

美国顶尖高校前沿学科项目

Deep Learning in Computer Vision

麻省理工深度学习应用于计算机视觉方向



深度学习的发展促进了计算机视觉的进步，在人脸识别，图像问答，物体检测和物体跟踪等方面，深度学习已经取得了非常好的效果，各类新的算法模型的运用也为实现计算机视觉提供了新的可能。本项目内容涵盖深度学习的经典算法模型以及模型应用于计算机视觉的热门方向，结合目标检测、图像语义分割、视频预测、自动驾驶等多应用场景的案例以及机器学习在医疗健康领域的应用等让学生了解深度学习的经典算法、相关前沿研究问题，以及各类神经网络模型中多个领域的最新应用，了解整个计算机视觉的发展过程中面临的挑战与机会。



■ 前沿大师课

10个模块核心课程囊括深度学习经典神经网络模型、计算机视觉的前沿应用，结合麻省理工学院人工智能与计算机实验室的最新研究成果，习得计算机视觉领域的前沿动态和挑战。

■ 学生可自行选择计算机视觉领域前沿研究方向作为课题

自动驾驶 End to end learning for self-driving cars

元宇宙室内三维图像 3D reconstruction in deep neural network

图像语义分割 Transformer network for image segmentation

街景视频图像理解 Video understanding

医学图像处理 Medical image processing in deep learning

■ 教授指导课题研究计划，快速构建学科知识体系

对每个方向所选课题进行对应的论文阅读指导、研究方法指导，帮助学生从大量的信息中快速构建学科知识体系，确立对研究项目全面的总体认知。

■ 课题项目实践

教授和助教将指导实践过程中的环境配置、参数调控优化等，对基础代码实现中的难题进行讲解。

■ 对论文的润色指导，优秀论文投稿支持瞄准顶级期刊

结合课程及项目研究，学生得项目论文将得到教授的指导及专业编辑的润色修改，优秀论文可获投稿学术支持，提供投稿建议及二次修改润色，提高SCI/EI等期刊对论文的接受率。

· 项目课时安排

项目日期：2023年1月8日-2月5日（春节假期调休）

时间安排：直播课程一般在北京时间上午9-12点

项目课时：60课时（每课时45分钟），包括核心课程、实践、科研写作等多个模块

论文产出：在项目结束后学生将获得项目证书，同时可结合项目课题产出一篇约5000词的科研论文（小组协作）

Master Class

10模块计算机视觉前沿课程；
一对一答疑指导；
5大前沿课题研究计划的指导以及论文选题建议；

Capstone

无人驾驶、元宇宙、图像语义分割、视频理解、医学图像实践案例；

Writing Booster

进阶写作科研论文指导

论文修改润色指导
投稿支持建议



Master Class 核心课程模块，总计25课时，周均6-7课时，由责任教授直播授课



Capstone Project 实践模块，总计15课时，周均3-4课时，由行业专家/博士后/助教直播授课及指导



Writing Booster 写作模块，总计20课时，周均3-4课时，由教授/期刊编辑直播授课及指导



论文指导、修改润色及投稿支持，每个小组论文完成后由教授/期刊编辑审核及指导

*本项目包含学术进阶写作模块，论文指导服务周期为项目结束后90天内。

- Introduction: Neural Networks and Convolutional Processing

In this module, we are going to realize the perceptron and feed-forward networks, image filtering and processing and mathematics of convolutions and feature learning. Students will learn how to stacking perceptrons to create networks, in order to learn it, training neural networks and convoluting processing is necessary. From convolutional layer to convolutional networks, we must know model local dependents on convolution, model strength dependents on hierarchy of features. Besides, we also will realize the image processing, which how to building block layer structure from 1D features to 2D processing.

- CNN Architectures

In this module, we will learn the foundational CNN (convolutional neural network) architecture. We are going to review LeNet: LeCun et al. 1998, AlexNet: Krizhevsky et al. NeurIPS 2012, GoLeNet/Inception: Szegedy et al. CVPR 2015, and VGGNet: Simonyan et al. ICLR 2015. During this period, we will realize the evolution of CNN, and find out the challenges in training CNNs. In addition to this, scaling CNNs also is the most part of the lecture and we will learn that from ResNet: He et al. CVPR 2016. The last part of this lecture is about the importance of data and how did datasets revolutionize computer vision.

- Sequential Image Processing

In this module, we will discover the neural networks and sequential image processing. This module is composed of three parts, realizing Recurrent neural networks (RNN), how does RNNs and CNNs apply for vision, and applications in vision. We are going to realize the requirements for sequential processing, models for sequential processing, and recurrent neural networks for sequential modeling. We will research how to use multimodal CNN + RNN architecture for captioning and how to use video data predict action.

- Generative Image Modeling

In this course, we will discover what is generative modeling and why generative modeling. This part includes density estimation, structured prediction and synthesis. Besides, realizing taxonomy of generative models also is the most part. We will find out that variational autoencoders will lead to reconstruction results. We also will discuss what is Generative adversarial networks (GANs) , compare traditional generation and StyleGAN, and find latest results about StyleGAN2.

- Neural Graphics and Rendering

In this module, we are going to discuss classical computer graphics, neural scene representations and implicit neural rendering. We will find out that rendering is a complex process and its differentiation is not uniquely defined, which prevents straightforward integration into neural networks. Differentiable rendering constitutes a family of techniques that tackle such an integration for end-to-end optimization by obtaining useful gradients of the rendering process.

- Neural Vision Applications

In this course, we will present some neural vision applications, including object detection, semantic segmentation, and self-supervised vision. The traditional object detection method uses the sliding window method to detect image region by region, but with the wide application of deep convolutional neural networks, Grishick et al. proposed R-CNN target detection framework. We will go deep on R-CNN. Semantic segmentation often requires the extraction of features and representations, which can derive meaningful correlation of the input image. This lecture will focus on the FCNN module.

- Interpretability and Uncertainty in Computer Vision

In this module, we are going to learn about uncertainty estimation in computer vision. We hope that the model can bring uncertainty and help people who use the model to make better decisions. Generally, there are two types of uncertainty, aleatoric uncertainty and epistemic uncertainty. We will research these two kinds of uncertainty in depth and realize the application of model uncertainty in life.

- Computer Vision Progress and Perspectives

In this course, we will research the progress and perspectives of computer vision in several areas, such as facial recognition, performance advances in generative models, AI-generated art, computer vision in robotics, autonomous vehicle training and control, and computer vision in safety-critical application. We will get better idea from specific case of image classification. Besides, we also will learn the bias in computer vision and how to debiasing it.

- **课程核心教授**



Dr. Alexander Amini

- 麻省理工学院计算机科学与人工智能实验室 (CSAIL) 研究科学家
- 麻省理工深度学习公开课课程负责人
- 美国国家科学基金会Fellow
- 研究领域包括计算机视觉、深度学习、无人驾驶等

- **课程日期** : 2023年1月8日-2月5日

- **课程时间** : 直播课程于每日9-11点 (北京时间)

L1	Summarize and Paraphrase	Emulating clarity, tone, clause types and combinations, language for stance and positioning, and information flow for crafting sophisticated sentences.
T1	Summarize and paraphrase practices	
L2	Quotation and Citation	Academic integrity and plagiarism, and citation systems such as MLA and APA.
T2	Citation tools in Microsoft Word	
L3	Thesis Development	Examining thesis sentences and moving through thoughtful strategies (specific, unique, debate causing, room for discussion) for complicating a thesis
T3	Construct thesis practices	
L4	Themed Seminar	The instructor will identify key concepts students can use to make sense of the complex communication experiences in their particular contexts.
T4	Handy tips in writing	
L5	Compose and Revise	Setting the context, stating the thesis, building points, developing ideas, and proofreading.
T5	Grading and peer review	

项目涵盖**计算机视觉、深度学习**等多个前沿学科、交叉学科和新兴学科方向，项目教学资料与顶尖高校该学期开设课程**同步更新**，与国外顶尖高校**保持教材一致以及研究数据同步**。

世界顶尖水平的师资教学团队全程指导，麻省理工学院**实验室科学家、深度学习领域公开课负责人**全程直播教学与答疑。项目配中文助教，提升学生学习效率。

项目中重要部分之一是应用实践，顶尖科技企业在当前热门领域的**应用案例**将作为实践教学案例，学生可根据深度学习的经典算法、相关前沿问题，研究其在多个领域的最新应用，将实践案例课题**进行实操分析及研究**。

学生将掌握计算机视觉的前沿热点，学会如何研究工具完成各类交叉学科研究，并基于此完成项目论文。在完成项目后将获得麻省理工学院实验室主任签发的**项目证书、支持网申要求的学术推荐信、论文期刊投稿支持**。

项目费用及奖学金

- 短期项目费用：16900元/人
- 项目费用包含课程、助教指导、项目服务管理、注册、资料费用
- 项目方为合作高校提供奖学金减免5000元/人，线上费用可抵扣2023年春季/暑期线下交流项目。

申请条件

- 本校全日制在读本科、研究生，符合本校国际交流派出要求
- 英文流利，能够接受全英文的教学环境
- 具备申请学科的基础知识
- 能保证项目时间投入

申请方式及咨询

- 扫描下方二维码，填写申请信息
- 通过初审后于3个工作日内收到项目录取邮件
- 更多问题欢迎咨询Cindy老师，微信tbstudy11
电话18917342671



项目申请二维码



咨询微信二维码