

# 法律声明

---

□ 本课件包括：演示文稿，示例，代码，题库，视频和声音等，小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

□ 课程详情请咨询

■ 微信公众号：小象

■ 新浪微博：ChinaHadoop



# 第1课 课程介绍

---

计算机视觉

Computer Vision

# 本章结构

---

- 什么是计算机视觉
- 主要研究问题
- 研究挑战
- 开源库介绍
- 环境搭建

# 什么是计算机视觉

---

研究理论和应用：

- 让计算机具有人类视觉的所有功能
- 让计算机从图像数据中，提取有用的信息，并解释
- .....

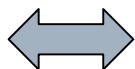
# 什么是计算机视觉

图像数据  
(图片、视频、深度图片等)



解释

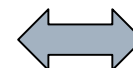
方向盘  
特斯拉  
地图  
无人驾驶  
.....



感知设备  
(摄像头)



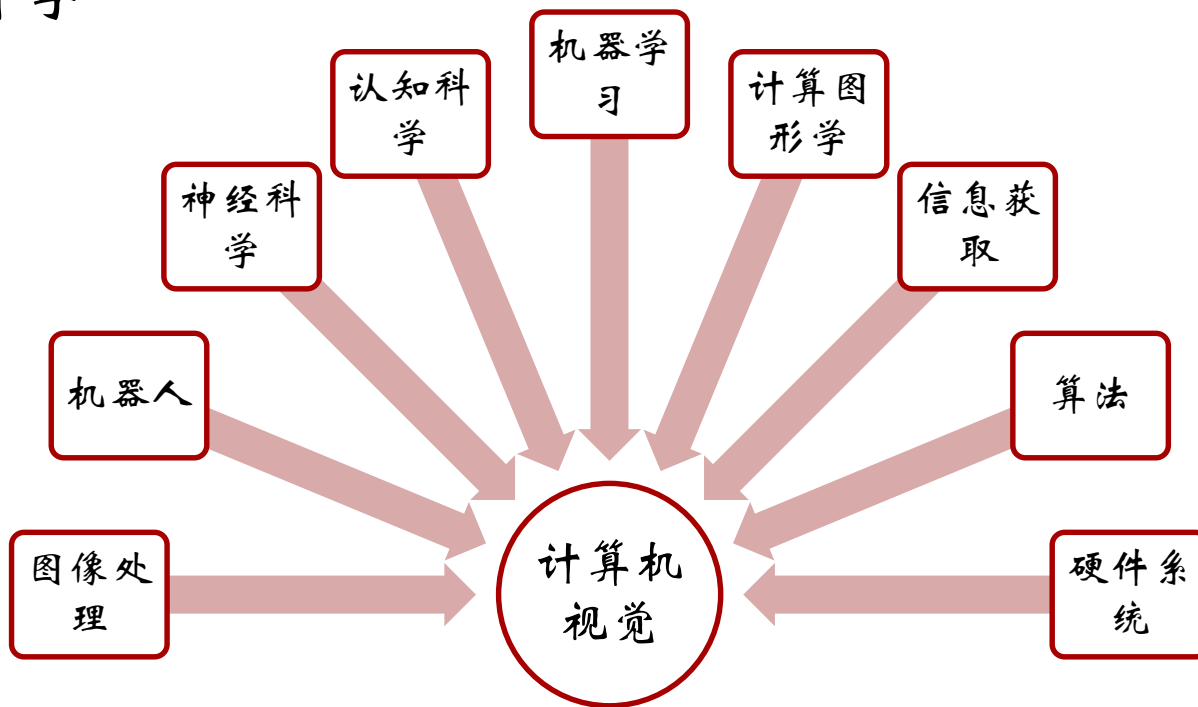
计算设备  
(CPU/GPU/FPGA)



# 什么是计算机视觉

## 高度复合学科

- 工程
- 计算机科学
- 数学
- 心理学
- 生物学
- ...

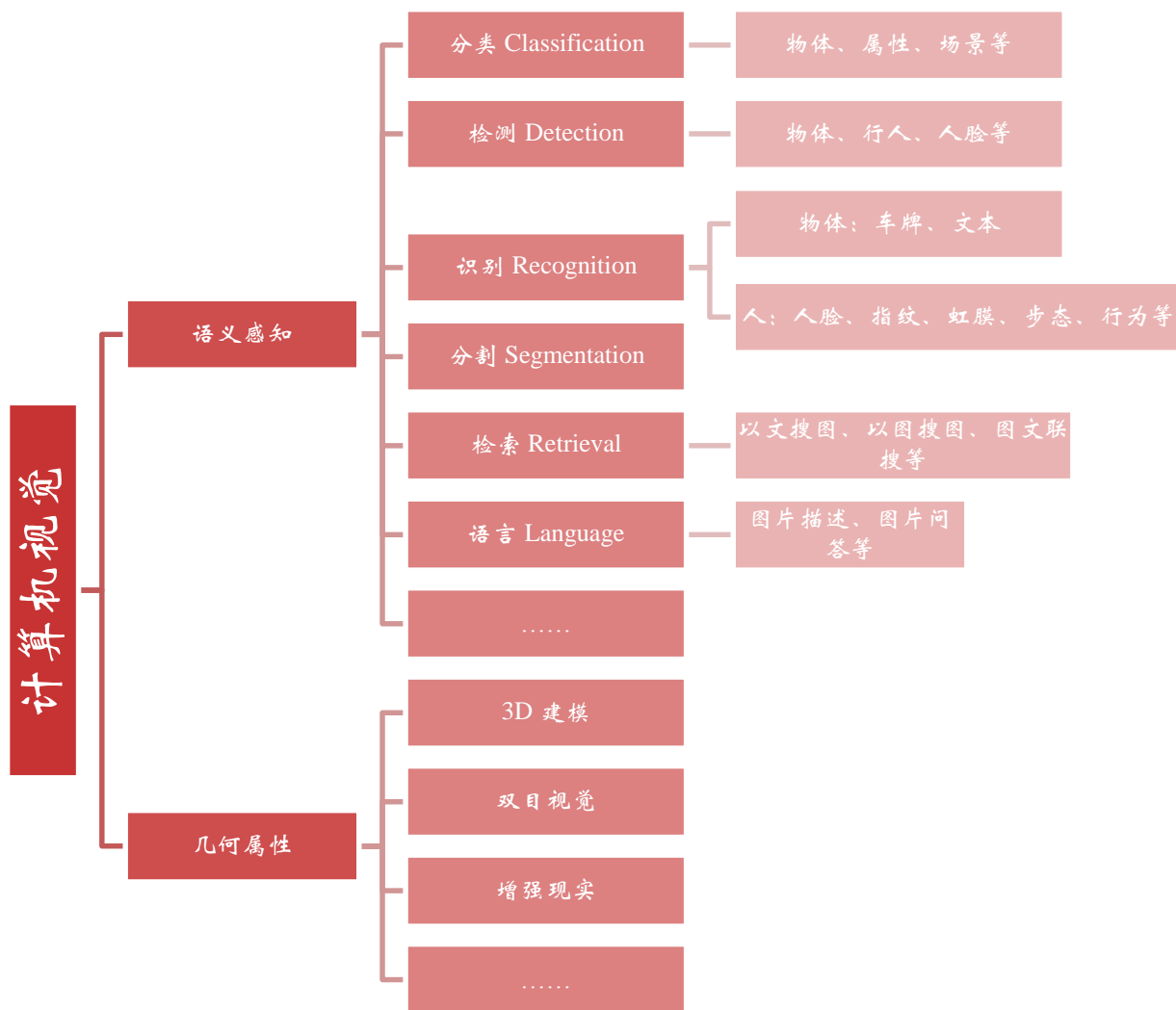


# 什么是计算机视觉

---

- 2个主要研究维度
  - 语义感知 (semantic)
  - 几何属性 (Geometry)
- 赶上人类视觉了吗
  - 迅猛发展 (特定领域已赶上)
  - 任重道远 ( not yet )

# 什么是计算机视觉



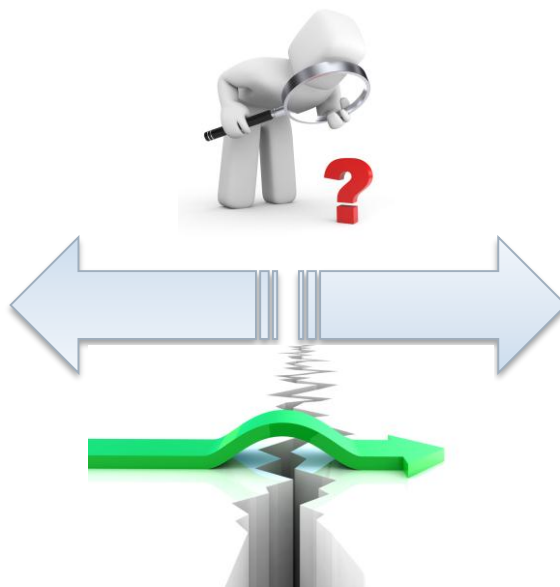


# 什么是计算机视觉

- 人工智能目标
  - 解决”像素值”与”语义”之间的差距 (Gap)



人眼所看到的



|   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|
| 0 | 3 | 2 | 5 | 4 | 7 | 6 | 9 | 8 |
| 3 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2 | 1 | 0 | 3 | 2 | 5 | 4 | 7 | 6 |
| 5 | 2 | 3 | 0 | 1 | 2 | 3 | 4 | 5 |
| 4 | 3 | 2 | 1 | 0 | 3 | 2 | 5 | 4 |
| 7 | 4 | 5 | 2 | 3 | 0 | 1 | 2 | 3 |
| 6 | 5 | 4 | 3 | 2 | 1 | 0 | 3 | 2 |
| 9 | 6 | 7 | 4 | 5 | 2 | 3 | 0 | 1 |
| 8 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 |

计算机所看到的

# 主要研究问题

- 图片分类→有没有？

Input Image

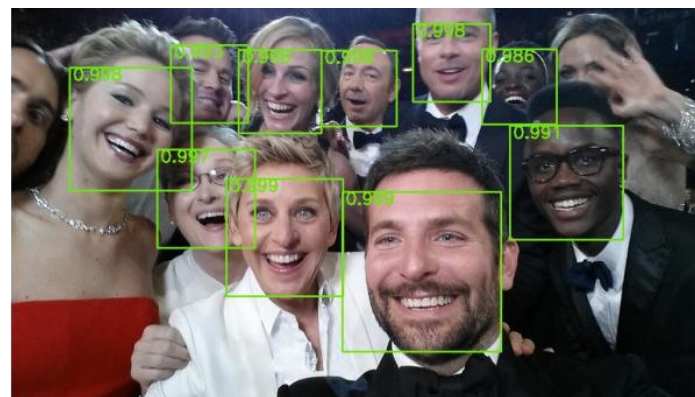
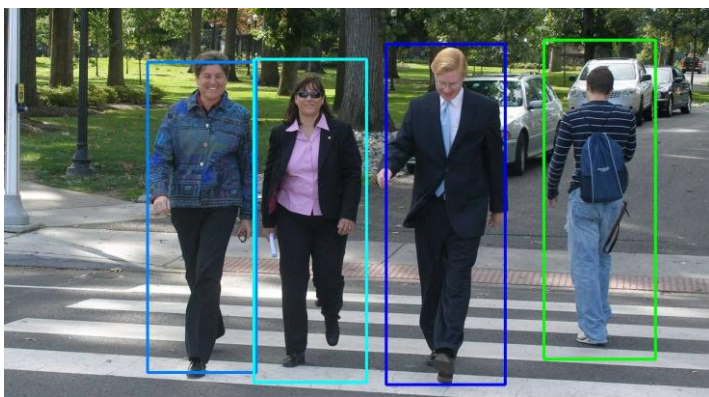
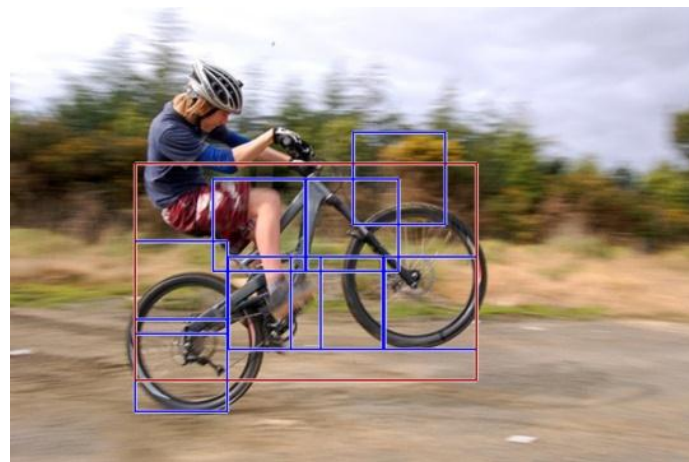


Thank you. After analysis, our system describes the image as follows:

|                      |         |
|----------------------|---------|
| butterfly            | 3.49354 |
| lepidopterous insect | 3.42100 |
| insect               | 2.50084 |
| arthropod            | 2.01440 |
| nymphalid            | 1.97341 |
| invertebrate         | 1.73042 |
| lycaenid             | 0.40975 |
| admiral              | 0.38683 |

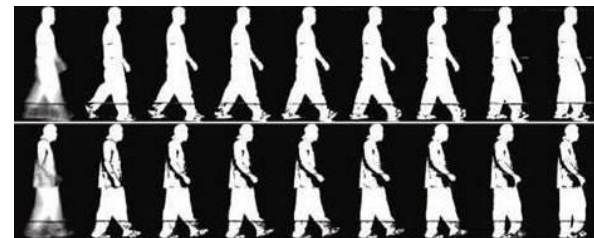
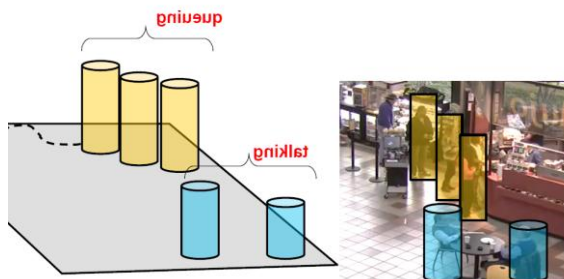
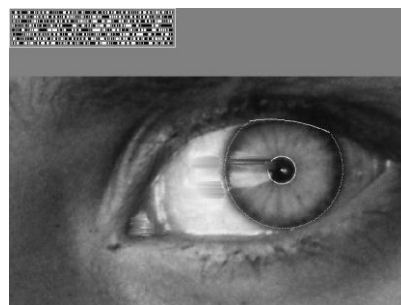
# 主要研究问题

- 检测 → 目标在哪儿?
  - 物体检测
  - 人脸检测
  - 行人检测



# 主要研究问题

- 识别 → 内容是什么？
  - 人脸识别
  - 字符识别
  - 车牌识别
  - 行为识别
  - 步态识别



# 主要研究问题

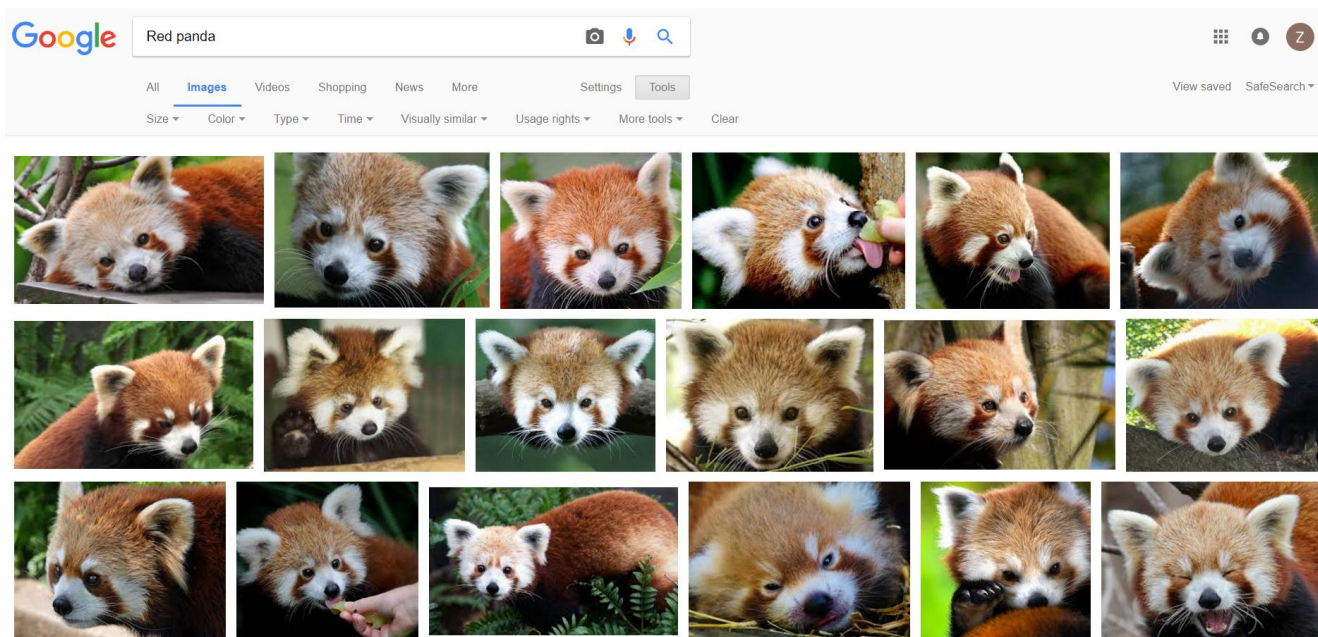
- 图片分割 → 区域轮廓？
  - 前景分割
  - 语义分割
  - .....





# 主要研究问题

- 图片检索→相似图片



# 主要研究问题



- 视觉&语言
  - 图片描述 (Captioning)
  - 图片问答 (Question Answering)
- Caption
  - 一个人在用自动驾驶系统开车
- QA
  - Q: 这个人为什么双手离开了方向盘
  - A: 因为他启动了无人驾驶功能

# 研究挑战

---

- 视角变化

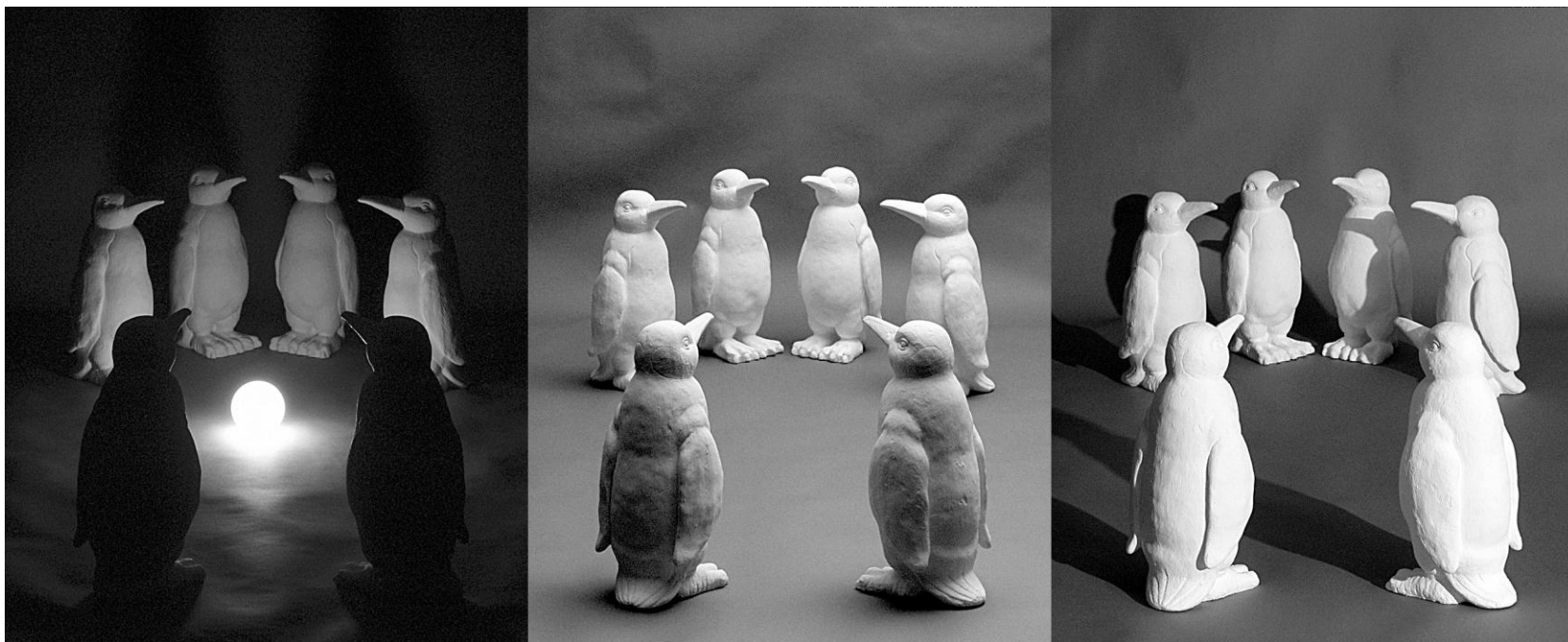




# 研究挑战

---

- 光照变化



# 研究挑战

---

- 尺度变化



# 研究挑战

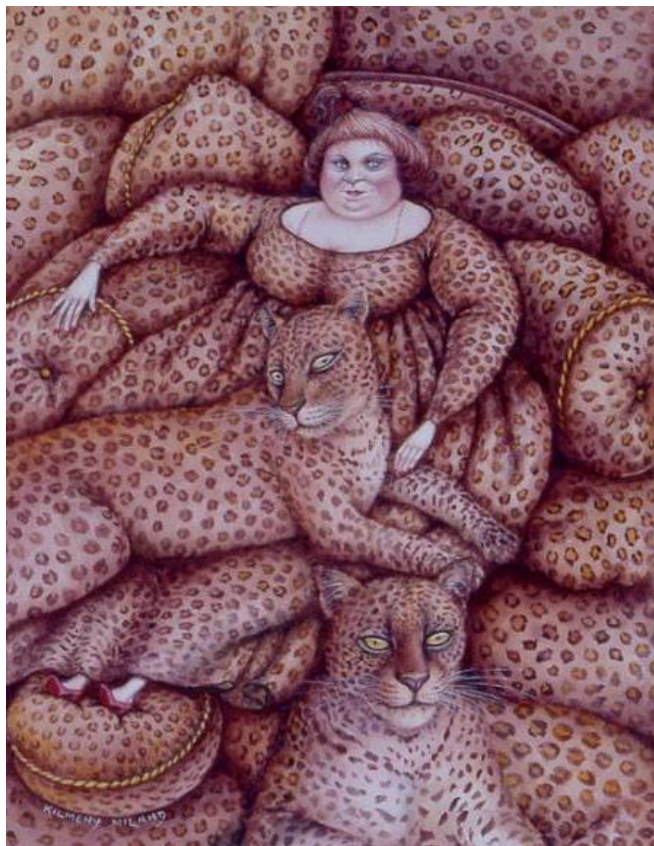
---

- 形态变化



# 研究挑战

- 背景混淆干扰

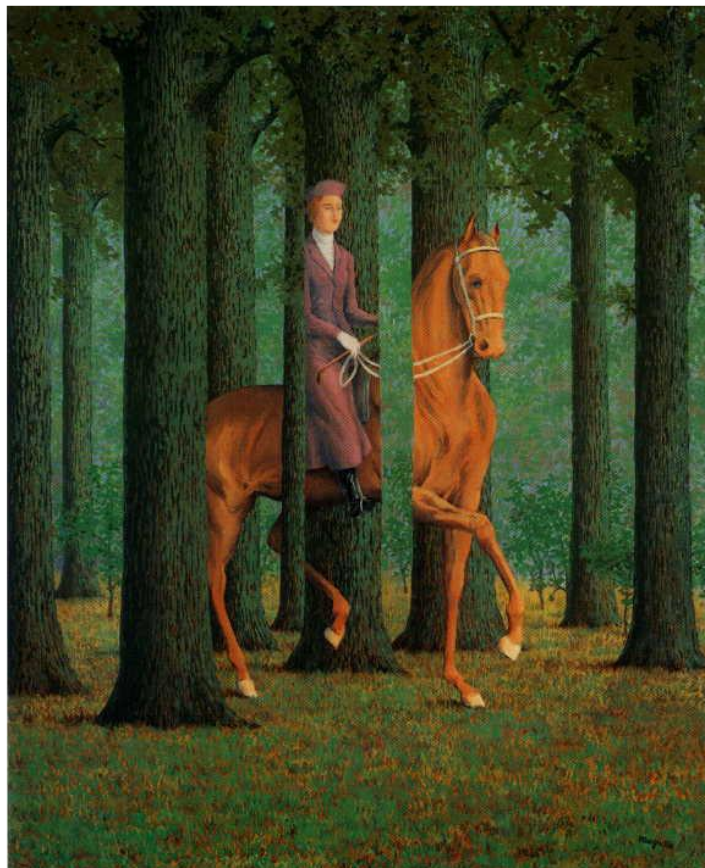




# 研究挑战

---

- 遮挡



# 研究挑战

---

- 类内物体的外观差异



# 开源库介绍

---

- OpenCV
  - 函数库：计算机视觉，机器学习
  - 优化算法：2500+
  - 编程语言：C, C++, Java, Python, Matlab
  - 系统：Windows, Linus, Mac OS, iOS, Android
  - 并行计算：CUDA, OpenCL
  - URL: <http://opencv.org/>

# 开源库介绍

---

- Caffe
  - 函数库：深度学习
  - 发布方：Berkeley Vision and Learning Center (BVLC)
  - 编程语言：C++, Python
  - 系统：Ubuntu, Red Hat, OS X , Windows
  - 并行计算：CUDA
  - URL: <http://caffe.berkeleyvision.org>



# 开源库介绍

---

- Theano
  - 函数库：深度学习
  - 发布方：LISA Lab, 蒙特利尔工程学院
  - 编程语言：Lua(Python)
  - 系统：Ubuntu, Red Hat, OS X, Windows
  - 并行计算：CUDA
  - URL: <http://deeplearning.net/software/theano/>

# 开源库介绍

---

- Torch(PyTorch)
  - 函数库：深度学习
  - 发布方：Facebook
  - 编程语言：Lua(Python)
  - 系统：Ubuntu, Red Hat, OS X
  - 并行计算：CUDA
  - URL: <https://torch.ch>(<http://pytorch.org/>)

# 开源库介绍

---

- Tensorflow
  - 函数库：深度学习
  - 发布方：Google
  - 编程语言：Python
  - 系统：Ubuntu, Red Hat, OS X
  - 并行计算：CUDA
  - URL: <https://www.tensorflow.org/>
  - 特色：
    - 多机多GPU分布式
    - Tensorboard

# 深度学习库对比

| 深度框架              | 编程语言<br>/灵活性             | 主要应用场景                          | 维护方           | Pre-trained<br>模型          | 多GPU<br>模型并行 | 多机<br>分布式  |
|-------------------|--------------------------|---------------------------------|---------------|----------------------------|--------------|------------|
| (Py)Torch         | Lua(Python)<br>/Yes +    | CNN/RNN(中等)/...<br>通用           | Facebook      | Yes ++                     | Yes          | No         |
| Theano            | Python<br>/Yes           | CNN/RNN(支持)/...<br>通用           | LISA Lab      | Yes<br>(Lasagne)           | Tesing       | No         |
| Caffe             | Python&C++<br>/No        | CNN<br>图像分类、检测、分割               | BVLC          | Yes ++                     | No           | No         |
| <b>TensorFlow</b> | <b>Python<br/>/Yes +</b> | <b>CNN/RNN(Best)/...<br/>通用</b> | <b>Google</b> | <b>Yes<br/>(Inception)</b> | <b>Best</b>  | <b>Yes</b> |

# 课程应用案例环境

---

- 编程语言
  - Python
- 函数库/计算框架
  - OpenCV
  - Tensorflow
  - Lire (Lucence Image Retrieval)

# 环境搭建opencv

---

- 官方教程
  - [http://docs.opencv.org/3.2.0/d6/d00/tutorial\\_py\\_roo\\_t.html](http://docs.opencv.org/3.2.0/d6/d00/tutorial_py_roo_t.html)
- Windows安装教程
  - [http://docs.opencv.org/3.2.0/d5/de5/tutorial\\_py\\_set\\_up\\_in\\_windows.html](http://docs.opencv.org/3.2.0/d5/de5/tutorial_py_set_up_in_windows.html)
- 开始演示

# 本章小结

---

- 计算机视觉&研究问题
  - 2D图片人工智能
  - 语言交互
- 开源框架
  - OpenCV
  - Tensorflow

# 疑问

---

□ 问题答疑：<http://www.xxwenda.com/>

■ 可邀请老师或者其他人回复问题

Q & A



# 联系我们

---

## 小象学院：互联网新技术在线教育领航者

- 微信公众号：小象
- 新浪微博：ChinaHadoop

