

## 1.3 深度学习的应用

### 1.3.1 计算机视觉

2015 年深度学习算法的错误率为 4%，已经成功超越了人工标注的错误率（5%），实现了计算机视觉研究领域的一个突破。

在 ImageNet 数据集上，深度学习不仅突破了图像分类的技术瓶颈，同时也突破了物体识别的技术瓶颈。物体识别的难度比图像分类高，需要给出所包含物体的具体位置。而且一张图片可能出现多个需要识别的物体。

在深度学习得到广泛应用之前，基于传统的机器学习技术并不能很好的满足人类人脸识别的精度要求。人脸识别的最大挑战在于不同人脸的差异较小，有时同一个人在不同光照条件、姿态或者表情下脸部的差异甚至会比不同人脸的差异更大。传统的机器学习算法很难抽象出足够有效的特征，使得学习模型既可以区分不同的个体，又可以区分相同个体在不同环境中的变化。深度学习技术通过从海量数据中自动习得更加有效的人脸特征表达，我怀疑是鼻子眼睛距离，大小等尺寸。当然要具体问题具体分析。在人脸识别数据集 LFW 上，基于深度学习算法的系统 DeepID2 可以达到 99.47% 的识别率。以前我有个本科学哥跟我是同一个毕业论文老师，做的 MNIST 手写体识别。

### 1.3.2 语音识别

谷歌、苹果、微软、IBM、百度等在国内外大型 IT 公司提供的语音相关产品，比如我们手机上那个 siri，谷歌的 Google Now、微软的 Xbox 和 Skype 等，都是基于深度学习算法的。

---

Google Now aims to answer a user's question before they need to ask it, by generating real-time information concerning what is important to each individual user. Google Now [mobile app](#) can tell you your local weather as you get dressed in the morning, how much traffic to expect during your commute to work, and when the next train will arrive on the platform you're waiting at, delivering information instantly to manage your day and make life a bit easier.

---

起初语音识别是用混合高斯模型，然后在 2012 年，深度学习的语音识别模式已经取代了混合高斯模型，并成功将谷歌语音识别的错误率降低了 20%。深度学习区别于混合高斯模型，因为混合高斯模型需要人工特征提取，但是深度学习可以自动地从海量数据中提取更加复杂且有效的特征。

### 1.3.3 自然语言处理

单词向量（word embedding）是深度学习解决很多自然语言处理问题的基础。近义词“狗”和“犬”。深度学习算法可以大幅度提高翻译的质量。情感分析最核心的问题是从一段自然语言中判断作者对评价主体是好评还是差评。工业界，用户评价，金融界用户对服务或者产品的态度。

