

计算机学院科研团队情况介绍表

团队名称	浙江大学创新软件研发中心 (Eagle-Lab)		团队负责人	卜佳俊	
联系人	马凌洲	Email	mlz@zju.edu.cn	电话	0571-87953955

主要情况介绍:

浙江大学创新软件研发中心（EAGLE-Lab）成立于 2004 年，依托于浙江大学计算机学院软件研究所。实验室致力于打造人工智能、大数据、物联网、机器视觉、信息无障碍等领域的知名实验室。实验室优秀校友在工业界有黄峥（拼多多创始人）、胡志鹏（网易副总裁）、陈刚（去哪儿 CEO）等，学术界有管子玉教授（2015 年度国家优秀青年科学基金获得者）、张利军副教授（2018 年首届达摩院青橙奖获得者）等。

EAGLE-Lab 拥有一支高质量的科研梯队。实验室始终坚持前沿科学研究，已在相关领域的顶级国际期刊与顶级国际会议上发表上百篇高质量学术论文，其中产生了 AAAI 最佳论文奖，系该会议创办 30 多年来中国大陆学者首次获奖论文。同时实验室也立足于科技落地，开展了多项国家级、省部级以及企业重大工程项目。荣获国家技术发明二等奖、国家科技进步二等奖，浙江省科学技术一等奖等国家和省部级奖励 10 余项。

在对外合作上，实验室与阿里巴巴、网易、拼多多、腾讯等国内外著名企业有着良好的合作关系，尤其是深度参与阿里巴巴-浙江大学前沿技术联合研究中心（AZFT）的合作，利用阿里巴巴提供的资金、数据、计算能力等优秀的研究资源，我们在国内国际专利、标准、学术论文、技术与商业结合等方面产生具有国际水平的创新成果。在社会公益上，实验室与中国残联已深度合作十多年之久，致力于助老助残事业，让特殊人群也能享受科技发展带来的福利成果，产生了良好的社会效益，得到了中国残联主席张海迪的肯定与支持。

此外，实验室与国外知名高校也建立了良好的合作关系，香港大学教授俞益洲博士（美国 UIUC 兼职教授）、加拿大 SFU 终身教授 Jian Pei 博士、Martin Ester 博士已成为实验室合作教授。

EAGLE-Lab 培养出的研究生及本科生中，有进入微软、Google、Facebook、阿里巴巴、华为、百度、腾讯、网易等国内外著名企业，也有进入美国 CMU、UIUC、UCLA、Columbia 等著名学府深造。其中包括在读期间获得 ACM 国际大学生程序设计竞赛全球总决赛冠军的队长和队员、大陆首次 AAAI 最佳论文奖获得者、竺可桢奖获得者等，使项目组成为一个培养人才的摇篮。

实验室网站: <http://eagle.zju.edu.cn>

团队主要成员

姓名	职称	研究方向	联系方式
卜佳俊	教授	信息无障碍、智能媒体计算、 大数据分析 & 挖掘	bjj@zju.edu.cn
宋明黎	教授	计算机视觉、计算机图形学、	brooksong@zju.edu.cn

		多模态模式识别	
董玮	教授	物联网, 智能边缘计算, 计算机网络	dongw@zju.edu.cn
朱建科	教授	计算机视觉、多媒体信息检索	jkzhu@zju.edu.cn
Jian Pei	教授	数据分析与挖掘、数据库技术	jpei@cs.sfu.ca
俞益洲	教授	计算机图形学	yzyu@cs.hku.hk
王强	副教授	计算机图形图像处理、计算机图形学	wangqiang@cs.zju.edu.cn
郭庆	副教授	Web 技术、移动多媒体、人工智能	13385718936@189.cn
高艺	副教授	传感网与物联网, 智能感知与移动计算, 网络测量	gaoyi@zju.edu.cn

目前承担的主要项目:

实验室已承担和完成国家 973、863、国家科技支撑计划、国家重点研发计划、国家自然科学基金和省部级重大科技攻关项目 30 多项, 国内外企业重大合作项目 20 多项。

主要研究成果:

一. 信息无障碍

浙江大学与中国残联合作成立了“中国残疾人信息和无障碍技术研究中心”(中国残联系统外第一个技术研究中心), 致力于该领域的人才培养、理论研究、系统研发和社会服务工作, 研发了中国盲人数字图书馆、无障碍(视频+文字)网络直播系统等一系列信息无障碍系统和平台, 利用互联网为全国 3600 多万在库实名残疾人提供便捷高效精准的信息无障碍服务。

实验室致力于人工智能助力信息无障碍的研究, 结合计算机视觉、语音识别、自然语言处理等前沿技术, 专注于无障碍智能媒体计算, 研究内容包括无障碍助盲眼镜、文字识别、网站无障碍自动检测等等。

同时, 实验室联合中国残联等多家单位成立信息无障碍技术标准工作联合组, 制定了多项已经实施的国家标准、行业标准, 包括《读屏软件技术要求》《信息技术 互联网内容无障碍可访问性技术要求与测试方法》《盲用数字出版物格式》《信息无障碍 视障者互联网信息服务辅助系统技术要求》等。目前, 已经初步形成了以互联网内容、辅助工具、内容加工与生成的无障碍标准体系基本框架, 并在浙江大学、北京航空航天大学等网站开展应用, 与阿里巴巴为代表的一流互联网公司合作, 推动互联网产品和业务的无障碍升级, 显著提升特殊用户的服务质量。

1. 在研项目: 无障碍助盲眼镜

智能助盲眼镜：研制出具有视觉感知定位、语音导航、视觉物体播报、文字朗读等功能的穿戴式智能助视设备。提升视障伤残人士环境感知能力、自主行动能力、与人沟通交流能力，使其基本实现生活自理和自助阅读学习，能参加更多工种的工作。本研究领域涉及的到关键视觉技术：人脸识别，文字识别 OCR，室内定位和导航，室外定位和导航，3D建模，障碍物检测和识别，交通标志识别，斑马线识别，盲道识别等关键技术。

2. 在研项目：智能手眼助盲眼镜（助盲生活辅助协同系统）

多重功能障碍者，尤其是双盲和双前臂截肢者，视觉和手部功能缺失给他们的生活、工作和学习带来了诸多不便，极大地降低了他们的生活质量。研究和开发先进的多功能生活辅助系统，能够帮助他们减轻或消除身体的功能缺陷，使其最大限度地恢复生活和工作能力，具有重要的社会意义。目前，国内外的义肢系统都需要借助使用者的视觉辅助完成义肢手的抓握、旋转，从而实现喝水，吃饭等基本日常活动。对于多重功能障碍者，由于丧失了视觉功能，佩戴义肢后几乎无法感知义肢手的空间位置和抓握的触觉信息。因此，佩戴义肢后患者几乎无法操控义肢完成相应的动作。本研究领域涉及的到关键视觉技术：人脸识别，目标识别，目标分类，目标空间定位，目标跟踪等。涉及到的 CNN, GCN 等人工智能深度学习和机器学习等关键技术。

3. 在研项目：信息无障碍自动检测技术研究

互联网的信息无障碍服务能力提升，离不开互联网内容的无障碍建设。大部分互联网内容顶层设计的时候往往缺乏信息无障碍方面的考虑，后续再增加或者改造的代价会比设计之初就考虑大很多。同时大部分内容发布人员不是特殊用户，因为无法很好的了解他们的特殊需求，因此需要一套检测方案，帮助他们更好的发现发布内容中的存在的障碍。然而，由于互联网上存在海量异构的多媒体数据，传统借助人力的检测方案存在效率低下、覆盖面窄、难以量化等缺陷，因此期望借助云计算、群智等理论，实现完整的机器自动+人工结合的检测机试，并研究各类平台终端自动检测方案，完成大批互联网内容的无障碍检测。

二. 数据智能

AI 已经成为了科技行业中“闪耀的新星”，并且在不断升温。当前市场上已经出现众多的 AI 应用。实验室在多个领域结合人工智能开展研究工作，包括医疗人工智能，5G 智能行业，智能图文信息分析等。

1. 在研项目：医疗人工智能

目前医疗健康领域，人工智能已经在疾病辅助诊断等许多方面得到广泛应用，尤其是人工智能与医学影像的结合，已经成为医生诊断、治疗工作的一种必备手段。人工智能 +

医学影像是计算机在医学影像的基础上，通过深度学习与大数据技术等，完成对影像视频中人体器官目标检测、目标分割分类工作，进而提出诊断报告，协助医生完成诊断、治疗工作的一种辅助工具。人工智能在医疗影像领域的应用场景重点是提升医生看病效率和诊断问题的准确率。

除了智能医疗影像识别，实验室还在人体运动分析（人体行为智能识别）、智能手术导航进行研究。基于人体骨架的行为分析方法在面对复杂背景的行为视频时具有较好的鲁棒性，因而这一方法是当下人体运动分析的研究热点，也是未来具有潜力的研究方向之一。深度摄像技术的发展使得人体运动的深度图像序列变得容易获取，结合高精度的骨架估计算法，能够进一步提取人体骨架运动序列。

外科手术进入智能机器微创手术阶段，骨科手术机器人在人工智能技术助力下在提高手术精度、优化手术操作、减少手术损伤等方面具有重要意义，是实现骨科手术微创化和智能化的重要途径。定位导航是现有骨科手术机器人的核心功能：定位是指利用空间定位算法计算手术靶点的空间坐标，定位算法的精度直接影响手术的精度和效果；导航是指利用合理的交互方式引导医生完成手术操作，导航模式的便利性直接影响系统的临床适应性和有效性。

2. 在研项目：智能图文信息分析

随着信息技术的快速发展，计算机视觉感知已成为人工智能时代下的关注核心，但有时图像中缺乏文字，对于图像含义的深入理解带来了巨大的困难。而当下大量的图文数据（rich-text image）中包含着图像载体形式的文字信息，对于视觉理解起到了巨大的辅助作用，充当着图文数据信息分析与理解的重要载体。

我们致力于探索图文大数据下的信息分析，通过场景文字（Scene Text）分析+文本结构内容化的研究思想，结合基础深度学习理论，创新定义多维度，多层次下的图文理解研究，在盲人无障碍图像辅助理解，通用图像问答及图像翻译等应用中提供技术与算法支持。

我们也与阿里巴巴达摩院视觉技术的 OCR 团队有着长期合作与学术交流，在电商，实拍文档，街景等图文数据场景中有着丰富的技术基础和经验积累；实验室也同步研发了一套基础的 OCR 线上实验平台并迭代更新，为相关研究与工程实践提供了良好的基础。具体的研究方向包括：OCR 基础算法研究，图文结构理解与内容化等。

3. 在研项目：5G 智能行业

实验室具有通讯研发 15 年以上的研发背景的 5G 专家做导师，并将在 5G+流程监控和品质控，5G+数据挖掘，5G+智能流程优化，5G+智能仓储，5G+远程维护，5G+医疗中研

究领域上进行突破性研究。

三. 物联网系统与网络

物联网是计算机学科相关研究中的一大热点，被誉为计算机、互联网之后的第三次世界信息发展浪潮，也是未来世界发展的主要走向。物联网正日益渗透到社会生活和国家发展的各个层面，已在信息化进程过程中发挥着越来越重要的作用。

EmNets 课题组的主要研究方向包括云-边-端一体的物联网系统、低功耗无线网络、边缘计算、智能传感和移动计算等。

课题组承担多项国家和省部级项目。在 ACM MobiCom, SIGMETRICS, UbiComp, IEEE INFOCOM, ICNP 等著名国际会议以及 IEEE/ACM Trans on Networking, IEEE Trans on Mobile Computing, IEEE Trans on Computers 等著名国际期刊上发表论文 100 余篇，其中 CCF A 类期刊/会议论文 50 余篇，IEEE/ACM Trans 长文 30 余篇。获 MSN' 16, SIGCOMM-China' 17 最佳论文奖，ACM MobiCom' 17 最佳视频展示奖。

1. 在研项目：基于 LinkLab 的物联网远程开发平台

物联网设备异构性加剧，物联网应用开发仍然十分困难。实验室与阿里云 IoT 事业部联合发起了 LinkLab 物联网远程开发平台 (<http://www.emnets.org/linklab>)。目前已在人工智能小镇部署了 150 余种物联网设备，用户只需通过浏览器 (WebIDE) 即可进行云-边-端一体的物联网应用开发。LinkLab 需要解决可靠设备管理、多环境高效在线编译、theia WebIDE 高效定制、大规模虚拟场景仿真等关键问题。

2. 在研项目：面向港珠澳大桥智能运维的边缘计算平台及应用

港珠澳大桥是我国首座集桥岛隧一体的现代跨海通道工程，为了“用好管好”这座世界级的现代跨海通道工程，迫切需要推动桥岛隧跨海通道工程运维管理技术创新。本课题为国家重点研发计划课题研究内容。

课题研究复杂环境下低功耗广域网传输模型，提出抗干扰、自适应性强的传输机制；研究边缘计算负载动态迁移技术、以及边缘侧网络资源感知传感协同技术，通过在边缘侧对不同业务进行预处理，提高数据处理效率和实时性，并降低带宽与能量消耗；研究边缘侧多集群管理技术和支持多 CPU 架构、多运行时环境的轻量级编排技术，提高系统稳定性；研究支持多源异构数据处理的边缘计算系统架构，构建多类型边缘计算系统平台，形成基于边缘计算的海量数据采传收一体化技术方案，为桥岛隧巡检、维保设备以及边缘视频处理提供模型与算力支持，在台风等极端天气时提升响应速度。

3. 在研项目：基于 AliOS Things 的微内核及 OTA 技术研发

伴随 5G 的商用，物联网应用和系统日趋复杂，将对当前的物联网操作系统 RTOS (如

freertos, zephyr, rtthread 等)在安全性, 健壮性, 低功耗, OTA(Over The Air 软件更新)稳定性等方面产生较大的冲击, 这就需要物联网操作系统有更先进的内核架构予以支持, 微内核架构是当前前沿的技术, 它能较好的解决 RTOS 面临的安全和健壮性问题; 同时在万物互联的时代下, OTA 能力已逐步发展为物联网设备的刚性能力需求, 它不仅能大大降低设备售后运维成本, 也能帮助客户缩短软件开发周期, 快速发布新功能, 抢占市场; 特别是基于微内核架构的差分 OTA, 具备在 rtos 上的高度差异化的优势, 可以实现更细颗粒度的单个独立应用或者驱动, 组件的差分升级, 最大限度的节省存储空间成本, 减少下载时间和设备功耗, 给用户以极致速度升级体验。本项目是实验室与阿里云 IoT 事业部的合作项目。

四. 视觉智能与模式分析

视觉智能与模式分析研究内容主要包含四个方面:

视觉感知增强方面, 我们主要针对高动态范围场景建模、颜色度量准则与管理、基于深度图像的计算机视觉等方面开展新的研究;

视觉嵌入方面, 我们研究物体的检测、追踪与识别, 图像/视频的语义分割与标注, 下一代编码/解码的生成技术;

基于视觉的人机交互方面, 我们志于研究人体行为与表情识别, 视线估计以及机器人模仿学习;

机器学习方面, 我们致力于搭建知识与数据鸿沟间的桥梁, 研究知识融合、自动化特征工程与图表征的新技术。

1. 在研项目: 视频 AI 分析平台——交互仿真虚拟人

视频 AI 分析平台面向广大视频行业用户, 为用户提供智能视频处理方案诸如视频素材提取、语义视频查询、视频 OCR 辅助。本平台以创新求实为准则, 提高视频处理的精度和实用性, 以最新的技术提供最高质量的服务。

实验室开发了一款高度模拟真人原型的虚拟人, 整合了语音识别、语音合成、机器问答、面部重塑等 AI 技术, 通过学习真人视频, 该款产品就能瞬间克隆出一个“数字化的人”。

2. 在研项目: 虚拟试装

服饰是淘宝最重要的商品类目, 为了帮助用户更快、更好、更直观地挑选服饰, 基于淘宝/天猫海量服饰图片, 本项目通过一系列检测、分割、形变算法, 将服饰真实地‘穿’到模特身上。用户可以直观地看到不同服饰的搭配以及真实的穿衣效果, 为用户提供更好的试穿体验。

提出基于类-规则树自适应虚拟试衣算法，通过将服饰划分为细类目，然后基于规则树实现自适应形变与轮廓匹配，经过细节与阴影优化，获得当前最好虚拟试衣效果，实现国内首个在线海量服饰图虚拟试衣。与已有同类方法相比，该算法不仅有更好的视觉效果，而且具有非常好的泛化性，可以简易得扩展到全类目服饰商品。提出基于自由度估计的 AR 试鞋算法，通过深度网络估计出视频图片帧中鞋子的六个自由度，将 3D 虚拟鞋模与所估六个自由度进行匹配，然后进行 3D 鞋模进行渲染，并进行抖动优化等后处理，获得 AR 试鞋效果。

3. 在研项目：助聋唇语识别

该项目需要研制出具有声源定位、声音类型识别、唇语辅助下的语音识别、交互场景下的手语识别等功能的便携式无障碍语言交流系统。帮助听障伤残人士基本无障碍感知环境、与人自然沟通，为他们提高生活和学习能力、获得更多工作机会提供支持。该项目要让听障残疾人获得与正常人相近的信息获取能力、与人沟通能力、环境感知能力，实现生活自理、跟班学习、获得工作机会，自主预防和摆脱贫困为目标。

无障碍语言交流系统主要包含四个分系统：多模态感知分系统（头戴）、嵌入式控制分系统（腰戴）、显示与输入分系统（腕戴）和云计算分系统。其中，多模态感知分系统通过各类感知设备采集并预处理音视频信息，主要负责整个系统的信号输入与传输以及端计算；嵌入式控制分系统将多模态感知系统传来的声音信号进行声源定位与声音类型识别，并提供无线传输功能；云计算分系统实现交互场景下的手语翻译和唇语辅助下的语音识别及语义理解等云计算；显示与输入分系统将最终的运算结果呈递给用户，供用户直观接收。

4. 在研项目：小样本机器学习

在实际生产中，为了获得高精度的分类模型，通常需通过标注大量的数据来支撑模型优化的目标。研究小样本学习算法，可以极大缓解深度模型对数据的依赖，从而大大提高识别模型的生产效率，解决视频监控、安全业务场景中小样本学习的问题，为小样本目标检测与识别提供技术支撑。

五. 数据挖掘与信息检索

在这个“大数据”的时代，数据中比以往任何时候都包含了更多有用的知识。我们渴望知识，然而他们却被海量的信息所淹没。数据挖掘与信息检索小组专注于数据挖掘、信息检索、机器学习等领域的研究工作，致力于从海量的数据中发现有用的模式和知识。我们目前的研究工作包括互联网多媒体信息检索、社交网络分析、个性化推荐算法与技术、机器学习与数据挖掘理论研究以及数据可视化技术的研究与分析等。近年来，我们在多个

TOP 会议和顶级国际 SCI 期刊上发表了高水平研究论文，包括 SIGIR、ACM MM、WWW、AAAI、IJCAI 等顶级国际会议和 TPAMI、TKDE、TIP、TC、TMC 等顶级国际期刊等。

在研方向：图数据与网络数据挖掘

图与网络广泛应用于社交网络、互联网、电网、引文网络等复杂关联与交互场景的建模。我们的研究工作主要从对象关系复杂、形态表现丰富的图数据与网络数据中发现信息交互、传播、表征等规律并支持各类工业应用。

1. 在研项目：网络表征学习

在大规模网络数据中，由于节点数量众多、特征丰富、网络结构多样化等问题，传统的数据分析和挖掘方法难以直接在原始数据上应用。为了提升图数据挖掘任务的性能，需要为复杂网络中的节点学习有效的低维向量表示。实验室目前已经在带属性网络、有向网络、异构信息网络等类型的复杂网络表征学习领域取得了多项研究成果，并将继续深化网络表征学习，并在网络结构学习，动态网络表征学习等领域实现突破。

2. 在研项目：复杂网络中的社区发现

社区发现是指通过对图数据进行分析，将图中内部节点之间联系紧密、与外部节点之间联系稀疏的部分找出来，这些内部稠密的部分就称之为社群。对网络进行社群发现具有极大的现实意义。例如，在交易网中，可以挖掘出由拥有相似兴趣偏好的用户组成的兴趣小组、由用户行为理解而形成的场景化的商品图谱、由黑产引导的刷单或“薅羊毛”的风险社群等。基于兴趣偏好相似的用户喜欢相似风格的商品，兴趣小组可用于基于兴趣社群的个性化推荐。风险社群则是反作弊业务中一种典型的作弊模式，商家通过进行作弊误导消费者、破坏平台商家之间的公平竞争，对平台带来了严重危害。

该项目利用图神经网络技术实现端到端的网络表征学习和社区发现框架，并在大规模社交网络，电商推荐，电商反欺诈领域实现落地应用。

3. 在研项目：跨模态检索与推荐系统

随着信息化水平的不断提高，用户可以接触和使用多种模态的数据，例如图片、文本、音频、视频等。多模态数据之间的语义鸿沟为跨模态检索与推荐任务带来了巨大的挑战。该项目结合单模态数据的语义理解，多模态数据的语义对齐以及个性化检索系统，为用户打造快速精准的跨模态检索与推荐系统。通过与落地应用场景结合，为用户提供了精准、个性化的检索与推荐结果。