



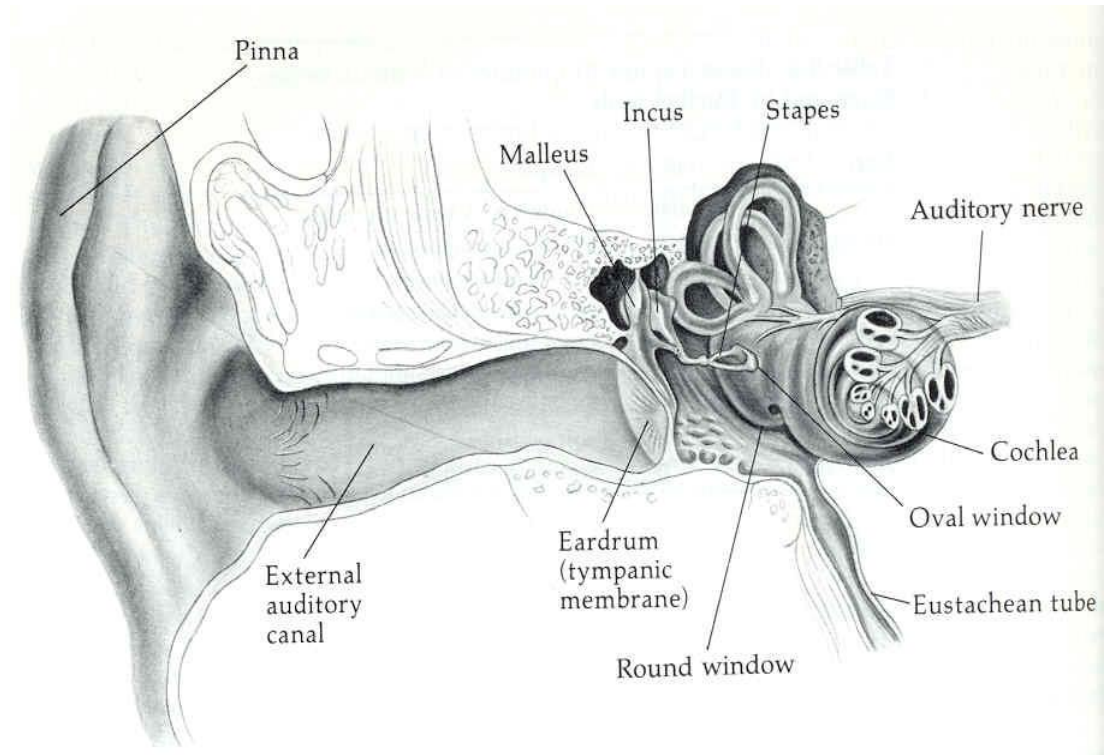
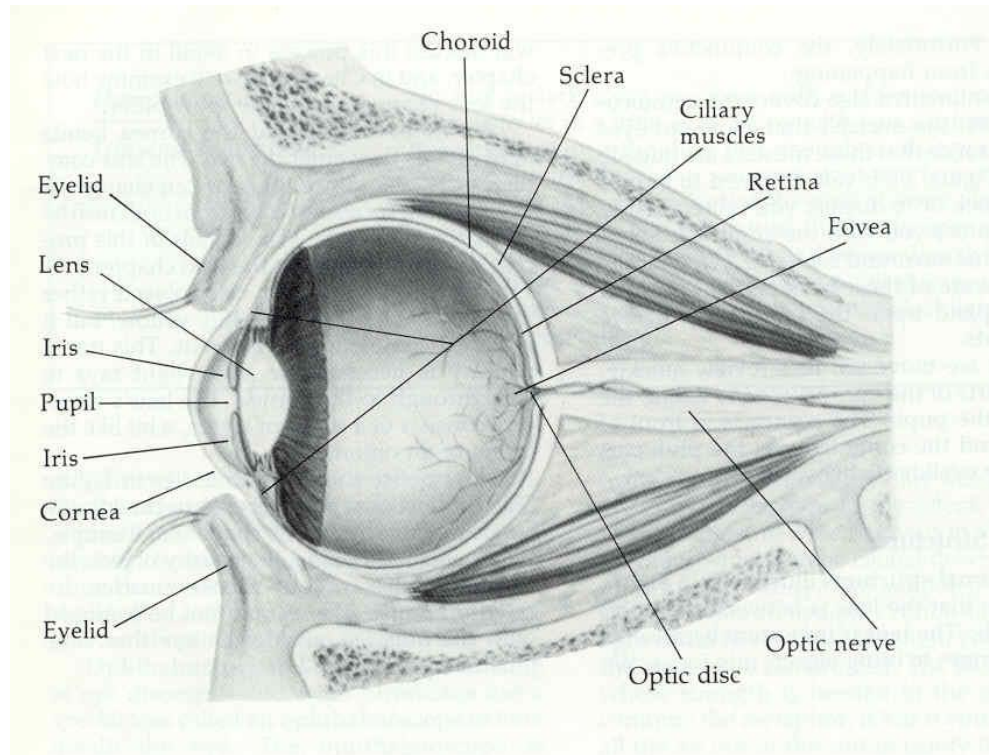
计算机视觉

Computer Vision

基础工业训练中心
清华大学

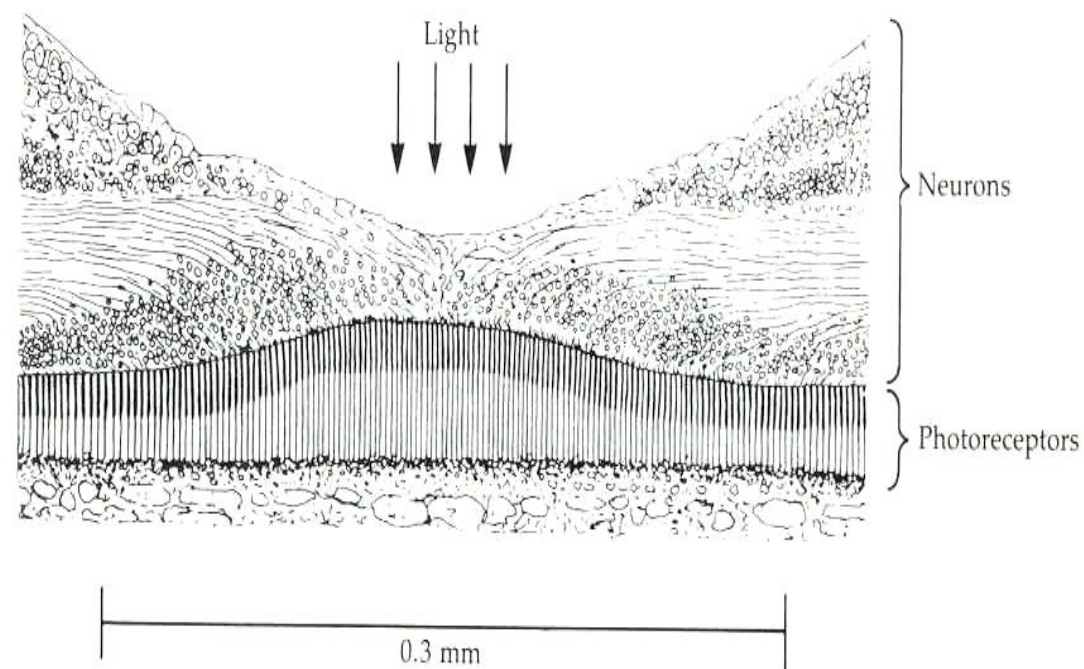
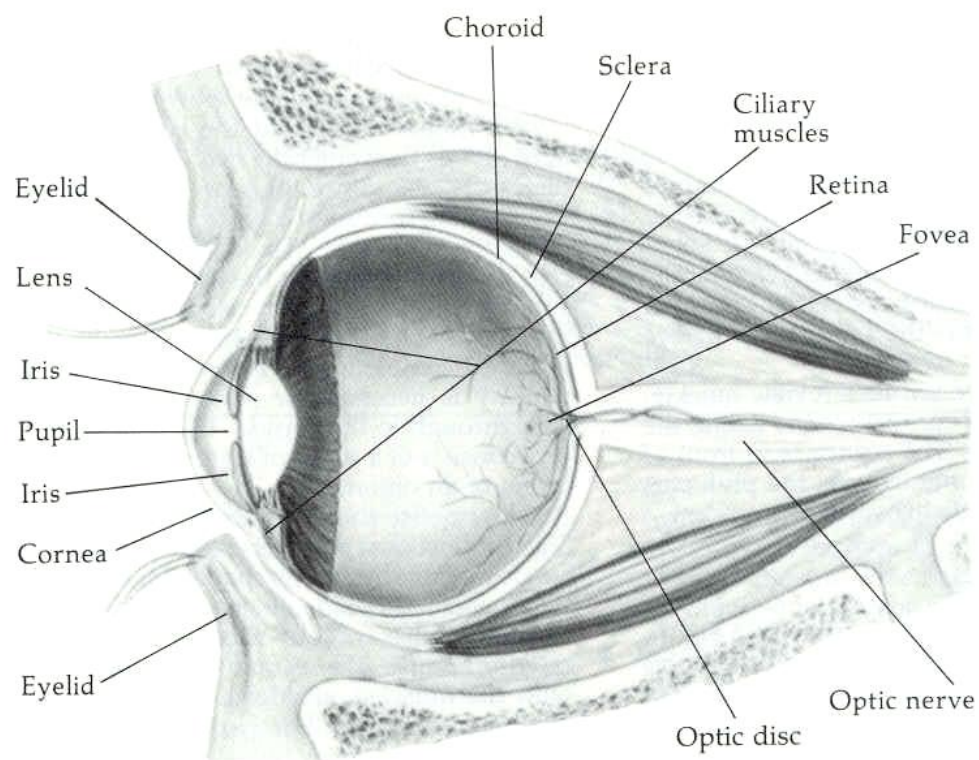
人类感知-视觉/听觉

- Human Perception: audiovisual information processing



人脑视觉处理系统

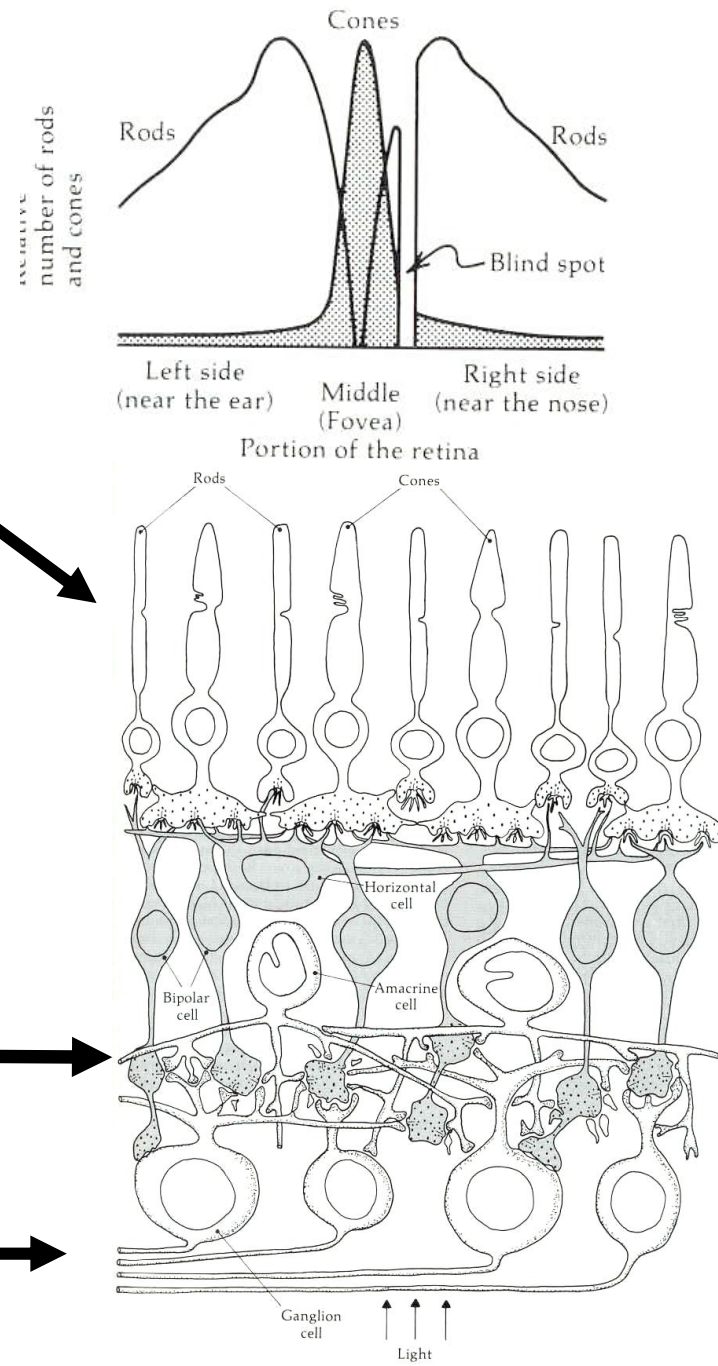
人眼视觉系统的处理从视网膜开始，有色彩处理和黑白光采集和处理系统



人脑视觉处理系统

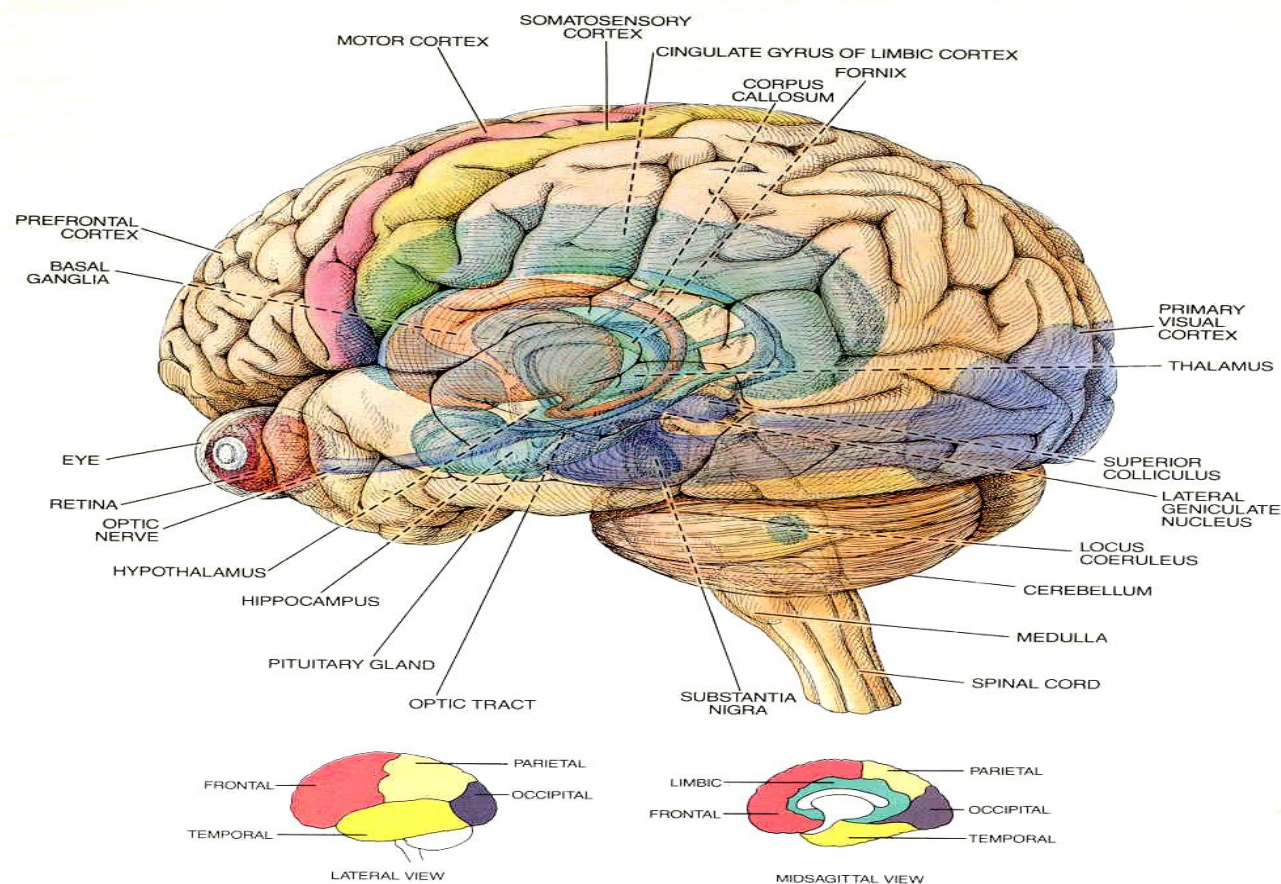
- 视杆细胞 (Rod cells) 负责明暗视觉, 视锥细胞 (Cone cells) 负责彩色视觉。
- 在非常低的光照水平下, 视觉是暗视觉 (Scotopic vision) ——光由视网膜上的视杆细胞检测。
- 在更明亮的光下, 比如白天, 视觉则是亮视觉 (Photopic vision) ——光由负责彩色视觉的视锥细胞检测。
- 色彩处理元素的数量比黑白色少很多 (1亿2千5百万: 6~7百万)。
- 从彩色和黑白受体的信号将进入视网膜神经节细胞。

输出连接



人类大脑的视觉通道

- 人的70%的信息来自于视觉
- 大脑皮层的不同区域
(Cortex)



The Brain: Organ of the Mind

计算机视觉

- 计算机视觉是什么？
- “What does it mean, to see? The plain man s' answer (and Aristotle's, too) would be, to **know what is where by looking.**” ---David Marr, Vision (1982)
- 计算机视觉定义
- Computer vision is the science (some say art) of programming a computer to process, and ultimately understand, images and video.

计算机视觉任务

- 分类Classification

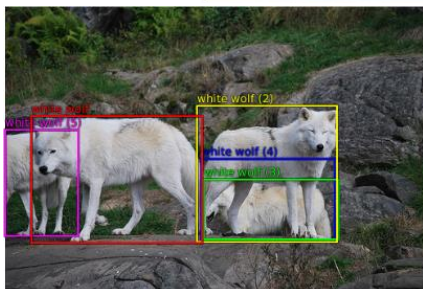


Top 5:
pencil sharpener
pool table
hand blower
oil filter
packet

Groundtruth:
pencil sharpener

ILSVRC2012_val_00010000.JPEG

- 定位localization



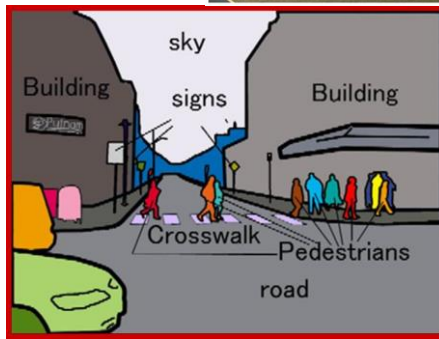
Groundtruth:
white wolf
white wolf (2)
white wolf (3)
white wolf (4)
white wolf (5)

- 检测detection



Groundtruth:
tv or monitor
tv or monitor (2)
tv or monitor (3)
person
remote control
remote control (2)

- 分区segmentation
- 场景解析与标记



困难

计算机视觉

- 计算机视觉Computer vision
 - 从图像中自动抽取信息
 - 模拟人的视觉行为。
- 机器人视觉Robot vision

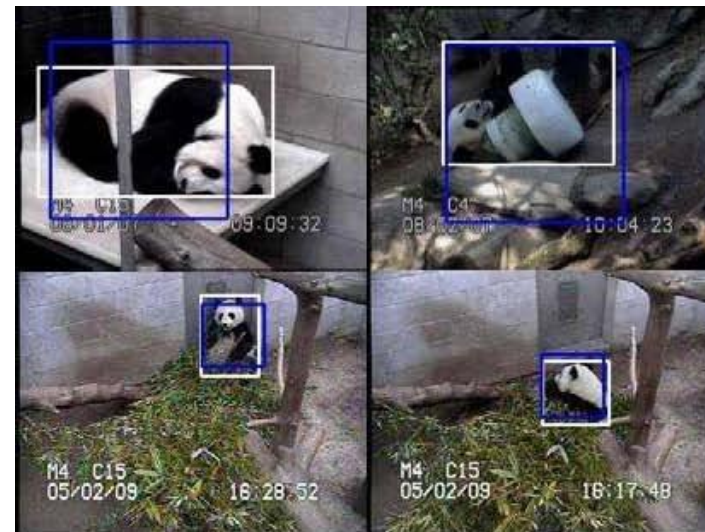
图像



计算机视觉的重要性

- 计算机视觉是实现人工智能的一个重要的挑战

- 机器人
- Robotics
- 智能监控
- Surveillance



OpenCV



- 开源的计算机视觉库， 目前版本为3.3
- <http://opencv.org/>
- 采用BSD协议， 跨平台Linux/Windows
- 应用广泛： 多点触摸， 摄像头互动， 体感交互， 人脸识别

Solem, Jan Erik. *Programming Computer Vision with Python: Tools and algorithms for analyzing images*. " O'Reilly Media, Inc.", 2012.

Howse, Joseph. *OpenCV Computer Vision with Python*. Packt Publishing Ltd, 2013.

OpenCV Library

- 最初由Intel开发，用于单核的 single-core x86 CPUs，目前应用广泛

IMAGE PROCESSING



General Image Segmentation
Processing



Machine
Learning,
Detection

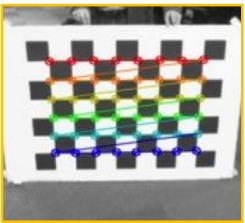


Image Pyramids Transforms

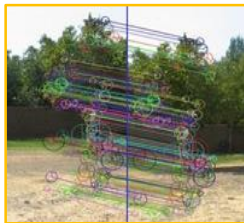


Fitting

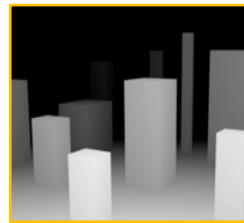
VIDEO, STEREO, AND 3D



Camera
Calibration



Features



Depth
Maps



Optical
Flow



Inpainting



Tracking

Jetson TX1 OpenCV4Tegra — Accelerated CUDA+NEON+GLSL+TBB multithreading

OPENCV示例-安装opencv与使用 (x86机器)

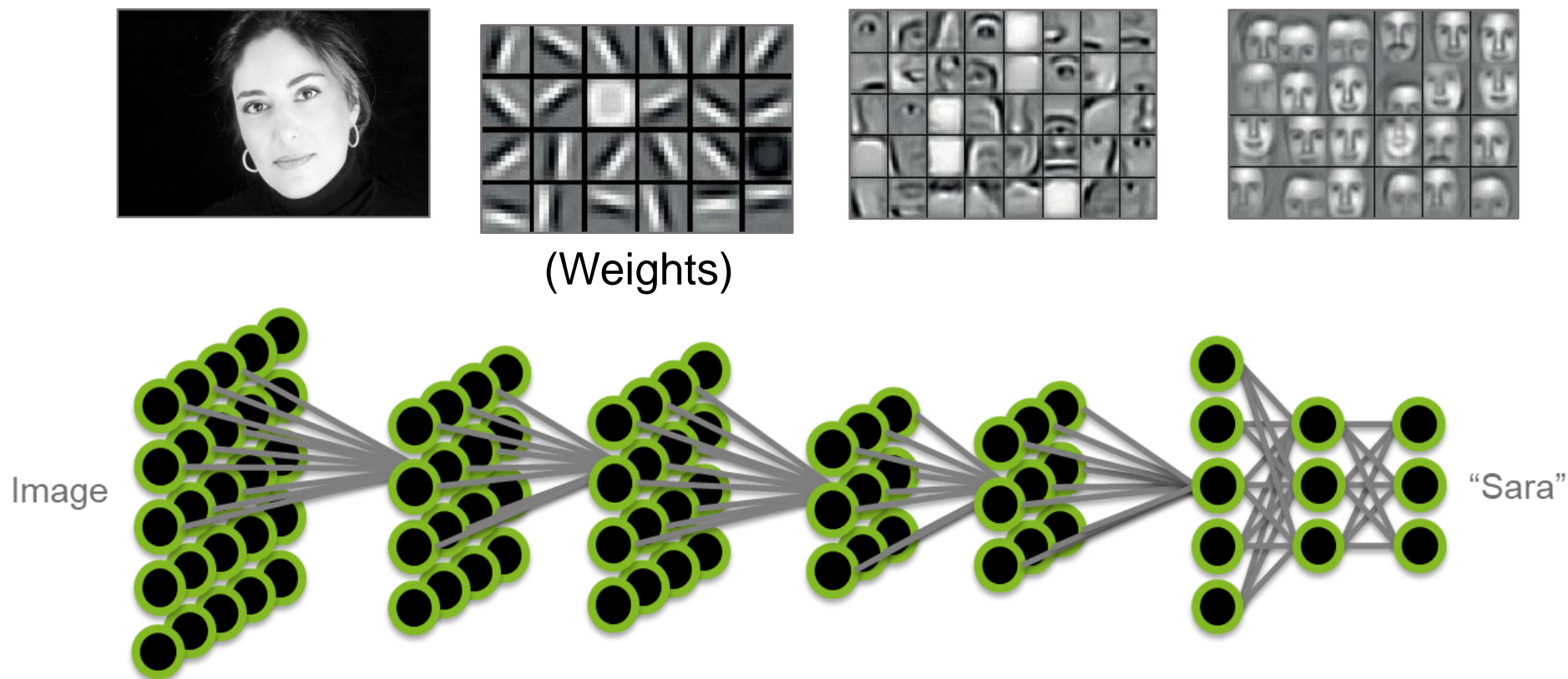
- 安装Anaconda 4.4
- TUNA: <https://mirrors.tuna.tsinghua.edu.cn>
- 安装 opencv 3.3
- `conda install -c conda-forge opencv`

- Windows 10平台为例。使用摄像头捕获视频并显示。

```
import cv2
import numpy
import matplotlib.pyplot as plot
imcap = cv2.VideoCapture(0)
while(1):
    # get a frame
    ret, frame = imcap.read()
    # show a frame
    cv2.imshow("capture", frame)
    if cv2.waitKey(1) & 0xFF == ord('q'):
        break
imcap.release()
cv2.destroyAllWindows()
```

深度神经网络 (Deep Neural Networks)

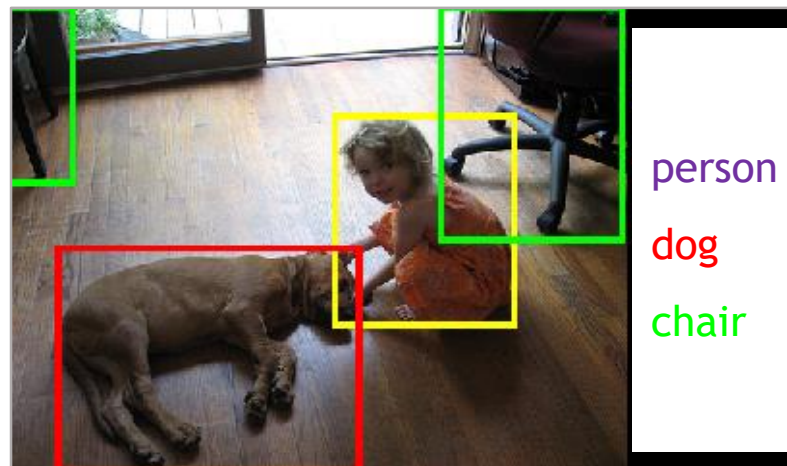
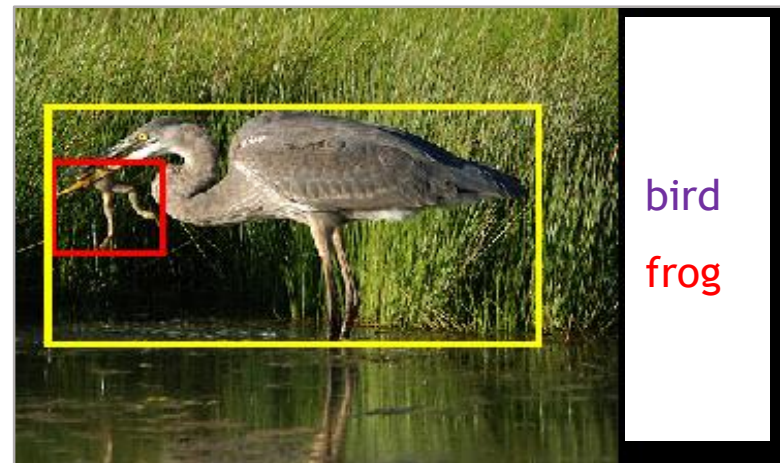
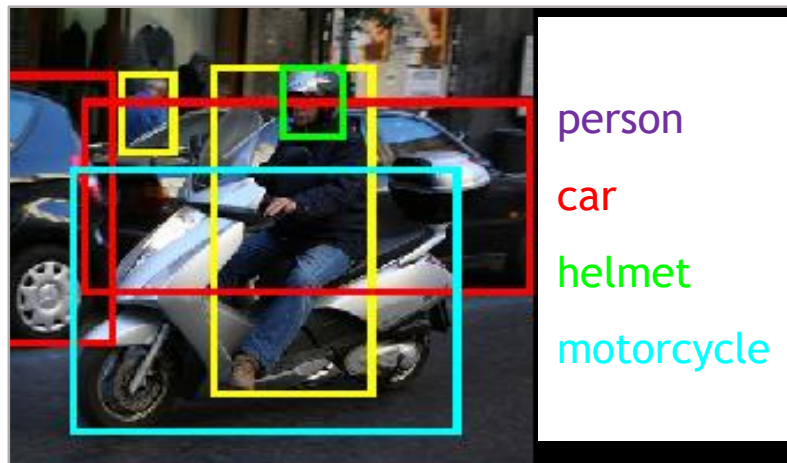
- Learning by abstraction



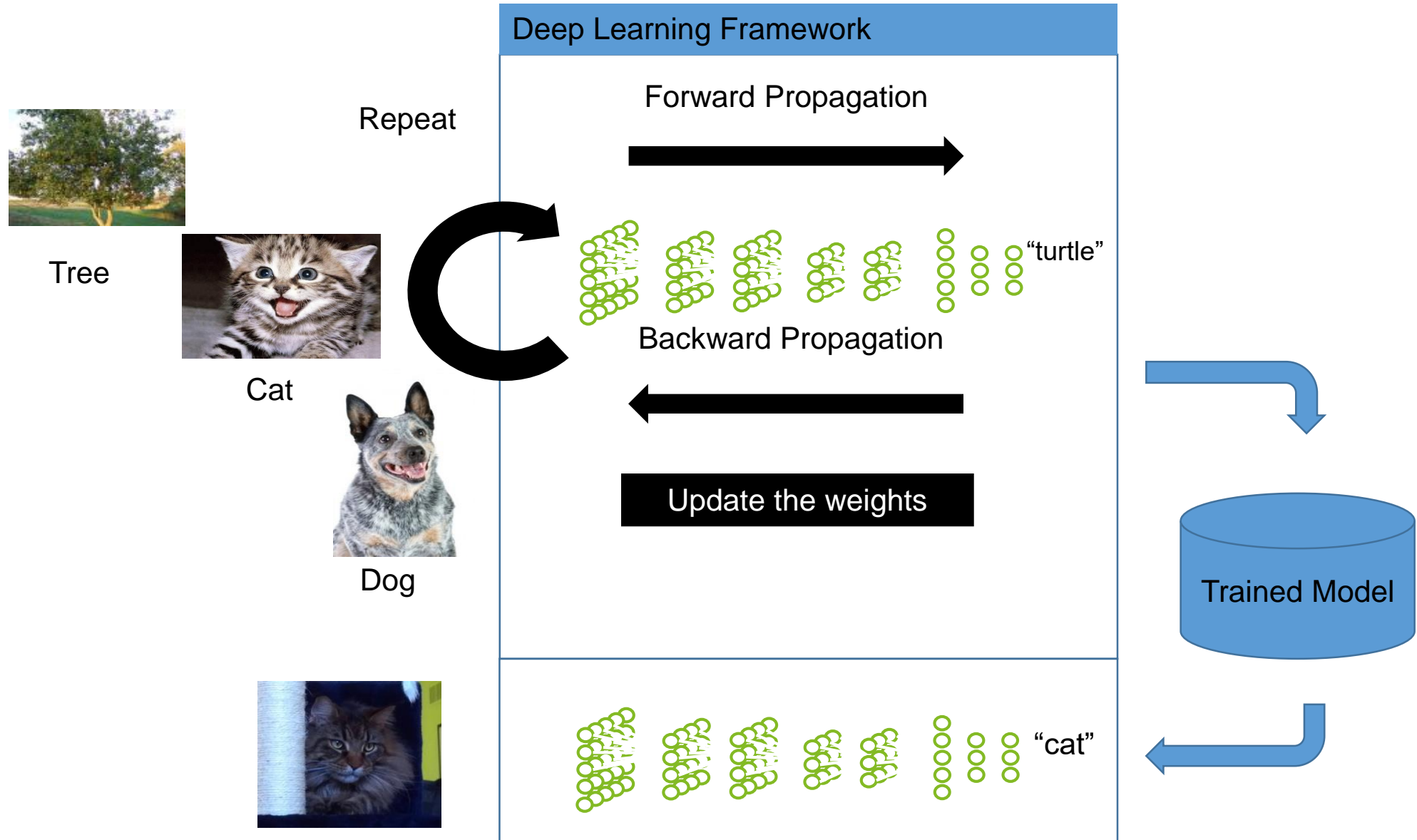
深度学习改变了计算机视觉的方法，大幅提升了准确度

ImageNet Recognition Challenge 2012

- 1.2M training images, 1000 object categories



深度学习框架 (Deep Learning Framework)



TensorFlow



- 开源的深度学习工具库
- <http://www.tensorflow.org>
- TensorFlow是一种元编程（meta programming），构建计算图（computational graph）的语言，编译系统生成实际的计算图
- TensorFlow是模型训练过程的数据流图，包括训练数据的读取和转换，队列，参数的更新以及周期性监测点生成。

TensorFlow-OD

- TensorFlow对象检测流水线(object detection pipeline)
 - 基于TensorFlow的对象检测模型。
 - 谷歌开发的一套框架，方便构建、训练和部署
- 当前版本包含：
 - 一个可训练性检测模型的集合，包括：
 - MobileNet模型与训练
 - 带有 MobileNets 的 SSD (Single Shot Multibox Detector)
 - 带有 Inception V2 的 SSD (Single Shot Multibox Detector)

https://github.com/tensorflow/models/tree/master/research/object_detection

无人驾驶或智能驾驶

- 自主（无人）驾驶（Self-Driving/Driverless），通过控制车辆的速度、方向刹车，来接替人类驾驶员，直接控制车辆。
- 辅助驾驶（ADAS）：利用车体加装的激光雷达、相机和GPS等传感器，观察周围环境，而后通过决策算法，提醒驾驶员注意道路状况的层次。
- 车载信息诊断（OBD）：利用车载诊断系统即OBD（Onboard Diagnostic）来监控并反馈车辆状态的层次，当车辆出现故障时，OBD会报警并记录数据。



<https://devblogs.nvidia.com/deep-learning-self-driving-cars/>

<https://blogs.nvidia.com/blog/2019/04/30/drive-labs-path-perception/>

无人机Droid

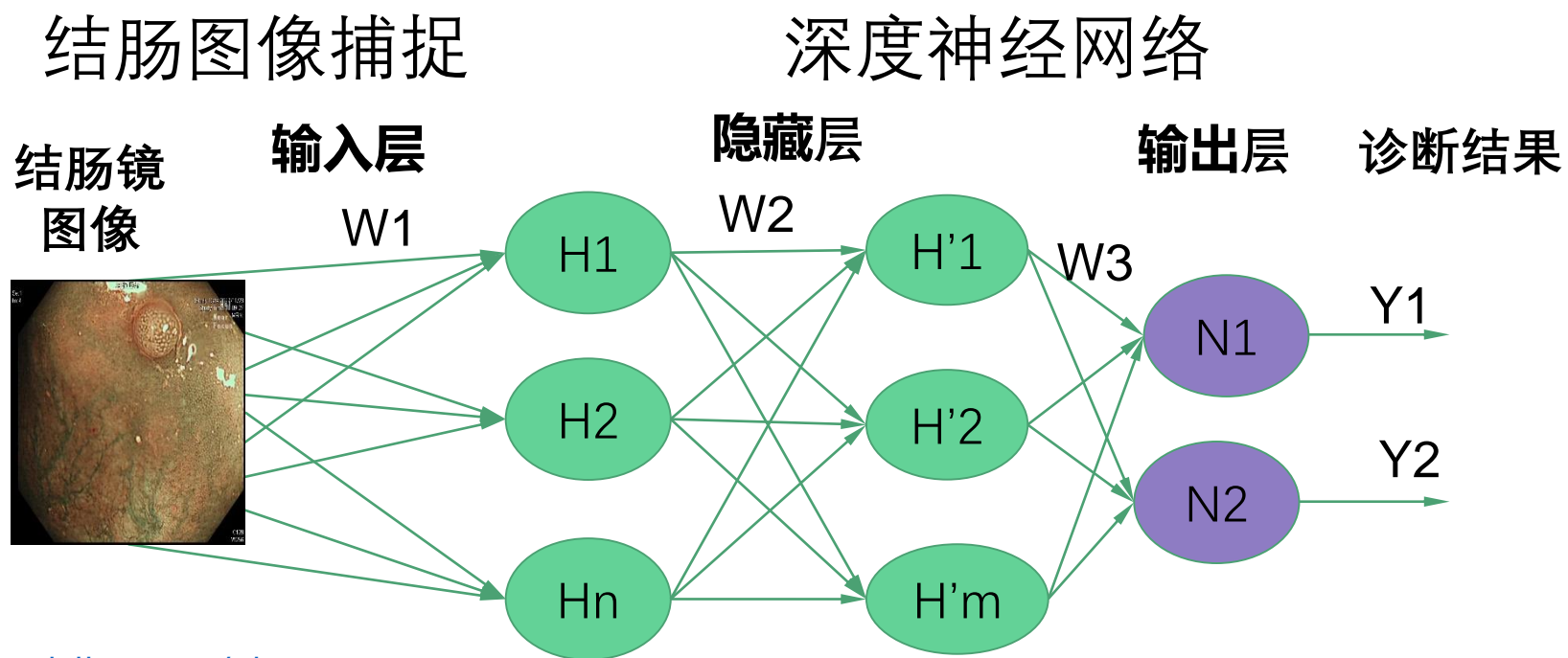
- 军事侦查
- 民用巡查



<https://github.com/NVIDIA-AI-IOT/redtail>

医学影像诊断

- 超声图像，计算机断层扫描和磁共振图像。
- 举例：结肠镜图像的病变动态捕捉



医学影像诊断

- 视网膜图像识别
- [1] Gulshan, V. et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA 316, 2402–2410 (2016).

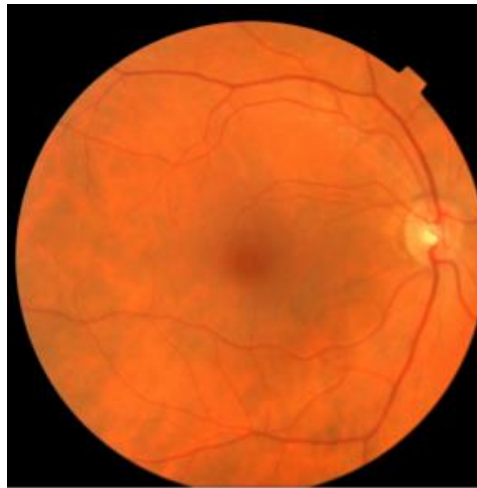
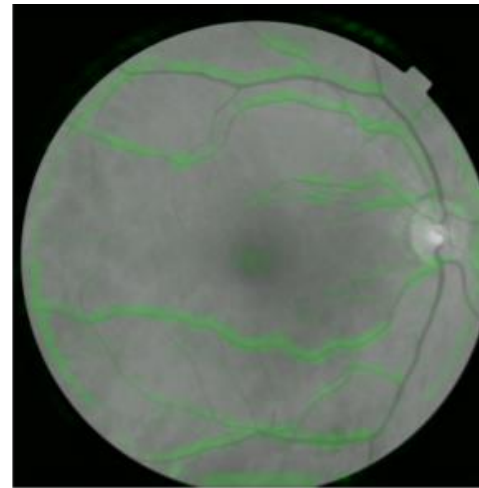
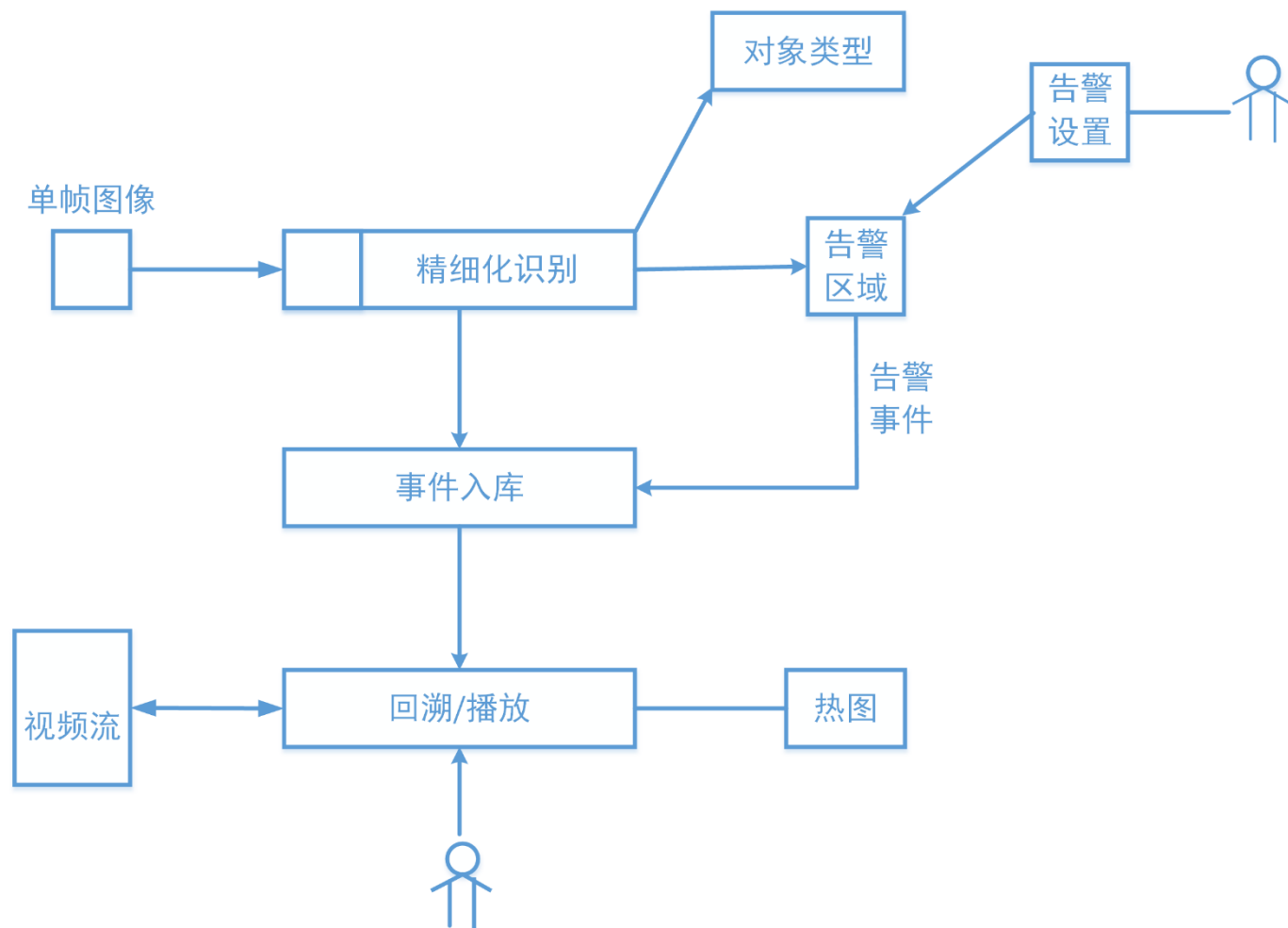


Image of retina



Blood pressure predictions
focus on blood vessels

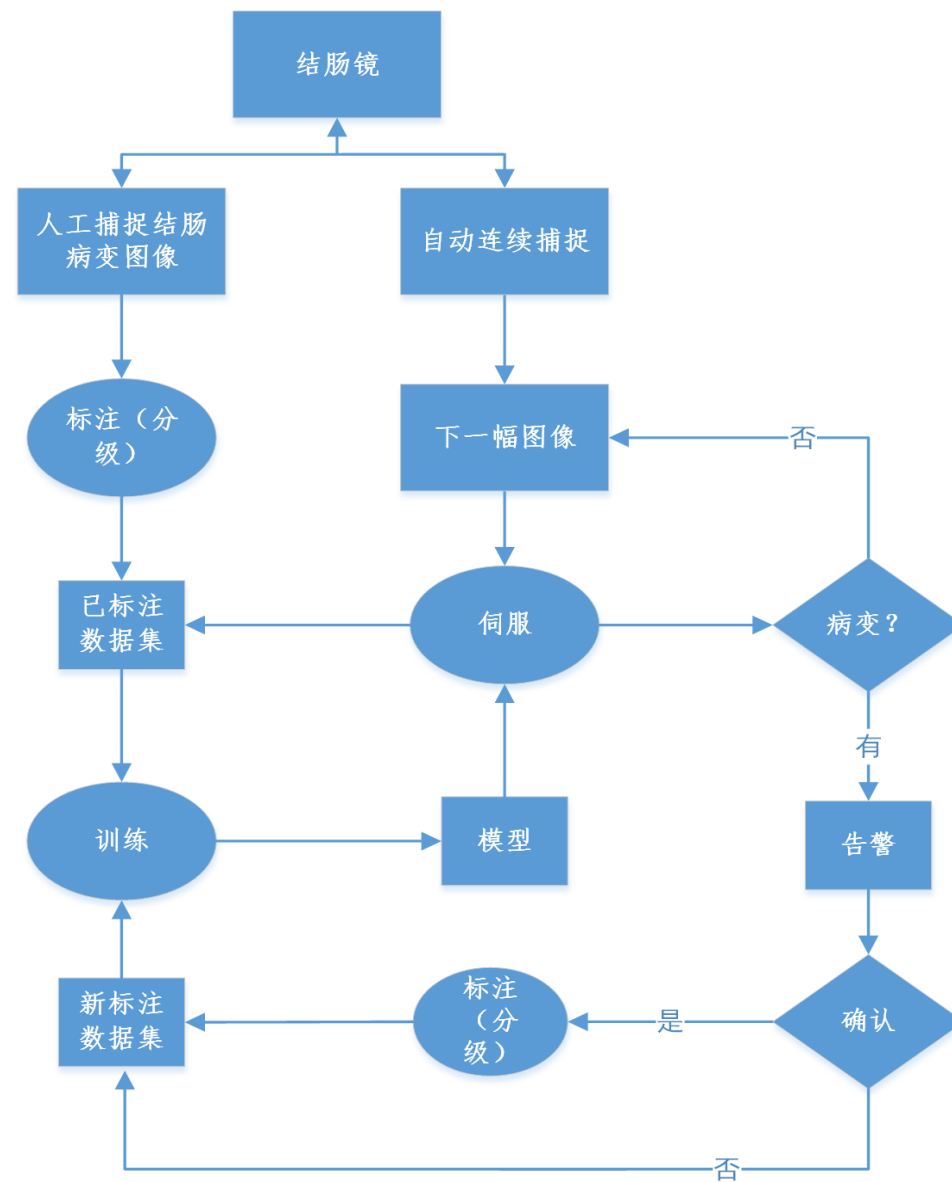
示例-基于深度学习的计算机视觉系统



- OpenCV（图像预处理），
- TensorFlow（图像识别）
- Flask（web框架），
- beanstalkd（事件队列），
- MySQL（档案化数据库部分），
- InfluxDB（实时数据库部分）。

精细化识别流程

- 数据集扩充
- 人工干预
- 识别告警



谢谢指正！

zhenchen@Tsinghua.edu.cn