

人工智能引论

视觉实践课初探

刘家瑛
北京大学





北京大学
PEKING UNIVERSITY

困惑一：

01

人工智能“热浪”引发的
课程建设

北京大学人工智能相关课程



北京大学
PEKING UNIVERSITY

课程设置

- 《智能科学技术导论》
(大一) 导论课程
- 《人工智能概论》
(大三) 系统性学习

VS

学生科研需求

- 对人工智能极大兴趣
- 学生大一/大二 进实验室的
茫然 ← 本科生科研实践

北京大学人工智能相关课程



其他相关课程设置 (信科)

课程号	课程名	院 系	课 类
04831250	机器智能实验	智能系	限选
04831270	智能信息系统	智能系	限选
04832220	智能机器人概论	智能系	任选
04832350	统计分析与商务智能	计算机系	任选
04833340	机器学习前沿：在线学习和优化	信科	公选
04831730	机器学习概论	信科	任选
04802020	机器学习原理	信科	任选
04804011	人工智能前沿与产业趋势	信科	任选
04833420	机器学习	信科	必修

北京大学人工智能相关课程



其他相关课程设置 (非信科)

课程号	课程名	院 系	课 类
00110950	人工智能	数院	任选
01235210	智能交通系统概论	地空	任选
00114250	机器学习	数院	任选
04630996	深度学习：算法与应用	元培	任选
00137130	深度学习：算法与应用	数院	限选
08408010	强化学习：理论与算法	数院	任选
00102892	统计学习	数院	任选
00134210	人工神经网络	数院	任选

北京大学人工智能相关课程



北京大学
PEKING UNIVERSITY

- 课程繁杂, 老师与学生一样没有头绪
- 时间分布不合理
- 课程同质化, 内容重复度高

面对人工智能“热潮”加大人工智能相关课程建设

→ 跟风开设新课 还是 深耕与探索已有课程



北京大学人工智能相关课程



北京大学
PEKING UNIVERSITY

课程定位

- 面向 计算机科学与技术, 智能科学与技术, 数据科学与大数据技术等专业的本科生
- 开设一门**基础性的必修课程**

课程特点

- **理论性强、涉及知识面广、实践要求高**
- 注重课程的基础性和启发性

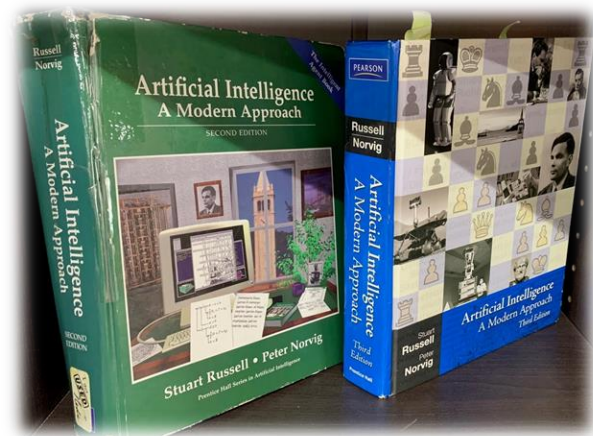
尝试：大小班相结合的课程

- 基于北京大学相关教师对人工智能的理解和在人工智能领域的研究成果
- 分析和借鉴国际一流大学 (CMU, MIT, Stanford, UCB等) 人工智能课程的授课内容
- 参考使用国际经典教材

Artificial Intelligence: A Modern Approach

Stuart Russell & Peter Norvig, Pearson

融入人工智能领域国际最新研究进展



注重理论学习和课程实践, 采用“大班课程和实践课程相结合”的教学方式

尝试：大小班相结合的课程



《人工智能引论》与实践课

- 建立课程团队, 根据课程不同内容, 采取一门课程多人讲授的方式
 - **理论讲授 (大班, 32学时)**
 - **前沿实践 (小班, 32学时)**
 - **交流展示 (分组, 4学时)**
- 下沉到大学本科一年级开设, 重新设计课程内容, 细致梳理逻辑结构

尝试：大小班相结合的课程



《人工智能引论》与实践课

- 建立课程团队, 根据课程不同内容, 采取**一门课程多人讲授**的方式
- 下沉到大学本科一年级开设, 重新设计课程内容, 细致梳理逻辑结构
- 逐步完善教案, 编写教材
- 搭建**课程教学平台和实验环境**
- 同步建设**网络在线课程**



北京大学
PEKING UNIVERSITY

困惑二：

02

各有所长，人工智能
大班课怎么上？

人工智能引论大班理论课



编号	课程内容	
1	Overview of AI	AI概述
2	Brain and Cognition	脑与认知
3	Intelligent Agents	智能体
4	Math Background	AI数学基础
5	Solving Problems by Searching	基于搜索的问题求解
6	Adversarial Search	对抗搜索
7	Constraint Satisfaction Problems	约束满足问题搜索
8	Machine Learning	机器学习
9	Artificial Neural Networks	人工神经网络
10	Deep Learning	深度学习
11	Logic and Reasoning	逻辑与推理
12	Knowledge Graph	知识图谱
13	Planning	规划
14	AI Ethics	AI伦理
15	AI Perspective	AI展望

人工智能引论大班理论课



北京大学
PEKING UNIVERSITY

2020年课程安排

1	AI概述
2	脑与认知
3	智能体
4	AI数学基础
5	基于搜索的问题求解
6	对抗搜索
7	约束满足问题搜索
8	机器学习
9	人工神经网络
10	深度学习
11	逻辑与推理
12	知识图谱
13	规划
14	AI伦理
15	AI展望

2021年课程安排

1	AI概述
2	智能体
3	脑与认知
4	约束满足问题搜索
5	对抗搜索
6	局部搜索
7	全局搜索
8	机器学习-基础
9	机器学习-进阶
10	人工神经网络
11	深度学习
12	逻辑与推理
13	知识图谱
14	机器人与AI伦理
15	AI展望

2022年课程安排

1	AI概述
2	AI数学基础
3	智能体
4	局部搜索
5	全局搜索
6	对抗搜索
7	约束满足问题搜索
8	机器学习-基础
9	人工神经网络
10	机器学习-进阶
11	深度学习
12	逻辑与推理
13	知识图谱
14	脑与认知
15	机器人与AI伦理

《人工智能引论》大班理论课 —— 分模块化

- 横向分模块化教学
- 发挥教学团队中, 各主讲教师自身
- 更贴近自身一线科研经历和所长
- 结合自身专长, 细化教学方式
- 学生可以在大班课堂感受不同老师的特长

《人工智能引论》大班理论课在线教学

- 多位教师上课的问题
 - 讲述具有差异性, 学生适应起来比较辛苦
 - 相互参考, 取长补短, 利用在线教学进行对齐与探讨
- ClassIn平台用于大班课程, 提前准备录播视频
 - 为开放平台课程打下基础
- 大班作业的布置与在各小班的对应follow相结合

尝试: 1门大班课 + X 门实践课

《人工智能引论》小班实践课

编 号	班级名称	
1	游戏AI	AI Game Playing
2	计算机视觉	Computer Vision
3	计算机听觉	Speech and Hearing
4	自然语言处理	Natural Language Processing
5	数据智能	Data Intelligence
6	智能机器人	Intelligent Robotics
7	可信机器智能	Trustworth Machine Learning

尝试: 1门大班课 + X 门实践课

《人工智能引论》小班实践课

编 号	班级名称	
1	游戏AI	AI Game Playing
2	计算机视觉	Computer Vision
3	计算机听觉	Speech and Hearing
4	自然语言处理	Natural Language Processing
5	数据智能	Data Intelligence
6	智能机器人	Intelligent Robotics
7	可信机器智能	Trustworth Machine Learning

人工智能引论大班课程



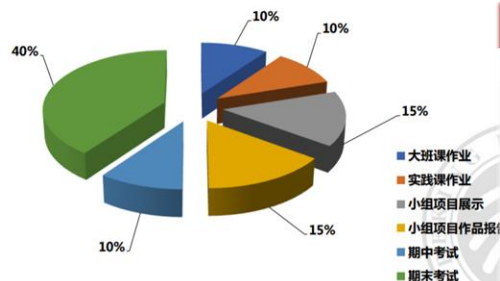
北京大学
PEKING UNIVERSITY

第一次大班课“闪亮登场”

- 介绍课程相关安排
- 介绍“1 + x”教学模式: 1大班课 + 6门实践课
- 6位老师分享介绍视频 ← 各显神通

人工智能概述

2. 课程设计 课程成绩评定方案



17

人工智能概述

• 作业 —— 《游戏AI》(李文新)

1. 程序阅读及阅读书面
 - 每个学生需要阅读一份Botzone上某个游戏的排名靠前的代码, 并撰写一份绝对报告。会提供报告模板。
2. 深度学习方法模拟某个已知算法
 - 应用深度学习的方法, 模拟某个游戏的某个算法, 并撰写相应的报告。会提供参考代码和对局数据。
3. 期末比赛 战旗(双人、同步游戏)
 - 写一个bot参加多轮比赛, 并在最后提交一个Poster。会提供Poster模板。

人工智能概述

• 实践课内容设计 —— 《智能机器人》(罗定生)

No.	Contents	Hours	Place	Date
Lecture 1	合并大班课	2	TBA	2020.02.20
Lecture 2	合并大班课	2		2020.02.27
Lecture 3	Introduction to Robotics	2		2020.03.05
Lecture 4	Robot Architecture	2		2020.03.12
Lecture 5	ROS & Simulation Platforms	2		2020.03.19
Lecture 6	Sensing and Sensors	2		2020.03.26
Lecture 7	Kinematics and Dynamics	2		2020.04.02
Lecture 8	Robot Motion Skills	2		2020.04.09
Lecture 9	Group Construction	2		2020.04.16
Lecture 10	Intelligent HRI	2		2020.04.23
Lecture 11	Embodied Intelligence	2		2020.04.30
Lecture 12	Developmental Robotics	2		2020.05.14
Lecture 13	AI Inspired from HI	2		2020.05.21
Lecture 14	Group Presentation (大班全体)	2		2020.05.28
Lecture 15	Summary (大班全体)	2		2020.06.04
Lecture 16	Final Exam (大班全体)	2		2020.06.08

29

人工智能概述

内容设计 —— 《计算机视觉》三教204 (刘家瑛)

引论实践课——计算机视觉

计算机视觉与机器智能
Computer Vision and Artificial Intelligence





北京大学
PEKING UNIVERSITY

困惑三：

03

人工智能实践小班课，
怎么玩才有兴趣？

计算机视觉相关课程



北京大学
PEKING UNIVERSITY

- [大二]: 04833510 视觉计算与深度学习
- [大三]: 04833970 计算摄像学: 成像模型理论与深度学习实践
- [大三]: 04833580 计算机视觉中的机器学习
- [大三]: 04814150 计算机视觉
- [研究生]: 04802019 计算视觉理论、模型与方法

有感性认知, 有收获有趣的课程

计算机视觉实践课 · 课程目标



北京大学
PEKING UNIVERSITY

- **发现自身的科研兴趣**
 - 喜欢 Vs. 不喜欢
- **培养科学素养**
- **锻炼科研方法**
 - 从兴趣入手
 - 系统性科研探索



计算机视觉实践课 · 科研兴趣



北京大学
PEKING UNIVERSITY

STRUCT Group

智能影像计算

北京大学 王选计算机研究所

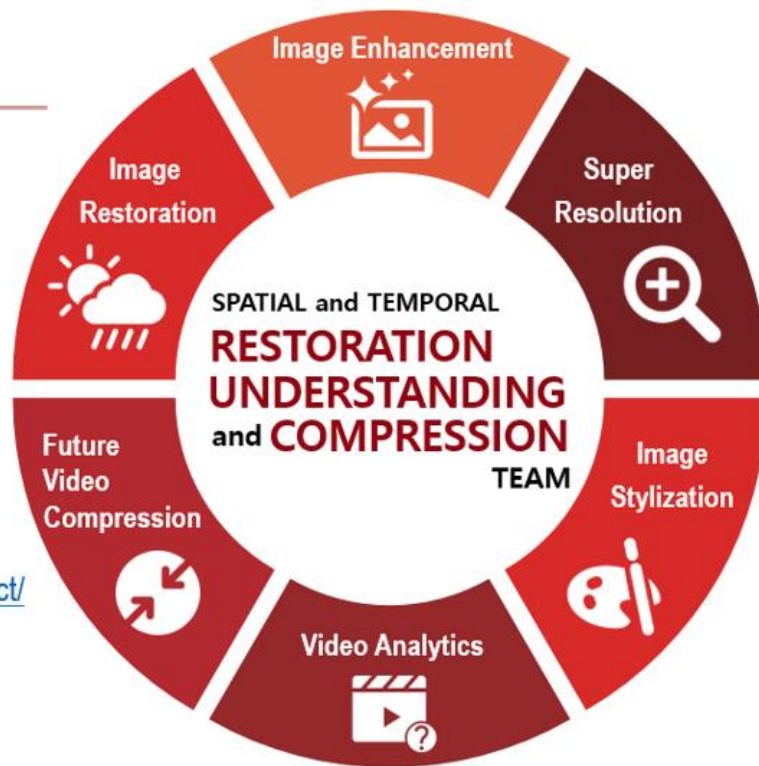
视频信息处理研究组

Spatial and Temporal Restoration,
Understanding and Compression Team

• PI: 刘家琪

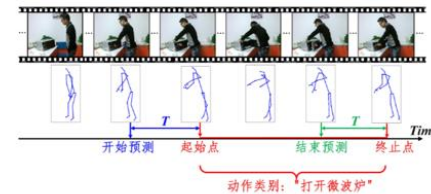
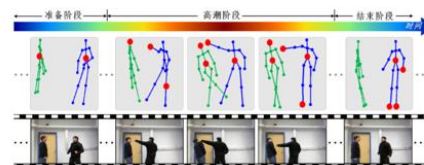
• 邮箱: liujiaying@pku.edu.cn

• 网页: <http://www.wict.pku.edu.cn/struct/>

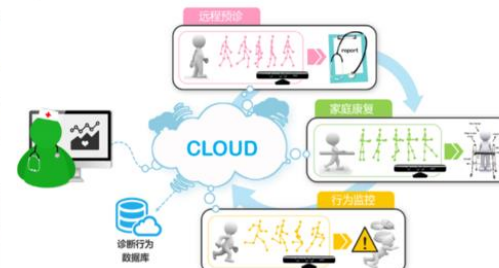


Intelligent Image/Video Editing

- Image / Typography Stylization
- Image / Video Restoration
- Image / Video Enhancement, i.e. derain, dehaze, contrast enhance



Video Action Analytics



计算机视觉实践课 · 课程安排

课 程	主讲内容	其他亮点
Lecture 1	Introduction & Perception	
Lecture 2	Image Transformations	
Lecture 3	Detector and Descriptors	
Lecture 4	Motion and Optical Flow	
Lecture 5	Object Detection	
Lecture 6	Image Restoration	STRUCT小组: 影像修复研究
Lecture 7	Image Stylization	STRUCT小组: 图文风格化研究
Lecture 8	Action Analytics	STRUCT小组: 多模态行为分析研究
Lecture 9	Applications	- 2019: 自动驾驶新技术 - 2020: 时尚与视觉 - 2021: 快手短视频中的影像编辑技术
Lecture 10	Project Presentation	

理论与实践相结合



北京大学
PEKING UNIVERSITY

- 每次实践课 分上下篇
 - 上篇讲**视觉技术**
 - 下篇讲**实践操作**
- 不求知识点满堂灌
- 只求点点如心

直方图均衡化

```
def hist_equ(gray):  
    '''Conduct histogram equalization for gray scale image'''  
    hist = np.histogram(gray, 256, [0, 256])  
    norm_hist = hist[0] / (gray.shape[0] * gray.shape[1])  
    integral = np.cumsum(norm_hist) # Cumulative Sum  
    integral = (integral * 255.).astype(np.uint8)  
    # now integral is a transformation function  
    result = integral[gray] # Pixel-Wise mapping  
    return result
```

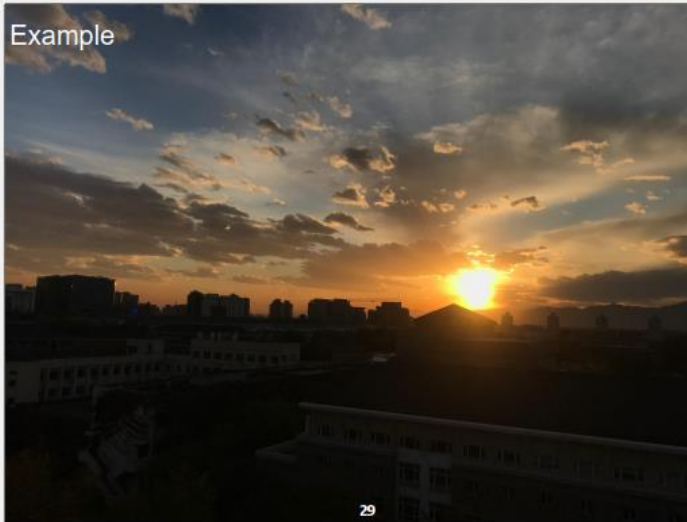
27

存在的问题：颜色异常

```
import cv2  
import numpy as np  
  
def hist_equ(gray):  
    ...  
  
img = cv2.imread('./example.jpg')  
b, g, r = cv2.split(img) # 默认按最后一维拆分  
nb, ng, nr = [hist_equ(gray) for gray in [b, g, r]]  
result = np.stack([nb, ng, nr], -1)  
cv2.imwrite('hist_equ.png', result)
```

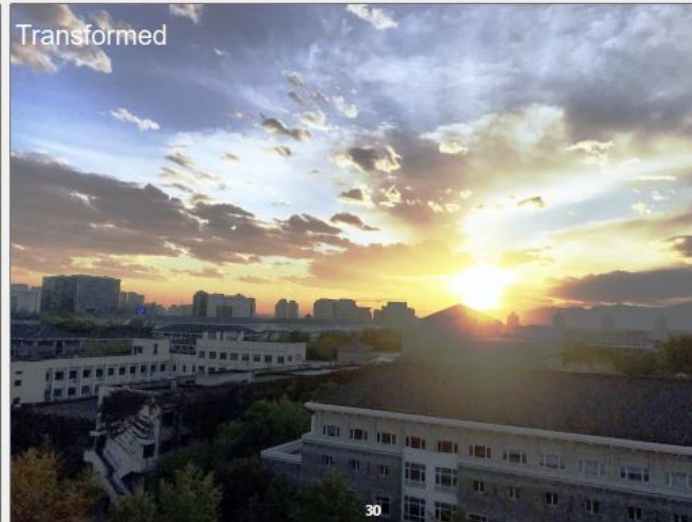
28

Example



29

Transformed



30

理论与实践相结合



北京大学
PEKING UNIVERSITY

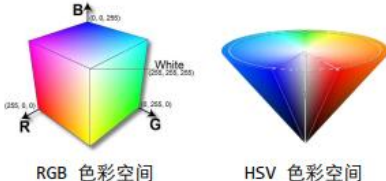
- 每次实践课 分上下篇
 - 上篇讲**视觉技术**
 - 下篇讲**实践操作**
 - 不求知识点满堂灌
 - 只求点点如心

上节课: OpenCV 基础图像处理

> OpenCV 色彩空间变换

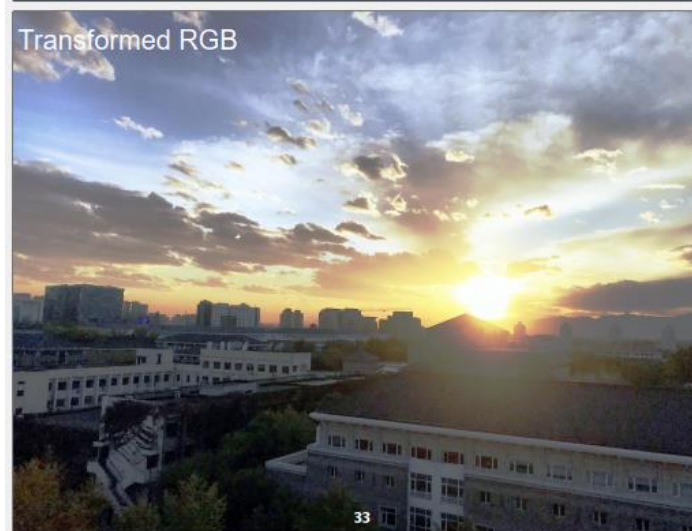
`cv2.cvtColor(src, code) → dst`

例子, 由RGB (OpenCV默认BGR) 转换到 HSV 空间
HSV: Hue, Saturation, Value
`hsv_img = cv2.cvtColor(img, cv2.COLOR_BGR2HSV)`



RGB 色彩空间 HSV 色彩空间

31



换用HSV色彩空间

```
import cv2
import numpy as np

def hist_equ(gray):
    ...

img = cv2.imread('./example.jpg')
hsv = cv2.cvtColor(img, cv2.COLOR_BGR2HSV) # to HSV
h, s, v = cv2.split(hsv)
new_v = hist_equ(v)
new_hsv = np.stack([h, s, new_v], -1)
result = cv2.cvtColor(new_hsv, cv2.COLOR_HSV2BGR) # to BGR
cv2.imwrite('hist_equ_hsv.png', result)
```

32



- 学会使用各种工具 和 阅读资料

→ 实践课程好的开始

安装 python 和 pip

Go to <https://www.python.org/downloads/>.

Download Python installation executables.

Python 3.6.8 Recommended

Then go to <https://pip.pypa.io/en/stable/installing/>.

Installing with get-pip.py

To install pip, securely download [get-pip.py](#). [1]:

```
curl https://bootstrap.pypa.io/get-pip.py -o get-pip.py
```

Then run the following:

```
python get-pip.py
```

Anaconda is also recommended.

安装 python 和 pip

Install `numpy` and OpenCV packages

```
> pip install numpy --user  
> pip install opencv-python --user  
> pip install scikit-image scikit-learn --user  
# READ DOCS https://pip.pypa.io/en/stable/installing/  
# Having Trouble ? Google IT!
```

Further Reading

Python & `numpy` Tutorial:
<http://cs231n.github.io/python-numpy-tutorial/#python-basic>

Official Python Tutorial:
<https://docs.python.org/3/tutorial/>

Official `numpy` Tutorial:
<https://docs.scipy.org/doc/numpy/user/quickstart.html>

有趣的小班作业



北京大学
PEKING UNIVERSITY

• 不走寻常路

小班课作业一

已附加文件: [rgb.zip](#) (25.76 MB)
[Day_and_Night_benchmark.zip](#) (38.481 MB)

第一次小班课编程作业, 截止时间为 3月29日 23:59

- 编程, 将附件中rgb.zip中的三张单通道图像, 拼合成一张正确的彩色图像并保存成文件, 提交代码和产生的图像。
- 寻找直方图和课件中展示的两张直方图相似的图像各一张, 提交图像和对应的直方图。
- (选做) 附件day_and_night_benchmark.zip中的benchmark.py程序提供了“日景-夜景”分类任务的评价程序, 阅读、修改完成该程序, 实现夜景和日景分类。

所有代码和图片打包成1.zip提交。



编程作业

- 编写程序: 程序读入目录下的 `r.png`, `g.png`, `b.png` 文件, 分别对应三个通道的信息, 将三个通道拼合成彩色图像, 输出为 `ans1.png`
- 以下给定了2个具有特点的直方图, 请你通过各种方式获取得到直方图相似的图像, 提交这些图像和对应的直方图。

小班课第一次作业

• 图像和直方图

有趣的小班作业



北京大学
PEKING UNIVERSITY

- 不走寻常路
- 作业来回做



小班课作业二

第二次小班课编程作业，截止时间为 4月5日 23:59

MNIST (手写字体数据集) 分类任务数据集

下载地址：<http://yann.lecun.com/exdb/mnist/>

要求：

1. 本次作业要求设计程序提取特征，然后使用分类器进行分类
2. 不允许使用神经网络模型
3. 请提交简要实验报告和代码，打包成zip 提交



小班课作业三

第三次小班课编程作业，截止时间为 4月19日 23:59

实现人工神经网络，求解MNIST任务

- 利用上一个作业构造的数据，实现数据加载器
 - 构造神经网络模型，包含卷积层和全连接层
 - 训练网络并且存储模型
 - 实现代码，利用存储的模型进行预测，计算准确率
- 请提交简要实验报告和代码，打包成zip 提交

项目大作业 · 举例



北京大学
PEKING UNIVERSITY

Project Assignment: Style Transfer

Class of Computer Vision

April 2019

1 导引

图像风格迁移 (Style Transfer) 旨在增强图片的艺术性。常见的应用有：将摄影图片转换成油画风格，将人物照片转换成素描图等。

Neural Style Transfer是Gatys等人在2016年提出的基于CNN的风格迁移方法。与Photoshop中的各种滤镜不同，Neural Style Transfer不是进行「摄影→油画」这种绘画大类的转换，而是将特定绘画作品的风格迁移到图像上，例如：将梵高的《星空》的风格迁移到一副摄影图片上，让摄影图片有和《星空》一致的黄蓝颜色和弯曲的笔触。



图1. 将梵高《星空》的油画风格迁移到一张月色湖水风景图上。

想要快速尝试图像风格迁移，可以尝试：Prisma APP (www.prisma-ai.com) (ios/安卓)



图2. Prisma APP 效果展示。

2. 任务一（基本任务）：将给定的未名湖摄影图片转换成水墨画风格（分数占比暂定80%）

给定未名湖摄影图片（下载地址：<http://39.96.165.147/TemporalItems/pku.jpeg>），要求将其转换成水墨画风格。



图3. 左：所给定的未名湖摄影图片。右：水墨画风格迁移后的结果（可供参考）。

不限制方法，可以自由选择：

1. 「推荐」Neural Style Transfer。
2. 查阅并参考其他方法。
3. 自己设计一套转换规则（例如先把图片从彩色变成黑白，然后再用一些滤波）。

项目要求

1. 熟悉一种图像风格迁移方法并实现（或自己设计一套转换规则），程序语言不限。
2. 必须保证风格转换全部通过程序实现。不允许使用Photoshop等美工软件进行辅助。
3. 如果借鉴了相关文献或者Github项目，须了解清楚其原理。
4. 所给摄影图片的分辨率为512x512，若计算资源不足，可以将分辨率降低为256x256。

评分标准

1. 美学主观质量（人眼观察）

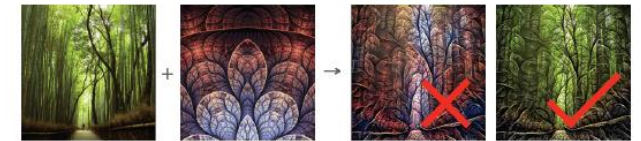
参考资料

1. Gatys, Leon A., A. S. Ecker, and M. Bethge. "Image Style Transfer Using Convolutional Neural Networks." Computer Vision & Pattern Recognition 2016.
2. Huang, Xun, and S. Belongie. "Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization." 2017.

2. 任务二：对已有方法进行改进，或与其他应用相结合（分数占比暂定20%）

对已有方法进行改进思路提供：

1. 尝试做到下面的效果：保持颜色不变，但是将纹理迁移。



2. 融合两种完全不同的风格，例如素描和油画。
3. 加速算法，实现实时风格转换。

将已有方法与其他应用相结合思路提供：

1. 视频风格转换
2. 交互网页
3. 扩展到移动端/Web

其他可选思路：

1. 针对其他画种改进转换效果，例如：卡通、插画等。

项目要求

1. 实现1-2种改进或与1种应用相结合。
2. 如果借鉴了相关文献或者Github项目，须了解清楚原理。

评价标准

1. 创新性
2. 美学主观质量（人眼观察）

参考资料

1. Gatys, Leon A., et al. "Preserving Color in Neural Artistic Style Transfer." (2016).
2. Ruder, Manuel, A. Dosovitskiy, and T. Brox. "Artistic style transfer for videos." (2016).
3. Jing, Yongcheng, et al. "Neural Style Transfer: A Review." (2017).

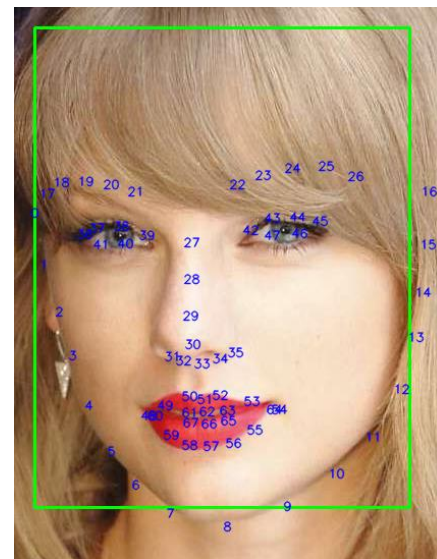
Project B 自动脸部贴纸



北京大学
PEKING UNIVERSITY

- 项目任务

- 1) 人脸检测：判断图像中是否存在人脸，找出人脸的位置
- 2) 人脸关键点检测：定位人脸中的关键点的物理位置
- 3) 贴图：在人脸图像上标定贴图位置，使贴图自然
- 4) (选做) 实时人脸贴图系统



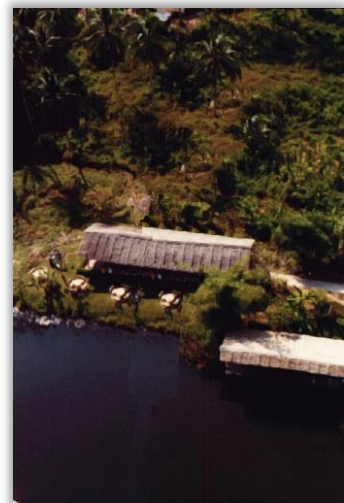
Project C 图像内容补全



北京大学
PEKING UNIVERSITY

- 项目任务

- 1) 给定图像和对应要去除区域的mask, 首先去除该区域所有像素, 然后根据周围信息补全这部分缺失的图像
- 2) 给定基本参考文献, 阅读并且实现描述的方法
- 3) 阅读文献, 改进方法, 或者实现一个完整演示系统



作品案例展示 - A



北京大学
PEKING UNIVERSITY



Original
Photo



Original
Neural Style Transfer



Neural Style Transfer
With Bilateral Filter



AdaIN



AdaIN
With Texture Multiplied



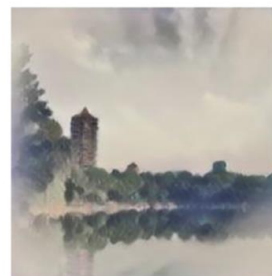
CIN
With Bilateral Filter



CIN
(Photoshop Processed)



CIN
With Masks



CIN
With Masks And Bilateral Filter

基于水墨风优化的 自由风格转换

周鑫渝, 张文博, 李臻

1700013222@pku.edu.cn, 1800012981@pku.edu.cn, 1800013025@pku.edu.cn

指导教师: 刘家瑛

北京大学, 人工智能引论课程

2018-2019, 春季学期

摘要

在此次项目中, 我们呈现了一种兼具相当自由度和较快速度的风格迁移方法。这个自由风格迁移神经网络建立在将任意风格图片映射到向量的 Style Prediction 网络和风格转换的 Style Transfer 网络上。

同时, 针对水墨画风格的转换, 我们在该模型的基础上尝试了多种优化方式, 尝试强化水墨画中墨水的特殊笔触效果和写意风格。

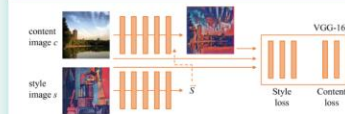
关键词: style transfer, arbitrary style, ink-painting style

引言

此前 Style transfer 的工作已经取得了长足的进步。最初我们尝试了 neural style transfer, 将图像的内容信息和风格信息分离, 从而实现风格转换。进一步扩展这个项目的想法, 我们引入一种快速、任意的艺术风格转换算法。通过 80000 幅绘画及纹理图片的预训练, 尝试捕获任意绘画或纹理的语义特征, 并以此完成风格转化。

此外, 不同于西方注重写实的绘画风格, 水墨画由于其独特的笔法和意境使得简单的风格转换并不能达到理想效果。因此, 我们通过使用一些图像处理手段在水墨画艺术范围进行了更多探索。

方法



丰富的文献表明, 如果图像识别系统提取的低级特征共享相同的空间统计信息, 则两个图像在样式上相似; 绘画风格可以被视作视觉纹理, 其重复图案可能具有低阶空间统计特征。通过用 Gram 矩阵表示图像风格信息, 图像识别网络的低级特征表征内容, 样式转换的完整优化目标可以表示为:

$$\min_c \mathcal{L}_c(x, c) + \lambda_s \mathcal{L}_s(x, s)$$

我们建立一个风格转换网络, 该网络通过提供代表任何绘画语义共享。

$$\hat{z} = \gamma_s \left(\frac{z - \mu}{\sigma} \right) + \beta_s$$

风格转换网络采用了一个预训练的 inception-v3 网络结构计算, 在其顶部放置两个完全连接的层, 以预测最终的向量。



另外, 在水墨画风格的转换中, 我们尝试对画面主体进行检测, 添加一个椭圆蒙版和图片主题的蒙版使得背景淡化, 减少背景特征在最终输出时的体现, 并且在输出图像上使用 Bilateral filter, 弱化笔触内部纹理。

实验



其他内容 / 风格图片的生成

参考文献

- [1] Gatys, Leon A., A. S. Ecker, and M. Bethge. "Image Style Transfer Using Convolutional Neural Networks." CVPR, 2016.
- [2] Huang, Xun, and S. Belongie. "Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization." ICCV, 2017.
- [3] Golnaz Ghiasi, Honglak Lee, Manjunath Kudlur, Vincent Dumoulin, Jonathon Shlens. "Exploring the structure of a real-time, arbitrary neural artistic stylization network." BMVC, 2017.

Demo



作品案例展示 - B



北京大学
PEKING UNIVERSITY



实时人脸贴纸：基于3D人脸模型

唐雯豪(1800013088), 韩宇栋(1800013097), 崔轩宁(1800013083)
联系邮箱: tangwh@pku.edu.cn, hanyd@pku.edu.cn, cxn@pku.edu.cn

指导教师: 刘家瑞

北京大学·人工智能导论课程
2018-2019, 春季学期

引言

Faceu和B612等相机APP深受青少年的欢迎, 原因之一便是其中丰富的贴纸选择。面部贴纸(facial filters), 是指通过计算机视觉方法将图案贴到人脸上。提出了一个基于3D人脸模型的人脸贴纸添加方案:

- ① 使用MTCNN进行人脸检测
- ② 使用3DDFA进行3D人脸关键点标注
- ③ 利用3D人脸信息添加贴纸, 并映射变换到原图片

此方案的优势在于:

- ① 使用3D信息, 贴纸添加在大角度侧脸上也有很好的表现。
- ② 拥有优秀的实时性能, 在GPU上可达200fps甚至更高, 在CPU上可达10fps。同时, 实现了一个带有GUI界面的面部贴纸添加Demo, 支持实时视频、拍照以及拖拽图片, 并且用户可以自由增减不同贴纸。

关键词: Face Detection, Facial Landmark Detection, 3DMM

方法 Part I

Face Detection:

- 人脸检测是计算机视觉的一部分, 经过对多种方法的测试, 选取了MTCNN作为人脸检测方法。

检测器	FPS
MTCNN	23.634
YOLO-V2	4.504
RFCh-RetNet101	4.564
TinyYOLO	16.911
SSD	6.902

不同方法测试结果 (测试环境: GTX1050Ti with Win-Q Design)

■ MTCNN检测精度和速度均优于其它方法, 结果如图:



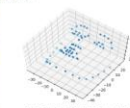
方法 Part II

Facial Landmark Detection:

- Motivation: 常见的2D平面关键点检测算法对于大角度侧脸鲁棒性不好
- 使用了3D Dense Face Alignment (3DDFA) 来检测人脸特征点标注
- 通过Cascaded Convolutional Neural Networks实现使用Dense 3D Morphable Model (3DMM) 人脸模型拟合人脸
- 3DMM人脸模型的公式描述:
$$S = S_0 + A_{10} \alpha_{10} + A_{100} \alpha_{100}$$
- 拟合过程使用了Cascaded Regression的方法, 其中每个Regressor是一个CNN, 使用了相同的网络结构, 回归过程公式如下:
$$\Delta p^i = \text{Net}^i(\text{PAF}(p^0, 1), \text{PNCC}(p^0, 1))$$
- 网络的损失函数使用了OW-PCD

Composing Filters with Images:

- 得到3DMM人脸模型之后, 建立一个以人脸为中心的三维坐标系, 并得到68个人脸关键点在三维空间中的坐标



- 在三维坐标中通过固定一个锚点和一个缩放尺度的方式确定和锚点4个点的坐标, 利用参数化函数将这四个点投影到相机坐标系中, 最后通过射影变换将人脸图片变换到原图平面上, 公式如下:
$$V(X) = f \cdot P \cdot R \cdot X + t$$
- 对于加速拟合过程时问题, 采取了一些代码层面的优化措施, 比如: 分离alpha通道, 只对非透明区域进行处理等。

结果



总结

提出了一个使用MTCNN检测人脸, 使用3DDFA得到3D人脸模型并添加人脸贴纸的方案流程, 并且实现了一个带有GUI界面的面部贴纸添加Demo, 支持实时视频、拍照以及拖拽图片添加贴纸, 用户可以自由增减不同贴纸。该方法有着良好的实时性能, 并且对于大角度的侧脸也有很好的鲁棒性, 最终的效果可以跟Faceu, B612等APP媲美。在这次项目中, 我们学会了很多人脸检测和人脸关键点检测的基于深度学习的方法与空间射影几何的基础理论, 同时编写代码和GUI界面的能力有了很大的提升。

下一步的改进有三个方向: 检测精度连续性的优化、性能提升和三维贴纸的添加。

参考文献

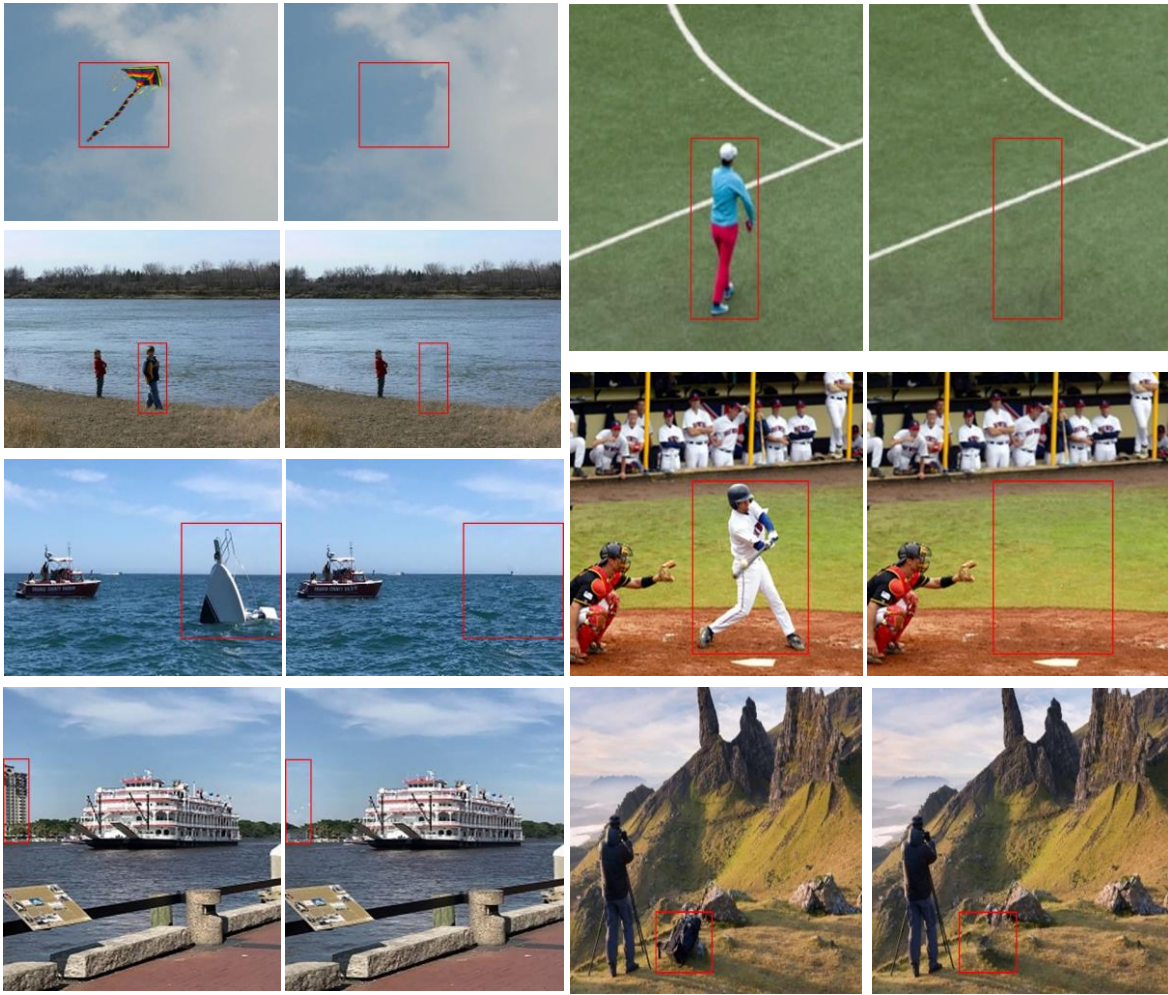
- [1] Zhang, Kaipeng, et al. "Joint face detection and alignment using multitask cascaded convolutional networks." IEEE Signal Processing Letters 23 (2016): 1499-1503.
- [2] Zhu, Xiangyu, et al. "Face alignment in full pose range: a 3D total solution." IEEE transactions on pattern analysis and machine intelligence 41.1 (2019): 78-92.



作品案例展示 - C



北京大学
PEKING UNIVERSITY



图像内容补全(Inpainting_Project)

郑元昊(1800013094), 杨炎峰(1700011436), 徐雨晨(1800013091)
联系邮箱: 1800013094@pku.edu.cn, 1700011436@pku.edu.cn, 1800013091@pku.edu.cn

指导教师: 刘家瑞

北京大学·人工智能引论课程
2018-2019, 春季学期

摘要

实验内容为给定图像和对应要去除区域的mask, 首先去除该区域所有像素, 然后根据周围信息补全这部分缺失的图像, 且在视觉上看似合理。我们的项目结合基于样本的纹理合成和填充图像中的细缝两种方法, 算法确定置信度优先级从而决定填充顺序, 通过基于样本块的图像修复实现了大的目标物体的移除, 合理而高效地还原了真实场景的线性结构和复合纹理结构, 并在实验速度和效果上进行了优化改进, 实现了即时交互与网页展示。

关键词: 物体移除, 图像修复, 区域填充

方法



区域填充算法:

1. 计算修补优先级。优先级决定填充顺序, 倾向于优先填充边缘和线性结构的块。
 $P(p) = C(p)D(p)$ $P(p)$: 优先级, $C(p)$: 置信度, $D(p)$: 数据项
 $C(p) = \frac{\sum_{q \in \Omega_p} \phi(q)}{|\Omega_p|}$ $D(p) = \frac{|\Omega_p|}{\alpha}$ α : 归一化因子
 $C(p) = 0, p \in \Omega$; $C(p) = 1, p \in \Omega$
2. 传播纹理和结构信息。从源区域 ϕ 提取优先级最高的补丁, 其中 d 为两个图像块中已填充像素的平方差之和(SSD)。
3. 更新置信度。

优化

随着大面积填充的进行, 原算法中中间区域的置信度迅速下降至0, 使得优先级项在后期填充的过程中不再可信, 容易出现中央区域纹理断层的现象。
改进方式:
 $P(p) = \alpha C(p) + \beta D(p)$
其中 α, β 为对应图片的超参数。对于每张图片利用穷举法得到最好的结果。
通过将乘法变换为加法, 避免了当置信度项趋于零时优先级中数据项的作用无法体现, 从而更好地保持原图的纹理信息。

实验

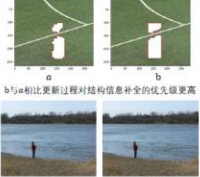


Case 1: 纹理还原

Case 2: 结构修复

Case 3: 结构+纹理补全

对比



b与a相比更新过程对结构信息补全的优先级更高
c与d相比对纹理(水波)信息还原的效果更好

总结

我们的项目较好地实现了图像内容补全的功能, 使用基于样本块的图像修复技术实现遮挡物移除的目标, 算法有效地确定目标区域的填充顺序, 可将线性结构和二维纹理传播到目标区域, 从合成图像到包含复杂纹理的图片, 均能有效补全。此外, 进行了方法优化、交互与网页展示。

未来展望:

- 对匹配块的搜索采用分块搜索的方式, 减少计算量的重复, 进而提升效率
- 对与目标块相似的一组匹配块进行线性组合, 得到更好的补全效果
- 鲜果为完整的应用

参考文献

[1] Y.A. Criminisi, P. Pérez, K. Toyama, "Region Filling and Object Removal by Exemplar-Based Image Inpainting", IEEE Transactions on Image Processing (TIP), September 2004, 13:1200-1212

网页展示



答辩现场



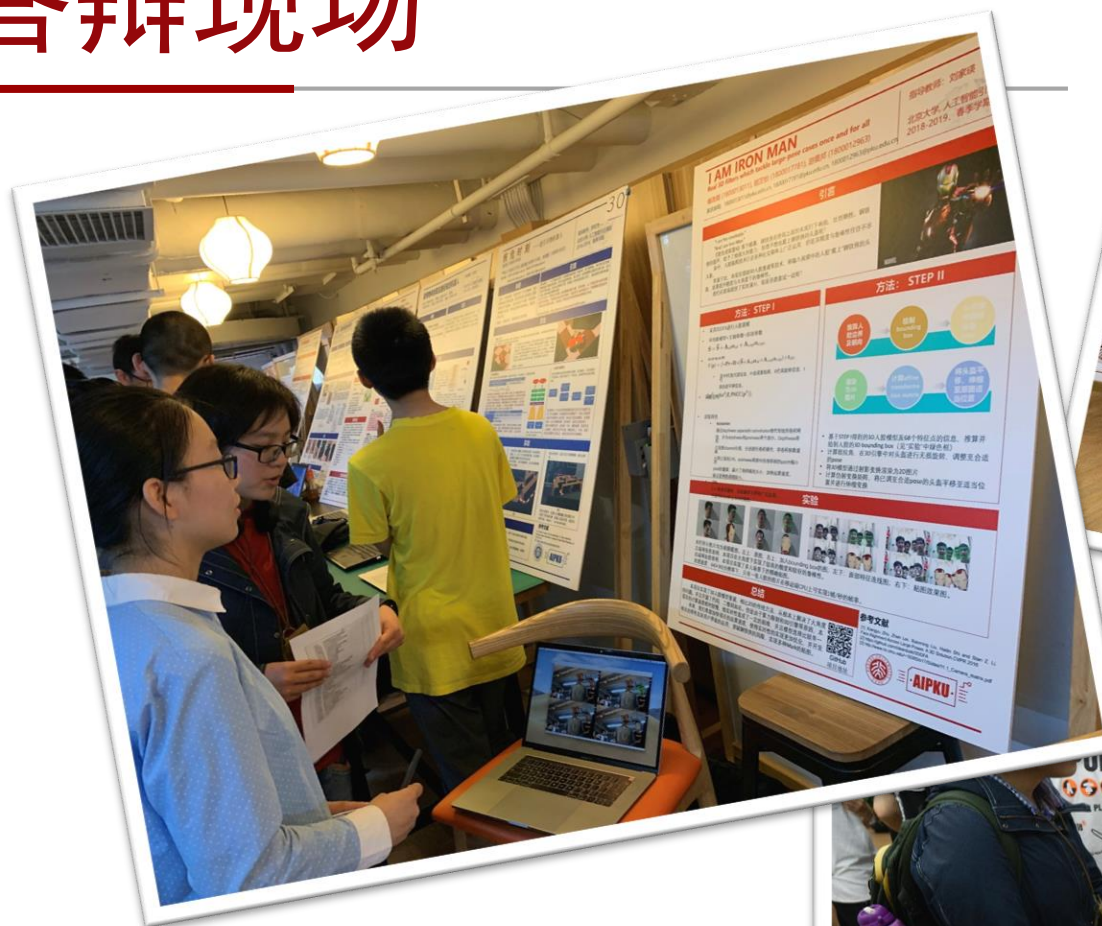
北京大学
PEKING UNIVERSITY



答辩现场



北京大学
PEKING UNIVERSITY

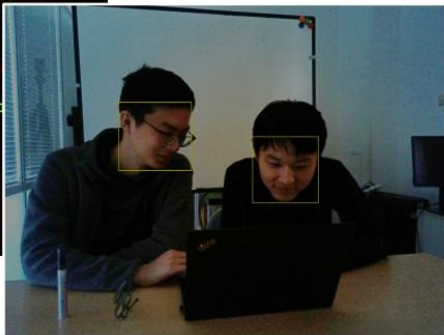
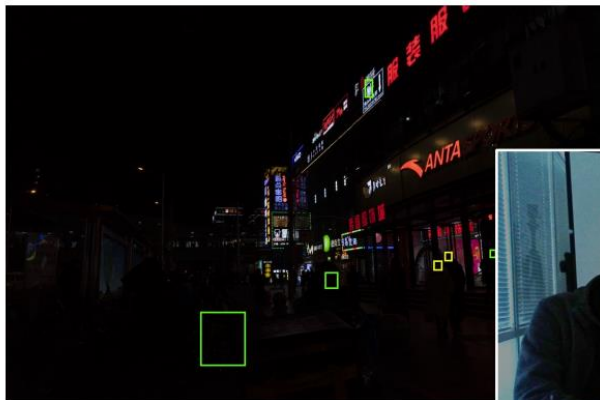


实战CV竞赛

- 用CVPR (UG2+) 举办多年的Challenge作为课程大作业
- 组队实战刷榜

Project: 低光人脸检测

- 检测低光照条件下的人脸



低光人脸检测数据集

- DarkFace
 - 链接: <https://flywh.github.io/CVPRW2019LowLight/>
 - 训练集: 6000张带标注的真实世界低光图像
 - 测试集: 4000张带标注的真实世界低光图像 (标注非公开)

DARK FACE: Face Detection in Low Light Condition
Spatial and Temporal Restoration, Understanding and Compression Team
Wangxuan Institute of Computer Technology, Peking University



实战CV竞赛



北京大学
PEKING UNIVERSITY

- 走通视觉研究全流程
- 数据收集 互助搭建新数据集
- 众筹标数据

低光人脸检测数据集

* DarkFace

* 链接: <https://flyywh.github.io/CVPRW2019LowLight/>

* 训练集: 6000张带标注的真人世界低光图像

* 测试集: 1000张带标注的真人世界低光图像

互助搭建新数据集

CVFace22

20

PKU-CVFace 数据情况

预计收集: $68 \times 20 = 1,360$

第一批合格数据: 964张 补交数据: 528张

共1,492张, 超额完成!

PKU-CVFace 数据情况

1,360

1,492

DARK FACE 数据集: 10,000张

数据准备: 低光数据拍摄

- * 小班作业2: 拍摄20张低光图像 (DDL: 4月17日)
- * 含有人脸内容
- * 避免重复内容 (不同图像间有80%以上的区域不重复)
- * 拍摄真实人脸: 避免海报、海报上的人脸
- * 避免人脸太小、太暗等问题
- * 注意肖像权

21

数据准备: 如何拍摄低光数据?

- * 设备: 手机或者相机均可
- * 实现方式
 - * 夜间/低亮度空间内直接拍摄
 - * 调整拍摄参数 (手机相机Pro模式)
 - * 调低EV值
 - * 调低ISO
 - * 调低快门时间
- * 图像格式
 - * RGB彩色图像 (非RAW格式)
 - * 推荐导出为PNG (JPG也可以)



22

数据准备: 低光数据拍摄

- * 一些样例



DarkFace

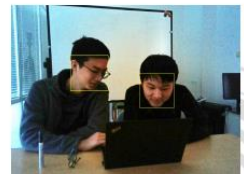


个人拍摄

23

数据准备: 低光数据标注

- * 在平台上对人脸bounding box进行标注
- * 届时在预提亮后的图像上进行标注
- * 后续会发布具体注意事项



24

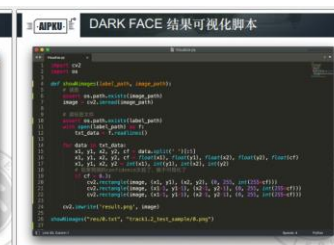
实战CV竞赛



北京大学
PEKING UNIVERSITY

• 走通视觉研究全流程

- 数据收集 互助搭建新数据集
- 众筹标注数据
- 和国际选手同台
- 细化完善自己的故事



课程解锁其他技能



北京大学
PEKING UNIVERSITY

- 查阅资料
- Latex使用 & Overleaf多人编辑
- Poster的设计与表示
- Oral和Spotlight的设计与表示
- CV研究领域的最新工作与会议
- 学会GitHub来组织多人工程



LaTeX 样例

已附加文件: [LaTeX样例] Intelli_Survey.zip (4.059 MB)

附件的压缩包是一个完整的 LaTeX 项目，其中包含了编译完成的 PDF 文件。

LaTeX 的本地配置可参考互联网上 TexLive 的安装文档。

Overleaf 提供了在线 LaTeX 编辑，共享和生成功能，本样例也在 Overleaf 上有提供：

<https://www.overleaf.com/read/tbkjhrtshwm>



北京大学
PEKING UNIVERSITY

04 | 困惑四： 除了视觉实践课， 我们还能玩儿点啥？

一次完美的碰撞

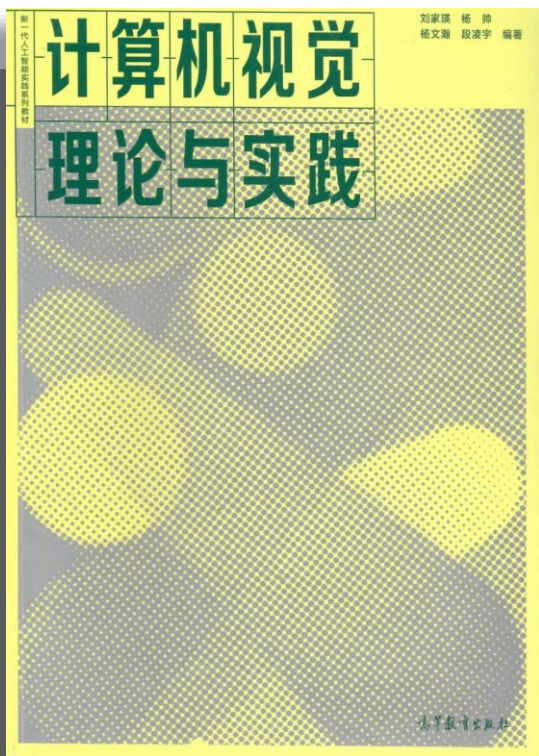


- 华为联合“新一代人工智能系列教材”编委会、教育部高等学校计算机类专业教指委、高等教育出版社共同出版
- 支撑高校人工智能相关专业的教学与实践
- 教材重视培养学生的应用和创新能力

视觉实践类教材编写



北京大学
PEKING UNIVERSITY



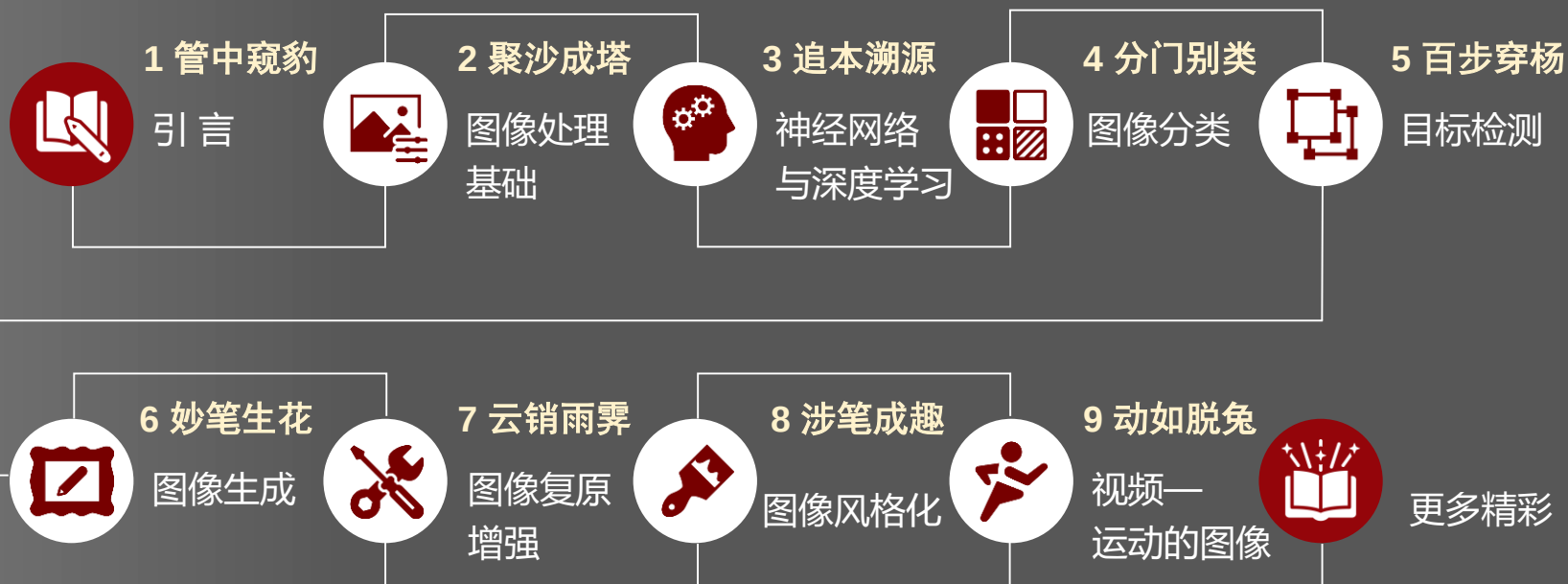
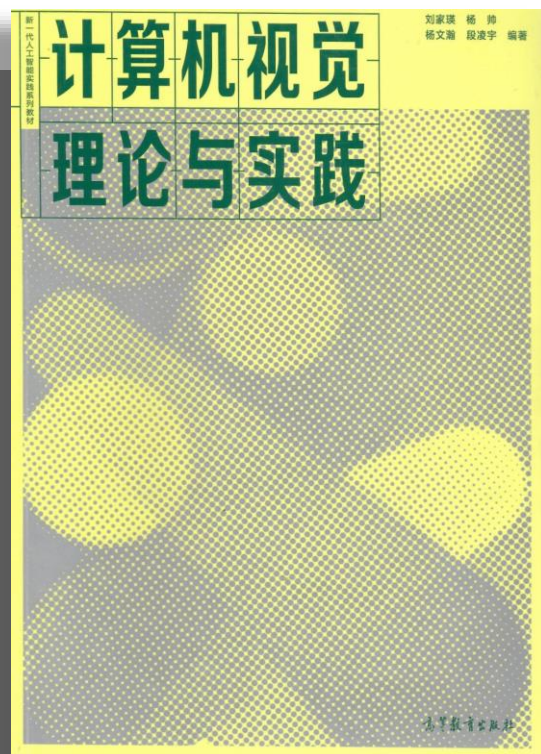
- 面向计算机视觉，聚焦**算法理论与代码实践**，简要介绍了计算机视觉的基本概念和计算机视觉研究的预备知识，分别讲述了图像分类、检测、生成、复原、增强、风格化与视频分析等计算机视觉领域的研究重点
- 深入浅出地介绍了现今已经成熟的算法和技术，并在各个章节辅以大量实例与核心代码详细展示如何使用主流框架搭建计算机视觉模型，对相关算法内容进行有效的补充和深入
- 通过理论与实践的结合，读者能够打下较为坚实的计算机视觉基础，熟悉经典算法框架，具备实现和改进成熟算法的能力，**为进一步在计算机视觉和相关领域提出新设想、开发新算法、解决新问题创造良好条件**

刘家瑛 杨帅 杨文瀚 段凌宇
高等教育出版社

视觉实践类教材编写



北京大学
PEKING UNIVERSITY



刘家瑛 杨 帅 杨文瀚 段凌宇
高等教育出版社

视觉实践类教材编写



北京大学
PEKING UNIVERSITY

教材配套网站

- 校内课程讲义
- 配套示例代码
- 勘误表



<http://39.96.165.147/Books/CV/main.html>



教材内容

序号	章节	校内课程讲义	示例代码	勘误表	助教
1	引言	讲义1			杨帅, 杨文翰
2	图像处理基础	讲义2	Practice_01_Non-Local-Means Practice_02_Canny_edge-detection Practice_03_Hough-transform Practice_04_OTSU		王德昭
3	神经网络与深度学习基础	讲义3	Practice_05_SimpleNN Practice_06_CustomizedLayer Practice_07_SimpleRNN		胡越予
4	图像分类	讲义4	Practice_09_mnist Practice_10_cifar100 Practice_11_Squeeze-and-Excitation Practice_12_Self-Adaptation		汪文靖
5	目标检测	讲义5	Practice_13_SVMDetector Practice_14_YOLOv3Detector Practice_15_FasterRCNNDetector Practice_16_DomainAdaptiveDetector	文档5	胡继章
6	图像生成	讲义6	Practice_17_VAE Practice_19_pix2pix Practice_20_CycleGAN		杨帅, 汪文靖
7	图像复原增强	讲义7	Practice_21_SRCNN Practice_22_JORDER Practice_23_RetinexNet		杨文翰
8	图像风格化	讲义8	Practice_24_NeuralStyleTransfer Practice_25_FastNeuralStyleTransfer Practice_26_TETGAN	文档8	杨帅
9	视频——运动的图像	讲义9			宋思捷

教材信息



书名: 计算机视觉理论与实践

作者: 刘家瑛, 杨帅, 杨文翰, 段凌宇

出版: 高等教育出版社

ISBN: 9787040574463

出版时间: 2022-10-01

《计算机视觉理论与实践》面向计算机视觉, 聚焦算法理论与代码实践, 简要介绍了计算机视觉的基本概念和计算机视觉研究的预备知识, 分别讲述了图像分类、检测、生成、复原、增强、风格化与视频处理等计算机视觉领域的研究重点, 深入浅出地介绍了现在已经成熟的算法和技术, 并在各个章节辅以大量实例与核心代码详细展示如何使用主流框架搭建计算机视觉模型, 对相关算法内容进行有效的补充和深入。通过理论与实践的结合, 读者能够打下较为坚实的计算机视觉基础, 熟悉经典算法框架, 具备实现和改进成熟算法的能力, 为进一步在计算机视觉和相关领域提出新设想、开发新算法、解决新问题创造良好条件。

《计算机视觉理论与实践》可作为人工智能专业和计算机相关专业本科生或研究生学习计算机视觉的教材, 也可作为对计算机视觉技术感兴趣的科研人员、计算机爱好者及各类自学人员的参考书。

更进一步：计算美学



北京大学
PEKING UNIVERSITY

- 视觉与艺术交叉新尝试



Overall Philosophy



- Computer vision is a huge field
- It can impact every aspect of life and society
- It will drive the next information and AI revolution
- An introductory course, we will ***not cover all topics of CV***
- Lectures/Book Chapters are mixture of details techniques and high level ideas
- ***Speak our “Language”***



北京大学
PEKING UNIVERSITY

谢谢！

敬请各位老师批评指导！