

计算机视觉的深度学习实践

【课程时长】

3 天 (6 小时/天)

【课程简介】

在所有人工智能的应用方向上，计算机视觉/机器视觉图像处理的落地应用最为成熟，近两年的技术发展是最为迅猛。

以互联网巨头为代表的企业正在开展如：基于识别的智能机器人，无人驾驶，人脸识别、图像搜索，图像处理，游戏界面等；专门从事视觉、图像处理的新兴公司正在发力于人脸识别，检测，跟踪，安防等领域。只有对计算机视觉这个领域有了一个初步的全面了解才能在这些领域进行研究，一步步深入下去。

【课程目标】

- 计算机视觉领域的重点研究问题。由浅入深得讲解数字图像的存储、预处理、特征提取，以及在深度学习兴起之前计算机视觉领域所取得的成就。
- 专门介绍深度学习的基础理论知识，包括神经网络的基本原理，以及深度学习对于传统神经网络的关键改进。
- 重点介绍深度学习模型在计算机视觉领域的应用。具体涉及在计算机视觉领域如何应用卷积神经网络 (CNN)、区域卷积网络 (R-CNN) 全卷积网络 (FCN)、循环神经网络 (RNN)、长短时记忆单元 (LSTM)、生成对抗网络 (GAN)等解决图像应用的难点 课程将使用 Python 语言及 Tensorflow、Keras 深度学习框架等进行案例实践教学。

【讲师介绍】

叶梓，博士、高级工程师。2005 年上海交通大学计算机专业博士毕业，主研方向为数据挖掘、机器学习、人工智能等。现为某大型上市软件企业的人工智能团队技术负责人。曾主持多项国家战略级人工智能项目，主持设计并搭建多个市级大数据平台，在大数据、人工智能应用等方面有着丰富的工程实践经验。先后在 SCI 或 EI 期刊上发表论文 4 篇，在中文核心期刊上发表论文近 20 篇，并被百度学术收录。2011 年获中国行业协会科技创新一等奖。

第一讲 课程概述

- 1、计算机视觉的研究意义
- 2、计算机视觉的难点
- 3、当前研究的主要热点问题 (分类、目标检测、实例分割、图说等)
- 4、本课程的主要内容介绍

- 5、相关开源库介绍 (OpenCV、Tensorflow、Keras、pyTorch 等)
- 6、应用案例：搭建 tensorflow+opencv 的环境

第二讲 图像预处理

- 1、图像平滑与去噪 (高斯滤波、中值滤波等)
- 2、基于直方图的对比度增强:CLAHE
- 3、边缘检测算子 (Sobel、拉普拉斯等)
- 4、形态学处理 (腐蚀、膨胀、开闭运算等)
- 5、高斯金字塔与拉普拉斯金字塔
- 6、频域分析及变换 (卷积计算、傅里叶变换、小波变换)
- 7、应用案例：平滑、边缘检测、CLAHE、FFT 等

第三讲 图像特征提取

- 1、颜色特征 (量化直方图、聚类直方图)
- 2、几何特征 (Edge、Corner、Blob 等)
- 3、Harris 角点与 FAST 角点
- 4、基于关键点的特征描述子 (SIFT、SURF、ORB)
- 5、其他特征提取 (LBP、Gabor)
- 6、应用案例：SIFT、图像拼接等

第四讲 未有深度学习之前

- 1、基于灰度的图像分割 (阈值分割、区域生长、分水岭等)
- 2、基于图论：graph-cut 与 grab-cut
- 3、用于人脸检测的 Haar-like 特征与级联分类器
- 4、用于行人检测的 HOG+SVM
- 5、用于行人检测的多尺度形变部件模型 (DPM)
- 6、应用案例：人脸识别、行人识别

第五讲 神经网络与误差反向传播算法

- 1、人工神经元及感知机模型
- 2、目标函数 (MSE)
- 3、激励函数 (sigmoid、tanh)
- 4、误差反向传播算法的推导
- 5、应用案例：可以手算的 BP 神经网络
- 6、深度学习与神经网络的区别与联系

第六讲 深度学习基础

- 1、深度学习中的目标函数与激励函数
- 2、深度学习中的求解方法 (Adagrad、RMSprop、Adam 等)
- 3、深度学习中的技巧 (dropout、BN、weights decay 等)
- 4、应用案例：利用 tensorflow 实现的手写数字识别
- 5、卷积神经网络介绍
- 6、卷积层的误差反向传播

7、池化层的误差反向传播

第七讲 图像分类

- 1、竞赛中的分类问题
- 2、CNN 的发展概述
- 3、开山之作：AlexNet
- 4、5 层变为 5 组：VGG
- 5、组合所有可能的模型：GoogLeNet
- 6、残差网络：ResNet
- 7、深与宽之外的改进方向：ResNext
- 6、应用案例：VGG、ResNet

第八讲 图像检索

- 1、检索特征（基于颜色，纹理，形状，局部特征）
- 2、特征相似度度量（EMD）
- 3、建立基于深度学习的检索索引
- 4、知识点：迁移学习的一种实现（fine-tune）
- 5、索引加速：KD-tree
- 6、大数据条件下的索引加速：Locality Sensitive Hash
- 7、应用案例：CBIR 的应用

第九讲 目标检测（上）

- 1、目标检测任务概述
- 2、区域卷积神经网络：R-CNN
- 3、共享卷积层与多尺度：SPP-Net
- 4、多任务的目标函数：Fast R-CNN
- 5、SS 改成 RPN：Faster R-CNN
- 6、其他数据集介绍：行人检测、人脸检测
- 7、应用案例：Faster R-CNN

第十讲 目标检测（下）

- 1、之前方法的总结
- 2、ROI-wise 子网继续共享：R-FCN
- 3、回归解决一切：YOLO v1
- 4、八大改进：YOLO v2
- 5、构建语义树：YOLO 9000
- 6、多尺度预测：YOLO v3
- 7、应用案例：Darknet 实现的 YOLO

第十一讲 通用场景下的图像分割

- 1、语义分割
- 2、全卷积网络语义分割：FCN
- 3、知识点：反卷积、转置卷积与空洞（膨胀）卷积

- 4、DeepLab v1 (含 CRF)
- 5、DeepLab v2 (多尺度)
- 6、DeepLab v3 与 v3+(多尺度级联)
- 7、PASCAL VOC、MS COCO、Cityscapes 等数据集介绍
- 8、应用案例：DeepLab、denseCRF

第十二讲 医疗影像分割

- 1、医学影像分析任务概述与数据集
- 2、U-Net
- 3、3D U-NET 与 V-Net
- 4、FC-DenseNet
- 5、病理切片分析任务概述与数据集
- 6、病理切片分析的实现
- 7、应用案例：利用 U-Net 实现的器官分割

第十三讲 图像描述 (图说)

- 1、深度学习的语言模型 (RNN)
- 2、知识点介绍；LSTM 与 GRU
- 3、图说模型原理与结构
- 4、模型增强：注意力机制
- 5、图说效果的评判标准
- 6、数据集介绍 (MS COCO, Flickr 等)
- 7、应用案例：RNN 简单示例，图像描述：show and Tell

第十四讲 图像生成

- 1、变分自编码器 (VAE)
- 2、生成对抗网络 (GAN)
- 3、知识点：KL 散度与 JS 散度
- 4、改进的 GAN：DCGAN
- 5、从根本上解决训练的困难：Wasserstein GAN
- 6、超分辨率问题：SRGAN
- 7、应用案例：GAN 与 DCGAN 等