

2025

寒假前沿学科  
科研实践项目

卡内基梅隆大学

# 《具身智能与机器人学习中的3D视觉革命》

Frontiers of 3D Vision Methods in Embodied AI and Robot Learning



## 卡内基梅隆大学

卡内基梅隆大学 (Carnegie Mellon University)，简称CMU。坐落在宾夕法尼亚州的匹兹堡(Pittsburgh)，是一所享誉世界的私立顶级研究型大学，在2023年U.S. News美国最佳大学排名第22，是“新常春藤”，全球大学校长论坛成员。

建校之初，卡内基技术学校是一个行业学校，主要提供2~3年制课程训练。经过一百多年的努力，今天的CMU凭借其在教学、科研和社会服务方面的卓越成就已成为美国研究型大学中的佼佼者，在信息技术、计算机科学、商业、公共政策与心理学等领域全球领先。至今，卡内基梅隆大学的教员和校友中共有20人获得诺贝尔奖，13人获得图灵奖，22人获评美国艺术与科学院院士，19人进入美国科学促进会，72人入选美国国家学院，114人获得艾美奖。

## 计算机学科

课程旨在具身智能和机器人学习的背景下探索和发展前沿3D视觉技术，是一门处于人工智能、计算机视觉和机器人学交叉点的前沿学科。它主要研究如何利用3D视觉技术提升机器人在现实世界中的感知、理解和交互能力。

随着深度学习的发展，3D视觉方法（如点云、深度图像等）变得更加成熟，使机器人能够更准确地识别和处理三维环境中的信息。该领域特别关注在嵌入式系统中实现高效的3D视觉处理，以应对计算资源受限的挑战。这需要开发轻量级的模型和优化算法，确保机器人能够在实际场景中进行实时感知和决策。与此同时，机器人学习方法（如强化学习和模仿学习）使得机器人能够通过与环境交互自主学习复杂任务，逐步提升其操作能力和适应性。

在应用层面上，3D视觉还促进了人机交互和协作的发展，使机器人能够更自然地理解人类的行为和意图，进而实现更有效的协同工作。学科发展前景广阔，将为智能机器人在现实世界中的应用奠定坚实基础。

# 项目介绍

**"具身智能与机器人学习中的3D视觉革命"**项目为期两周，将在美国卡内基梅隆大学举行。项目将专注于探索具身智能与机器人学习领域中3D视觉技术的前沿应用。项目旨在通过理论与实践相结合的方式，帮助学生深入理解和掌握这些先进技术，特别是在智能系统与物理世界交互中至关重要的3D感知、建模与导航能力。

项目包含学科前沿专题讲座、产业界研讨、项目实践和跨文化交流等多个模块。课程部分由来自卡内基梅隆大学发世界顶尖水平教授讲解，深入探讨3D视觉技术如何与深度学习和机器人技术相结合，构建具备自主感知与操作能力的智能系统。理论学习将与产业界实际案例演练紧密结合，通过实际案例演练和项目实践让学生深入了解和学习人工智能在不同的应用场景中的机遇和挑战。学生将以小组为单位，选择具身智能或机器人学习的实际应用场景，如机器人导航、操作或人机交互，设计并开发相应的解决方案并进行汇报。这种融合理论与实践的学习方式将帮助学生更好地理解前沿技术的实际应用，并提升他们的创新能力和团队合作精神。

项目包括参访卡内基梅隆大学的前沿实验室及当地高科技企业，让学生能够直观了解3D视觉技术在不同领域的实际应用，并与行业专家进行深入交流，获取宝贵的实践经验和职业发展建议。项目还将为学生提供深度体验美国顶尖大学校园生活和跨文化交流的机会，他们将了解卡内基梅隆大学等世界顶尖名校的申请和就读体验，参观和游览匹兹堡当地的著名博物馆、美术馆和历史遗迹，通过体育赛事了解美国的竞技体育文化，体验匹兹堡的独特文化魅力，从而增强他们的国际化视野和跨文化交流能力。



## 卡内基梅隆大学适配学科

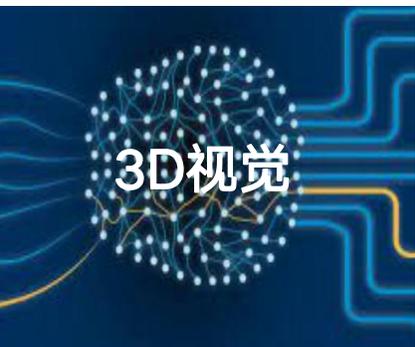
学院	适配学科和专业
计算机与人工智能学院	计算机、人工智能、软件工程、大语言模型
智慧城市与交通学院	智能建造、智能制造工程

## 项目模式

学生将沉浸于最前沿的3D视觉与机器人学习的学术研究氛围中。这一项目不仅是知识的传授，更是思想的碰撞。学生将接触到最前沿的研究课题，包括如何在三维环境中增强机器人的感知与决策能力。通过与世界级学者的交流，学生将挑战传统的理论框架，探索新的研究视角。项目中的课程设计旨在打破学科壁垒，将3D视觉技术与深度学习、人工智能的最新成果融入到具身智能的研究中。学生将在高度互动的课堂讨论、学术研讨会和实验项目中提升自己的科研能力，并学会如何将学术研究与实际应用结合，推动前沿领域的发展。

# 学术 + 产业

在产业模式中，学生将专注于3D视觉技术在真实世界中的具体应用场景，特别是在智能机器人领域的创新应用。学生将分析和研究高科技行业中最前沿的技术方案，探讨如何将3D视觉技术与深度学习算法相结合，以提升机器人的环境感知能力和决策准确性。通过案例研究，学生将学习如何解决实际应用中遇到的挑战，例如在动态环境中实现高精度的三维重建与实时响应。这些技术层面的探讨将帮助学生理解3D视觉在产业中的应用方式，增强他们对技术前沿的洞察力，并为未来在相关领域中的研究与开发做好准备。



# 项目成果

## 1. 理解与阐述

清晰地理解并准确阐述3D视觉的核心原理，以及这些原理在具身智能和机器人学习中的关键作用。

## 2. 分析与评估

深入分析并评估最先进的3D视觉技术，包括深度感知、以物体为中心的推理方法，以及复杂环境中的场景表示与建模。

## 3. 开发与实施

开发并有效实施用于机器人导航、操作和交互任务的3D视觉算法，提升机器人在现实环境中的应用能力。

## 4. 应用与优化

灵活应用深度学习技术，增强机器人在3D空间中的感知能力和决策精度，为实际任务提供智能解决方案。

## 5. 批判性评估

批判性地审视当前3D视觉在具身智能与机器人研究中的局限性，深入思考并预测该领域的未来发展方向。

## 1. 具身智能中的3D视觉导论

- 3D视觉系统和传感器
- 具身智能与机器人技术的关联
- 机器人3D感知的挑战与机遇

## 4. 端到端学习 (End2End Learning)

- 3D物体中心化表示
- 视觉语言模型 (VLMs) 在机器人学习中的应用
- 技能发现与学习

## 7. 实践项目与展示

- 现实问题中的3D视觉技术应用
- 项目结果展示与同侪评审

## 2. 深度感知与3D传感

- 立体视觉与深度估计
- 激光雷达、RGB-D相机及其他深度感知技术
- 深度信息的AI系统集成

## 5. 3D环境中的学习与交互

- 深度强化学习
- 仿真到现实 (Sim2Real) 转移技术
- 多智能体系统与协作机器人

## 3. 3D重建与场景理解

- 3D点云和网格生成
- 体积方法和SLAM技术 (同步定位与地图构建)
- 3D语义场景理解

## 6. 高级主题与案例研究

- 3D视觉与具身智能最新研究前沿探讨
- 机器人导航、操作与人机交互案例分析

# 实践课题项目

## 室内导航机器人场景表示项目

### 目标

- 设计并实现一款能够自主导航室内环境的机器人，借助3D视觉技术进行环境感知与路径规划。该机器人应具备障碍物检测与规避、特定物体或地标识别，以及实时地图构建的能力。

### 关键组件

- 3D视觉系统：采用立体相机或LiDAR传感器，捕捉并处理环境的三维数据，用于识别障碍物并生成环境地图。
- 场景表示：实现当前最先进的3D场景表示技术，如NeRF或3DGS，帮助机器人深入理解室内环境。
- SLAM（同步定位与地图构建）：应用SLAM算法，确保机器人能够实时定位并理解其在环境中的位置。
- 路径规划：开发算法，使机器人能够规划并执行路径，成功避开障碍物并到达指定目的地。
- 物体识别：集成识别模块，用于检测和与特定物体互动，如定位并搬运物品。
- 用户界面：设计直观的用户界面，方便用户输入目的地或指定任务。

### 实践步骤

- 规划与准备：确定导航环境和项目目标，划分为3D视觉系统、场景表示、SLAM、路径规划、物体识别、用户界面六个模块。
- 视觉与场景表示：使用立体相机或LiDAR采集环境数据，并实现NeRF或3DGS场景表示。
- SLAM与路径规划：实现SLAM算法用于定位和地图构建，并开发路径规划算法，实现避障与导航。
- 物体识别与交互：开发物体识别模块，并设计简单的用户界面，用于输入导航任务。
- 系统集成与测试：整合各模块，测试并优化机器人在模拟环境中的导航与物体识别能力。
- 总结与展示：撰写项目报告，准备展示，演示机器人在室内环境中的自主导航与物体识别。

### 预期成果

通过该项目，学生将设计出一款能够在室内环境中自主导航的机器人，具备识别特定物体并与之互动的能力

# 实践课题项目

## 基于物体特征的机器人抓取与分类策略项目

### 目标

- 开发一套机器人手臂系统，利用3D视觉技术识别、抓取并根据物体的形状、大小或颜色进行分类。

### 关键组件

- 3D视觉系统：集成3D相机，提供机器人工作区域内物体的深度信息，确保系统准确识别物体的形状和位置。
- 物体检测与分类：开发计算机视觉算法，根据物体的3D形状和大小进行检测与分类，并可增加颜色识别功能。
- 机器人手臂控制：实现逆运动学算法，精确控制机器人手臂，使其能够准确抓取并放置物体。
- 抓取策略：设计并优化多种抓取策略，确保机器人能有效处理多种类型的物体。
- 分类逻辑：开发分类算法，引导机器人根据分类结果将物体放置在对应的位置或容器中。

### 实践步骤

- 规划与准备：确定需要处理的物体类型及分类标准，划分为3D视觉、物体检测、手臂控制、抓取策略、分类逻辑五个模块。
- 视觉与物体检测：使用3D相机获取物体数据，并开发算法进行物体识别与分类。
- 手臂控制与抓取策略：实现逆运动学控制手臂，并设计有效的抓取策略。
- 分类逻辑与测试：实现自动分类功能，整合系统并测试抓取与分类效果。
- 总结与展示：撰写报告，展示机器人手臂的物体分类与抓取能力。

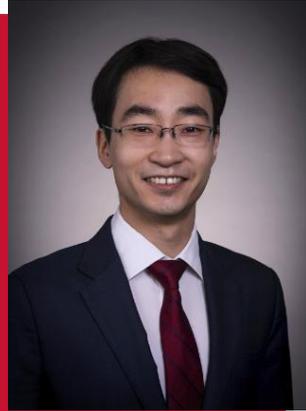
### 预期成果

学生将完成一个高度自动化的机器人系统，能够根据物体的形状和大小进行分类，展示其在3D视觉、机器人控制和实时处理领域的高级技能。

## Ding Zhao

### Associate Professor and Dean's Early Career Fellow Professor

Mechanical Engineering Department | Computer  
Science Department  
Robotics Institute | CyLab Security & Privacy  
Institute | Wilton E. Scott Institute for Energy  
Innovation



Ding Zhao教授是自动驾驶、智能交通系统和机器学习领域的知名学者。他的研究重点在于开发安全可靠的自动驾驶技术，尤其是通过仿真和学习技术来提高自动驾驶系统的安全性和鲁棒性。Zhao教授的工作涵盖了从车辆控制算法、风险评估、到多智能体系统的协同工作等多个方面。

Zhao教授目前在卡耐基梅隆大学担任机械工程系的教授，并且是该校安全、自动化与计算实验室（Safe AI Lab）的主任。他的实验室致力于通过先进的仿真技术和数据驱动的方法来开发更加安全、智能的自动驾驶系统。Zhao教授的研究不仅局限于学术界，他还积极与工业界合作，推动自动驾驶技术的实际应用和发展。

## Shubham Tulsiani

### Assistant Professor

The Robotics Institute  
Carnegie-Mellon University

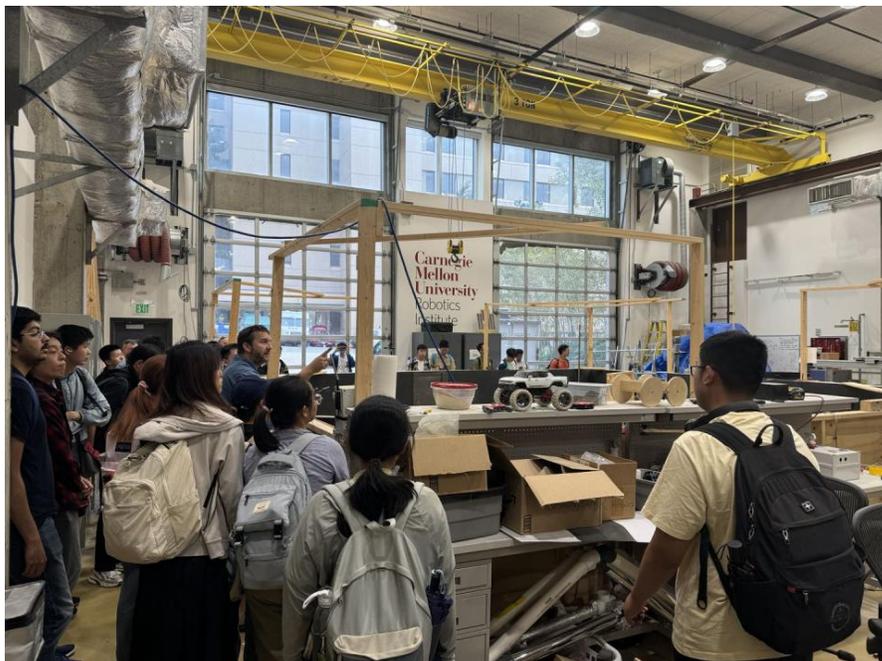


Shubham Tulsiani教授是计算机视觉和人工智能领域的知名学者。他的研究主要集中在三维视觉、场景理解、生成建模和自监督学习等方面。他的工作在计算机视觉社区中得到了广泛认可，尤其是在如何从二维图像推断出三维结构的研究上。Tulsiani教授在其研究生涯中发表了多篇高影响力的论文，涵盖了深度学习与计算机视觉的多个前沿领域。他还曾在多个顶级国际会议上展示他的研究成果，如CVPR（计算机视觉与模式识别会议）和NeurIPS（神经信息处理系统会议）。

除了学术研究，Tulsiani教授还对推动实际应用感兴趣，他的研究成果在自动驾驶、机器人和虚拟现实等领域有着广泛的应用潜力。他目前担任卡耐基梅隆大学的助理教授，并在此继续推动视觉与学习技术的进步。

## 实验室参访

### Robotics Institute



Robotics Institute是全球领先的机器人研究机构之一，致力于推动机器人技术的发展与应用。成立于1979年，研究所汇聚了顶尖的学术专家和研究人员，专注于机器人感知、控制、规划和智能系统等领域。其研究涵盖从自主车辆和人机交互到医疗机器人和工业自动化等多个应用领域。Robotics Institute以其创新的研究成果和对机器人技术的深远影响而闻名，推动了相关技术在实际应用中的延续和发展。

参访Robotics Institute将为学生们提供宝贵的学习机会。通过现场参观和专家讲解，学生们将深入了解前沿的机器人技术和研究成果；通过观察各种先进的机器人系统和实验设施，获得对机器人感知、控制和规划技术的直观理解。此外，学生们还将与领域内的专家交流，了解了最新的研究趋势和实际应用，拓宽视野，并激发对未来研究和职业发展的兴趣。

## Carnegie Museum of Art & Natural History



Carnegie Museum of Art & Natural History 是位于美国匹兹堡的著名文化和科学机构，由安德鲁·卡内基于1895年创立。博物馆分为两个主要部分：艺术馆和自然历史馆。艺术馆收藏了广泛的艺术作品，从古典到现代艺术应有尽有，是世界上最早关注当代艺术的博物馆之一。自然历史馆则以其丰富的恐龙化石展品、矿物收藏和生物多样性展览而闻名，为参观者提供了深入探索地球历史和生命演化的机会。博物馆不仅是文化和科学教育的重要场所，也是学术研究的中心。

## 钢铁河国家遗产区 River of Steel Carrie Furnace Blast



River of Steel Carrie Furnace Blast 是匹兹堡地区一座具有历史意义的工业遗址，曾是美国钢铁业的核心部分。学生将在向导的带领下探索这座遗址，了解钢铁生产的历史、工艺以及它在工业革命中的重要角色。导览过程中，学生将参观巨大的高炉和其他保存完好的工业设备，体验昔日工人的工作环境。此外，Workshop环节将让学生亲身体验金属工艺，如铸造和锻造，进一步理解钢铁生产过程。

## Senator John Heinz History Center

Senator John Heinz History Center 是宾夕法尼亚州西部最大的历史博物馆，也是匹兹堡的文化瑰宝。这个富有活力的博物馆将历史与现代相结合，通过丰富的展览和互动体验，生动地讲述了从早期移民到工业巨变，再到当代的区域故事。馆内展览包括美国工业的起源、钢铁业的辉煌历史，以及匹兹堡在体育、艺术和创新领域的卓越贡献。除了令人印象深刻的永久收藏外，Heinz History Center 还定期举办独特的特别展览和活动，使每次参观都能带来新的发现。这个博物馆不仅是历史爱好者的圣地，也是一处激发好奇心和灵感的文化中心。



阿克里舒尔体育场 (Acrisure Stadium) 位于匹兹堡市的奥黑尔区。阿克里舒尔体育场于2001年开放，除了是钢人队的主场外，还承办匹兹堡大学的美式足球比赛。球场的设计能够容纳约68,400名观众，以其现代化设施和优越的视野而受到球迷的喜爱。阿克里舒尔体育场的独特之处在于其沿河的美丽景观，以及在比赛日时充满活力的氛围。由于钢人队的忠实球迷，球场常常被称为“北方的最热血场地”。而匹兹堡钢人队不仅在NFL历史上具有重要地位，而且也是匹兹堡城市精神的象征之一。

## 阿克里舒尔体育场 (Acrisure Stadium)



## 线下课程安排

- 项目计划日期为2025年2月2日-2月15日，项目共包含32个学术课时，2次+企业参访，4次+文化活动。

### 线下课程参考行程

	时间	Day1	Day2	Day3	Day4	Day5	Day6	Day7
第一周	08:00-09:00	接机&办理入住	早餐	早餐	早餐	早餐	早餐	早餐
	09:00-12:00		项目介绍	学术课程	学术课程	学术课程	学术课程	文化活动
	13:00-15:00		文化活动	学术课程	学术辅导课	企业/机构参访	学术辅导课	
	15:00-17:00		文化活动	学术辅导课	文化活动		文化活动	
	时间	Day8	Day9	Day10	Day11	Day12	Day13	Day14
第二周	08:00-09:00	早餐	早餐	早餐	早餐	早餐		
	09:00-12:00	城市自由探索	学术课程	学术课程	学术课程	小组讨论&团队作业展示	送机&离开	到达国内
	13:00-15:00		企业/机构参访	学术课程	团队作业讨论	结业仪式暨颁发证书和成绩单		
	15:00-17:00			文化活动				

\*本日程仅为参考行程，最终安排可能会有调整。

## 项目预期成果



### ✓ 计算机视觉学术领域前沿

卡内基梅隆大学作为人工智能学科的四大名校之一在计算机视觉领域专注于大语言模型图像生成，视频理解和机器人视觉等方面的研究，项目中将学习到这些领域的经典理论和前沿动态。



### ✓ 实践课题应用与指导

分小组进行实践课题的演练和实操，融合理论与实践的学习方式将帮助学生更好地理解前沿技术的实际应用，并提升他们的创新能力和团队合作精神。



### ✓ 跨文化体验&Dream School

零距离体验美国高校校园生活，参加当地学生的交流活动，参访顶级学府，体验当地人文特色，全面客观了解美国留学生活。



### ✓ 官方证书、学术推荐信

提供官方教学团队签发的项目证书，项目表现优秀者有机会获得教授签署的学术推荐信。



### ✓ 产业实践行业评价

深入国家机器人中心、谷歌、亚马逊等科研企业和平台，参与项目的产业实践，结合行业评价持续地对实践项目进行优化。



### ✓ 学术人脉拓展

与CMU学生社团同台竞技，结识世界顶尖水平的院士、知名教授、学术权威零距离交谈，参观世界顶尖的大学、拓展学生人际网络。

# 项目申请及咨询服务

**项目费用：5950 USD/人**

项目费用包含：

- 1.核心课程费用；
- 2.实践和文化活动参访费用；
- 3.住宿、早餐、接送机、当地交通费；
- 4.个人境外旅行意外保险费；
- 5.美国签证咨询及协助申请服务；
- 6.项目管理费用。

## 项目申请条件：

- 1.满足学校国际交流派出要求；
- 2.本科一年级至博士生三年级， 年满18岁；
- 3.具备一定的计算机、编程等基础课程知识， 各项目专业基础课程要求详询Olivia老师；
- 4.具备一定的学术英语能力、海外生活能力、开放积极的交流心态， 参与项目期间遵纪守法， 尊重项目组安排。

## 申请流程：

- 1.填写报名提交材料
- 2.等待审核结果
- 3.收到录取通知后签署项目合约
- 4.完成缴费
- 5.获得官方邀请函
- 6.办理签证
- 7.购买往返机票
- 8.参加线上/线下行前培训
- 9.出境

注：申请过程中我们将为学生提供全程的指导服务。

### 项目申请链接



### 项目咨询Olivia老师





谢谢审阅！