

模式识别与计算机视觉

课程简介 &
模式识别系统简介

吴建鑫
南京大学人工智能学院，2026

课程介绍

本学期课程安排

✓ 智汇南雍 <https://lms.nju.edu.cn/>

- 作业在智汇南雍提交

- PPT及其他: 见本课程主页

- ◆ https://cs.nju.edu.cn/wujx/teaching_PR.html

- 交互、答疑、作业批改疑问等

- 课间或通过email联系我: wujx2001@nju.edu.cn

- 或联系助教: 童锦 (1901134679@qq.com)

- 课堂小测验

- ◆ 在智汇南雍在线进行,请准备好手机或电脑, 在课上完成

- 成绩占比高, 请重视!

- 以《模式识别》教材为主

- 将补充部分对计算机视觉基础知识介绍

关于选课、补选

- ✓ 是人工智能学院的必修课程
 - 无力继续加大课程规模
 - 按学校规定正常退选、补选
 - 请勿私人联系选课名额事项
- ✓ 未能选课，但是感兴趣怎么办？
 - 欢迎跟班学习
 - 可以学到知识，只是不能获得学分！

课程目标

- ✓ 理解、记忆模式识别中的基本概念、步骤和方法
- ✓ 对重要方法，要能实际应用，并能理解其前提条件、应用范围、应用注意事项和方法的原理及推导
- ✓ 对讲授的其他方法，要能理解其含义和使用环境
- ✓ 要对模式识别的前沿领域有感性的认识
- ✓ 计算机视觉基础知识介绍
- ✓ 提高目标
 - 进一步能通过独立阅读和实践掌握较深入的问题和方法，并能应用到学习、研究遇到的问题中

模式识别

✓ 前置课程(prerequisite)

- 程序设计、高等数学（数学分析和线性代数）、**概率**和数理统计
- 一般来说，最早2年级下半学期才可以选

✓ 有若干次作业

- 包含纸上解题和一些小的编程(需要Linux软件安装)

✓ 成绩考核包括：期末考试、作业、课堂小测试

- 期末考试：50%，作业+**课堂测试**：50%
- 期末考试形式：**开卷**考试，**不划**重点

教材

✓ 《模式识别》

- 吴建鑫著，罗建豪、张皓译，吴建鑫审校
- 机械工业出版社，书号: ISBN 978-7-111-64389-0
- https://cs.nju.edu.cn/wujx/PR_Book_CN/PR_Book_CN.htm

✓ 英文

- Essentials of Pattern Recognition – An accessible approach, Jianxin Wu
- 剑桥大学出版社, 书号: ISBN 9781108483469
- https://cs.nju.edu.cn/wujx/PR_Book_EN/PR_Book_EN.htm

✓ 中英文教材欢迎纠错，email联系，我确认后可加分

Matlab / python

- ✓ 本课程中用到的编程，很多在Matlab环境中最方便
 - <http://cn.mathworks.com/help/matlab/>
 - 学校已购买Matlab，本校师生均可免费使用
 - 学会随时调用Matlab的帮助（或在线帮助）
- ✓ 如果没有Matlab环境
 - 有很多免费软件，实现Matlab的功能足够本课程使用
 - 语法和编程语言与Matlab一致
 - 如GNU Octave: <https://octave.org/>
- ✓ 也可以用Python/numpy代替

数学背景

- ✓ 会简要介绍/回顾需要的最基本知识
 - 前置课程，（如较好的掌握了）就已经够本课程使用了
 - 要用到的更多的数学知识会在授课中随时补充
- ✓ 对研究或工作有帮助
 - 特别是对有志CS/AI研究的（不限于模式识别），有较大帮助
- ✓ 如对进一步的知识感兴趣
 - 每章会提供关于进一步阅读的资源指南
 - 通用资源：
 - 教材
 - 参考书（DHS, PRML）
 - 维基（Wiki）
 - Wolfram MathWorld (<http://mathworld.wolfram.com/>)

模式识别简介

什么是模式识别?

- ✓ “**Pattern recognition** is a field in machine learning.” (??)
 - 取自2013年12月25日，2016年2月9日仍然如此
 - Pattern recognition (**psychology**), identification of faces, objects, words, melodies, etc.
 - [http://en.wikipedia.org/wiki/Pattern_recognition_\(disambiguation\)](http://en.wikipedia.org/wiki/Pattern_recognition_(disambiguation))
- ✓ In machine learning, pattern recognition is the **assignment** of a **label** to a **given input value**.
 - http://en.wikipedia.org/wiki/Pattern_recognition
- ✓ PRML: **automatic discovery** of **regularities** in **data** ... and with **the use of these regularities** to take actions such as classifying the data into different categories

模式识别的特征

✓ 输入：数据

- data, input value, ...

✓ 输出：“模式” (Pattern)

- faces, objects, words, melodies, (classification) label, categories, regularity, ...

✓ 输入转换到输出：

- 由计算机系统完成(automatic, algorithm, assignment ...)
- 是一个困难但有趣的过程(discovery)

模式识别举例：体感游戏



✓ 输入？

视频来源

✓ 输出？

<https://www.bilibili.com/video/av930246515/>

✓ 困难？

模式识别举例：语音助手

✓ 例子？

✓ 输入？

✓ 输出？

✓ 困难？

模式识别举例：Autopilot



- ✓ 输入？
- ✓ 输出？
- ✓ 困难？

视频来源

<https://www.tesla.cn/>

FSD主页

https://www.tesla.com/en_hk/fsd

从数据到模式 From data to pattern

传感器

- 产生物理读数
- 物理学、声学、电子学、光学、触觉、味觉、...

数字信息

- 文本、图像、音频、视频、三维点集、...

模式

- 物体检测和识别、音频转化为文本、行为识别、姿态识别、...

应用

- 决策、计划、处理、...

(虚拟例子) 性别识别: 特征

✓ 选取什么 **数据** 来判别一个人的性别?

- 长相? 喉结?
- 身高?
- 体重?
- 声音?
- 行为?
- ...



✓ 被 **使用** 的数据或从原始输入数据(raw input data)中提取(extract)的数据称为 **特征**(feature)

- 为什么要选取、提取、或者 **学习** 特征?

(虚拟例子) 性别识别: 评估

- ✓ 怎么知道模式识别的结果好坏?
 - 评估或评价(evaluate)
 - 通常是把模式识别系统的输出(称为预测, prediction)与真实值(groundtruth)进行比较
- ✓ 那么, 真实值从哪里来?
 - 女性的性染色体 $XX \rightarrow X$ (来自母亲) + X (来自父亲)
 - 男性的性染色体 $XY \rightarrow X$ (来自母亲) + Y (来自父亲)
- ✓ 如何进行比较?
 - 以后会讲。现在先想一想?

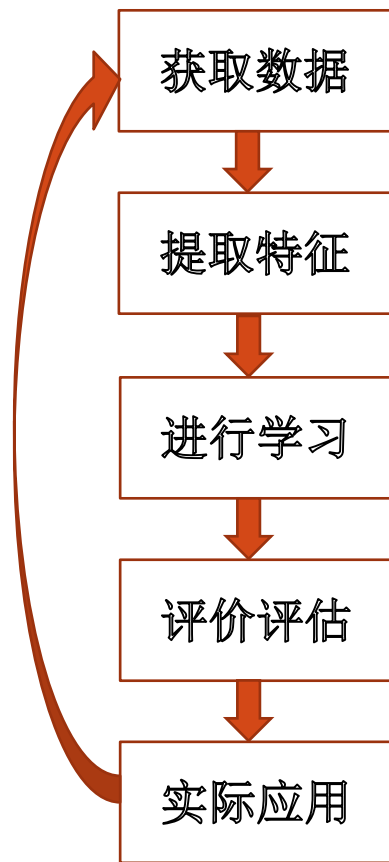
(虚拟例子) 性别识别：转换

- ✓ 如何构建一个模式识别系统将输入转换为输出
 - 常见方法：机器学习(machine learning)
 - 从过去的经验学习 Learning from experience
 - 训练集(training set): 已搜集的数据
 - 模型(model): 从训练集得到的规律，表现形式多样
 - 测试(prediction, testing): 将这些规律应用到新的例子以得到针对该例子的结果（模式识别的输出）
 - 测试集(testing set): 为了评估这些规律搜集的数据
 - 那么，对训练集和测试集应该有什么要求？

（虚拟例子）性别识别：应用

- ✓ 得到输出后，有什么用？
 - 通常在更大的系统中起作用
 - 自动驾驶、智能语音、体感游戏， ...
- ✓ 那么，模式识别系统要做到多好呢？
 - 由那个“更大的系统”决定
 - 是不是越高越好？

小结：模式识别的步骤



问题：

- ✓ 是这样吗？
 - 需要多次反馈 (feedback)、尝试、修改
- ✓ 什么步骤最重要
 - ??
- ✓ 深度学习在图中什么位置？

模式识别为什么困难？

✓ 多种原因

- 语义鸿沟semantic gap,如,一副 2000×1000 的图像是什么？
 - 对人的眼睛和大脑：教师、桌椅、讲台、黑板、人、...
 - 对计算机： $3 \times 2000 \times 1000 = 6,000,000$ 个数字
- 计算能力computational power
 - 很多算法在台式机上要数年或更长时间
 - 大数据，存储和计算都成为问题
- 数据的获取data acquiring
 - 很多时候数据难以获取(如，医学图像medical imaging)
 - 或者虽然容易获取，但是难以标注(如，对Google图像的精确标注annotation)

一只狐狸掉进湖面的冰窟窿里面



OpenAI Sora



Seedance



各步骤的（实践）重要性

1. 提取特征

- 设想一下，告诉你一个人的年纪，需要判别其性别？
- 或者，告诉你一个人的染色体组成，需要判别其性别？

2. 数据获取

- 设想一下，根据图像判断性别，所有训练图像都只包含男子、而不包含任何女子？
- 或者，图像中有男有女，但是所有都标记为女？

3. 机器学习

- 有很多理论上重要的问题，对实践效果也非常重要

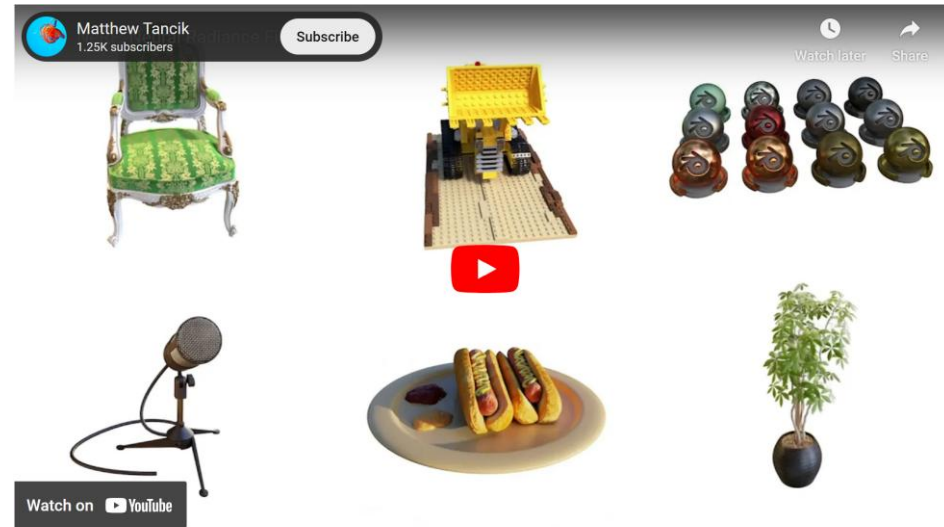
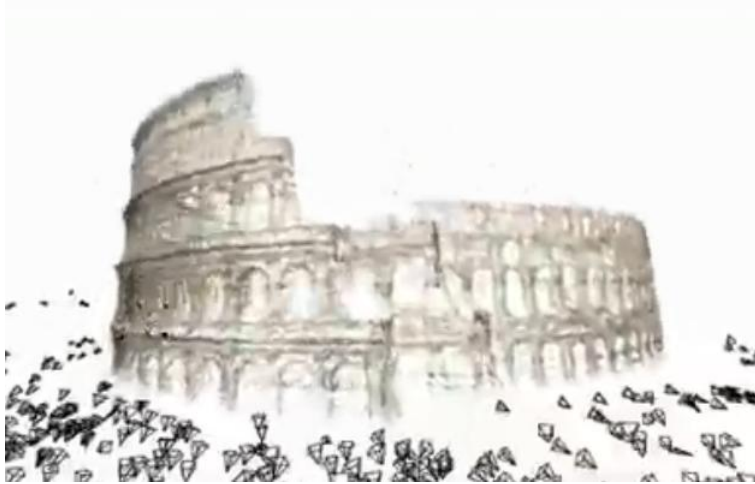
和其他领域的关系

模式识别 vs. 机器学习

- ✓ 机器学习在模式识别中有非常重要的作用
- ✓ 但是，模式识别具有更多的“系统”性
 - 数据获取
 - 提取特征
 - ...
- ✓ 主要的区别是：“数据” vs. “特征”
 - 但是，在深度学习中，这个区别不那么明显了

模式识别 vs. 计算机视觉

- ✓ 模式识别与计算机视觉(computer vision)的研究和应用有非常多的重合
 - 识别recognition是计算机视觉中最重要的问题之一
 - 模式识别中很大部分输入是图像
- ✓ 模式识别包含很多视觉以外的问题
 - 音频、雷达、文本、...
- ✓ 计算机视觉包括很多识别以外的问题
 - 如，超分辨率super-resolution
 - 三维重建3D reconstruction



ECCV Technical Talk Video

http://v.youku.com/v_show/id_XMTUwMDg00TQw.html
<http://grail.cs.washington.edu/rome/>

<https://www.matthewtancik.com/nerf>

模式识别与多媒体(multimedia)

- ✓ 都可能牵涉多种媒体
- ✓ 但多媒体更具有“系统”性
 - 多媒体不特别强调单个模块的性能
 - 更注重整个大系统的成功
 - 比模式识别更强调多种媒体之间的配合
 - 就算每个模块都采用了已有的技术，但是一个科学利用现有技术和多种媒体的系统仍然可以是很大的成功

小结：模式识别vs.其他学科

- ✓ 模式识别与很多学科有广泛的联系
 - 机器学习、计算机视觉、多媒体、可视化(visualization)、图形学(graphics)、...
- ✓ 需要很多学科的支持
 - 物理学、电子学、...
- ✓ 模式识别与相关学科是有区别的
 - 但是，现在各个学科呈现越来越融合的趋势
 - 所以，不要强调区别，让我们更注重共性、合作、以及更重要的一**解决问题**！

进一步的阅读

- ✓ 各相关领域的最新进展可以参考其重要会议和期刊
 - 机器学习: JMLR, ICML, NeuRIPS, ICLR
 - 计算机视觉: TPAMI, IJCV, CVPR, ICCV, ECCV
 - 多媒体: ACM Multimedia, TMM
- ✓ 有用的资源:
 - 深度学习软件包（百花齐放）！
 - 计算机视觉、图像处理OpenCV: <http://opencv.org/>
 - Matlab: 若干工具箱toolbox

你适合本课程吗?

- ✓ 拿到学分不容易
 - 需要投入时间和精力,仅课堂听一下不能过关

分数段	90-100	80-89	70-79	60-69	0-60
人数	34	45	8	9	7
比例	32.69%	43.27%	7.69%	8.65%	6.73%

- ✓ 尤其适合于对研究感兴趣的同学
 - 不一定毕业以后要做研究
 - 你觉得解决一个模式识别/计算机视觉问题有意思
 - 想想我们举的几个例子
 - 你觉得模式识别/计算机视觉本身挺有意思
 - 或者其他让你觉得有趣的东西
- ✓ 如果只是为了混学分,请退选!

智慧南雍与课堂小测试

- ✓ 智汇南雍 <https://lms.nju.edu.cn/>
- ✓ 移动端：南京大学APP，搜索“智慧教学平台”
- ✓ 下周对智慧南雍系统中的课堂小测试功能进行试用
 - 请务必提前装好南京大学APP

本课程 vs. 深度学习？

- ✓ 在现在深度学习如火如荼的情况下，我们为什么还要学习这么课程？

习题课 vs. 计算机视觉简介

- ✓ 本学期只有**16周**！
 - 课时非常紧张
 - “习题课”与“计算机视觉简介”只能二选一