处理都在云端进行，那么网速就会成为瓶颈，一两秒的延迟可能会让追踪对象逃离视线。相反的，**前置计算让前端设备成为数据采集设备和数据处理单元的合体，一方面提升了处理的速度，另一方面还可以处理云端难以解决的问题。**

以安防领域为例，对识别速度的需求已催生出智能前端化的安防摄像机产品。该摄像机在机身集成了人脸分析算法，在相机内部即可进行大量运算，实现独立的人脸识别能力，提升监控效率。此外，该相机还能进行曝光调节自适应，避免因光线问题导致的识别难点的产生。随着人工智能技术的推进，更多的智能前端化产品将不断出现。

• 软硬件一体解决方案更具竞争力：

目前，人脸检测技术是计算机视觉领域发展最成熟、创业公司也是最多的一个领域。然而，大部分公司仅仅着力软件，提供简单场景的人脸检测服务。**这类技术壁垒较弱，竞争激烈，进而造成企业议价能力低，难以实现自负盈亏。**

计算机识别技术的引爆点在于具体行业的解决方案，即所提供的产品应解决应用场景的复杂问题，比如基于人脸检测的客流分析，基于机器人或智慧家电的视觉系统。而对于该类解决方案，**单一的软件技术是无法契合用户需求的，高集成化、软硬件一体的解决方案是关键。**

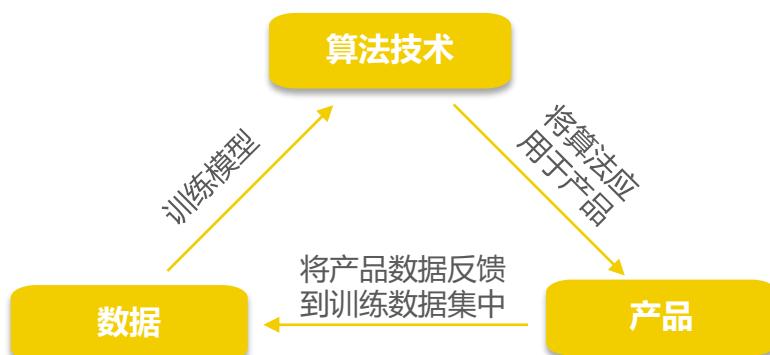
此外，**计算机视觉在消费领域落地的一个障碍是支持高性能运算的低功耗低价位芯片选择太少。**从低功耗，高运算能力的芯片出发，结合先进的算法，这类企业将在计算机视觉领域拥有核心竞争力。

- 3.1 产业链
- 3.2 全景图
- 3.3 商业模式
- 3.4 行业竞争
 - 市场格局
 - 创业公司分析
- 3.5 行业思考**
 - 商业模式
 - 技术趋势

行业思考 · 商业模式

数据是争夺要点，应用场景是着力关键

在人工智能领域，对数据量的重视不能更多。假如把算法比作计算机视觉的发动机，数据量比作燃料，那么基于应用场景的大数据集则是优质燃料。



计算机视觉的研究虽然始于学术界，但最终作为一种商业应用，能解决现实世界中的实际识别问题，才是视觉识别公司的核心竞争力。也就是说，**这里的数据绝不能只是一些标准数据集，应用场景的数据对模型的训练是非常重要的**。一些公司可以在LFW或者FDDB数据集上显示很好的结果，然而实际用于商业环境时精准度却差强人意。

当一家公司先天能够获得大量连续不断的优质场景数据，又有挖掘该数据价值的先进技术时，商业模式、数据模式上就能形成协同效应。

典型的一个案例就是Google。它从搜索引擎开始，积淀了非常多的优质数据，而先进的技术让其成为世界上最理解数据，最会应用数据的AI公司之一。

例如，在美国，医疗影像识别领域所面临的一个问题就是大量数据，即训练数据的图片被Google 和Facebook 垄断，因而小的创业公司在计算机视觉发展上会受限。

我们认为，在未来，**创业公司要么通过自有平台获取数据，要么选择与拥有数据源的大公司进行合作，同时选择一个商业落地的方向，实现快速的数据循环**。

3.1 产业链
3.2 全景图
3.3 商业模式
3.4 行业竞争
市场格局
创业公司分析
3.5 行业思考
商业模式
技术趋势

行业思考 · 技术趋势

知晓机器如何思考为AI重要议题，届时技术将迎来大的飞跃

“在人工智能和机器学习上我们已经取得巨大的进步，但是今天就宣布取得了机器学习的成功就像是「我们爬上了树梢却宣布自己登上了月球」。”

----- (CEO of Allen Institute) Oren Etzioni

目前计算机视觉在识别物体上已经有一定的成效。然而，**尽管对算法的上层逻辑和原理人是清楚的，但具体到某种数据的训练和学习过程，人们是无法预测无法知晓的。**即，机器如何认知仍是一个黑匣子。当我们知晓机器是如何思考时，计算机视觉技术将迎来大的飞跃。

• 深度神经网络很好骗？

计算机视觉的经典方法是运用神经网络，即深度学习的方式。而深度学习就是执行算法---即执行如何完成任务的指令。**由于机器认知事物的规则是被设定的，而这种规则是不能穷举的，因而机器有时会给出非常诡异的，人极少几率会犯的错误结果，比如，把蓝天背景下的白色卡车看成蓝天下的白云（Tesla自动驾驶的首次事故）。**

• 机器如何认知长久以来都被称为“黑匣子”

对于机器是如何思考的，至今对我们来说是个黑匣子。在一个典型的神经网络中，人们能够轻松辨认的层级只有输入层和输出层，而在输入层与输出层之间的，虚拟神经元处理信息并互相连接的层级中，我们无法确认这个系统是如何运行的。因为，“要理解某个具体的、经过训练的神经网络究竟如何工作非常困难，有太多相互作用的、非线性的部分了。”**目前科技水平的发展现状是，我们知道机器的神经网络是和人类不同的，但我们不知道有何相同，也不知道有何不同。**

• 更深入的了解机器认知机理成为下一个重要议题

随着人们对人工智能的依赖程度越大，这种无法了解原理的不安也会越大，因为我们不知道机器何时会犯错，何时不会犯错。如果不明白机器是如何给出一些怪异的结果，就无法避免这种事情的再次发生。因此，了解机器神经网络每一层的机理就成为人工智能的下一个重要议题。因为只有弄明白每一次逻辑推算的方式，才能确保机器的行为具有可预测性。**而届时，人工智能，包括计算机视觉的技术发展将会发生一次大的飞跃。**

分析师声明

作者具有专业胜任能力，保证报告所采用的数据均来自合规渠道，分析逻辑基于作者的职业理解，本报告清晰准确地反映了作者的研究观点，力求独立、客观和公正，结论不受任何第三方的授意或影响，特此声明。

免责声明

36氪不会因为接收人接受本报告而将其视为客户。本报告仅在相关法律许可的情况下发放，并仅为提供信息而发放，概不构成任何广告。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议。在法律许可的情况下，36氪及其所属关联机构可能会持有报告中提到的公司的股权，也可能为这些公司提供或者争取提供筹资或财务顾问等相关服务。

本报告的信息来源于已公开的资料，36氪对该等信息的准确性、完整性或可靠性不作任何保证。本报告所载的资料、意见及推测仅反映36氪于发布本报告当日的判断，本报告所指的公司或投资标的的价格、价值及投资收入可升可跌。过往表现不应作为日后的表现依据。在不同时期，36氪可发出与本报告所载资料、意见及推测不一致的报告。36氪不保证本报告所含信息保持在最新状态。同时，36氪对本报告所含信息可在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。



为创业者提供最好的产品和服务