



LiveVideoStackCon

聚音视 研修不止于形

2017年10月20日-21日

北京.丽亭华苑酒店

增强人类智慧- 利用AI技术实现智能媒体分析

Transforming business with artificial intelligence

Leon Liang 梁健
资深技术顾问
微软大中华区创新技术合作事业部 - CSE
leliang@microsoft.com





推理

Learn and form conclusions
with imperfect data



理解

Interpret meaning of data
including text, voice, images



自然人机交互

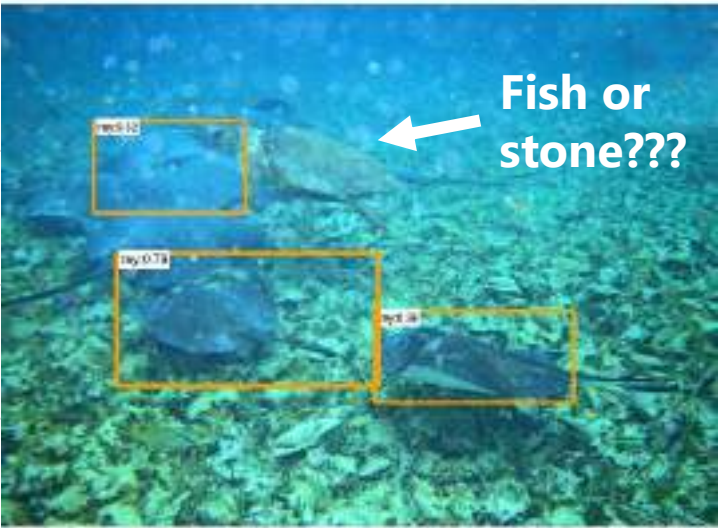
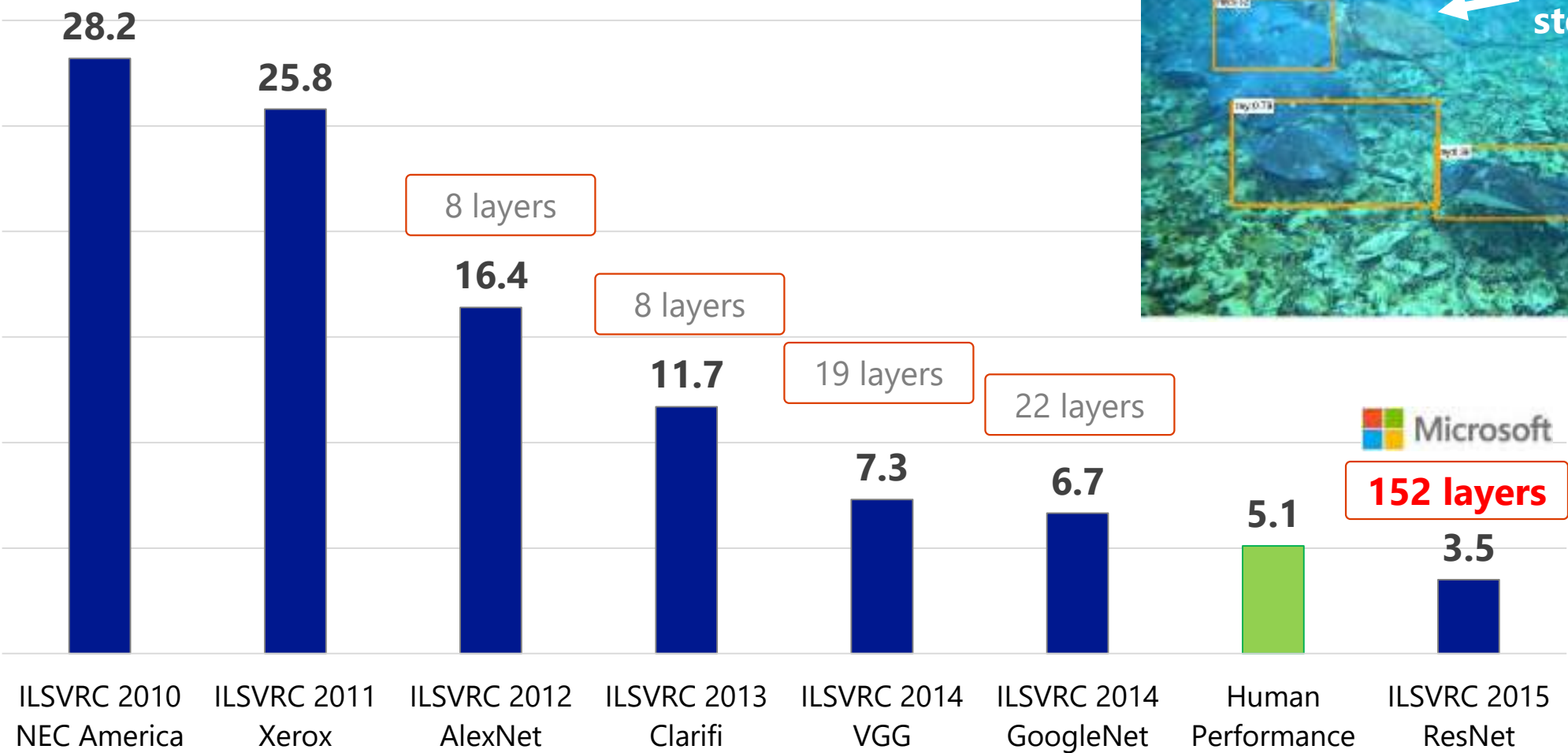
Interact with people
in natural ways

AI 增强人类的智慧

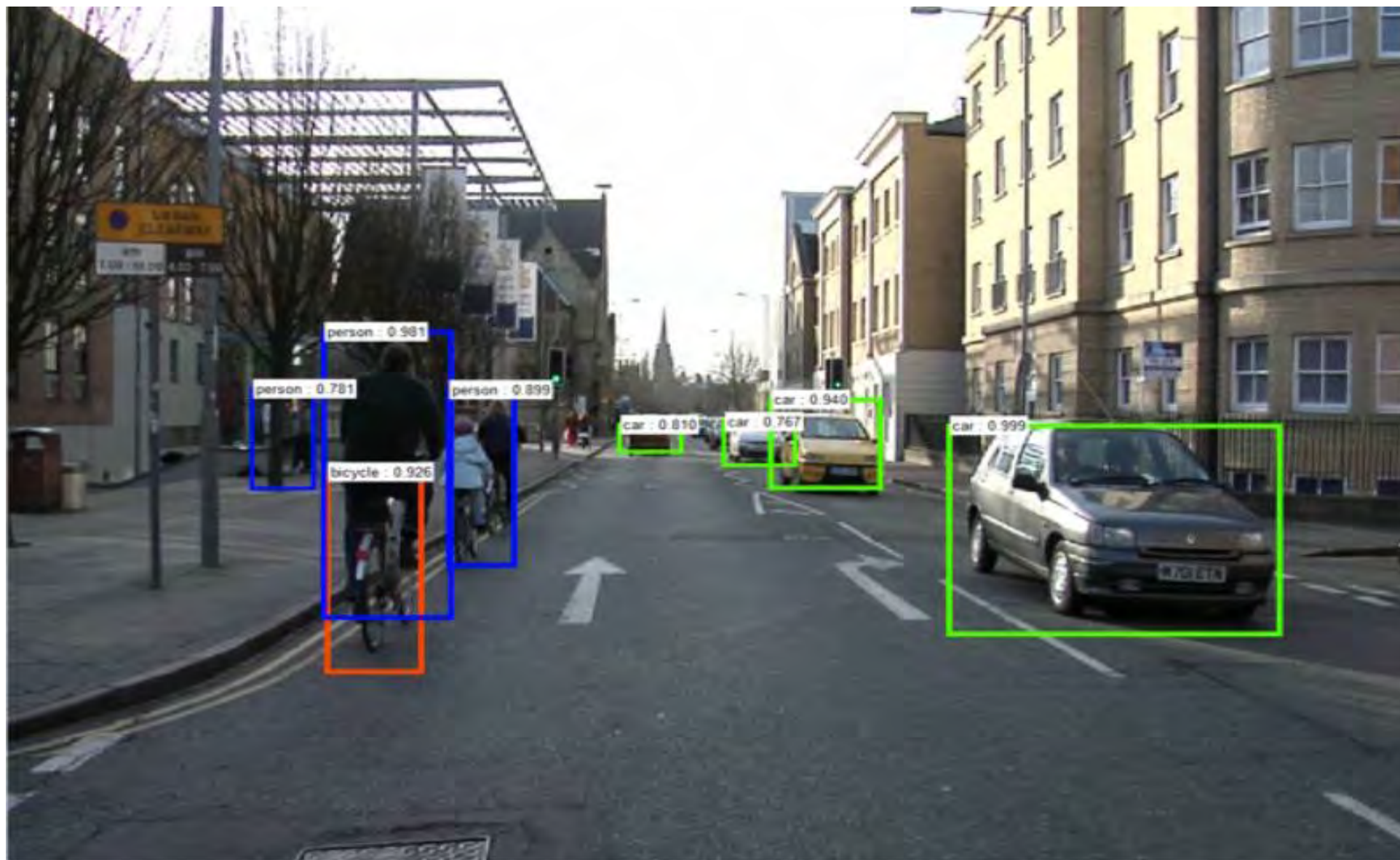
深度学习算法层级革命

ImageNet Large Scale Visual Recognition Challenge (ILSVRC)

ImageNet Classification top-5 error (%)



例子 – 时时视频分析



Media Analytic 在实际生产场景中的应用



在实现更智能服务
的过程中不断突破



Big Data



Powerful
Algorithms



Cloud
Computing

每一个开发者都
可以开发和使用
AI技术



Media AI

您的客户需要什么

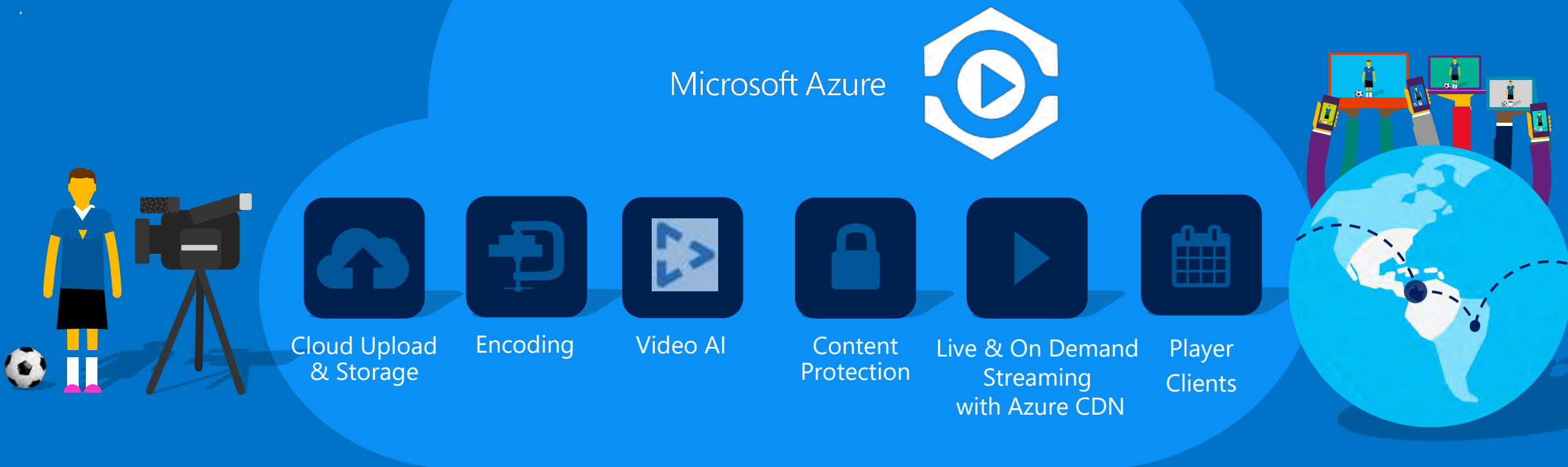
Video AI is the key to solving these challenges....

- 让内容更容易被找到
- 提升内容的价值和吸引力
- 个性化的观看体验
- Uncover Hidden Content Insights
- 降低人工成本
- 创造利润 / 提高收视率
 - 定向的广告投放
 - 预测模型和智能推荐



Azure Media Services (媒体服务)

Azure Media Services is an extensible, multi-tenant platform that enables end-to-end video workflows in the Azure public cloud.



Azure Media Scenarios

Cloud based Video Streaming

Video On-Demand Portals

- Sports
- Entertainment
- Education
- News

Live Broadcast & Events

- Sports Broadcasters
- News Channels
- Music Concerts
- Awards Ceremonies
- Town Halls / Conferences

Digital Marketing / Advertising

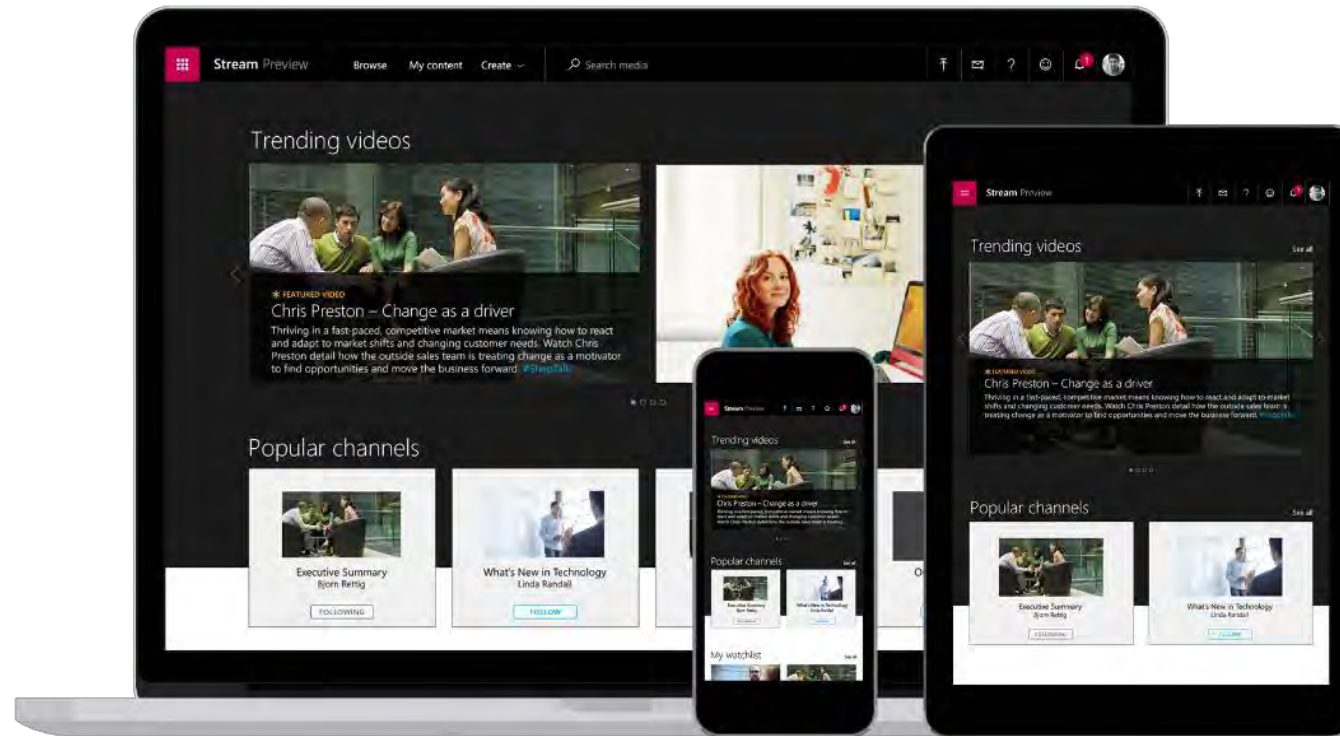
- Retail
- Finance / Banking
- Brand Awareness
- Sales Support

Enterprise / Collaborative

Corporate

CDN Use Cases

- Websites



Microsoft Stream

Industries that can benefit with Video AI...



News



Entertainment



Sports



Retail



Government



Surveillance



Public Safety

More...

...they all have a lot of video content!

微软认知服务

Microsoft

Cognitive Services

Give your apps
a human side



视觉

从面部感官到感觉，让您的对话机器人了解图像、视频和情绪



语音

把语音转换为文本或把文本转换为语音；
了解您的意图，翻译语言，过滤噪音以及识别说话者



语言

教您的对话机器人理解自然语言指令，解析复杂的文本以及了解用户情绪



知识

从网络、学术界或您自己的数据积累中融合丰富的知识



搜索

通过必应API的强大功能访问数十亿网页、图片、视频和新闻



实验室 - Labs

对一些认知服务里的还未正式发布的前沿技术，进行发现，尝试，同时，提供宝贵的反馈

微软认知服务



视觉



语言



语音



搜索



知识图谱

Computer vision

Face

Emotion

Content Moderator

Video

NEW

Video Indexer

NEW

Cognitive Services Labs

Text analytics

Spell check

Web language model

Linguistic analysis

Translator

Speaker recognition

Speech

Web search

Image search

Video search

News search

Autosuggest

Academic knowledge

Entity linking service

Knowledge exploration

Recommendations

QnA maker

NEW

Custom

Vision Service

Custom

Language
Understanding

Custom

Speech Service

NEW

Custom

Search

NEW

Custom

Decision Service



计算机视觉 API

图像分析

基于所输入图像的视觉内容分析出图像的视觉特征

光学字符识别

在图像中检测并识别文本

生成缩略图

生成高品质和高存储效率的缩略图



图片分析



图片类型:

Clip Art Type	0 Non-clipart
Line Drawing Type	0 Non-Line Drawing
Black & White Image	False

图片内容:

Adult Content	False
Adult Score	0.18533889949321747
Faces	[{ "age": 27, "gender": "Male", "faceRectangle": { "left": 472, "top": 258, "width": 199, "height": 199 } }]
Tags	[{ "name": "water", "confidence": 0.9996442794799805 }, { "name": "sport", "confidence": 0.9504992365837097 }, { "name": "swimming", "confidence": 0.9062818288803101, "hint": "sport" }, { "name": "pool", "confidence": 0.8787588477134705 }, { "name": "water sport", "confidence": 0.631849467754364, "hint": "sport" }]

图片色值:

Dominant Color Background	White
Dominant Color Foreground	Grey
Dominant Colors	White
Accent Color	



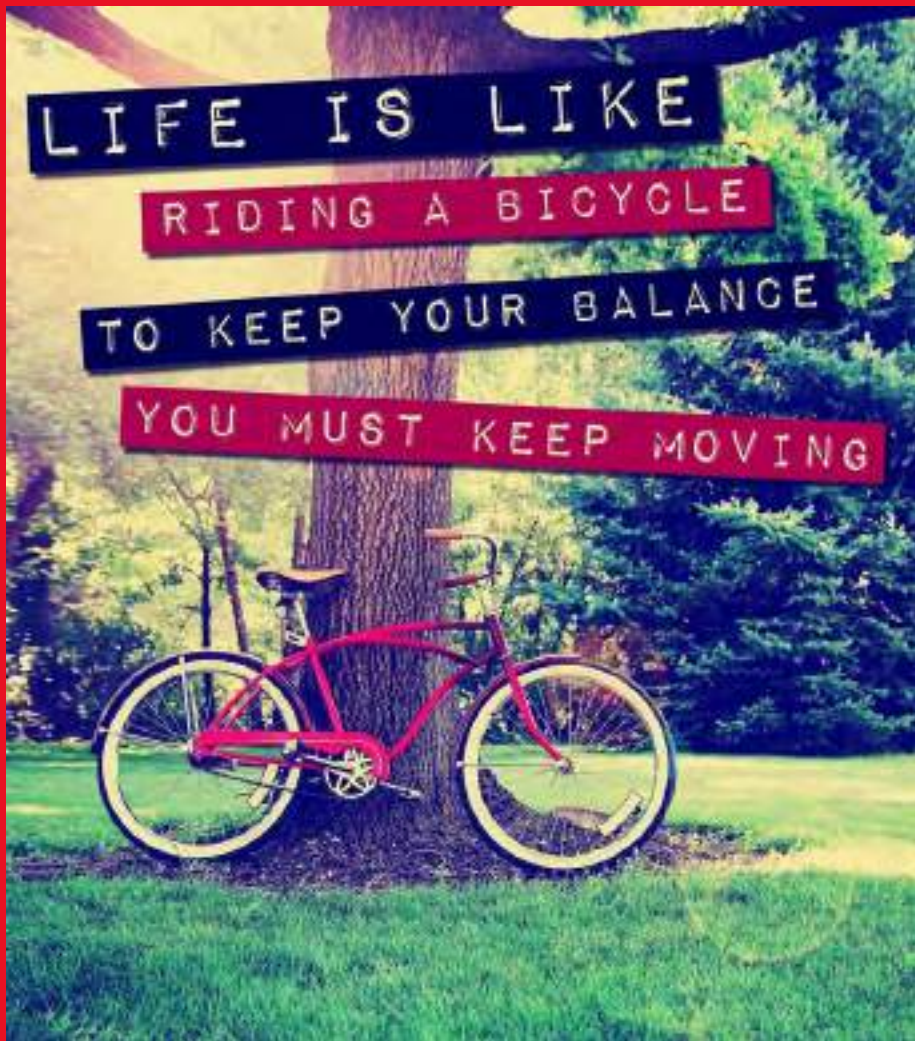
光学字符识别

已支持功能:

- 扫描图片
- 含有文本的图片
- 纹理致密的位置信息

有待增强功能

- 机动车牌照
- 手写文本
- 超大文本识别





智能缩略图



开启智能缩略功能：



Microsoft Azure

[为什么选择 Azure?](#) [解决方案](#) [产品](#) [文档](#) [价格](#) [培训](#) [应用商店](#) [合作伙伴](#) [博客](#) [资源](#) [支持](#)

免费帐户



Welcome to the Intelligent Kiosk! Here you will find several demos showcasing the Perceptual Intelligence power of Azure Machine Learning, the Cortana Intelligence Suite and the Microsoft Cognitive Services.

The kiosk is an example of an intelligent application, one that can understand human faces and predict their age, gender and emotion. It can do so from live images from a web cam, Bing Images or local photos. It also allows:

Microsoft Azure

Through

[为什么选择 Azure?](#) [解决方案](#) [产品](#) [文档](#) [价格](#) [培训](#) [应用商店](#) [合作伙伴](#) [博客](#) [资源](#) [支持](#)

免费帐户



Want

Find it

[https://](#)

Want

Download

[https://](#)

Need

Contact

For more

[Azure](#)

[Cortana](#)

[Cortana](#)

[Microsoft](#)

[Microsoft](#)

[Microsoft](#)

[Microsoft](#)

[Microsoft](#)

[Microsoft](#)

[Microsoft](#)

[Microsoft](#)

[Microsoft](#)

[Microsoft](#)

[Microsoft](#)

分析视频，在视频中使用高精度面部位置检测，可检测最多 64 张面孔。

在实际操作中查看



表

JSON

时间 (秒)	面部 ID	X、Y	宽度、高度
0	0	0.59, 0.23	0.09, 0.16
0	1	0.38, 0.15	0.07, 0.12
1	0	0.54, 0.25	0.09, 0.15
1	1	0.23, 0.18	0.07, 0.12



人脸识别 API

人脸检测

检测图片中的人脸并识别丰富的属性，**同时支持64张脸**

人脸验证

核对两张人脸是否属于同一个人

相似人脸搜索

从多张人脸中找出与所查人脸相似的人脸

人脸分组

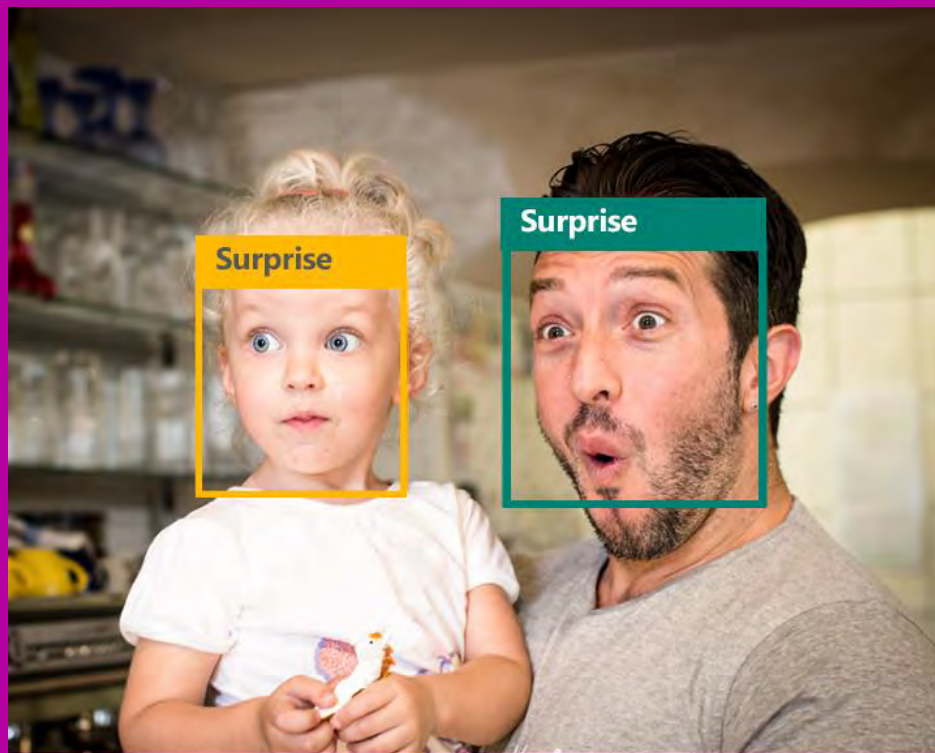
根据识别的人脸进行分组 **Max 64 组，每组10K person，248 训练图片/Person**

人脸辨识

查询某张人脸，与提供的数据中的何人相匹配



情感识别 API



人脸检测

```
"faceRectangle": {"width": 193, "height": 193, "left": 326, "top": 204}
```

...

情感指数

```
"scores": { "anger": 5.182241e-8,  
            "contempt": 0.0000242813,  
            "disgust": 5.621025e-7,  
            "fear": 0.00115027453,  
            "happiness": 1.06114619e-8,  
            "neutral": 0.003540177,  
            "sadness": 9.30888746e-7,  
            "surprise": 0.9952837}
```