简报

目录

最新动态 .......................................................................................................................... 1
实验室与英国剑桥大学共建中英人工智能伦理与治理研究中心 .................................. 1
实验室 2 项成果分获第九届吴文俊人工智能科学技术奖一、二等奖 ......................... 1
实验室获 2019 年 NeurIPS 模型压缩与加速竞赛双料冠军 ....................................... 2
实验室在 ICCV2019 竞赛中揽获 5 项冠军 ................................................................... 3
喜报：实验室蒋田仔研究员当选 2020 IEEE Fellow ..................................................... 6
喜报：实验室陈然研究员入选为第五批青年创新促进会优秀会员 .............................. 6
喜报：实验室两位老师入选中国图象图形学学会首批会士 ......................................... 6
喜报：实验室宗成庆研究员获得宝钢优秀教师奖 ......................................................... 7
喜报：实验室何世柱等 3 人入选为 2020 年青年创新促进会会员 ................................ 7
科研进展 .......................................................................................................................... 7
基于时空图结构学习的交通预测 ..................................................................................... 7
基于自适应上下文网络的场景解析 ............................................................................... 8
基于图关联的无监督行人再识别 ..................................................................................... 8
基于对抗视觉特征残差的零样本学习 ............................................................................ 9
基于快速核相关滤波的目标跟踪 ................................................................................... 10
GSLAM：一个通用的 SLAM 框架和基准 .................................................................... 10
基于图神经网络的视频精彩片段检测 ........................................................................... 11
基于条件门循环单元的文字风格模仿 ........................................................................... 12
基于交互式学习的同步语音识别与语音翻译 ................................................................. 12

实验室网址：http://www.nlpr.ia.ac.cn/ 实验室微博：http://weibo.com/u/5146164852
基于多角度图编码的答案句选择 ................................................................. 13
基于脑功能影像时间序列的多尺度卷积循环神经网络模型 .............................. 14
学术交流 ......................................................................................................... 15
法国贡比涅技术大学 Thierry Denoeux 教授访问模式识别国家重点实验室 ......... 15
澳大利亚墨尔本大学 Tim Baldwin 教授访问模式识别国家重点实验室 .............. 16
模式识别大讲堂：深度学习之路在何方？未来人工智能之路是否还依赖深度学习 16
模式识别大讲堂：智能感知与计算团队研究方向与最新进展 ............................ 17
模式识别青年学术沙龙：人工智能结合真实场景在垂直领域中的应用 ............... 18
项目立项 ......................................................................................................... 19
实验室第四季度新建立课题 65 项 .................................................................. 19
文体活动 ......................................................................................................... 21
智能感知中心“云蒙山秋游”活动 .................................................................. 21
《模式识别国家重点实验室简报》编委会

内容审核
刘成林  陶建华

编辑小组
组长：陈玉博  刘斌
成员：张一帆  原春锋  徐士彪  樊彬  张煦尧  崔海楠
       杨阳  李瑾  黄岩  钱胜胜  赵微  曹迪

实验室网址：http://www.nlpr.ia.ac.cn/  实验室微博：http://weibo.com/u/5146164852
实验室与英国剑桥大学共建中英人工智能伦理与治理研究中心

为推动人工智能伦理与治理研究与国际协作，2019年11月4日，中英人工智能伦理与治理研究中心（China-UK Research Centre for AI Ethics and Governance，下称研究中心）在京成立。中方依托单位为中国科学院自动化研究所，英方依托单位为剑桥大学。我所模式识别国家重点实验室研究员、类脑智能研究中心副主任、科技部国家新一代人工智能治理委员会委员曾毅研究员出任中方主任，剑桥大学未来智能研究中心主任Huw Price教授出任英方主任。研究中心结合北京市相关优势力量，邀请了北京智源人工智能研究院、北京大学、中国社科院等单位相关学者和机构加入协作网络，目标是建成以北京和剑桥为中心，融合东西方学术思想、面向人工智能伦理、风险、治理的综合性、融合自然科学与社会科学的跨学科、跨机构的研究中心。揭牌仪式后召开了第一届学术研讨会，由中科院自动化所副所长、模式识别国家重点实验室主任刘成林代表中方单位发言。科技部国际合作司欧洲处张冷炀、科技部新一代人工智能发展研究中心李修全副主任、北京智源人工智能研究院副院长曹岗、北京大学哲学系副主任、人工智能学院副院长刘哲、英国驻华大使馆科技处Anthony Myers、Lori Han、剑桥大学、牛津大学等相关学者代表出席并发言。中英人工智能伦理与治理研究中心将与合作机构共同推动围绕人工智能伦理与治理相关的联合研究课题、青年学者互访项目，推动中英双方在人工智能及其社会影响、伦理与治理的跨文化认知、技术模型、人工智能技术风险与伦理评估、通用人工智能不同实现技术途径及其长远风险等方面开展深入合作。

实验室2项成果分获第九届吴文俊人工智能科学技术奖一、二等奖

近日，中国人工智能学会公布了《关于授予81项成果2019年度“吴文俊人工智能科学技术奖”的决定》，实验室2项成果分获第九届吴文俊人工智能科学技术奖一、二等奖。实验室脑网络研究中心为第一完成单位和天津医科大学总医院、首都医科大学宣武医院合作完成的项目成果“脑网络信息处理规律及其遗传
机制研究”获得吴文俊人工智能自然科学奖一等奖。实验室智能感知计算与研究中心为中国唯一单位与中科院、信工所、华北电力大学和山东科技大学合作完成的项目成果“基于图像视频分析的智能监测技术与应用”获得吴文俊人工智能技术发明二等奖。吴文俊人工智能科学技术奖是为加快实施国家创新驱动发展战略，深入贯彻《新一代人工智能发展规划》，更好地推进科教兴国、人才强国的政策落实，弘扬“尊重劳动、尊重知识、尊重人才、尊重创造”的良好风尚，通过选优优秀的智能科学技术成果及应用项目，充分调动和激发广大智能科技工作者投身建设创新型国家的积极性和创造热情，大力提升我国智能科学技术与产业应用的创新发展水平，隆重表彰在智能科学技术领域取得重大突破，做出突出贡献的单位和个人。

实验室获 2019 年 NeurIPS 模型压缩与加速竞赛双料冠军

2019 年 12 月 8 日至 14 日，神经信息处理系统大会（NeurIPS 2019）在加拿大温哥华举行，实验室图像与视频分析课题组程健研究员率队在本次大会的神经网络压缩与加速竞赛（MicroNet Challenge）中获得双料冠军！今年的 MicroNet Challenge 神经网络压缩与加速竞赛由 Google、Facebook、OpenAI 等机构在 NeurIPS2019 上共同主办，集结了 MIT、KAIST、华盛顿大学、浙大、北航等国内外著名前沿科研院校，同时还吸引了 ARM、IBM、高通、Xilinx 等国际一流芯片公司的参与。MicroNet Challenge 竞赛包括 ImageNet 图像分类、CIFAR-100 图像分类、WikiText-103 语言模型三个子任务，以模型的压缩率和加速比的比拼为主要内容，目的是通过算法达到最优的运算效率，牵引更高效的神经网络处理器架构设计，并最终实现神经网络软件、硬件加速的协同设计。程健研究员团队参加了竞争最激烈的 ImageNet 和 CIFAR-100 两个子赛道比拼，并最终包揽了图像类的两项冠军。团队结合极低比特量化技术和稀疏化技术，在 ImageNet 任务上相比主办方提供的基准模型取得了 20.2 倍的压缩率和 12.5 倍的加速比，在 CIFAR-100 任务上取得了 732.6 倍的压缩率和 356.5 倍的加速比，遥遥领先两个任务中的第二名队伍。团队同时受邀在大会上以“A Comprehensive Study of Network Compression for Image Classification”为主题详细介绍了相关的量化和稀疏化压缩和加速技术。针对比赛任务，团队在报告中给出解决办法：采用量化和稀疏化技术，将深度学习算法模型进行轻量化和计算提速，以大幅降低算法模型对算力、功耗以及内存的需求，让低端设备实现人工智能方案。团队成员冷聪副研究员表示，量化及稀疏化技术也是人工智能深度学习的软、硬件协同加速方案突破口。通过将其与人工智能硬件架构设计紧密结合，可以进一步降低人工智能技术落地难度，让 AI 更为易得易用。
实验室在 ICCV2019 竞赛中揽获 5 项冠军

在 2019 年国际计算机视觉大会（ICCV 2019）举办的多项竞赛中，实验室成员提交的算法表现优异，获得 5 项竞赛冠军。

实验室智能感知与计算研究中心俞宏远联合微软亚洲研究院在 VOT2019-RGBD 跟踪挑战赛中夺得了冠军。 Visual-Object-Tracking Challenge (VOT) 是国际目标跟踪领域最权威的评测平台，旨在评测在复杂场景下单目标跟踪的算法性能。其中 VOT-RGBD 竞赛所使用的数据集是目前国际公开数据集中挑战最大的 RGBD 跟踪数据集，该数据集中包含了跟踪目标形状变化、遮挡、消失后再出现以及黑暗环境下的目标跟踪等诸多挑战。我们使用的 SiamDW-D 跟踪模型，融合了 RGB 和深度两种模态的数据，赋予了传统跟踪模型处理多种模态数据和在线更新的能力。因此，在跟踪过程中能更好地适应跟踪环境和跟踪目标的变化，实现了在复杂场景下的高精度鲁棒跟踪。
实验室智能感知与计算研究中心牛凯等人在 WIDER Face and Person Challenge 2019 基于自然语言信息的行人检索 Person Search by Language 任务中获得冠军。基于自然语言信息来检索大规模数据集中的行人图片是一项重要且极具挑战性的任务。在很多场景下，我们并不能得到可靠的待搜索对象的视觉信息。比赛所用测试数据与训练数据来自完全不同的真实监控场景，将该任务向实际应用推进了一大步，同时也对模型提出了更高的挑战。我们的模型不仅能够很好的处理自然语言和视觉信息间巨大的模态差异进行准确检索，同时能够克服真实监控场景下的跨领域难题，极大的提升了检索的准确性，在获得竞赛冠军的同时得分远超其他参赛队伍，在智能安防、零售、智慧城市等场景中均具有巨大的应用前景。

实验室视频内容安全团队张子琦、史雅雅、魏久桐、原春锋、李兵、胡卫明在视频描述竞赛 VATEX Captioning Challenge 中获得中英文双料冠军。本次比赛由国际计算机视觉大会 ICCV 和美国加州大学圣巴巴拉分校 UCSB 联合举办，分为中文描述和英文描述两个赛道。视频描述涉及到计算机视觉和自然语言处理两个领域，而且视频中有表观、运动、语义属性甚至语音等多个模态。因此如何更好地融合多模态信息，成为本次比赛的最大挑战。团队在没有大量样本训练的前提下，采取恰当的训练方式，依靠坚实技术积累在中英文描述两个赛道均获冠军。

实验室智能感知与计算研究中心牛凯等人在 WIDER Face and Person Challenge 2019 基于自然语言信息的行人检索 Person Search by Language 任务中获得冠军。基于自然语言信息来检索大规模数据集中的行人图片是一项重要且极具挑战性的任务。在很多场景下，我们并不能得到可靠的待搜索对象的视觉信息。比赛所用测试数据与训练数据来自完全不同的真实监控场景，将该任务向实际应用推进了一大步，同时也对模型提出了更高的挑战。我们的模型不仅能够很好的处理自然语言和视觉信息间巨大的模态差异进行准确检索，同时能够克服真实监控场景下的跨领域难题，极大的提升了检索的准确性，在获得竞赛冠军的同时得分远超其他参赛队伍，在智能安防、零售、智慧城市等场景中均具有巨大的应用前景。

实验室视频内容安全团队张子琦、史雅雅、魏久桐、原春锋、李兵、胡卫明在视频描述竞赛 VATEX Captioning Challenge 中获得中英文双料冠军。本次比赛由国际计算机视觉大会 ICCV 和美国加州大学圣巴巴拉分校 UCSB 联合举办，分为中文描述和英文描述两个赛道。视频描述涉及到计算机视觉和自然语言处理两个领域，而且视频中有表观、运动、语义属性甚至语音等多个模态。因此如何更好地融合多模态信息，成为本次比赛的最大挑战。团队在没有大量样本训练的前提下，采取恰当的训练方式，依靠坚实技术积累在中英文描述两个赛道均获冠军。
实验室生物识别与安全技术团队刘浩、朱翔昱、雷震等人在 Light Weight Face Recognition Challenge（LFR）大规模图像人脸识别赛道中获得冠军。经过近几年深度学习的发展，人脸识别技术已经在学术界众多人脸测试集上的性能达到饱和，但是在实际应用场景中仍有许多问题亟待解决，尤其是大规模人脸识别和视频人脸识别问题。ICCV2019 Light Weight Face Recognition Challenge（LFR）竞赛便是针对这些问题举办的人脸识别挑战赛。在大模型图像人脸识别赛道（DeepGlint-Large）上，团队使用改进的 SE-AttentionNet-IR 作为基础架构，使用 CosFace 作为损失函数并重点调整了 margin，最终在有限的计算资源下战胜了众多参赛企业取得该赛道的冠军，并受邀到会议现场作题为 “High Performance Face Recognition without Bells and Whistles”的口头报告。

实验室图像与视频分析团队程健、冷聪、张一帆等在轻量级人脸识别比赛 Light Weight Face Recognition Challenge 和快速人脸识别 Face Detection Runtime 中分别获得亚军。Light Weight Face Recognition Challenge 和 WINDER CHALLENGE Face Detection Runtime Challenge 两项比赛是针对真实应用场景，在要求模型准确度同时，还分别重点关注模型的体积和检测速度。比赛所用数据集是目前国际公开的数据规模最大、真实性最贴近实际（数据来自真实场景）、检测难度最高的人脸识别检测数据集。我们的人脸识别模型，在模型精度、体积和速度上取得了突破，不但达到了高精度，还极大地降低了模型计算量和模型体
积，实现了超高性能，在安防、金融、物联网等各种实际人脸识别场景，尤其在AIoT边缘计算场景中有着广阔的应用潜力。

喜报：实验室蒋田仔研究员当选 2020 IEEE Fellow

近日，实验室蒋田仔研究员收到通知，当选 2020 IEEE Fellow！IEEE 是目前全球最大的非营利性专业技术学会，其会员人数超过 40 万人，遍布 160 多个国家，在电气与电子工程、计算机与控制、通信等领域具有广泛的影响力。IEEE Fellow（IEEE 会士）是 IEEE 授予会员的最高荣誉，被学术界认定为权威的荣誉和重要的职业成就，其评选过程非常严谨，每年由同行专家在做出突出贡献的会员中评选出，每次获选人数不超过会员总数的 0.1%，在世界学术界及工程技术界竞争非常激烈。

喜报：实验室赫然研究员入选为第五批青年创新促进会优秀会员

近日，院人事局发布通知，经多轮遴选，实验室赫然研究员入选为第五批青年创新促进会优秀会员。特向赫然同志表示祝贺！依据《中国科学院青年创新促进会章程》和《中国科学院青年创新促进会管理办法》，中国科学院青年创新促进会会员执行期四年，资助期满后对表现优异、成绩突出的应届会员开展优秀会员评选活动，对入选者给予 200 万-400 万追加经费支持，资助其持续地开展原创性研究工作和学术活动，加速成长为学术青年领军人才。优秀会员评选按照青促会理事会通过的《中国科学院青年创新促进会优秀会员评审细则》进行，主要程序包括本人申请、单位审核、秘书处形式审查、初评、复评、拟定入选公示等。

喜报：实验室两位老师入选中国图象图形学学会首批会士

中国图象图形学学会于 2018 年决议通过《中国图象图形学学会会士条例》。会士是学会会员的最高学术荣誉，授予在图像图形领域取得卓越成绩和为学会作出突出贡献的会员。12 月 2 日，首批会士名单公布。实验室卢汉清研究员、谭铁牛研究员入选。
喜报：实验室宗成庆研究员获得宝钢优秀教师奖

经国科大推荐，宝钢教育基金会审议，实验室宗成庆研究员获得 2019 年度宝钢优秀教师奖。宝钢优秀教师奖旨在奖励长期坚持工作在教学第一线，在教学内容、教材、方法、手段改革方面取得显著成果的优秀教师。本年度国科大仅有 2 名教师获评该奖项。

喜报：实验室何世柱等 3 人入选为 2020 年青年创新促进会会员

根据《中国科学院青年创新促进会管理办法》（科发人字〔2019〕25 号），经单位推荐、院审核，院人事局发布了 2020 年青年创新促进会会员人选名单。本次遴选共产生新会员 458 人，实验室何世柱、高晋、王伟 3 位青年科研人员成功入选。

科研进展

基于时空图结构学习的交通预测

交通预测是智能交通系统中不可或缺的一个部分。该任务受限于以下三个方面：第一，交通数据关联于道路网络，因此它属于图结构数据而不是标准的栅格矩阵；第二，交通数据中呈现出很强的空间依赖关系，这表明交通图中的节点之间通常拥有复杂且动态的关系；第三，交通数据展现出很强的时间依赖关系，因此交通时间序列建模是很重要的。为了解决上述问题，实验室时空数据分析师团队的博士生张奇、向世明研究员和潘春洪研究员等人提出新一个结构学习卷积框架，它可以将传统卷积扩展到图域，进而学习交通预测的图结构。该方法显式地将结构信息建模到卷积操作中；所提方法中利用了两个结构学习卷积模块，分别捕获
全局和局部的结构：然后将这两个模块整合到一个端到端的交通预测网络中。此外，该方法利用伪三维卷积抓取交通数据的时间依赖性。在六个公开数据集上的对比实验表明了该方法的有效性。相关工作已被国际会议 AAAI 2020 接收。

图 1. 基于时空图结构学习的交通预测模型框架

基于自适应上下文网络的场景解析

场景解析是计算机视觉领域中重要且具有挑战的方向，旨在预测出图像中的像素点属于某一个目标类或场景类。图像场景的复杂多样（光照，视角，尺度，遮挡等）对于场景的理解和像素点的判别造成很大困难。有效的上下文信息对于场景解析十分重要。传统方法在融合全局上下文和局部上下文到每个像素点时都是同等对待的，忽略了不同像素点对于全局上下文和局部上下文的需求差异性。针对这一问题，实验室图像与视频分析团队博士生付君，刘静研究员等人提出基于自适应上下文网络的场景解析方法。该方法首先利用全局上下文模块，根据每个像素点自适应地融合全局上下文信息。然后，利用局部上下文模块为每个像素点自适应地融合局部上下文信息。最后进行局部上下文和全局上下文的选择性融合，实现自适应融合像素感知的上下文信息捕获。在多个国际公开数据集上的实验结果证明该方法的有效性，相关工作发表在国际会议 ICCV 2019 上。

图 2. 基于自适应上下文网络的场景解析模型框架

基于图关联的无监督行人再识别

行人再识别旨在从多摄像机拍摄到的画面中检索到同一个行人的图像。该任
务在公共安全领域有重要的应用。由于应用场景的差异化，需要花费大量的时间和人力标注数据以提高行人再识别模型的性能。因此，使用无标注数据训练的无监督行人再识别算法，成为该领域的一个重要研究方向。针对这一问题，实验室生物识别与安全技术课题组博士生吴锦林，杨阳博士和雷震研究员等人提出了一种基于图关联的无监督行人再识别方法。该方法首先采用多任务训练的策略，使模型专注于学习单个视角内的行人图像的特征表示，然后使用图关联的方法挖掘潜在的正样本，使模型专注于学习跨视角行人图像的特征表示。该方法不仅能够建立一个较为精确的跨视角行人轨迹关联图，而且能够降低跨视角学习时噪声样本带来的干扰。在七个国际公开数据集上的实验结果证明了该方法的有效性。相关工作发表在国际会议 ICCV 2019 上。

图 3. 基于图关联的无监督行人再识别模型框架

基于对抗视觉特征残差的零样本学习

当前主流的物体识别算法的学习过程需要为每个物体类别提供大量的标注样本，且学习得到的模型鲁棒性不足。受限于大规模的样本标注，零样本学习成为了计算机视觉领域一个重要的研究方向。然而，零样本学习目前仍然存在着巨大挑战，传统的法通常是建立一种视觉特征与语义特征之间的映射关系，然而这种映射关系在未见类别上的泛化能力有限。针对这一问题，实验室机器人视觉课题组博士生刘博，董秋雷研究员和胡占义研究员提出了一种基于视觉特征残差的对抗特征生成网络；该网络将视觉特征看作视觉特征原型和特征残差的叠加，使用条件对抗生成网络生成视觉特征残差，使用回归模型预测视觉原型，然后将特征残差与视觉原型叠加，得到样本的视觉特征。为了增强视觉特征与语义特征的一致性，该方法基于回归误差进行特征选择，最终得到更优的语义一致性的对抗视觉特征。所提方法在国际公开数据集 AWA1、AWA2、SUN、APY、CUB 和 NABird 上达到了最优的结果，证明了该方法的有效性。相关研究工作已经被国际会议 AAAI 2020 接收。
基于快速核相关滤波的目标跟踪

近年来，基于相关滤波的目标跟踪算法在定位精度上不断取得突破，基于核相关滤波的目标跟踪算法通过缓解边界效应以及利用高维的卷积神经网络特征提高算法的定位精度。尽管这类方法显著提升了核相关滤波的定位精度，它们均难以实时运行。针对这一问题，图像与视频分析课题组博士生郑林宇、唐明研究员、王金桥研究员等人提出了基于快速核相关滤波的目标跟踪方法。该方法首先通过真实且稠密的采样方式在线获取训练样本，然后在时域对偶空间构建基于核岭回归的优化问题用于求解跟踪器模型，因此该方法不受边界效应影响。此外，该方法利用稠密样本特有的数据高度重叠的特点，并使用基于动态规划的方法过滤传统核矩阵构建过程中的冗余计算，进而降低算法的时间复杂度。该方法在求解线性方程组的过程中，使用上一帧的最优解作为本帧的初始解，并利用基于高斯-赛尔德的迭代方法高效求解跟踪器模型，进一步提高计算效率。在国际公开数据集上的实验结果证明了该方法的有效性，相关工作发表在国际会议 ICCV 2019 上。

<table>
<thead>
<tr>
<th></th>
<th>rdKCF*</th>
<th>C-COT</th>
<th>ECO</th>
<th>GPRT</th>
<th>HCF</th>
<th>SRDCF</th>
<th>decon</th>
<th>deep</th>
</tr>
</thead>
<tbody>
<tr>
<td>mOP-13</td>
<td>0.884</td>
<td>0.821</td>
<td>0.871</td>
<td>0.841</td>
<td>0.741</td>
<td>0.785</td>
<td>0.799</td>
<td>0.779</td>
</tr>
<tr>
<td>mOP-15</td>
<td>0.828</td>
<td>0.816</td>
<td>0.842</td>
<td>0.791</td>
<td>0.661</td>
<td>0.728</td>
<td>0.759</td>
<td>0.765</td>
</tr>
<tr>
<td>mPN-13</td>
<td>0.846</td>
<td>0.782</td>
<td>0.832</td>
<td>0.818</td>
<td>0.783</td>
<td>0.748</td>
<td>0.772</td>
<td>0.740</td>
</tr>
<tr>
<td>mPN-15</td>
<td>0.820</td>
<td>0.805</td>
<td>0.819</td>
<td>0.793</td>
<td>0.735</td>
<td>0.713</td>
<td>0.753</td>
<td>0.755</td>
</tr>
<tr>
<td>mFPS-15</td>
<td>24</td>
<td>3</td>
<td>6</td>
<td>5</td>
<td>11</td>
<td>8</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

表 1. 在数据集 OTB-2013 和 OTB-2015 上定位精度和速度对比结果。mOP, mPN, mFPS 分别为平均定位精度，平均归一化定位精度，平均速度。

GSLAM：一个通用的 SLAM 框架和基准

SLAM 技术目前取得了较大的成功并吸引了各个领域的关注。但是，如何统一已有算法接口、如何有效地确定速度、鲁棒性和可移植性的基准是当前急需解决的问题。针对上述问题，实验室三维可视计算团队徐士彪副研究员等人提出一种新型通用 SLAM 平台：GSLAM。该平台不仅能够支持性能评估功能，还提供相关工具包以便利研究者快速开发他们自己的 SLAM 系统。GSLAM 的核心贡献在于它提供一种通用的、跨平台和完全开源的 SLAM 接口，方便不同用户的研
究或商业需求。通过该平台，用户可以实现他们自己的功能函数作为插件以取得更好的性能，有效推动 SLAM 应用到实用化阶段。该研究成果发表在国际会议 ICCV 2019 上，代码开源网址: https://github.com/zdzhaoyong/GSLAM

基于图神经网络的视频精彩片段检测

随着视频应用的迸发，用户每天都会接触到大量的视频，浏览整个视频费时又乏味。视频精彩片段检测旨在提取视频的精华片段，从而缓解这种情况。现有视频精彩片段检测方法存在两个问题。首先，大多数现有方法仅专注于学习视频的整体视觉表示，但忽略了视频中物体及其交互对精彩部分的影响。其次，当前最好的方法通常采用成对排序损失的策略，没有使用全局信息。针对上述问题，实验室多媒体计算团队博士生张莹莹、徐常胜研究员等人提出了基于图神经网络的视频精彩片段检测方法。该方法构造一个对象感知图并从全局建模对象之间的关系。为了降低计算成本，将视频建模成两种类型的图：1）空间图，用于捕获每一帧中物体的复杂相互作用；2）时间图：获得每个帧的物体信息表示并捕获全局信息。在此基础上，该方法还利用图神经网络操作来学习视频片段的表示及它们之间的关系。此外，该方法采用多阶段损失来优化模型，第一阶段计算每个视频片段的得分，并使用分类损失优化；第二阶段根据前一阶段的得分得到难分样本对，再使用成对排序损失对模型进行优化。在两个国际公开数据集上的实验结果证明了该方法的有效性。相关研究工作已被国际会议 AAAI2020 接收。
基于条件门循环单元的文字风格模仿

笔迹风格模仿是一项十分具有挑战性的工作，它可以看成是一种风格迁移方法，即将一部分文字上包含的书写风格迁移到其它文字中。针对这一问题，实验室智能交互团队博士生赵博程、杨明浩副研究员、陶建华研究员等人设计了一种文字模仿算法，该算法能够根据 10 个左右的输入汉字提取手写体特征并输出模仿的手写体文字。该方法通过使用深度卷积神经网络提取作者手写风格特征，通过高维空间聚类后得到风格化编码；在此基础上将条件 GRU 网络改进为双条件 GRU 网络，利用该网络对于风格特征与文字信息的拼接编码进行处理，从而生成基于输入文字风格特征的模拟文字。相关工作已经被国际期刊 Pattern Recognition 接收。

基于交互式学习的同步语音识别与语音翻译

语音翻译是指将一种语言的语音翻译到另一种语言的语音或文本。现有方法多采用语音识别加文本翻译的级联系统实现，但存在着时间延迟、参数冗余和错误传递的缺陷。端到端的语音翻译模型在理论上可以缓解这些缺陷，但是受限于训练数据量和学习困难度，模型的性能仍然不佳。针对此问题，自然语言处理团
队博士生刘宇宸、张家俊研究员、宗成庆研究员和百度公司合作提出基于交互式学习的同步语音识别与语音翻译模型。该模型包含一个语音识别解码器和一个语音翻译解码器，两者通过一个交互注意力模块相连在训练过程中实现交互式学习。在解码的每一步，两个解码器不仅可以关注到已识别词语的信息，也可以关注到已翻译词语的信息，实现同步与交互式的语音识别与语音翻译。实验表明该模型不仅能够有效提升机器翻译性能，也能够得到更好的语音识别结果。相关研究成果已被国际会议 AAAI 2020 接收。

基于多角度图编码的答案句选择

问答任务是检验机器语言理解能力的一个重要手段，同时问答系统的研究也能够直接推动社交、医疗、金融等多个领域的智能化进程。答案句选择是问答任务的重要子任务，同时也是问答系统中的一个重要组成部分。该任务通常会给定一个问题，以及多个候选句子，要求系统能够从多个候选句中找出能够回答给定问题的答案句。传统方法大多忽略了候选句之间的语义关联，对每个候选句单独建模，并单独计算其作为答案句子的概率。然而，候选句子之间的关系是普遍存在的，尤其是在所有的候选句都来自同一篇章的场景下。忽略这种关系，可能会导致对每个候选句的语义理解不充分，进而使得答案句选择错误。针对这一问题，实验室自然语言处理团队博士生田志兴、张元哲博士、刘康研究员、赵军研究员等人提出了基于多角度图编码的答案句选择模型。该模型使用图结构从多个角度建模句间关系。构图策略分为静态和动态两种。其中，静态构图策略根据候选句原始文本，从实体共现、句间距离、语义相似度三个角度构建固定的句子关系图；动态构图策略根据候选句的初始表示，对节点应用自注意力机制构建一种动态优化的句子关系图。对句间关系的多策略多角度的图结构建模使得该模型更加充分的理解每个候选句的语义，进而更准确的选择答案句。相关工作已经被国际会议 AAAI 2020 接收。
图 9. 基于多角度图编码捕获句间关系的答案句选择模型

基于脑功能影像时间序列的多尺度卷积循环神经网络模型

功能磁共振（fMRI）在认知科学以及精神疾病研究中被广泛应用。但是基于 fMRI 的脑疾病分类问题通常具有高维小样本的特点，常用的解决方案难以充分利用时序动态变化信息。针对这一问题实验室脑网络组团队隋婧研究员等人提出了多尺度卷积循环神经网络算法，该算法能够直接在时间序列上使用 RNN 模型自动学习 fMRI 的时间和空间特征，突破了现有 fMRI 分析“先借助时间序列构建功能连接矩阵再分类”的传统框架。在实验室建立的千例级精神分裂症脑影像学数据库上的实验结果证明了该方法的有效性，并且在实验中通过该方法发现了区别诊断贡献最显著的脑区集中于纹状体和小脑，提高了算法的可解释性，有助于挖掘疾病相关的影像学标志物。相关研究工作已发表在 The Lancet 旗下国际期刊 EBiomedicine。

图 10. 多中心集成以及跨站点分类结果对比
学术交流

法国贡比涅技术大学 Thierry Denoeux 教授访问模式识别国家重点实验室

2019 年 11 月 12 日上午，法国贡比涅技术大学 Thierry Denoeux 教授应邀访问模式识别国家重点实验室，并做了题为《On New Connections between Deep Learning and The Theory of Belief Functions》的学术前沿报告。Thierry Denoeux 教授介绍了 Dempster-Shafer 理论中的置信函数在不确定性建模和推理上的应用，特别是介绍了 Dempster-Shafer 理论与 Logistic 回归以及多层神经网络回归模型的等价性。在本次报告中，Thierry Denoeux 教授首先介绍了 Dempster-Shafer 理论和基于该理论的建模方法。该方法主要是将独立的观测证据用置信函数表示，然后利用 Dempster’s rule 置信函数的结合律来结合所有的观测证据。在此之后，他以 Logistic 回归为例回顾了机器学习模型。对于多层神经网络，可以将第一层至倒数第二层网络看作是一个特征提取器，而最后一层可以看作传统机器学习模型中的分类器。接下来，他介绍了 Dempster-Shafer 理论与 Logistic 回归的等价关系。具体来说，每一个特征表示的输入可以看作 Dempster-Shafer 理论中的一个证据。每一个证据对特定的类别的 mass function 是特征信号的线性组合。在假设每个证据是相互独立的情况下，利用 Dempster 结合律，所得到的归一化的置信函数与 Logistic 回归模型完全相同。类似的方法也可以扩展到其他机器学习模型中的分类算法中，如支持向量机等。在最后 Thierry Denoeux 教授总结道相较于概率的模型，基于 Dempster-Shafer 理论的模型在量化不确定性方面更有优势。在报告期间和报告之后，与会师生对 Dempster-Shafer 理论在深度学习的应用等方面与 Thierry Denoeux 教授进行了广泛的交流。探讨了该理论在开放性、不确定的环境中模式识别方面的应用，深入讨论了在处理高置信度的分类问题、增量学习上 Dempster-Shafer 理论的优势等。
澳大利亚墨尔本大学 Tim Baldwin 教授访问模式识别国家重点实验室

2019 年 11 月 22 日上午，澳大利亚墨尔本大学 Tim Baldwin 教授应邀访问模式识别国家重点实验室，并做了题为《Memory-Augmented Neural Networks for Better Discourse Understanding》的学术前沿报告。Tim Baldwin 教授首先介绍了记忆网络（Memory Network）的两种类型：静态记忆网络和动态记忆网络。静态记忆网络是指一旦构建完毕后后续不再变化；而动态记忆网络则指网络内容会随着模型处理需要不断变化。然后，Tim Baldwin 教授利用非常形象的例子说明命名实体识别和对话状态跟踪等自然语言处理任务都需要篇章级别的理解，而记忆网络便是对篇章信息进行建模的一个有效手段。于是，Tim Baldwin 教授介绍了如何针对自然语言处理中的任务优化记忆网络，并分别针对静态记忆网络和动态记忆网络提出了切实有效的优化模型，在命名实体识别和对话状态跟踪等自然语言处理任务上取得了当时最好的性能。最后，Tim Baldwin 教授还提到记忆网络研究的一些未来方向。在报告结束后，与会师生与 Tim Baldwin 教授在记忆网络的设计、效率和优化方法进行了深入探讨。同时，针对记忆网络与当前流行的 BERT 等预训练模型的关系也进行了开放性讨论。

模式识别大讲堂：深度学习之路在何方？未来人工智能之路是否还依赖深度学习

2019 年 11 月 25 日下午，模式识别国家重点实验室青年学术沙龙活动在智能化大厦学术报告厅举行，模式学术交流小组有幸邀请到字节跳动人工智能实验室总监李航博士、北京大学王立威教授、中国科学院自动化所曾毅研究员以及陶建华研究员四位企业界和学术界一线专家与大家分享以“深度学习之路在何方？未来人工智能之路是否还依赖深度学习？”为主题的学术报告。本次活动由郭建伟副研究员主持，学术报告厅座无虚席，近 200 位所内外的老师和同学参加了本次活动。李航博士的报告题目是“AI 的未来：如何超越深度学习？”。李航博士首先表达了对 AI 的基本看法，认为当前 AI 的主要目的是为人类提供智能工具，提高人类的工作效率与生活舒适度。然后，从类脑计算、神经符号处理、合理思
考的机器三个方向对当前的深度学习理论进行讲述。最后，进一步深入介绍了如何超越深度学习。王立威教授报告的主题是“Thoughts on the future of deep learning”。王教授以两个问题为开篇：1.深度学习未来在哪？2.人工智能如何发展？然后从深度学习的任务、学习范式、模型以及学习算法四个方面对深度学习的未来作主题报告。王立威教授认为当前深度学习算法距离Human-level的任务还很遥远，而从人类进化的角度来看，监督学习占得比例很少而非监督学习占的比重很大。曾毅研究员报告的主题是“类脑人工智能：愿景、挑战与进展”。曾研究员首先介绍了行为建模与机制理解之间的差异，指出类脑人工智能是一项跨学科的研究工作，需要一个人同时具备神经科学和人工智能科学两种技术的人才能进行探索和研究。然后介绍了人工智能当前全球各个国家的不同愿景，以及面临的技术挑战和进展。

模式识别大讲堂：智能感知与计算团队研究方向与最新进展

2019年10月23日，中科院自动化研究所模式识别国家重点实验室副主任、智能感知与计算研究中心常务副主任王亮研究员在智能化大厦第1会议室模式识别国家重点实验室学术交流会作关于“智能感知与计算团队的研究方向和最新的进展”的报告。郭建伟副研究员主持了本次报告，张晓鹏研究员、张家俊副研究员、万军副研究员等老师和研究所诸多学生参加了报告会。近一个小时的报告结束后，与会人员对王亮研究员的报告意犹未尽，热烈进行提问和讨论，活动得到了很好的反响。王亮研究员首先对智能感知与计算研究中心的基本情况进行简单的介绍，包括中心成立背景、定位、组织机构等，中心致力于研究泛在智能感知理论与技术以及与之相伴的海量感知数据的智能分析与处理。接着，重点介绍了智能感知中心的研究方向与研究内容，主要包括智能感知基础理论、生物启发的智能计算、生物识别与安全及多模态智能计算四大方向。智能感知基础理论旨在基于认知机理和信息理论，研究面向感知问题的机器学习理论和优化方法，智能采集、处理、分析符合人类视觉认知的数据生成和模拟方法，从而赋予机器类人表达能力与想象能力。生物启发的智能计算方向以卷积神经网络为代表的计算机视觉技术近年来在性能上取得突破，在部分任务上甚至超越人类感知水
模式识别国家重点实验室简报

2019年第4期

平，但现有的视觉感知与深度学习的局限依旧明显，模拟大脑已经成为人工智能发展趋势。生物识别与安全在人脸识别核心算法方面，从人脸检测、关键点定位、人脸识别、特征提取、特征比对，活体检测全流程提出了一系列创新理论和方法；另外，中心近几年在远距离虹膜识别、移动端虹膜识别、动物虹膜识别等领域取得了新的应用突破。多模态智能计算涵盖三种理论方法，其一深度认知神经网络在深度认知神经网络模拟认知过程中的选择性注意、长短时记忆等机制扩展到现有深度神经网络的认知功能；其二视觉语言交互主要研究视觉与语音、文本等不同模态数据之间的匹配、融合、生成等问题；其三跨模态学习主要研究缩小语义鸿沟、处理混合小样本、提升模型泛化性等。

模式识别青年学术沙龙：人工智能结合真实场景在垂直领域中的应用

2019年10月25日下午，模式识别国家重点实验室青年学术沙龙活动在智能化大厦学术报告厅举行。本次活动邀请了滴滴AI Lab图像组负责人沈海峰博士、墨奇科技CTO汤林鹏博士和旷世科技研究院基础模型组负责人张祥雨博士三位企业界一线专家为大家讲述人工智能结合真实场景在垂直领域中的具体应用。本次活动由万军副研究员主持。沈海峰博士的报告题目是“复杂出行环境下的人脸识别技术”。沈博士主要从滴滴AI Lab图像组介绍，应用场景和市场规模分析，复杂人脸分析特点出发，详细介绍滴滴在人脸识别中的工作，比如人脸检测优化加速，人脸识别目标函数构造，3D姿态估计等。汤林鹏博士的报告题目为“大规模、非结构化数据搜索”，介绍利用人工智能技术处理和搜索大规模（二百亿级别）非结构化指纹数据。汤博士从传统特征与深度特征融合的多尺度特征提取，基于多尺度、多样化特征表示以及GPU+CPU混合的异构高性能比对，基于高性能特征存储和容错、多中心系统的大规模分布式计算三个方面进行详细介绍。张祥雨博士的报告题目是“Recent Advances in AutoML”。首先，张博士回顾自2017年以来神经网络架构搜索（NAS）技术，说明基于NAS的深度学习模型已经受到学术界和工业界越来越多的关注，如在分类、检测、分割等取得了超越手工设计模型的性能，然后张博士介绍了旷世科技在NAS方面的研究成果，
如利用 Single Path One-shot NAS 技术在“性能”、“效率”和“灵活性”上如何取得平衡，最后介绍了 NAS 在模型剪枝、检测等方面的工作。

项目立项

实验室第四季度新建立课题 65 项

实验室 2019 年第四季度新建立课题 65 项，总科研经费 5110.1 万元，其中国家任务 38 项、中科院任务 3 项、企业委托 24 项，部分项目如下:

<table>
<thead>
<tr>
<th>项目名称</th>
<th>项目负责人</th>
<th>项目类型</th>
<th>经费（万元）</th>
<th>执行期</th>
</tr>
</thead>
<tbody>
<tr>
<td>星载图像处理软件</td>
<td>潘春洪</td>
<td>企业委托</td>
<td>160</td>
<td>2019-7-31 至 2020-7-31</td>
</tr>
<tr>
<td>短视频文字提取软件开发</td>
<td>杨沛沛</td>
<td>企业委托</td>
<td>30</td>
<td>2019-8-7 至 2020-8-7</td>
</tr>
<tr>
<td>基于音视频的精神状态检测技术研究</td>
<td>刘斌</td>
<td>高技术应用项目</td>
<td>45</td>
<td>2019-6-1 至 2019-12-30</td>
</tr>
<tr>
<td>重点任务图区域信息采集</td>
<td>赫然</td>
<td>国家重点研 发计划</td>
<td>205</td>
<td>2019-2-1 至 2022-2-28</td>
</tr>
<tr>
<td>知识关联与事件推理类问题求解关键技术与系统</td>
<td>刘康</td>
<td>国家重点研 发计划</td>
<td>237.51</td>
<td>2019-5-1 至 2022-10-31</td>
</tr>
<tr>
<td>智能语音评价方案（语音合成方向）</td>
<td>陶建华</td>
<td>企业委托</td>
<td>19.5</td>
<td>2019-8-20 至 2020-8-20</td>
</tr>
<tr>
<td>基于知识挖掘提升 GAN 稳定性的技术研究</td>
<td>赫然</td>
<td>企业委托</td>
<td>95</td>
<td>2019-9-1 至 2020-8-31</td>
</tr>
<tr>
<td>快速高精度的大场景三维建模</td>
<td>崔海楠</td>
<td>联合资助基础</td>
<td>40</td>
<td>2019-1-1 至</td>
</tr>
</tbody>
</table>

19
<table>
<thead>
<tr>
<th>课题名称</th>
<th>负责人</th>
<th>项目类型</th>
<th>学校</th>
<th>课题编号</th>
<th>开始日期</th>
<th>结束日期</th>
</tr>
</thead>
<tbody>
<tr>
<td>面向小数据语音建模的跨语言迁移学习研究</td>
<td>易江燕</td>
<td>青年科学基金项目</td>
<td>哈尔滨</td>
<td>26</td>
<td>2020-1-1</td>
<td>2022-12-31</td>
</tr>
<tr>
<td>跨场景的人体目标轨迹分析</td>
<td>姚涵涛</td>
<td>青年科学基金项目</td>
<td>哈尔滨</td>
<td>29</td>
<td>2020-1-1</td>
<td>2022-12-31</td>
</tr>
<tr>
<td>基于场景精细感知对抗的虚假图像合成研究</td>
<td>彭勃</td>
<td>青年科学基金项目</td>
<td>哈尔滨</td>
<td>28</td>
<td>2020-1-1</td>
<td>2022-12-31</td>
</tr>
<tr>
<td>基于视-听觉认知的对话类视频压缩编码</td>
<td>刘雨帆</td>
<td>青年科学基金项目</td>
<td>哈尔滨</td>
<td>28</td>
<td>2020-1-1</td>
<td>2022-12-31</td>
</tr>
<tr>
<td>基于射影不变量的射影空间曲面匹配问题研究</td>
<td>韩杰</td>
<td>青年科学基金项目</td>
<td>哈尔滨</td>
<td>18</td>
<td>2020-1-1</td>
<td>2022-12-31</td>
</tr>
<tr>
<td>语言表征机理及受脑启发的文本表示模型研究</td>
<td>王少楠</td>
<td>青年科学基金项目</td>
<td>哈尔滨</td>
<td>27</td>
<td>2020-1-1</td>
<td>2022-12-31</td>
</tr>
<tr>
<td>面向网络短视频的暴力内容识别与语义描述</td>
<td>李文娟</td>
<td>青年科学基金项目</td>
<td>哈尔滨</td>
<td>25</td>
<td>2020-1-1</td>
<td>2022-12-31</td>
</tr>
<tr>
<td>深度神经网络无监督定点量化方法研究</td>
<td>王培松</td>
<td>青年科学基金项目</td>
<td>哈尔滨</td>
<td>27</td>
<td>2020-1-1</td>
<td>2022-12-31</td>
</tr>
<tr>
<td>文本信息抽取与问答系统</td>
<td>刘康</td>
<td>优秀青年科学基金项目</td>
<td>哈尔滨</td>
<td>130</td>
<td>2020-1-1</td>
<td>2022-12-31</td>
</tr>
<tr>
<td>图像语义理解</td>
<td>刘静</td>
<td>优秀青年科学基金项目</td>
<td>哈尔滨</td>
<td>130</td>
<td>2020-1-1</td>
<td>2022-12-31</td>
</tr>
<tr>
<td>基于多视角卷积网络的稠密SLAM场景重建与识别</td>
<td>徐士彪</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>58</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
<tr>
<td>基于空间采样与深度感知融合策略的语音分离方法研究</td>
<td>梁山</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>59</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
<tr>
<td>成年人脑老化全阶段的磁共振影像学研究</td>
<td>左年明</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>59</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
<tr>
<td>面向轮式机器人应用的在线实时目标跟踪研究</td>
<td>高晋</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>61</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
<tr>
<td>深度神经网络高效计算方法研究</td>
<td>程健</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>60</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
<tr>
<td>基于RGB-D视频的多模态人体行为识别研究</td>
<td>原春锋</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>61</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
<tr>
<td>城市场景序列化图像稠密重建与高精语义模型构建</td>
<td>张晓鹏</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>60</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
<tr>
<td>基于对抗机制的深度几何变换学习及其应用研究</td>
<td>孟高峰</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>61</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
<tr>
<td>基于双样本学习的目标跟踪</td>
<td>唐明</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>62</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
<tr>
<td>融合异构知识的交互式对话模型关键技术研究</td>
<td>赵军</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>60</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
<tr>
<td>基于注意记忆计算建模的模型可解释性研究</td>
<td>王威</td>
<td>面上项目</td>
<td>哈尔滨</td>
<td>60</td>
<td>2020-1-1</td>
<td>2023-12-31</td>
</tr>
</tbody>
</table>
模式识别国家重点实验室简报

基于深度学习的生物荧光显微图像分割关键技术研究与应用

杨戈

面上项目

60

2020-1-1 至2023-12-31

医疗对话文本中的信息抽取关键技术研究

张元哲

青年科学基金项目

25

2020-1-1 至2022-12-31

城市空间智能建模与虚实融合

张晓鹏

国家重点研发计划

145

2019-7-1 至2022-6-30

大数据技术分析

陶建华

企业委托

1000

2019-9-6 至2022-9-5

基于实况的机器学习预报及偏差订正技术探索

向世明

企业委托

40

2019-9-24 至2021-12-31

面向分类用户个性化需求的科技大数据精准服务技术

张兆翔

国家重点研发计划

199

2019-7-1 至2022-6-30

CX-6(01)图像数据单元

潘春洪

企业委托

310

2019-11-1 至2020-12-31

数据驱动的多维度视频编辑与评价

严冬明

企业委托

49.68

2019-6-1 至2020-11-30

多模态人体特征识别算法与系统

万军

企业委托

17.45

2019-4-4 至2022-4-3

人脸识别技术研发

雷震

企业委托

20

2019-9-9 至2020-6-30

面向压缩域的视频内容分析与高效比对

刘雨帆

企业委托

100

2019-11-2 至2021-12-31

深度学习模型小型化和硬件加速技术研究

程健

企业委托

135

2019-8-20 至2022-12-31

室内室外大场景鲁棒的重定位技术研究

吴毅红

企业委托

177.41

2019-8-20 至2022-12-31

乡村住宅设计逻辑提取与空间布局生成研究

郭建伟

企业委托

20

2019-9-1 至2020-4-30

神经网络机器翻译关键技术研究

张家俊

企业委托

30

2019-1-1 至2020-7-31

文体活动

智能感知中心 “云蒙山秋游” 活动

天高云远，秋色宜人。10 月 15 日，智感中心工会联合中心党支部及研究生会组织百余名职工和研究生共赴密云云蒙山地质公园，共同领略这一可观山观水观长城、赏云赏瀑赏石松的京华胜境的奇美风光。云蒙山位于密云西部，集泰山之雄、华山之险、黄山之奇、峨眉之秀于一体，素有“北方小黄山”之称。大家徒步几小时，在磨练自己意志的同时，也陶冶了情操，增进了友谊。