**一、个人简介**

余宙，讲师，男，浙江杭州人，1988年生。分别于2010年和2015年在浙江大学获得数字媒体技术专业学士学位和计算机科学与技术专业博士学位，并于2015年至2016年期间就职于阿里巴巴集团搜索事业部任资深算法工程师。自2016年9月起任职于杭州电子科技大学计算机学院，是媒体智能实验室（Media Intelligence Lab, MIL）和复杂系统建模与仿真教育部重点实验室成员。

研究领域：主要研究兴趣包括使用机器学习和深度学习算法对计算机视觉和自然语言的跨媒体表达，涉及视觉自然描述、视觉问答系统、跨媒体检索、高维数据哈希索引。现已发表国内国际期刊会议论文10余篇，其中在CCF A类会议和IEEE期刊上发表论文7篇，主持国家自然科学基金青年项目1项。担任IEEE TIP, TMM, TCSVT, INS, PR, NEUCOM, SIGPRO, CVPR, AAAI, IJCAI, ACMMM等多个国际顶级期刊和会议的审稿人。

指导学生：指导的本科生多次参加大学生创新创业、新苗计划、互联网+等竞赛并获得资助；指导的研究生参加国际的学术竞赛，如VQA Challenge 2017，获得第二名；实验室毕业的研究生就业BAT等公司，也推荐优秀的本科生到BAT等大公司实习。

个人主页：<http://mil.hdu.edu.cn/people/zhou_yu/index.html>

联系方式：[yuz@hdu.edu.cn](mailto:yuz@hdu.edu.cn)

**二、项目规划**

**课题方向：深度学习+跨模态表达与交互（计算机视觉+自然语言处理）**

### 专业方向：计算机科学与技术、物联网工程、软件工程本科生

**整体规划**：研究“跨媒体表达”方向上的相关问题，涉及计算机视觉、自然语言处理、机器学习（尤其是深度学习）的基本理论。并以项目驱动形式，对特定课题进行深入研究，掌握理论、方法与系统等方面的研究进展及前沿成果，发现存在的问题，探索可行的解决方案，并进行实验验证，力争取得兼具实用性与创新性的成果。在此过程中，锻炼学生的实践能力和创新能力。

**项目设计：**

1. **项目1：视觉描述系统**

对视觉对象，如图像、视频，让算法自动生成自然语言的描述。涉及图像、文本特征表达、机器学习、深度学习等。最后完成算法+加系统实现（Web/App/树莓派）。

1. **项目2：视觉问答交互系统**

给定一个视觉对象（图像、视频）和一个自然语言描述的问题，让算法同时理解视觉对象和问题，在进行深度推理后使用自然语言回答相应的问题。涉及图像、文本特征表达、机器学习、深度学习等。最后完成算法+加系统实现（Web/App/树莓派）。

1. **项目3：跨模态搜索引擎**

在海量的跨媒体数据中，实现简易的搜索引擎进行高效地跨媒体检索（如，使用输入文本检索相关图像，输入图像检索相关文本）。涉及图像、文本特征表达、机器学习、深度学习等。最后完成算法+加系统实现（Web/App）。

1. **项目4：交互式视觉定位系统**

给定一个视觉对象（图像、视频）和一个自然语言描述的物体，让算法在图像中进行物体精确定位该任务。涉及图像、文本特征表达、机器学习、深度学习等。最后完成算法+加系统实现（Web/App/树莓派）。

**进度及内容安排：**

1. **第一学期：机器学习、计算机视觉、自然语言处理基础**

Python编程基础：环境构建、基本语法、数据处理、工具包使用等

机器学习基础包括：回归、分类, 决策树、支持向量机、神经网络、降维、强化学习和最优化方法。

计算机视觉基础包括：图像表示、图像特征、图像识别、图像分割、物体检测、视频理解；阅读项目相关经典论文，了解经典方法并进行实现。

自然语言处理基础包括：文本表示、文本特征（词袋模型、词向量、TF-IDF、LSI、pLSA、LDA等主题建模模型）；阅读项目相关经典论文，了解经典方法并进行实现。

实践任务：包括机器学习任务（井字棋胜负判断、客户聚类）、计算机视觉任务（MNIST手写数字识别、CIFAR-10图像分类）和自然语言处理任务（文档分类）。需要组合不同的机器学习方法、图像特征、文本特征完成任务，进行性能比较和分析，并撰写规范的技术报告。

1. **第二学期：深度学习基础及项目初步实现**

学习深度学习的基础理论和方法，包括卷积神经网络、循环神经网络。阅读项目相关论文，了解相关进展。

熟练掌握Python下深度学习得工具包使用，如Tensorflow、PyTorch、caffe等。能使用工具包复现相关论文的代码。

1. **第三学期：项目的改进**

针对选定的项目，分析现有算法的缺陷，设计对应的解决策略，构建新的模型，并进行实验验证。

1. **第四学期：项目的完善**

针对项目进展中出现的问题进行完善，构建演示性系统或撰写学术论文，作为对整个工作的集成性总结。