

学AI用AI公益活动（第7次）



# 计算机视觉

王红梅

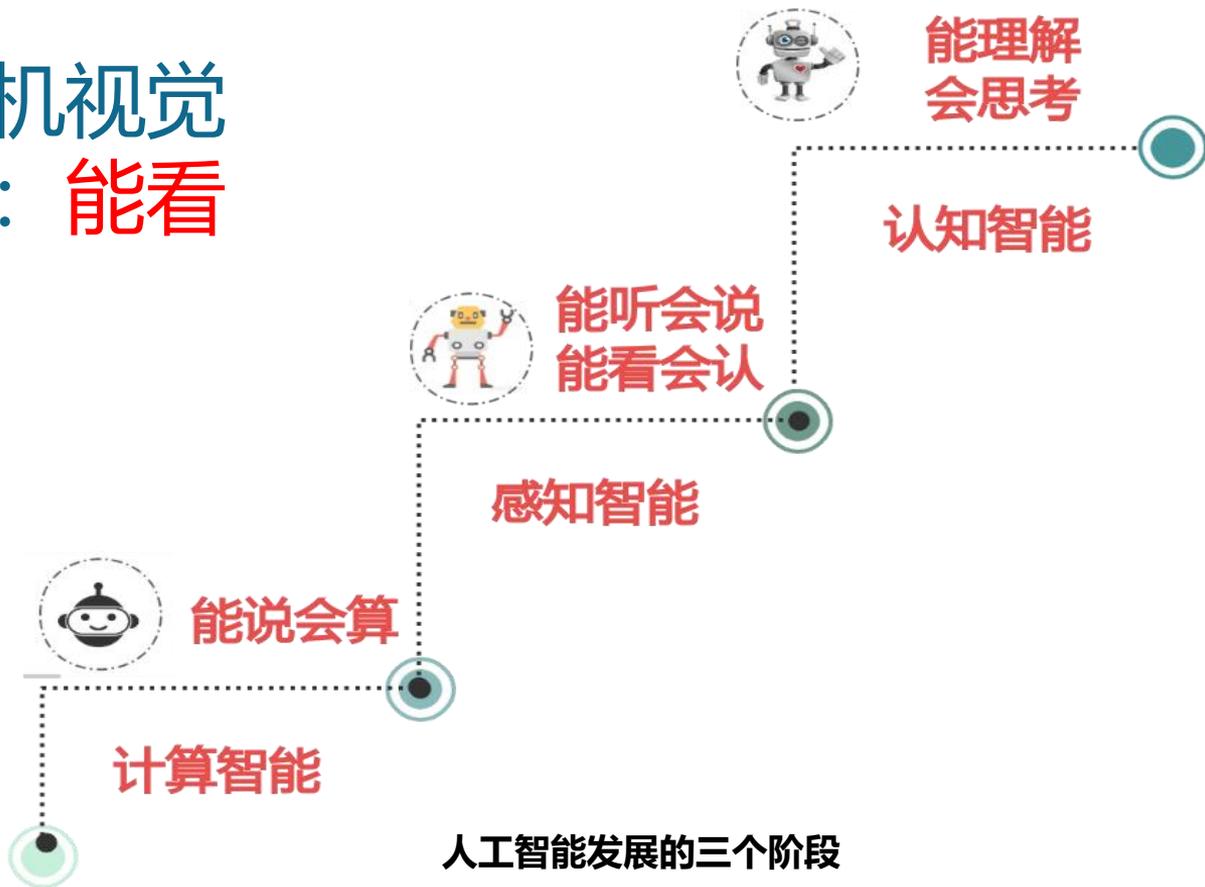
郑州航空工业管理学院

2026-03-06



- 一、计算机视觉概述**
- 二、计算机视觉的处理方法**
- 三、计算机视觉的任务、分类及应用**
- 四、计算机视觉实践**

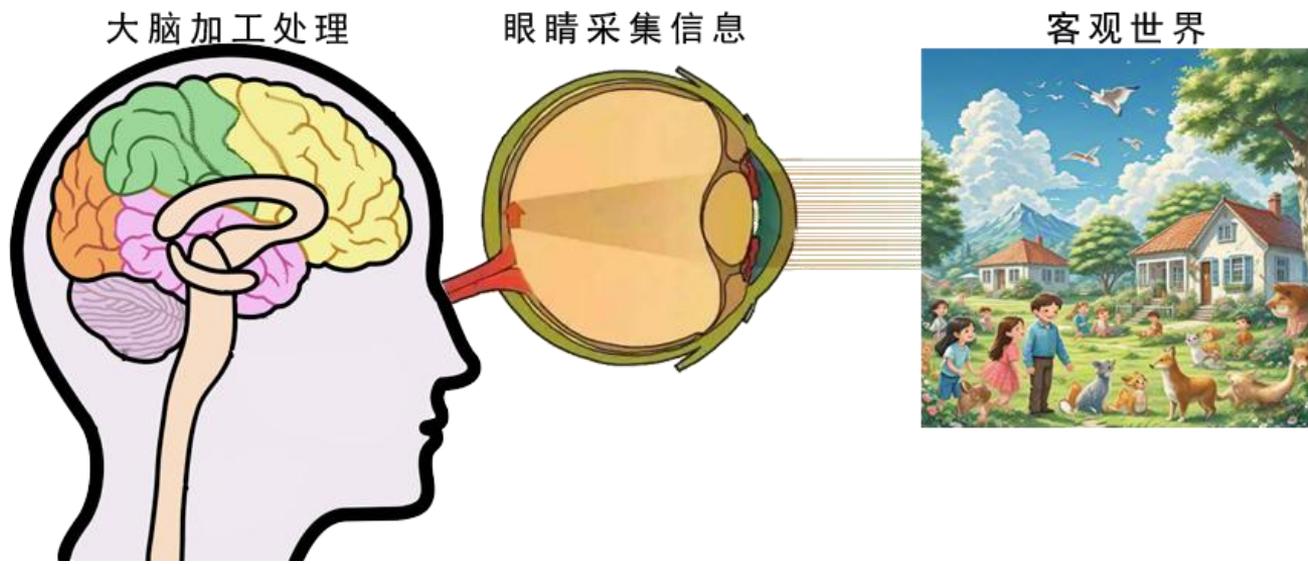
计算机视觉  
负责：**能看**  
**会认**



人工智能发展的三个阶段

讨论：大家觉得眼睛采集的信息，有谁来负责处理？

在这个非常漫长的过程，“视觉”从眼睛采集数据开始，而大脑才是它呈现意义的地方，



**讨论：大家觉得眼睛采集的信息，有谁来负责处理？**

**摄像头采集信息，视觉模型才是加工和处理信息的地方，**



类似摄像头等传感器采集数据，视觉模型才是加工信息的场所。

## (一) 出现有标注的高质量数据集

**2009年，李飞飞教授等发布了ImageNet数据集。**



类似摄像头等传感器采集数据，视觉模型才是加工信息的场所。

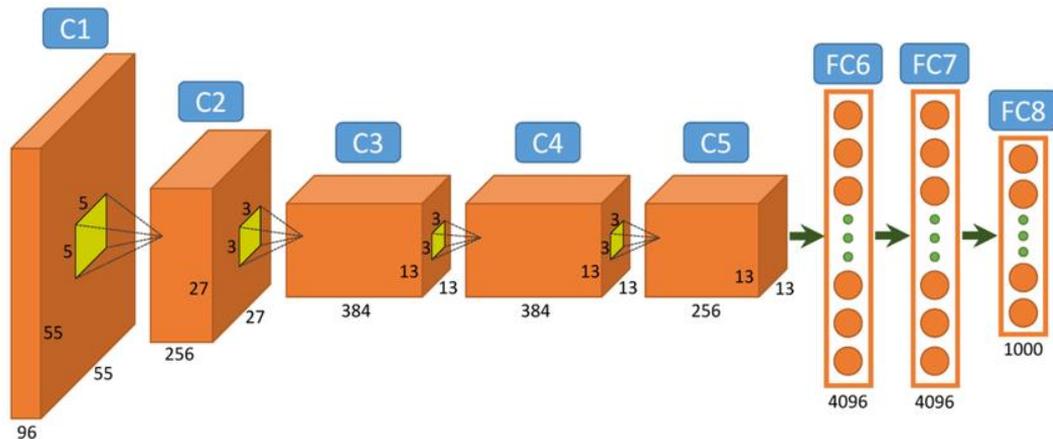
### (一) 出现有标注的高质量数据集

ImageNet数据集包含了**1400多万幅图片**。**2万多个类别**的大型数据库，这个数据集是一个用于**图像识别和分类**的大型数据库，**包含了大量的标注图片，用于训练和测试图像识别算法**。**ImageNet数据集的规模和多样性对于推动计算机视觉领域的发展起到了重要作用**，使得研究人员能够开发出更加准确的**图像识别和分类算法**。

类似摄像头等传感器采集数据，视觉模型才是加工信息的场所。

(二) 2010年-至今 深度学习在视觉中流行，在应用上百花齐放

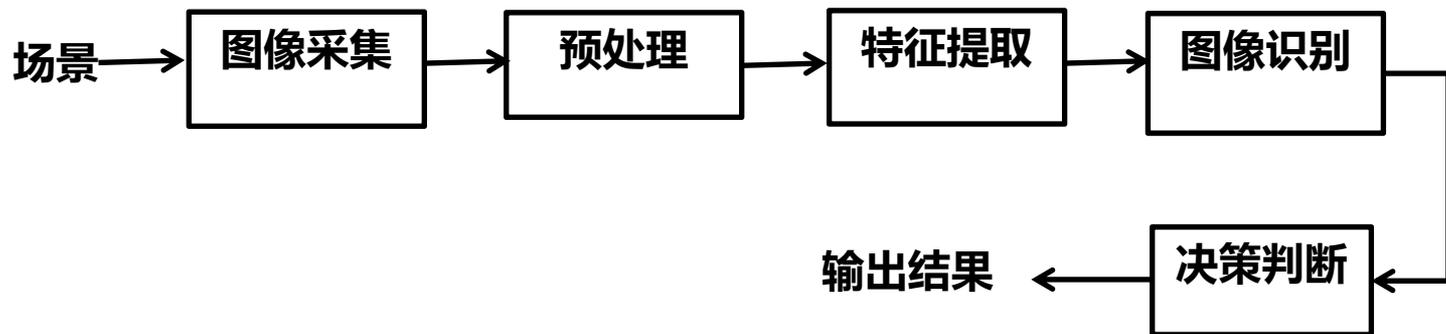
2012年，Alex Krizhevsky、Ilya Sutskever 和 Geoffrey Hinton 创造了一个“大型的深度卷积神经网络”，也即现在众所周知的 **AlexNet**。英伟达的**GPU**也发展的能支持AlexNet的运行。



# 一、计算机视觉概述

类似摄像头等传感器采集数据，视觉模型才是加工信息的场所。

(三) 计算机视觉应用系统主要包括**图像采集**、**预处理**、**特征提取**、**图像识别**和**决策判断**等模块。

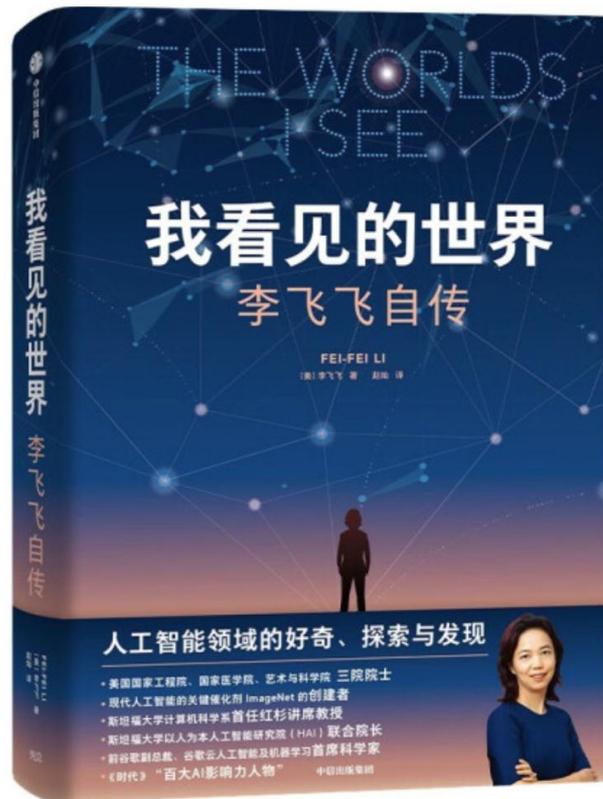




### 图书推荐

- (1) 讲述李飞飞自身成长的励志书
- (2) 讲述人工智能发展
- (3) 讲述计算机视觉的发展

...



计算机视觉关于图像识别主要有**四大任务**，分别为**分类**、**定位**、**检测**和**分割**，不同任务解决不一样的问题。

#### 分类



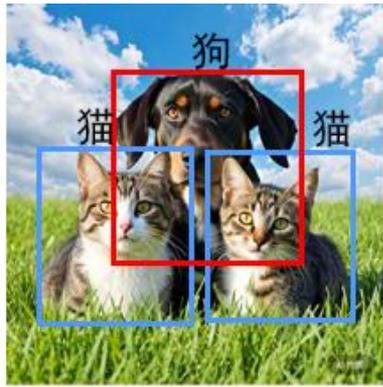
是什么？

#### 定位



在哪里？

#### 检测



是什么？ 在哪里？

#### 分割



怎么分离出来？

计算机视觉关于图像识别主要有**四大任务**，分别为**分类**、**定位**、**检测**和**分割**，不同任务解决不一样的问题。

#### 1. 分类 (Classification)

问题：**图片里是什么？**

例子：给你一张图，**判断是“猫”还是“狗”**。

技术本质：**算法给整张图打一个标签**，

实际应用：**相册自动分类（猫/狗/风景）、垃圾邮件过滤（正常/垃圾）**。

关键特点：**只关心“是什么”，不关心“在哪里”或“有多少”**。



计算机视觉关于图像识别主要有**四大任务**，分别为**分类**、**定位**、**检测**和**分割**，不同任务解决不一样的问题。

## 2. 定位 (Localization)

问题：目标**在哪里**？（通常已知类别）

例子：已知图中有猫，**用框标出猫的位置**。

技术本质：**在分类基础上加一个边界框**。

实际应用：人脸识别时框出人脸、自动驾驶中标记车辆位置。

关键特点：通常针对**单个目标**，既要**分类也要定位**。



计算机视觉关于图像识别主要有**四大任务**，分别为**分类**、**定位**、**检测**和**分割**，不同任务解决不一样的问题。

### 3. 检测 (Detection)

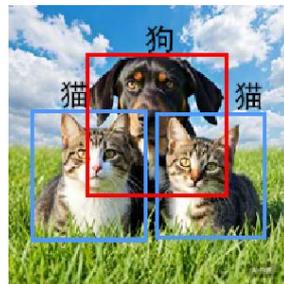
问题：图中**是什么**？它们**在哪里**？

例子：一张街景图中找出所有的车、行人、红绿灯，并分别标出位置和类别。

技术本质：**多目标定位+分类**。

实际应用：自动驾驶、安防监控（识别多个人或物体）。

关键特点：**处理多个目标**，输出一堆边界框和类别标签。



计算机视觉关于图像识别主要有**四大任务**，分别为**分类**、**定位**、**检测**和**分割**，不同任务解决不一样的问题。

#### 4. 分割 (Segmentation)

问题：**每个像素归属问题或怎么把图像分离出来？**

例子：把照片中的猫的每一根毛发、脚爪的像素都精确标出来，其他背景排除。

技术本质：给每个像素分类。

细分类型：语义分割和实例分割：

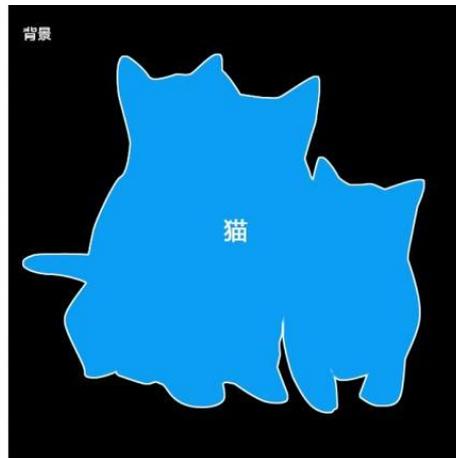
实际应用：医学图像（标记肿瘤区域）、地图分割等。

关键特点：**像素级精度**，比检测的边界框更精细。



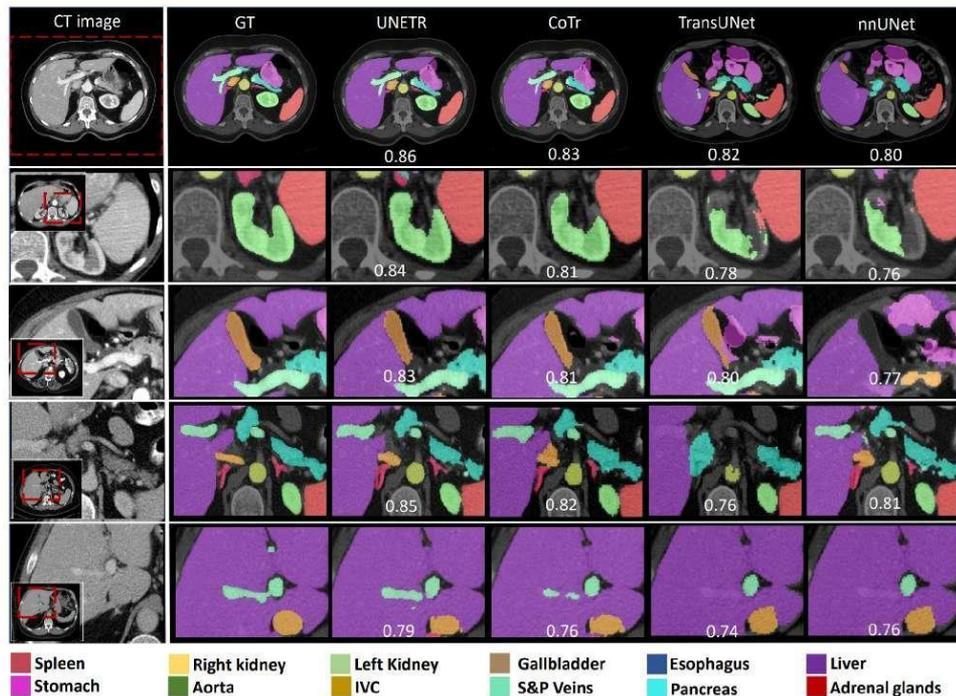
## 4.1 语义分割

**语义分割旨在将输入图像中的每个像素标记为属于哪个语义类别。**与目标检测和图像分类不同，语义分割不仅可以识别图像中的物体，还可以为每个像素分配标签，从而提供更详细和准确的图像理解。



## 4.1 语义分割

**主要应用：适用于对图像进行精细分割和像素级分类的场景，例如自动驾驶中的道路分割、医学图像中的病灶分割、地理信息系统中的土地分类等。**

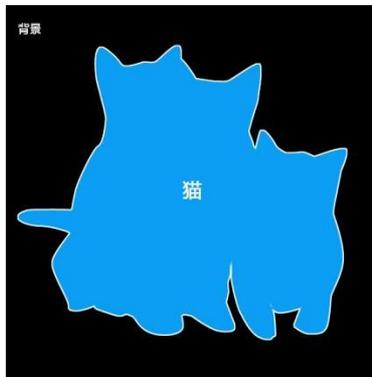


## 4.2 实例分割

实例分割是结合目标检测和语义分割的一个更高层级的任务，旨在检测图像中的物体的同时将**每个物体分割成精确的像素级别的区域**。



(a) 原图



(b) 语义分割



(c) 实例分割



#### 讨论：语义分割和实例分割的区别

实例分割和语义分割都属于计算机视觉任务，而且都属于预测单个像素点的类型。给定一张人群的图片，两者的区别在于，**语义分割只对类别进行了区别，不管有几个人，都会预测为一个类别。而实例分割不仅预测类别，还要识别每个不同的个体，每个个体会被预测为不同的对象。有点像结合了目标检测的语义分割。**



计算机视觉关于图像识别主要有**四大任务**，分别为**分类、定位、检测和分割**，不同任务解决不一样的问题。

技术关联：

这些任务通常共用底层技术（如**卷积神经网络CNN**）来完成，**四个任务难度递增**：

- (1) 分类是基础，
- (2) 分类+定位是对一个对象
- (3) 检测=分类+定位的多次应用，
- (4) 分割需要更精细的处理。



### 计算机视觉类别和任务对应一览表

类别 \ 任务	分类	定位	检测	分割	备注
图像分类	√				
目标检测	√	√	√		
目标跟踪	√	√	√		持续的目标检测
语义分割	√	√	√	√	只考虑类型
实例分割	√	√	√	√	考虑类型和每一个对象



**环境安装需要Python+第三方库**



群聊：郑州航院人工智能通  
识课9群



该二维码7天内(3月13日前)有效，重  
新进入将更新

保存图片



群聊：郑州航院人工智能通  
识课10群



该二维码7天内(3月13日前)有效，重  
新进入将更新

保存图片

没有在群的老师可  
以扫描加群，但不  
用重复加群