

## Chapter 1-1

# 计算机视觉导论

An introduction to Computer Vision

分享人

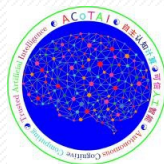
ACoTAI Lab, Dalian Maritime University

安泓郡

ICDC department, Dianhang Association

[an.hongjun@foxmail.com](mailto:an.hongjun@foxmail.com)

## 分享人简介



### **HONGJUN AN, Researcher**

Autonomous Cognitive Computing & Trustworthy Artificial Intelligence Laboratory

**ACoTAI Lab**, College of Artificial Intelligence, Dalian Maritime University.

**R.D.** Computer Vision, Deep Learning, Machine Learning

**Bilibili:** <https://space.bilibili.com/425295059>

**GitHub:** <https://github.com/BestAnHongjun>

**Email:** [an.hongjun@foxmail.com](mailto:an.hongjun@foxmail.com)



## 分享计划

章节	课题名称	课程目标	预计时间	主讲
第一章：计算机视觉基础课程	计算机视觉导论	1.定性认识计算机视觉的基本任务 2.定性认识计算机视觉的主要方法学	2022.2.14 周一 20:00	安泓郡
	计算机图形学基础	1.了解图像在计算机中的存储与计算方法 2.掌握图像的基本处理方法与常用算法	2022.2.23 周三 20:00	
	计算机图形学基础实验	1.认识OpenCV工具 2.编程实现图像处理的常用算法	2022.2.27 周日 20:00	
第二章：计算机视觉 - 图像多分类任务课程	卷积神经网络初步	1.进一步明确图像分类任务的目标 2.定性认识卷积神经网络是如何提取图像特征的，学习卷积、池化等基本算子 3.常用卷积神经网络介绍	2022.3.6 周日 20:00	臧景奇
	数据集读取与切分	1.学习设计适合网络训练的DataLoader	2022.3.13 周日 20:00	
	图像分类网络搭建	1.初步认识Pytorch 2.学习如何使用Pytorch封装的卷积、池化、全连接模块 3.学习使用torchvision.models中的预训练模型 4.学习fine-tune预训练模型	2022.3.20 周日 20:00	
	图像分类模型评估	1.学习模型的主要评估指标	2022.3.27 周日 20:00	

## 分享计划

章节	课题名称	课程目标	预计时间	主讲
第三章：计算机视觉 - 目标检测任务初步课程	目标检测导论	1.进一步明确目标检测任务的目标 2.定性认识目标检测任务的主要方法学，包括朴素DPM目标检测系统、基于RPN的two-stage目标检测系统、以YOLO为主流的one-stage目标检测系统	2022.4.3 周日 20:00	安泓郡
	YOLO目标检测算法初探	1.了解YOLO目标检测算法的核心思想 2.了解YOLO v1目标检测算法的损失函数设计与head设计	2022.4.10 周日 20:00	
	YOLO目标检测算法深入	1.学习YOLO v2-v5的主要更新特性	2022.4.17 周日 20:00	
	YOLO在服务器端的部署	1.学习借助预训练模型fine-tune自己的YOLO目标检测模型 2.学习YOLO目标检测算法在服务器端的部署	2022.4.24 周日 20:00	
	YOLO端侧部署	1.学习YOLO算法在k210单片机的端侧部署	2022.5.1 周日 20:00	黄子豪

# Chapter 1-1 目录

## CONTENTS

- 1/ 计算机视觉的基本任务
- 2/ 计算机视觉的基本方法学
- 3/ 计算机视觉导论 - 总结

# 第一部分

## BASIC TASK

### 计算机视觉的基本任务

# 什么是计算机视觉?

计算机视觉的概念，历史上给出了许多种不同的定义。目前一种相对权威的说法如下：

Computer vision is a field of artificial intelligence (AI) that enables computers and systems to derive meaningful information from digital images, videos and other visual inputs — and take actions or make recommendations based on that information. If AI enables computers to think, computer vision enables them to see, observe and understand.<sup>[1]</sup>

实际上，对计算机视觉的定义，不同研究者都会给出不同的说法。并且，随着人们对该领域研究的深入和时代背景的变化，人们会不断地更新对这一概念的认识。但普遍来说，从最初到现在，人们研究计算机视觉，是期望构建一种信息处理系统，使其具备类似人类视觉的功能，完成一些高级的任务，把人们从一些单一复杂的工作中解放出来，提高人类社会生产力。

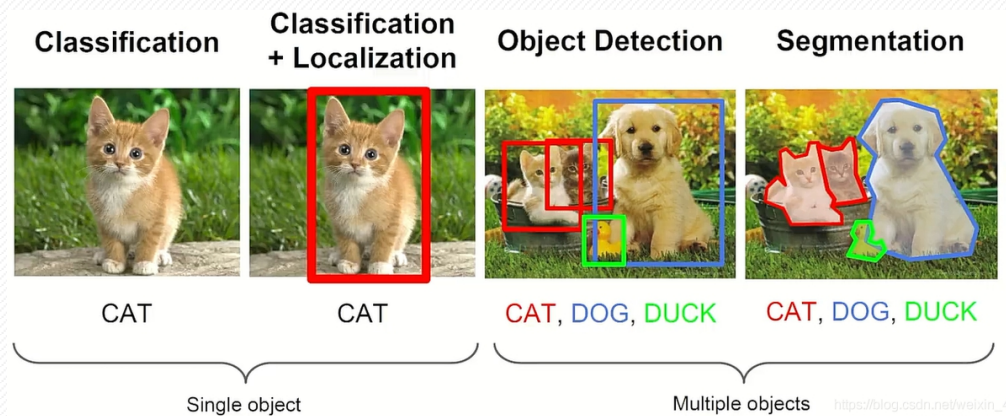
Reference:

[1] <https://www.ibm.com/topics/computer-vision>

# 计算机视觉的基本任务

计算机视觉有四大基本任务<sup>[1]</sup>:

- 分类(Classification): 解决“是什么”的问题, 即给定一张图片或一段视频判断里面包含什么类别的目标。
- 定位(Location): 解决“在哪里”的问题, 即定位出这个目标的位置。
- 检测(Detection): 解决“是什么、在哪里”的问题, 即定位出这个目标的位置并且知道目标物是什么。
- 分割(Segmentation): 解决每一个像素属于哪个目标物或场景的问题。



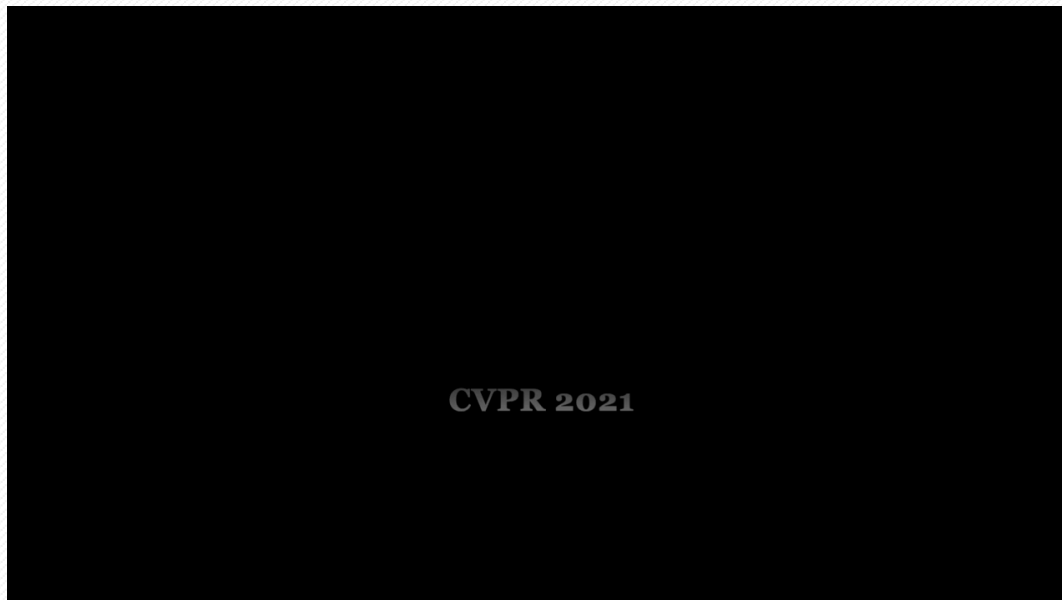
Reference:

[1] [https://blog.csdn.net/weixin\\_43558052/article/details/107692797](https://blog.csdn.net/weixin_43558052/article/details/107692797)



# 计算机视觉的高级任务

目标跟踪与分割 示例: TraDes (CVPR 2021)<sup>[1]</sup>



Reference:

[1]Wu, Jialian and Cao, Jiale and Song, Liangchen and Wang, Yu and Yang, Ming and Yuan, Junsong.Track to Detect and Segment: An Online Multi-Object Tracker. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2021.

# 计算机视觉的高级任务

行人轨迹预测 示例: Social-VRNN<sup>[1]</sup>

## Social-VRNN: One-Shot Multi-modal Trajectory Prediction for Interacting Pedestrians

Bruno Brito, Hai Zhu, Wei Pan and Javier Alonso-Mora

Cognitive Robotics  
Delft University of Technology

CoRL 2020



Reference:

[1] Bruno Brito and Hai Zhu and Wei Pan and Javier Alonso-Mora. Social-VRNN: One-Shot Multi-modal Trajectory Prediction for Interacting Pedestrians. arXiv, 2010.09056.2020

## 第二部分

## METHODOLOGY

### 计算机视觉的基本方法学



## 计算机视觉的基本方法学



由几个简单的例子，定性了解并  
体会计算机视觉的基本方法学

## 思考1：这是什么？



[1]

苹果

Reference:

[1] <https://www.vcg.com/creative/1002835253>

## 思考2：你怎么知道这是个苹果？



[1]

苹果

Reference:

[1] <https://www.vcg.com/creative/1002835253>

### 思考3：能否设计一个识别器，识别图像中的苹果？



[1]

苹果

Reference:

[1] <https://www.vcg.com/creative/1002835253>



## 能否将形状作为识别苹果的特征?



[1]

特征提取

圆形

决策

苹果

Reference:

[1] <https://www.vcg.com/creative/1002835253>



## 以形状为决策特征的识别器是否足够健壮？

[1]



特征提取

圆形

决策

苹果

[2]



特征提取

圆形

决策

苹果

梨

Reference:

[1] <https://www.vcg.com/creative/1002835253>[2] <http://www.zwboshi.com/10901.html>

## 在形状特征的基础上加上颜色特征?

[1]



特征提取

圆形  
红色

决策

苹果

[2]



特征提取

圆形  
绿色

决策

不是  
苹果

Reference:

[1] <https://www.vcg.com/creative/1002835253>[2] <http://www.zwboshi.com/10901.html>

## 以形状、颜色为决策特征的识别器足够健壮吗？

[1]



特征提取

圆形  
红色

决策

苹果

[2]



特征提取

圆形  
红色

决策

苹果

樱桃

Reference:

[1] <https://www.vcg.com/creative/1002835253>[2] <https://m.huabaike.com/qita/9520.html>

## 再加入纹理特征?



[1]

特征提取

圆形  
红色  
条状

决策

苹果



[2]

特征提取

圆形  
红色  
光滑

决策

不是  
苹果

Reference:

[1] <https://www.vcg.com/creative/1002835253>[2] <https://m.huabaike.com/qita/9520.html>

## 再来点别的数据怎么办？



[1]

苹果



[2]

苹果

Reference:

[1] <https://www.tcn-aomoriapple.com/variety.html>

[2] <http://616pic.com/tupian/pingguologo.html>

早期的研究者也经历过和我们类似的绝望。

The common and almost despairing feeling of the early investigators like B.K.P Horn and T.O. Binford was that proctically anything could happen in an image and furthermore that practivally everything did.<sup>[1]</sup>

但我们刚才的思维过程并非完全没有借鉴意义。这至少给了我们两点启示：

- 似乎存在一种可行的识别器框架（方法学）：特征提取 -> 决策
- 特征的质量，决定着识别器的性能。

人们开始试图寻找一些抽象的、较为健壮的特征用于表示图像。**特征工程，模式识别**等领域兴起。

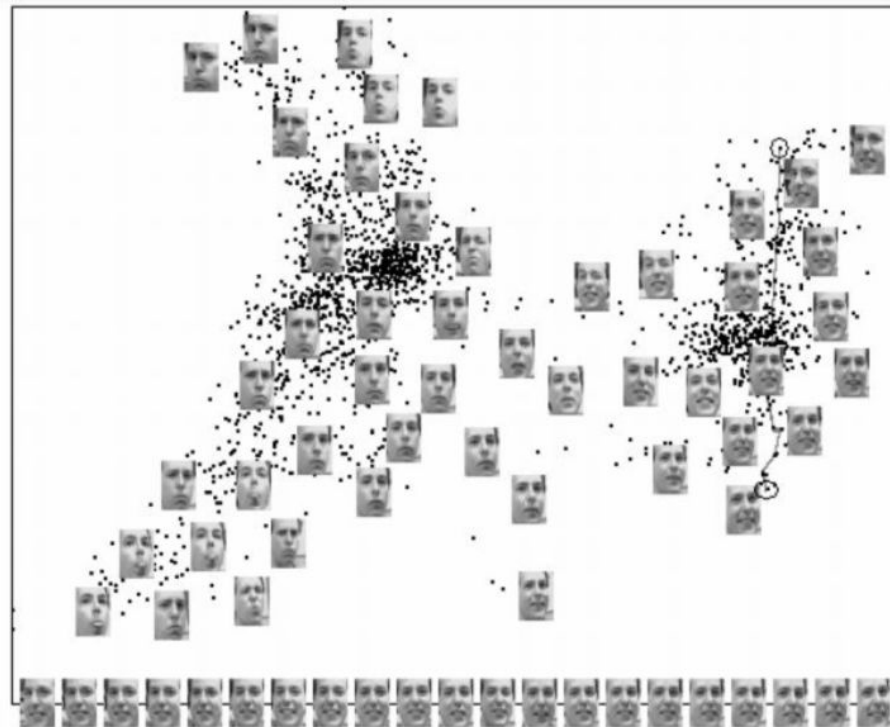
Reference:

[1] David Marr. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information

经过科学家的不懈努力，设计出了许多经典的图像数字特征。

代表：HOG特征、SIFT特征、拉普拉斯特征等。

这些特征提取算法，能将高维的图像映射到低维空间的一些离散点，然后就可以借助SVM、MP等经典学习算法进行决策分类。



[1]

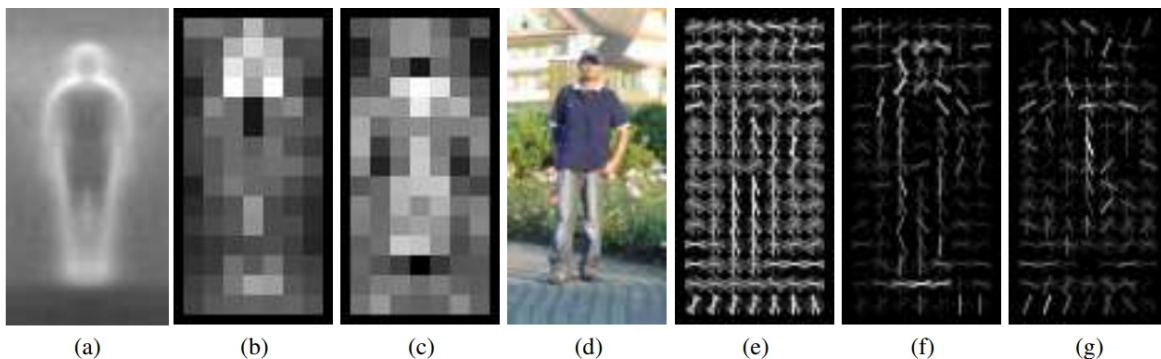
Reference:

[1] He, Xiaoferi, et al. Face recognition using Laplacianfaces. Pattern Analysis and Machine Intelligence, IEEE Transactions on 27.3 (2005): 328-340.



## 定性认识几种经典的特征提取算法

方向梯度直方图(Histogram of Oriented Gradient,HOG)是用于在计算机视觉和图像处理领域, 目标检测的特征描述子。



[1]

Figure 6. Our HOG detectors cue mainly on silhouette contours (especially the head, shoulders and feet). The most active blocks are centred on the image background just *outside* the contour. (a) The average gradient image over the training examples. (b) Each “pixel” shows the maximum positive SVM weight in the block centred on the pixel. (c) Likewise for the negative SVM weights. (d) A test image. (e) It's computed R-HOG descriptor. (f,g) The R-HOG descriptor weighted by respectively the positive and the negative SVM weights.

### Reference:

[1]Dalal, N. and B. Triggs. "Histograms of Oriented Gradients for Human Detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1 (June 2005), pp. 886–893.

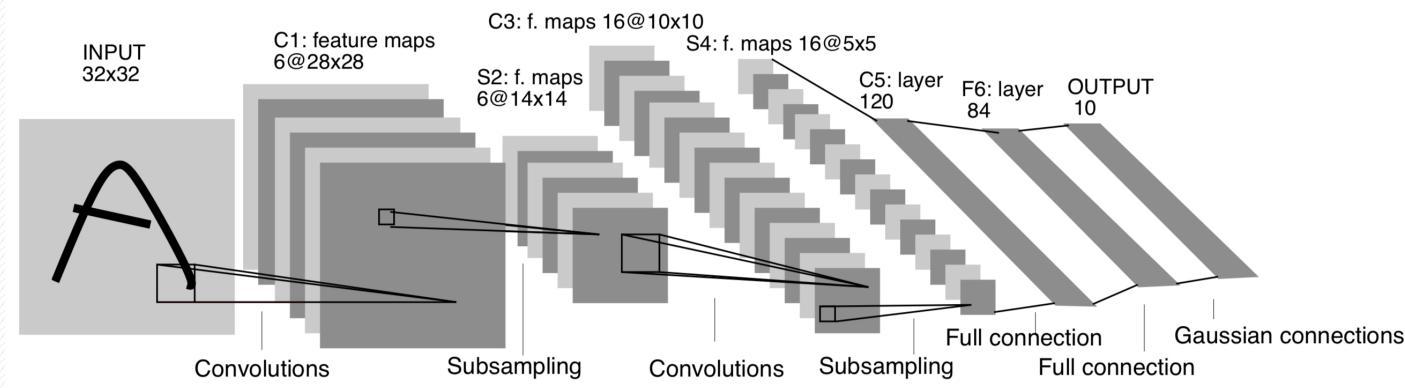


这些经典的特征提取算法虽然解决了很大一部分问题，但都存在一个普遍的缺陷：不够健壮。主要原因是这些特征都是人们“手工设计”的，是受制于人类主观意识的。

**能否设计一种算法，让计算机自动提取一个适合决策的特征？**

1988年，LeNet出现，标志着卷积神经网络(CNN)正式步入计算机视觉的历史。

[1]



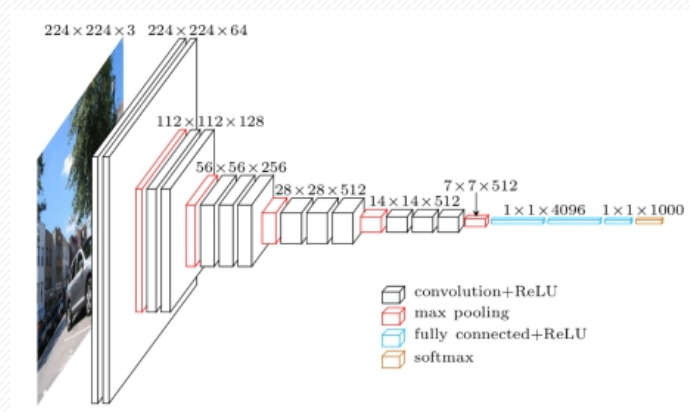
Reference:

[1] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.

早期卷积神经网络在许多图像问题上大放光彩。如在手写数字识别、邮票识别等任务上准确率能达到99%以上。

受制于网络深度影响，网络无法提取到足够抽象的特征，在一些复杂问题上仍然表现不佳。受制于当时计算机硬件的算力，网络很难做到很深。

2010年左右，计算机硬件尤其是显卡算力大幅提升，深度神经网络爆火。得益于深层的网络模型，能够提取到稳定的、足够抽象的特征，使得以深度学习为基础的计算机视觉系统在许多问题上取得了极佳的成绩。



[1]

Reference:

[1] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in International Conference on Learning Representations, May 2015.

## 第三部分

# Summary

## 计算机视觉导论 - 总结

## 总结

模型复杂度越来越高，特征越来越抽象，健壮性越来越强，可解释性越来越差。

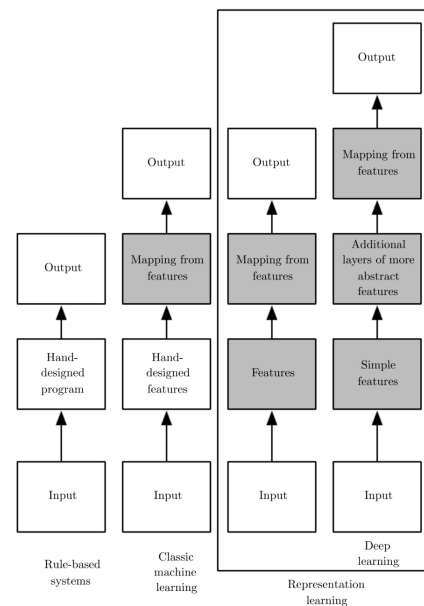
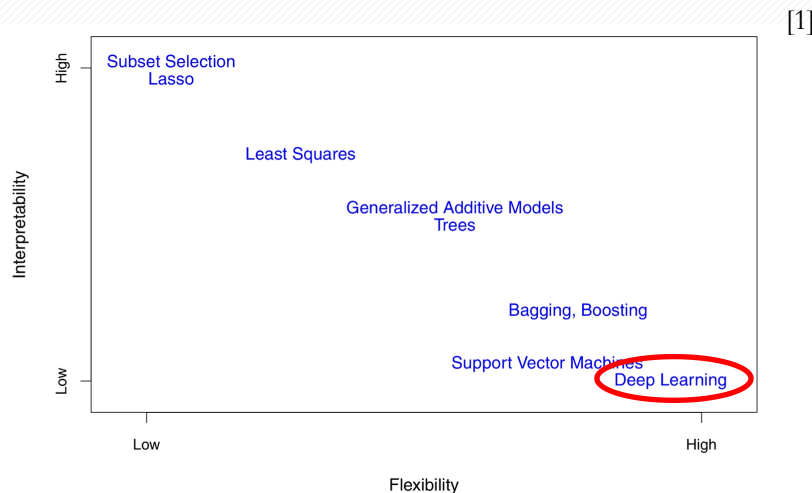


Figure 1.5: Flowcharts showing how the different parts of an AI system relate to each other within different AI disciplines. Shaded boxes indicate components that are able to learn from data.

Reference:

[1] Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani. An Introduction to Statistic Learning, with Applications in R.

[2] Ian Goodfellow, Yoshua Bengio and Aaron Courville. Deep Learning.

## 课程资源

课程PPT，代码等将于Github同步更新：

<https://github.com/BestAnHongjun/ICDC-Department-training>

课程回放将于直播结束三天内上传至bilibili平台：

<https://space.bilibili.com/425295059>

学习交流、问题指正请联系：

[an.hongjun@foxmail.com](mailto:an.hongjun@foxmail.com)



*bilibili*主页

# THANKS!

分享人

ACoTAI Lab, Dalian Maritime University

安泓郡

ICDC department, Dianhang Association

[an.hongjun@foxmail.com](mailto:an.hongjun@foxmail.com)