

CCF

计算机视觉 专委会简报

COMPUTER VISION NEWSLETTER

2019/03 期
总第 19 期

专委动态

走进高校系列活动

科技前沿

研究热点追踪

专题综述



主 办： CCF 计算机视觉专业委员会

主 编： 王 亮

执行主编： 李实英

网 址： <http://ccfcv.ccf.org.cn>

E m a i l： ccfcvn@gmail.com

COMPUTER VISION NEWSLETTER

计算机视觉

专委简报编委会

主 编 王 亮 中国科学院自动化研究所

执行主编 李实英 上海科技大学

专委动态

主 编 毋立芳 北京工业大学

编 委 马占宇 北京邮电大学

王瑞平 中国科学院计算技术研究所

虞晶怡 上海科技大学

科技前沿

主 编 王金甲 燕山大学

编 委 崔海楠 中国科学院自动化研究所

邓 成 西安电子科技大学

任传贤 中山大学

任桐炜 南京大学

苏 航 清华大学

杨巨峰 南开大学

委员风采

主 编 余 焯 合肥工业大学

编 委 黄 岩 中国科学院自动化研究所

刘海波 哈尔滨工程大学

张汗灵 湖南大学

资源平台

主 编 李 策 兰州理工大学

编 委 樊 鑫 大连理工大学

贾 同 东北大学

蹇木伟 山东财经大学

金 鑫 北京电子科技学院

刘 丽 国防科学技术大学

沈沛意 西安电子科技大学

COMPUTER VISION NEWSLETTER

CONTENTS

目录

专委动态

走进高校	走进高校系列活动	04
走进企业	走进企业系列交流会	05
视界无限	视界无限系列研讨会	06
专委新闻	CCF-CV 常委年度工作会议	08
会议推介	PRCV 2019 工作进展	09
	PRCV 2019 诚招赞助	10
	PRCV 2019 诚邀申请讲习班	12
	PRCV 2019 诚邀申请成果展示	13
	PRCV 2019 博士生论坛征集通知	14

科技前沿

专题综述	生成式对抗网络年度进展概述	16
热点追击	基于情感区域自动挖掘的视觉情感预测	22
	多长度哈希联合学习方法	23

委员风采

委员访谈	南京大学吴建鑫教授访谈	24
委员好消息		26

资源平台

开源代码	从集成聚类到大规模聚类	28
数据集	交通场景动态理解	31
招聘信息		35
征文通知		40

CCF-CV 走进高校系列报告会

第 69 期 青海大学



2019年3月28日，第六十九期“CCF-CV”走进高校系列报告会活动在青海大学科技馆二楼举行。此次报告会以“计算机视觉前沿技术及应用”为主题，会议邀请了中国科学院自动化所王亮研究员、浙江大学潘纲教授、西安电子科技大学苗启广教授三位学者作特邀报告。报告会由计算机技术与应用系副主任刘志强及副主任张玉安共同主持，来自青海大学、青海师范大学、青海民族大学这三所高校的近200名师生及相关企业同行们聆听了报告。

第 70 期 太原理工大学



2019年4月16日，第七十期走进高校系列

报告会在明向校区大数据学院9层报告厅举行。本期报告会邀请了中科院自动化所王亮研究员，武汉大学夏桂松教授，北京大学刘家瑛副教授，上海交通大学马超助理教授四位专家学者做特邀报告。大数据学院陈泽华教授、李东喜副教授和郭学俊老师担任本次报告会的执行主席。

第 71 期 国防科技大学



2019年4月21日，第七十一期走进高校系列报告会活动在国防科技大学103教学楼举行。本期报告会邀请了中山大学林惊教授，清华大学鲁继文副教授，复旦大学姜育刚教授，微软亚洲研究院王井东研究员，华中科技大学白翔教授，厦门大学纪荣嵘教授，浙江大学章国锋教授，中国科学院大学叶齐祥教授，中国科学院自动化研究所赫然研究员，清华大学孙富春教授十位专家学者做特邀报告。国防科技大学系统工程学院刘丽副教授、郭得科教授和智能科学学院胡德文教授担任本次报告会的执行主席。报告会吸引了来自国防科技大学、中南大学、湖南大学等高校约200人聆听了报告。

(责任编辑：马占宇)

CCF-CV 走进企业系列交流会

第 17 期 高新兴科技集团

时间：2019 年 4 月 26 日

中国计算机学会计算机视觉专委会（CCF-CV）走进企业系列交流会第十七期活动——走进高新兴，在广州市黄埔区高新兴科技集团股份有限公司成功举办。这是 CCF-CV 专委会开展走进企业系列交流会以来，首次走进广州企业，领域内多位著名学者齐聚一堂，针对计算机视觉前沿技术及应用展开了深度交流与探讨。



在活动开始前，高新兴科技集团董事长刘双广在致辞中首先向 CCF-CV 以及到场业界专家和嘉宾表达了感谢。刘双广表示，高新兴是拥有 20 多年历史的智慧城市物联网产品和服务提供商，主要从事基于物联网、人工智能等核心技术、产品及解决方案的研发和应用。期待未来同各位专家学者开展合作，共同促进计算机视觉领域的产学研融合。在热烈的掌声中，交流会正式开始。

首先，高新兴集团中央研究院副院长毛亮博士展示了高新兴中央研究院的技术实力和在人工智能领域取得的成果，全面介绍了高新兴围绕人脸识别、行为分析、车辆识别、视频内容分析

等方向的布局和算法落地情况，并分析了工程实践中面临的难点和算法需求。

接下来，上海交通大学林巍晓教授带来了面向监控视频的行为识别与理解的技术分享报告。林教授在报告中介绍了行为识别技术的发展和主要类别，之后介绍了其课题组的一些工作进展，主要包括多目标轨迹鲁棒提取、基于轨迹信息的行为识别、基于时空序列的行为识别等。

随后，中国科学院自动化研究所程建研究员分享了边缘端视觉计算技术的精彩主题报告。在介绍边缘端上的密集计算后，程建研究员重点讲解了多种轻量化的解决方案，并对未来轻量化的发展给出了独到的见解。

中山大学郑伟诗教授带来弱监督行人重识别的专题报告。郑教授介绍了无监督行人重识别的方法，并重点讲解了弱监督行人重识别技术，针对 RGB-IR 行人重识别、Attribute-Image 行人重识别等问题进行了深入探讨。

北京科技大学殷绪成教授分享了面向边缘计算的文字识别与人脸识别技术的专题报告。殷教授就面向边缘计算/AI 芯片的人工智能技术在车牌检测与识别、文本检测与识别、人脸检测与识别等几个领域的应用作了详细的介绍。

主题报告结束后，进入深度交流环节。参会人员纷纷与嘉宾展开互动，问答环节精彩纷呈，现场气氛非常热烈。活动结束后，高新兴为每位参会人员赠送了纪念徽章和讲义汇编，大家一同合影留念，祝贺本次交流活动圆满成功。

（责任编辑：王瑞平）

CCF-CV 视界无限系列研讨会

第 2 期 浙江大学



2019年3月31日，由中国计算机学会计算机视觉专委会主办、浙江大学CAD&CG国家重点实验室承办的第2期CCF-CV“视界无限”系列活动——“同时定位与地图构建（SLAM）的前沿进展与未来趋势”研讨会在杭州启真酒店成功举行。本期研讨会由计算机视觉专委会委员、浙江大学章国锋教授组织和主持。研讨会邀请了北京大学查红彬教授、中科院自动化所吴毅红研究员、上海交通大学邹丹平副教授、西蒙弗雷泽大学谭平副教授做主题报告，浙江大学鲍虎军教授、刘勇教授、阿里巴巴李名杨研究员参加了深度研讨环节。本次活动报名非常火爆，收到了200余份参会申请，经过认真遴选，来自全国各高校和企业的80名参会者汇聚杭城，就SLAM的理论与应用进行了热烈的交流和深入细致的探讨。

会议议程依次包括引导发言、主题报告、海报/系统展示、深度研讨四部分。首先，浙江大学鲍虎军教授代表研讨会承办单位欢迎大家来到杭州参加“视界无限”活动。鲍老师指出，这

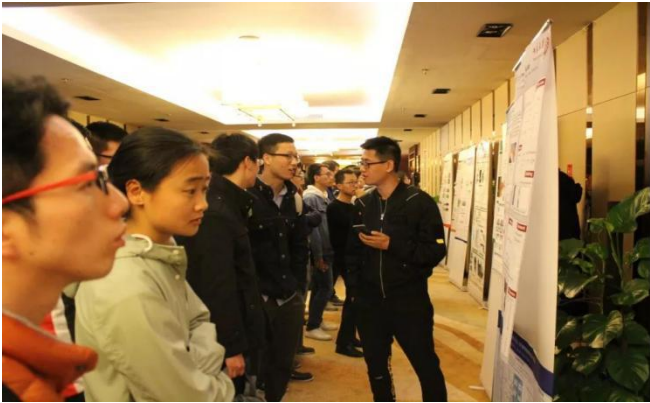
种小规模深入讨论对SLAM的发展有很大的推动作用，期待听到专家的精彩报告，并祝愿研讨会圆满成功。

章国锋教授的引导发言题目是“面向增强现实的单目视觉惯性SLAM算法评测”。章国锋教授指出，尽管视觉惯性SLAM已经取得了很大的成功，但由于缺乏合适的基准，目前还比较难从增强现实的角度来定量地评估各种SLAM系统的定位结果。现有SLAM数据集/基准一般只提供相机位姿的精度估计，而且相机运动类型比较简单，与移动增强现实中的常见运动情况并不吻合。为了解决这一问题，构建了一个新的视觉惯性数据集及相应的面向增强现实的评测标准，对现有的单目VSLAM/VISLAM方法进行了细致的分析和比较，并从中选出几个代表性的方法/系统在新基准上进行定量评估。

查红彬教授的引导发言题目是“基于数据流处理的SLAM技术”。传统的SLAM技术在实际应用中，仍存在重建误差积累严重、计算成本高昂等问题，影响了机器系统的在线响应速度以及对复杂环境的自适应能力。针对这些问题，应该最大限度地利用传感数据的时空一致性与三维地图的几何不变性，在现有多视点几何计算的基础上，强化SLAM算法的系统性与泛化能力，进一步改善其基本性能。

在主题报告环节，吴毅红研究员的报告题目是“Marker SLAM及FMD SLAM”。首先给出了目前marker SLAM存在的问题：几乎所有的marker SLAM都需要与空间点进行对应，然后采用PnP

与 RANSAC 进行定位。在运动模糊、较远距离、较大噪声情况下都容易失败。由于点个数较少，即使采用 RANSAC 也难以剔除误匹配。为此提出了一种新的基于圆形的 marker SLAM。接下来，吴毅红研究员介绍了自然场景下的 FMD SLAM，其中将特征法与直接法进行融合，在前端采用直接法进行跟踪，在后端采用特征法对地图进行优化和更新，达到了速度与精度都可以兼顾的目的。



邹丹平副教授的报告题目是“无人系统中的视觉 SLAM-融合环境与载体信息的方法”，首先回顾了环境结构特性的描述方法与观测模型，以及移动机器人与旋翼无人机两种典型载体的运动特性，在此基础上介绍了针对应用场景深度订制的视觉 SLAM 思路。该思路从环境结构特性与载体运动特性两方面入手，并将两者融入到视觉 SLAM 系统中以提升实际应用中的精度与可靠性。

谭平副教授的报告题目是“A Novel Approach Toward Dense Monocular SLAM”。视觉 SLAM 已有多年的研究历史，且有 ORB-SLAM 和 DSO 等成功的应用。这些方法通常只重构场景的稀疏图，即，用一组稀疏点云描绘场景结构。然而在 AR/VR 情境下，很多应用需要稠密的 3D 点云以实现避障或用户交互。Dense SLAM 旨在追踪相机运动轨迹的同时建立稠密 3D 图，其中单目 Dense SLAM 仍是一个亟待解决的问题。谭平副教授介绍了两种最新的解决方案：利用偏振相机实

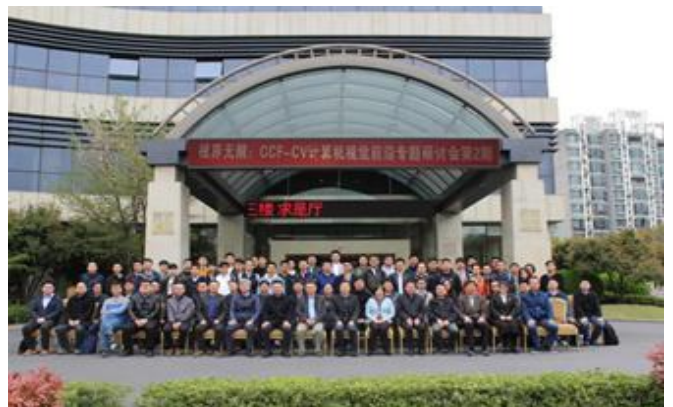
现立体重建或使用深度学习学习形状先验信息以实现立体重建。

本次研讨会新增了海报/系统展示环节，共有来自企业和高校的 14 篇论文海报和 4 个系统进行了演示，为与会者提供了更多细致交流和深入探讨问题的机会。



在 Panel 环节，与会嘉宾与现场师生就“SLAM 的未来发展之路在哪里？”、“SLAM 的 killer app 是什么？”、“SLAM 的产学研如何做？”、“怎么看 Event Camera SLAM？”等问题展开热烈讨论，参会者也就自己关心的问题与嘉宾交换了观点。

最后，计算机视觉专委会副主任查红彬教授总结了本次会议取得的成果并对承办方浙江大学 CAD&CG 国家重点实验室的师生表示了感谢，第 2 期“视界无限”研讨会在热烈的掌声中圆满结束。



(责任编辑：毋立芳)

CCF-CV 常委年度工作会议

2019年4月19日，中国计算机学会计算机视觉专委会（CCF-CV）常务委员会2019年度工作会议于北京京仪大酒店圆满召开。

本次常委会工作会议系统梳理了一年以来专委各项工作的进展、存在的不足，并针对专委后续发展的重点，特别是为过去几年中开创的若干品牌学术活动的继续创新突破指明了方向。会议首先由专委主任谭铁牛院士做主题发言，他充分肯定了专委会过去一年取得的各项进展，并对后续工作提出了多项切实可行的发展建议。

接下来，由PRCV2019大会组织委员会主席、西北工业大学王鹏教授代表大会组委会向常委会汇报了“2019第二届中国模式识别与计算机视觉大会（PRCV2019）”的各项筹备情况。作为本年度专委最重要的学术活动，PRCV2019将于11月份在西安举行，由中国四大一级学会（CCF、CAA、CSIG、CAAI）联合主办，由西北工业大学、西安交通大学、西安电子科技大学以及陕西师范大学等单位承办。会议目前进入论文的征稿环节，国内相关领域师生投稿活跃，一方面反映了国内模式识别与计算机视觉领域研究的繁荣态势，另一方面也体现了PRCV作为领域国内旗舰会议的品牌影响力。本届大会的论文集将由Springer出版社出版，并被EI和ISTP检索；部分优秀论文经扩展后将被推荐至国内外高质量期刊的特刊（包括Pattern Recognition、Neurocomputing等）发表。此外，大会的四位Keynote讲者也已全部确定，分别是Marr奖获得者、加拿大多伦

多大学Kyros Kutulakos教授，IEEE Fellow、IET Fellow、西安电子科技大学焦李成教授，IEEE TPAMI 副主编、比利时鲁汶大学 Tinne Tuytelaars 教授，IEEE TPAMI 副主编、首尔国立大学 Lee Kyoungmu 教授。会议的其它各项组织工作目前正有条不紊地推进中。

之后，常委会针对秘书处所整理的专委发展相关议题展开讨论，并逐一形成了具体可行的指导性建议，为专委下一步发展明确了方向。议题涉及专委年度学术大会 RACV、专委奖励、专委工作条例和换届选举、常规学术活动、日常工作等5个方面近20项议题。其中，专委发展新举措包括：改革专委学术大会 RACV 的举办目标与形式，会议英文名称确定为 Recent Advances in Computer Vision，定位为国内计算机视觉领域的小规模精品研讨会；举办计算机视觉暑期学校，推广领域前沿并惠及广大研究生和青年学者等。



会议最后由谭主任作了总结发言。会议结束后，全体与会委员合影留念。

（责任编辑：王瑞平）

PRCV2019 工作进展

中国模式识别与计算机视觉大会 (Chinese Conference on Pattern Recognition and Computer Vision) 是由中国模式识别学术大会 (CCPR) 和中国计算机视觉大会 (CCCV) 合并而来, 定位国内顶级的模式识别和计算机视觉领域学术盛会。首届中国模式识别与计算机视觉大会 (PRCV2018) 于 2018 年 11 月 23-26 日在广州成功召开。

第二届中国模式识别与计算机视觉大会 (PRCV2019) 将于 2019 年 11 月 8-11 日在西安举行。PRCV2019 由中国计算机学会 (CCF)、中国自动化学会 (CAA)、中国图象图形学学会 (CSIG) 和中国人工智能学会 (CAAI) 联合主办; 由西北工业大学、西安交通大学、西安电子科技大学以及陕西师范大学等单位联合承办。

论文投稿情况

PRCV 2019 投稿已于 4 月 30 日截止, 共收到 425 篇稿件, 其中有效稿件 408 篇。

特邀讲者



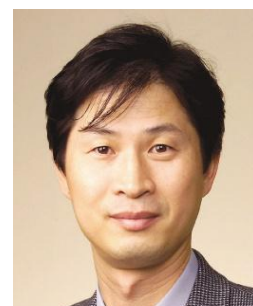
Kyros Kutulakos
加拿大多伦多大学教授
Marr 奖获得者



焦李成
西安电子科技大学教授
IEEE Fellow、IET Fellow



Tinne Tuytelaars
比利时鲁汶大学教授
IEEE TPAMI 副主编



Lee Kyoung Mu
首尔国立大学教授
IEEE TPAMI 副主编

已获赞助

• 钻石赞助



• 金牌赞助



• 银牌赞助



•铜牌赞助



会议邮箱

prcv2019@nwpu.edu.cn

会议网址

<http://www.prcv2019.com>

PRCV2019 诚招赞助

中国模式识别与计算机视觉大会 (Chinese Conference on Pattern Recognition and Computer Vision) 是由中国模式识别学术会议 (CCPR) 和中国计算机视觉大会 (CCCV) 合并而来, 由中国计算机学会 (CCF)、中国自动化学会 (CAA)、中国图象图形学学会 (CSIG) 和中国人工智能学会 (CAAI) 联合主办, 是模式识别与计算机视觉领域国内最高级别的学术盛会。首届中国模式识别与计算机视觉大会 (PRCV2018, <https://prcv-conf.org/2018>) 于 2018 年 11 月 23-26 日在广州举行, 由中山大学承办, 大会注册规模达到 1600 人, 被国内 10 多家媒体报道, 共获 25 家业界知名公司赞助, 有力促进了本领域国内学术水平的提高, 学术界与工业界之间的融合与交流。

第二届中国模式识别与计算机视觉大会 (PRCV2019) 将于 2019 年 11 月 8-11 日在西安举行, 由西北工业大学、西安交通大学、西安电子科技大学以及陕西师范大学等单位联合承办。本届大会将秉承提升我国模式识别与计算机视觉科学研究水平的发展理念, 着眼国内、面向国际, 为业内人士倾力奉献精彩的学术盛宴, 共同分享我国模式识别与计算机视觉领域的最新理论和技术成果。大会将在全面继承 PRCV2018 成

功经验的基础上, 进一步完善大会程序, 提高学术论文水平, 优化赞助商权益, 将大会报告、专题论坛、论文交流、企业宣讲、展台展示、前沿技术讲习班和行业应用竞赛等精彩内容与我国在本领域的学术进展和应用需求紧密结合, 以期更好地服务于我国在本领域的科技进步和发展需要。在当前人工智能技术蓬勃发展的大背景下, PRCV2019 预计参会人数将超过 1800 人, 可进一步落实国内本领域顶级盛会的地位, 并积极向国外拓展影响。

为成功举办 PRCV2019 大会, 现诚征赞助企业。通过本次大会, 赞助企业可有效扩大在国内模式识别和计算机视觉领域的知名度和影响力, 挖掘合作伙伴, 并促进产学研紧密合作。

大会根据赞助额度设钻石、金牌、银牌和铜牌等赞助级别。赞助级别和享有赞助权益如下:

钻石赞助

- 赞助金额: 15 万元
- 赞助服装: 含 PRCV2019 标志的企业冠名志愿者服装 (20 件)
- 赞助回报: 10 个全注册名额, 大会特邀演讲 20 分钟 (共 4 名, 以赞助合同签订时间排序, 并至少提前两周与赞助主席沟通演讲主题和讲者

人选)或1个专题论坛演讲30分钟(排名5及以后),1个专题论坛企业冠名,晚宴冠名权,1个晚宴颁奖人。1分钟企业宣传视频循环播放,1个展位(含2名参展人员)。在网站、会议手册和背板上显示企业LOGO,会务资料企业彩页,晚宴颁发致谢奖牌。

特注:对于赞助高于15万的企业,为了更好保障赞助权益,组委会考虑采取一事一议。

金牌赞助

- 赞助金额:8万元
- 赞助回报:8个全注册名额,1个专题论坛演讲20分钟,1个奖项冠名,1个茶歇冠名权,1个晚宴颁奖人。1分钟企业宣传视频循环播放,1个展位(含2名参展人员)。在网站、会议手册和背板上显示企业LOGO,会务资料企业彩页,晚宴颁发致谢奖牌。

银牌赞助

- 赞助金额:6万元
- 赞助回报:6个全注册名额,1个专题论坛演

讲15分钟。1分钟企业宣传视频循环播放,1个展位(含2名参展人员)。在网站、会议手册和背板上显示企业LOGO,会务资料企业彩页,晚宴颁发致谢奖牌。

铜牌赞助

- 赞助金额:4万元
- 赞助回报:4个全注册名额,1个专题论坛演讲10分钟。1分钟企业宣传视频循环播放,1个展位(含2名参展人员)。在网站、会议手册和背板上显示企业LOGO,会务资料企业彩页,晚宴颁发致谢奖牌。

注:PRCV2019将对2017年以来持续赞助PRCV/CCCV/RACV/CCPR的相关企业授予中国模式识别与计算机视觉大会特别赞助奖。

赞助主席

王涛, wtao@qiyi.com, 爱奇艺公司
杨金锋, jfyang@vip.163.com, 中国民航大学
赵歆波, xbozhao@nwpu.edu.cn, 西北工业大学
(责任编辑:毋立芳)

PRCV 2019

诚邀申请讲习班

中国模式识别与计算机视觉大会(Chinese Conference on Pattern Recognition and Computer Vision)是由中国模式识别学术大会(CCPR)和中国计算机视觉大会(CCCV)合并而来,定位国内顶级的模式识别和计算机视觉领域学术盛会。首届中国模式识别与计算机视觉大会(PRCV2018)于2018年11月23-26日在广州成功召开。

第二届中国模式识别与计算机视觉大会

(PRCV2019)将于2019年11月8-11日在西安举行。PRCV2019由中国计算机学会(CCF)、中国自动化学会(CAA)、中国图象图形学学会(CSIG)和中国人工智能学会(CAAI)联合主办;由西北工业大学、西安交通大学、西安电子科技大学以及陕西师范大学等单位联合承办。

现诚挚邀请广大模式识别与计算机视觉领域的从业者在PRCV2019会议上组织与承办计算机视觉、模式识别与机器学习方面的各种讲习班

活动。讲习班针对模式识别、计算机视觉与机器学习领域的某一特定方向（举例：稀疏编码、行为识别、目标检测、图像细分类、跨媒体分析、生成对抗学习等），由若干名讲者从不同维度、不同视角或者不同侧面等对特定方向的理论、方法与应用进行详细讲解。

有意愿在 PRCV 2019 组织或承办讲习班活动者，请于 2019 年 8 月 31 日前将讲习班申请书以电子邮件的形式提交给讲习班主席。申请结果将会在 2019 年 9 月 15 日通知组织者申请是否被大会采纳。所有被采纳的讲习班将会在 PRCV 2019 网站上公布。

PRCV 2019

诚邀申请成果展示

中国模式识别与计算机视觉大会 (Chinese Conference on Pattern Recognition and Computer Vision) 是由中国模式识别学术大会 (CCPR) 和中国计算机视觉大会 (CCCV) 合并而来，定位国内顶级的模式识别和计算机视觉领域学术盛会。首届中国模式识别与计算机视觉大会 (PRCV2018) 于 2018 年 11 月 23-26 日在广州成功召开。

第二届中国模式识别与计算机视觉大会 (PRCV2019) 将于 2019 年 11 月 8-11 日在西安举行。PRCV2019 由中国计算机学会 (CCF)、中国自动化学会 (CAA)、中国图象图形学学会 (CSIG) 和中国人工智能学会 (CAAI) 联合主办；由西北工业大学、西安交通大学、西安电子科技大学以及陕西师范大学等单位联合承办。

为促进模式识别与计算机视觉领域先进科研成果转化，打造实用性创新型研究成果的展示

讲习班申请书须包括以下信息

1. 讲习班主题
2. 讲习班简介、目的、意义
3. 拟邀请讲习班讲者及个人简介

讲习班主席

彭宇新, pengyuxin@pku.edu.cn, 北京大学
聂飞平, feipingnie@gmail.com, 西北工业大学
孟德宇, dymeng@mail.xjtu.edu.cn, 西安交通大学

(责任编辑: 毋立芳)

平台与窗口，加强学术界与工业界的沟通与交流，现面向广大高校、科研院所和企业公开征集模式识别与计算机视觉领域的优秀成果展示项目。展示内容包括但不限于模式识别与计算机视觉领域非商用的新型算法、模型、技术、系统原型、产品等，欢迎推选相关优秀 PR&CV 成果到大会进行宣传 and 展示。

有意参与 PRCV2019 成果展示者需填写申请表并发邮件至 xirunping@126.com，提供展示内容名称、作者、单位、成果简介、图片示例等信息。申请展示的成果将由组委会组织专家进行评选，大会将为每个入选的展示成果提供展位一个，每个展位需要至少全注册两名参会人员。

重要日期

- 申请截止: 2019 年 9 月 15 日
- 评选结果: 2019 年 9 月 30 日

- 会期展示：2019年11月8-11日

展示环境及要求

- 展位(会务提供)：标准展位(含一张展台、两把凳子和电源等)
- 展示设备(自备)：笔记本、显示器及其它设备等

- 成果简介：制作展示期间用海报1页，推荐尺寸：宽80cm，高120cm。(参展单位也可以根据展位大小结合需求自备展架或自行布展)

联系方式

郝润平, xirunping@126.com, 西北工业大学

PRCV 2019 博士生论坛征集通知

中国模式识别与计算机视觉大会(Chinese Conference on Pattern Recognition and Computer Vision)是由中国模式识别学术大会(CCPR)和中国计算机视觉大会(CCCV)合并而来,定位国内顶级的模式识别和计算机视觉领域学术盛会。首届中国模式识别与计算机视觉大会(PRCV2018)于2018年11月23-26日在广州成功召开。

第二届中国模式识别与计算机视觉大会(PRCV2019)将于2019年11月8-11日在西安举行。PRCV2019由中国计算机学会(CCF)、中国自动化学会(CAA)、中国图象图形学学会(CSIG)和中国人工智能学会(CAAI)联合主办;由西北工业大学、西安交通大学、西安电子科技大学以及陕西师范大学等单位联合承办。

论坛概况

本论坛为最近或即将获得博士学位的学生提供与领域内经验丰富的研究者沟通交流的机会。每位学生都会被分配一位领域内的资深成员作为导师,具体分配方案会基于学生的个人偏好或两者研究兴趣的相似程度而制定。投稿PRCV2019且论文被录用的学生可以优先选择导

师,获得跟导师面对面一对一交流的机会。导师团队包括但不限于如下著名专家学者:

谭铁牛 中科院自动化所研究员
王亮 中科院自动化所研究员
陈熙霖 中科院计算所研究员
赖剑煌 中山大学教授
王涛 爱奇艺公司首席科学家
查红彬 北京大学教授
白翔 华中科技大学教授
纪荣嵘 厦门大学教授
贾云得 北京理工大学教授
林宙辰 北京大学教授
刘青山 南京信息工程大学教授
卢湖川 大连理工大学教授
鲁继文 清华大学副教授
苗启广 西安电子科技大学教授
山世光 中科院计算所研究员
王蕴红 北京航空航天大学教授
虞晶怡 上海科技大学教授
张艳宁 西北工业大学教授
高新波 西安电子科技大学教授
黄庆明 中国科学院大学教授

赵 耀 北京交通大学教授
 薛建儒 西安交通大学教授
 杨小康 上海交通大学教授
 刘成林 中科院自动化所研究员
 王井东 微软亚洲研究院研究员
 姜育刚 复旦大学教授
 程明明 南开大学教授
 高盛华 上海科技大学助理教授、研究员
 程 健 中科院自动化所研究员
 彭宇新 北京大学教授
 孟德宇 西安交通大学教授
 左旺孟 哈尔滨工业大学教授
 欧阳万里 悉尼大学副教授
 董伟生 西安电子科技大学教授
 方玉明 江西财经大学教授
 丛 杨 中科院沈阳自动化所研究员
 毋立芳 北京工业大学教授
 韩军伟 西北工业大学教授
 耿 新 东南大学教授
 刘 偲 北京航空航天大学副教授

所有的学生和导师都会被邀请参加博士生论坛，使学生有机会与导师探讨其目前研究进展和未来职业规划。同时每位学生都需向其他参会者和导师展示一张海报，内容为介绍博士学位论文成果或最近发表的一篇论文成果。除海报展示，组委会还将综合申请材料择优评选出部分参会学生进行口头报告。2019 年度评选出的 CCF-CV 学术新锐奖获得者将被直接邀请发表口头报告。

参会资格

参会学生投稿时须为在读博士研究生，且研究方向为计算机视觉或模式识别领域。

申请步骤

满足参会资格的学生需把申请材料发送至

论坛主席邮箱，申请材料须为单个 pdf 文件或单个 zip/rar 压缩文件（内含多个 pdf 文件），材料中须包括内容如下：

1. 个人简历；
2. 1-2 页研究说明，总结申请者目前的研究进展；
3. 所展示海报的题目和作者列表；
4. 5 位候选导师的姓名和邮箱列表，按照偏好程度降序排列；

注：请不要列出自己的博士导师作为论坛导师。请确保申请材料中包含以上信息，不完整的材料将会被拒绝。

审核流程

如果因条件所限无法全部邀请满足要求的申请者，论坛组委会将基于上述申请材料及申请者的毕业时间进行选择，且会考虑覆盖尽量多的研究机构。

海报格式

海报格式为宽 240cm，高 120cm 的矩形。参会时请在指定房间中自行选择可用的展示板进行张贴，论坛组委会不指定海报序号。

重要时间节点

- 申请材料提交截止日期：2019 年 8 月 31 日
- 审核结果公布日期：2019 年 9 月 15 日

论坛主席

有任何问题，请联系以下论坛主席：

韩军伟, junweihan2010@gmail.com, 西北工业大学

耿 新, xgeng@seu.edu.cn, 东南大学

刘 偲, liusi@buaa.edu.cn, 北京航空航天大学

（责任编辑：毋立芳）

生成式对抗网络年度进展概述

哈尔滨工业大学 刘铭 颜肇义 左旺孟

一、引言

生成式对抗网络 (Generative Adversarial Network, GAN)^[1] 是近年来深度学习研究的热点之一, 在计算机视觉领域中的底层视觉、图像生成和迁移学习等任务中获得了广泛的关注与应用。GAN 最初关注的是图像生成, 采用对抗的方式同时学习一个生成网络 G 和一个判别网络 D 。 G 的输入是零均值标准方差的噪声向量 $\mathbf{z} \in \mathbf{R}^{n \times 1}$ ($z_i \sim \mathbf{N}(0,1)$), 旨在输出高质量的生成图像 $G(\mathbf{z})$ 。 D 是一个二值判别器, 用于判断输入的是真实图像 \mathbf{x} 还是生成图像 $G(\mathbf{z})$ 。生成式对抗网络优化的目标函数为,

$$\min_G \max_D \mathbb{E}_{\mathbf{x} \sim p_{data}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]. \quad (1)$$

由上式可见, 判别器的目标在于学习一个最优的二值分类器 D 以尽可能区分真实图像和生成图像, 而生成器 G 的目标在于学习一个最优的生成器以尽可能欺骗判别器 D 。通过以对抗学习的方式交替训练 D 和 G , 模型在理想情况下最终将达到纳什均衡, 此时 \mathbf{x} 和 $G(\mathbf{z})$ 的分布将完全相同, $G(\mathbf{z})$ 因而可以产生与真实样本难以区分的生成样本。

在 GAN 模型的各种推广形式中, 条件生成式对抗网络 (cGAN) 是最有价值 and 最具代表性的方式之一。不同于标准的 GAN, cGAN 的输入在随机变量 \mathbf{z} 之外, 还包括一个条件变量 \mathbf{c} 。CGAN 的生成器因而可以写为 $G(\mathbf{z}, \mathbf{c})$ 。条件变量 \mathbf{c} 的引入不但有助于改善 GAN 训练的稳定性 (如: InfoGAN、ACGAN), 还可进一步拓展 GAN 的应用范围。如图像翻译、智能图像填充、人脸属性编辑等任务都可以表示为条件图像生成问题。

然而, 基于公式(1)的 GAN 或 cGAN 训练往往会遇到梯度不稳定以及等问题而容易致使模式

坍塌 (Mode Collapse), 导致生成结果的多样性有限。针对这一问题, 国内外学者近年来开展了大量的研究工作并提出了一些有效的改进方法, 并推动了 GAN 在图像生成与翻译等任务中的广泛应用。

二、GAN 研究进展

1. 分布差异度量

直观地说, GAN 网络通过生成器与判别器的对抗学习, 使生成器能够产生判别器难以区分真假的数据样本, 这个过程本质是不断拉近两个 (或多个) 数据分布之间的距离。因此, 分布之间的差异度量是生成式对抗网络的学习中的一个核心问题。

如公式(1), Goodfellow 等^[1] 利用 KL 散度 (Kullback-Leibler divergence) 与 JS 散度 (Jensen-Shannon divergence) 衡量真实数据分布与生成数据分布之间的差异。然而, 生成式对抗网络将低维噪声映射到图像空间, 其实质是图像空间中的低维流形, 因此, 真实数据分布与生成数据分布之间的重叠往往可以忽略。此时, KL 散度与 JS 散度趋近于正无穷或一个常数, 无法准确反映两个分布之间的差异程度 (距离), 进而导致生成式对抗网络存在训练不稳定、梯度消失、模式坍塌等问题。

Arjovsky 等^[2] 对原始生成式对抗网络的不足进行了理论分析, 据此针对性地提出使用 Wasserstein 距离 (又称 Earth-Mover 距离) 对数据分布的差异进行度量, 并在此基础上提出 Wasserstein GAN (WGAN)^[3], 理论上解决了生成式对抗网络梯度消失等问题。Wasserstein 距离可以转换为基于 Lipschitz 约束的形式进行求解, WGAN 通过将判别器参数截断到一定范围内保证 Lipschitz 约束。然而, 该方法会导致参数二值化的现象, 使得判别器的判别能力大大下降。由

于 Lipschitz 约束 $\|T\|_L \leq 1$ 可以由 $\|\nabla T\| \leq 1$ 保证, Gulrajani 等^[4]进一步提出了改进的 WGAN-GP 模型, 通过梯度惩罚代替截断操作保证 Lipschitz 约束, 同时采取真伪样本随机插值的方法尽量保证 Lipschitz 约束在整个样本空间的成立。Mescheder 等^[5]通过对生成式对抗网络训练稳定性的数学推导, 提出了相比于 WGAN-GP 更为简单有效的梯度惩罚项。Wu 等^[6]提出了满足对称性约束的 Wasserstein 散度的 WGAN-div, 避免了对 Lipschitz 约束的要求。

2. 特征规范化

深度网络的规范化是指对于某层的特征或者权重进行某种形式的规范化, 从而改善网络学习的稳定性和训练效率。目前常见的有批规范化 (Batch Normalization, BN)、正交正则 (Orthogonal Regularization) 以及谱规范化 (Spectral Normalization) 等。其中批规范化^[7]在网络训练过程中, 对每个 mini-batch 的网络特征首先进行规范化然后进行线性变换, 通常能有效改善学习算法的收敛和稳定性。

除了对网络学习的特征进行规范化, 我们也可以对网络的权重进行约束。在生成式对抗网络中, 由于网络训练的不稳定性, 一种较好的初始化网络权重的方法显得尤为重要。为此, [8] 提出用正交初始化方法。在生成网络中, 如果卷积的权重是正交的, 那么特征经过正交的矩阵变换之后可以保持秩不变, 从而在较大程度上避免出现梯度消失或爆炸的情况。正交初始化方法如下:

$$L_{ortho} = \sum \left(\|WW^T - I\| \right), \quad (2)$$

其中 Σ 表示对卷积求和, W 是卷积权重, I 是单位阵。然而, 论文[9]指出这个初始化约束太强, 所以在[10]中提出一种新的正交规范化变体来最小化卷积核内部的余弦相似度以提升稳定性,

$$R_\beta(W) = \beta \|W^T W \otimes (1 - I)\|_F^2. \quad (3)$$

正交正则通过将所有的权重的奇异值设置为 1, 从而破坏了权重的频谱信息。为此人们提出谱归一化方法, 只对权重的谱进行了放缩, 保证其最大奇异值为 1, 因此并不会破坏网络的矩

阵结构。此外, 谱归一化能够保证网络的 Lipschitz 连续性, 从而稳定网络的训练。对于每层变换 $g: h_{in} \mapsto h_{out}$, Lipschitz 范数的定义是 $\sup_h \sigma(\nabla g(h))$, 其中 $\sigma(A)$ 是权重矩阵 A 的谱范数, 进而有:

$$\sigma(A) = \max_{h: \|h\|_2=1} \frac{\|Ah\|_2}{\|h\|_2} = \max_{\|h\|_2 \leq 1} \|Ah\|_2. \quad (4)$$

由上式可见, 每层网络的 Lipschitz 范数等价于其权重矩阵 A 的最大奇异值。一方面, 对于线性层 $g(h) = Wh$, 其 Lipschitz 范数为 $\|g\|_{Lip} = \sup_h \sigma(\nabla g(h)) = \sup_h \sigma(W) = \sigma(W)$ 。另一方面, 对于激活函数也同样符合 Lipschitz 连续性。比如, 对于 Lipschitz 范数 $\|a_i\|_{Lip} = 1$ 的激活函数 a_i (如 ReLU), 可以利用不等式 $\|g_1 \circ g_2\|_{Lip} \leq \|g_1\|_{Lip} \cdot \|g_2\|_{Lip}$ 。因此, 约束网络每层权重 W^l , 使得其谱范数不大于 1, 即可保证整个网络的 Lipschitz 连续性。通过对网络权重 W 进行谱规范化, 即有:

$$\bar{W}_{SN}(W) := W / \sigma(W), \quad (5)$$

则可以保证 $\|f\|_{Lip}$ 的上界为 1。

然而, 谱规范化操作 $\sigma(W)$ 如果直接进行奇异值分解, 那么计算的开销是巨大的。作者提出用幂迭代^[11]方法来减少计算开销。首先, 对于每个权重进行随机初始化向量 \tilde{u} , 迭代构造左奇异向量 \tilde{v} 和右奇异向量 \tilde{u} :

$$\tilde{v} \leftarrow W^T \tilde{u} / \|W^T \tilde{u}\|_2, \tilde{u} \leftarrow W \tilde{v} / \|W \tilde{v}\|_2, \quad (6)$$

最终估计出 $\sigma(W) \approx \tilde{u}^T W \tilde{v}$ 。

对于 WGAN-GP 来说, 其惩罚采样点 \tilde{x} 的梯度只能是对于当前生成网络的分布进行正则, 并不能对生成网络分布的支集也进行正则。另一方面, 生成网络的分布及其支集会随着训练不断改变, 从而这种方式会造成正则本身的不稳定性。谱规范化对于 WGAN-GP 来说, 对于更高的学习率具有更强的稳定性。实际情况下, 谱规范化可以和 WGAN-GP 联合使用以获得更好的训练效果。

3. 多尺度生成

随着生成图像空间分辨率的提升, 图像生成

任务的问题维度与难度也随之上升。多尺度生成作为降低高分辨率图像生成问题难度的一种有效策略，被广泛应用于生成式对抗模型。

Denton 等^[12]提出 LapGAN 模型，通过在拉普拉斯金字塔框架内将多个 GAN 模型级联的策略，由粗到精地生成目标图像。其中，最低分辨率的 GAN 模型以随机噪声作为输入，其后每个 GAN 模型都是以更低分辨率模型输出和随机噪声作为输入的条件 GAN 模型，通过残差学习生成更高分辨率的目标图像。Wang 等^[13]在图像翻译任务中采用类似的思想，将多个生成器进行嵌套，提出了 Pix2pixHD 模型，生成 2048×1024 分辨率的目标图像。同时，该模型采用嵌入物体边缘信息、多尺度判别器等方法进一步提升生成图像质量。

然而，上述方法采用多个生成器与判别器，导致网络结构较为复杂，需要消耗大量的计算资源。Karras 等^[14]提出 Progressive GAN (PGGAN 或 ProGAN) 模型，通过模型结构的动态调整实现多尺度生成。具体来说，模型从 4×4 像素的生成器与判别器出发，当前分辨率达到稳定状态时，在生成器末端与判别器前段分别增加新层以提升二者的分辨率，同时采用逐步提升新增层权重的方式进行平滑过渡，减少增加新层带来的不稳定性。该模型一方面充分利用不同分辨率生成难度的差异，降低了高分辨率模型训练的难度，另一方面由于模型的动态结构调整，显著减少了模型训练所需要的计算资源和时间。

Brock 等^[10]在 BigGAN 中进一步探究了通道数、批样本数量等对生成式对抗网络生成质量的影响，提出增加通道数、扩大批样本数量可以明显提高生成图像的质量，但同时，二者的增加带来了稳定性问题。因此，BigGAN 采用正交正则化等多种策略提升模型训练的稳定性，并使用噪声截断技巧对生成结果的真实性和多样性进行平衡。此外，BigGAN 采用共享嵌入的方式，将噪声向量嵌入到生成器的多个层，使得不同分辨率的特征直接受到噪声向量的影响。这些策略的应用使得 BigGAN 的图像生成质量得到极大的提升，同时基于共享嵌入的多尺度生成策略也对模型的收敛

起到了加速作用。

4. 应用：多样化图像翻译

多样化图像翻译指的是，给定一张图像，可以生成多张互异的目标域的图像。比如，给定一张夏天的图像，网络可以生成多张不同的冬天图像。一种直接的解决方案是在网络(如 Pix2Pix)的输入加入噪声，然而这样做往往会让网络在训练过程中直接忽略所加入的噪声，进而在测试时无法获得多样的输出结果。

为了解决上述问题，[15]提出了一种条件变分自动编码对抗网络 (cVAE-GAN) 与条件隐变量回归对抗网络 (cLR-GAN) 相结合的多样化图像翻译解决方法。对于成对的图像 (A, B) ，cVAE-GAN 将目标图像 B 先送入编码网络 E 中进行编码，得到编码向量 $E(B)$ ，再将 $E(B)$ 与输入图一起输出到生成网络 G 中，并最终获得重构图像 \hat{B} 以定义重建损失和对抗损失。进而，引入 KL 散度损失，使得 $E(B)$ 近似服从高斯分布。

不同于 cVAE-GAN，cLR-GAN 则直接将输入图 A 与随机噪声 z 一起输入到生成网络 G 中，得到 \hat{B} ，再将 \hat{B} 输入 E 中以还原隐变量 z 。在此基础上，定义 L_{GAN} 以保证 \hat{B} 的无法与目标图 B 无法区分， L_1^{latent} 以保证 E 可以重建隐变量 z 。由于 E 需要重建 z ，因而可以保证 \hat{B} 的生成不会直接忽略 z ，而是将 z 看成生成 \hat{B} 的关键条件。通过将 cVAE-GAN 和 cLR-GAN 进行结合，得到最终 Bicycle-GAN 的模型。

虽然 Bicycle-GAN 可以生成多样化的图像，然而需要大量的样本对数据。最近，MUNIT^[16]和 DRIT^[17]通过将不同域的图像可以分解为域不变内容空间 (domain-invariant content space) 以及域相关属性空间 (domain-specific attribute space) 两个部分，研究了基于非样本对数据的多样化图像翻译问题。由于 MUNIT 与 DRIT 的思想以及做法均相似，这里将以 DRIT 为例加以介绍。

DRIT 网络由内容编码器 $\{E_x^c, E_y^c\}$ ，属性编码器 $\{E_x^a, E_y^a\}$ ，生成器 $\{G_x, G_y\}$ 以及域判别器 $\{D_x, D_y\}$ 以及内容判别器 D_{adv}^c 。以源域为例，

DRIT 会首先将 X 映射到两个空间,即域不变内容空间 ($E_x^c: X \rightarrow C$) 以及域相关属性空间 ($E_x^a: X \rightarrow A_x$)。 G_x 会根据送入的内容以及属性 ($G_x: \{C, A_x\} \rightarrow X$), 生成相应的图像 u , 此时域判别器 D_x 判别 u 和原始图像 X 。另一方面, D_{adv}^c 用于判别分别从 X 和 Y 提取的内容, 使其无法分辨, 即有内容对抗损失 $L_{adv}^content$ 。为实现基于非样本对的图像翻译模型学习, 进一步引入循环一致性损失 $L_1^c(G_x, G_y, E_x^c, E_x^a, E_y^a)$ 。此外, 网络还加入了自身重建损失 L_1^{rec} 、属性编码器输出的 KL 损失 L_{KL} 、以及隐变量重建损失 L_1^{latent} 。

5. 应用: 多域图像翻译

相比于双域图像翻译工作^[18, 19], 多域图像翻译利用单个模型完成多个领域间的相互映射, 一方面避免了双域图像翻译存在的所需模型数量随领域个数的增加指数增长的问题, 另一方面可以充分利用不同领域图像数据进行模型学习。如图 1 所示, Perarnau 等^[20]提出 IcGAN 模型, 利用单个模型生成多个不同领域(属性)的目标图像, 但由于生成网络中解码器与编码器分别进行训练, 使得 IcGAN 在人脸属性编辑任务中难以保留人脸图像的内容和身份。

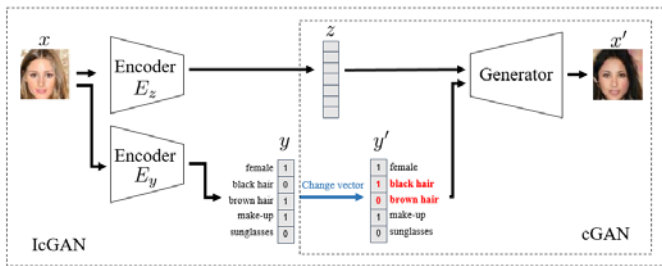


图 1 IcGAN 模型^[20]

Choi 等^[21]与 He 等^[22]分别提出 StarGAN (图 2)与 AttGAN 模型(图 3), 将属性识别网络与判别网络进行融合, 通过端到端学习约束生成网络产生具有特定属性的目标图像。其中, StarGAN 将目标属性与源图像同时输入生成网络, 通过循环一致性损失约束网络保留源图像内容, 而 AttGAN 将目标属性与隐层特征结合, 通过重建损失约束源图像信息的保留。

如图 4 所示, Liu 等^[23]提出使用差值属性向量(即目标属性向量域源属性向量的差值)代替

目标属性向量嵌入生成网络, 显式地区分需要修改的属性与无需修改的属性, 同时引入选择性转移单元, 将源图像特征选择性地传递到目标图像。相比于 StarGAN 与 AttGAN, STGAN 有助于减少属性的误修改, 有效提升了生成图像的属性编辑准确率和视觉质量(图 5)。

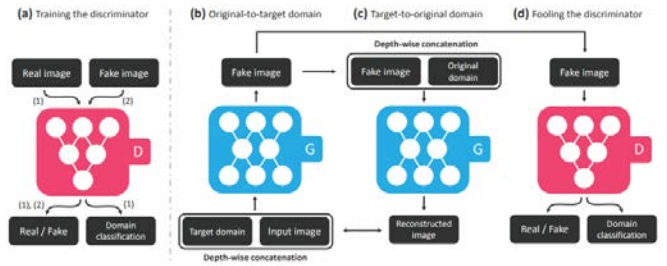


图 2 StarGAN 模型^[21]

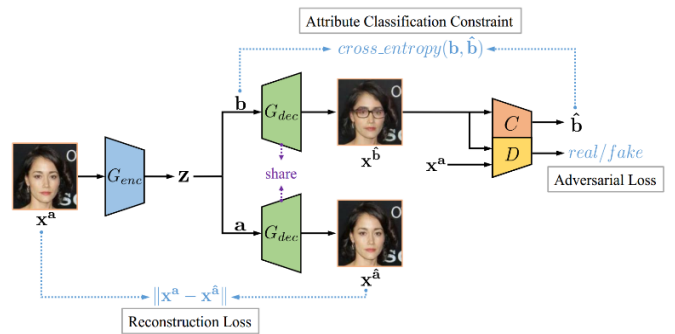


图 3 AttGAN 模型^[22]

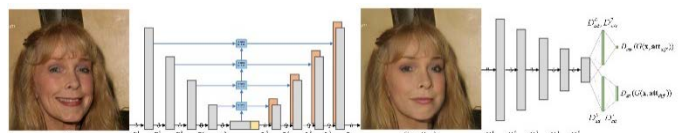


图 4 STGAN 模型^[23]



图 5 基于 STGAN 的人脸属性编辑

三、未来发展方向

本文回顾了生成式对抗网络学习与训练方面的一些主要进展, 包括从稳健损失出发的 WGAN 和 WGAN-GP, 以及从网络参数或特征规格化角度出发的正交规格化和谱规格化。此外, 多尺度学

习作为一种常用的改善生成质量策略，也在 Pix2pixHD、Progressive GAN 和 BigGAN 等得以使用与进一步推广。针对 GAN 的应用，本文以图像翻译为例，回顾了多样化和多域图像翻译方面的主要进展。

虽然生成式对抗网络近年来取得了较大的进展，但在学习的收敛性和稳定性等方面仍然缺乏足够的理论与技术支持。最近的 BigGAN 和 StyleGAN^[24] 虽然能够生成包含丰富纹理细节的

高清图像，但往往伴随着一定程度的伪影和瑕疵，以及难以克服全局不一致性问题。此外，相对于训练图像，基于 GAN 生成的图像往往并不能带来额外的多样性，也是许多 GAN 应用中需要解决的一个重要问题。在应用方面，除图像翻译外，如何将生成式对抗网络更好地应用于迁移学习、视频生成以及其它低层视觉任务也是非常值得关注的研究方向。

(责任编辑：苏航)

参考文献

- [1] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *NIPS*, 2014.
- [2] Martin Arjovsky and Léon Bottou. Towards principled methods for training generative adversarial networks. In *ICLR*, 2017.
- [3] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *ICML*, 2017.
- [4] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville. Improved training of wasserstein gans. In *NIPS*, 2017.
- [5] Lars Mescheder, Andreas Geiger, and Sebastian Nowozin. Which training methods for gans do actually converge? In *ICML*, 2018.
- [6] Jiqing Wu, Zhiwu Huang, Janine Thoma, Dinesh Acharya, and Luc Van Gool. Wasserstein divergence for gans. In *ECCV*, 2018.
- [7] Sergey Ioffe and Christian Szegedy. Batch normalization: accelerating deep network training by reducing internal covariate shift. In *International Conference on Machine Learning*, 2015.
- [8] Andrew Brock, Theodore Lim, James M Ritchie, and Nick Weston. Neural photo editing with introspective adversarial networks. *arXiv preprint arXiv:1609.07093*, 2016.
- [9] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. In *ICLR*, 2018.
- [10] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large scale gan training for high fidelity natural image synthesis. In *ICLR*, 2019.
- [11] Frank Lin and William W Cohen. Power iteration clustering. In *ICML*, 2010.
- [12] Emily L Denton, Soumith Chintala, Arthur Szlam, and Rob Fergus. Deep generative image models using a laplacian pyramid of adversarial networks. In *NIPS*, 2015.
- [13] Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. High-resolution image synthesis and semantic manipulation with conditional gans. In *CVPR*, 2018.
- [14] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen. Progressive growing of gans for improved quality, stability, and variation. In *ICLR*, 2018.
- [15] Jun-Yan Zhu, Richard Zhang, Deepak Pathak, Trevor Darrell, Alexei A Efros, Oliver Wang, and Eli Shechtman. Toward multimodal image-to-image translation. In *NIPS*, 2017.
- [16] Xun Huang, Ming-Yu Liu, Serge Belongie, and Jan Kautz. Multimodal unsupervised image-to-image translation. In *ECCV*, 2018.
- [17] Hsin-Ying Lee, Hung-Yu Tseng, Jia-Bin Huang, Maneesh Singh, and Ming-Hsuan Yang. Diverse image-to-image translation via disentangled representations. In *ECCV*, 2018.
- [18] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *CVPR*, 2017.

- [19] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *ICCV*, 2017.
- [20] Guim Perarnau, Joost van de Weijer, Bogdan Raducanu, and Jose M Álvarez. Invertible conditional gans for image editing. 2016.
- [21] Yunjey Choi, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, and Jaegul Choo. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In *CVPR*, 2018.
- [22] Zhenliang He, Wangmeng Zuo, Meina Kan, Shiguang Shan, and Xilin Chen. Arbitrary facial attribute editing: Only change what you want. *arXiv preprint arXiv:1711.10678*, 2017.
- [23] Ming Liu, Yukang Ding, Min Xia, Xiao Liu, Errui Ding, Wangmeng Zuo, and Shilei Wen. Stgan: A unified selective transfer network for arbitrary image attribute editing. In *CVPR*, 2019.
- [24] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *CVPR*, 2019.



左旺孟

哈尔滨工业大学教授，博导。主要研究方向为的图像增强与复原、图像生

成与编辑、视觉跟踪、物体检测与图像分类等计算机视觉任务。

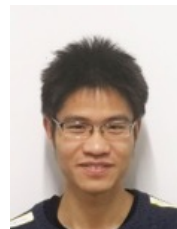
Email: wmzuo@hit.edu.cn



刘铭

哈尔滨工业大学博士生，主要研究方向为图像生成与编辑。

Email: csmliu@outlook.com



颜肇义

哈尔滨工业大学博士生，主要研究方向为智能图像填充、人流密度估计。

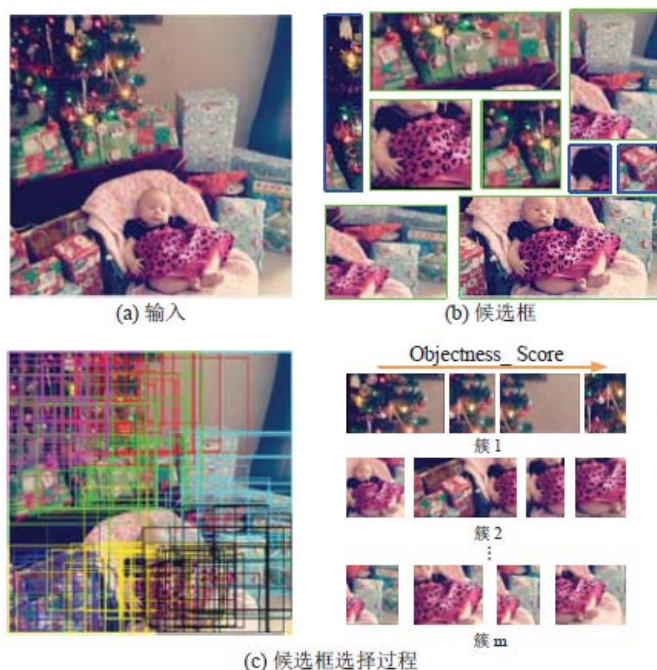
Email: yanzhaoyi@outlook.com

基于情感区域自动挖掘的视觉情感预测

南开大学 杨巨峰 折栋宇

随着社交网络的不断普及，越来越多的网络用户倾向于用不同的媒介去表达他们的观点，识别其中所蕴含情感对于理解这些用户行为很有帮助，尤其是理解图片视频等视觉媒体内容中的情感已经在相关领域引起了越来越多的注意。该种分析算法的潜在用途是非常广泛的，主要包括情感图片检索、美学质量分类、意见挖掘、评论助手等应用。

我们研究涉及高度抽象概念的视觉情感分析问题。现有的大多数方法都只关注于提升从图像全局角度捕获特征的表达能力，但通过观察可以发现图像的整体和局部区域都可以传达出重要的情感信息，受到该启我们提出了一个框架来有效利用图片中表达情感的区域。我们首先利用现有的目标检测工具生成候选区域，再使用候选区筛选算法去除冗余和具有干扰的目标区域，然后利用卷积神经网络计算每一个候选区域的情感得分，通过同时考虑物体得分和情感得分自动发掘表达情感能力较强的区域。最后，将整张图片和局部区域的卷积神经网络输出相结合，产生最后的预测结果。由于标注情感区域非常主观并且费力，我们的框架只需要图像级别的标签，可以显著地减少在训练中需要的标注负担，这对于



情感分析尤其重要。实验表明我们提出的算法在八个主流的基准数据集上的结果均超过了目前最先进的方法。

以上系列工作发表于 IEEE TMM 2018、IEEE CVPR 2019、AAAI 2017、IJCAI 2017。

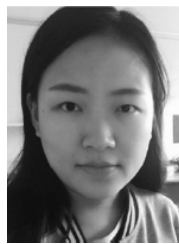
(责任编辑：王金甲)



杨巨峰

南开大学副教授，主要研究方向为计算机视觉、机器学习、多媒体计算。

Email: yangjufeng@nankai.edu.cn



折栋宇

南开大学计算机视觉实验室硕士研究生，主要研究方向为计算机视觉、情感计算。

Email: sherry6656@163.com

多长度哈希联合学习方法

山东财经大学 聂秀山 山东大学 刘兴波

随着互联网和信息技术的发展，网络上数据呈爆炸性增长，如何高效的存储、管理、检索、分析大数据成为学术界和产业界关注的热点。哈希学习 (Hash Learning) 正是解决上述问题的可行方案之一。哈希学习结合数据自身分布和内在特性，利用机器学习方法，把高维的数据转化为二值码表示，同时尽可能保持数据在原特征空间的相似性，哈希学习对数据的二值表示形式，可以显著的节省存储空间，提升数据处理速度，因此在大数据学习中占有重要的地位。另一方面，哈希码也可以看作是一种重要的二值特征表示形式，而二值特征表示在识别、分类、匹配、搜索等计算机视觉和多媒体信息处理领域也取得了较好的应用效果。因此哈希学习越来越受到研究者的重视。

现有的哈希学习算法中，大多在模型训练之前，需要预先设定一个哈希码长度 (例如 48 或 128 比特等)，然后利用相关模型训练参数，当哈希长度发生变化时，需要重新运行模型来得到新长度的哈希码。因此，如何在一个框架下同时得到多个不同长度的哈希码，供用户在不同情境下选择使用是一个值得研究的问题。

另一方面，如果把哈希码看作数据的二值特征表示的话，不同长度的哈希码本质上也是数据的多种特征表示。模式识别和计算机视觉领域相关研究表明，充分利用多特征之间的关联关系对特征表示和数据分析性能的提升具有较好的促

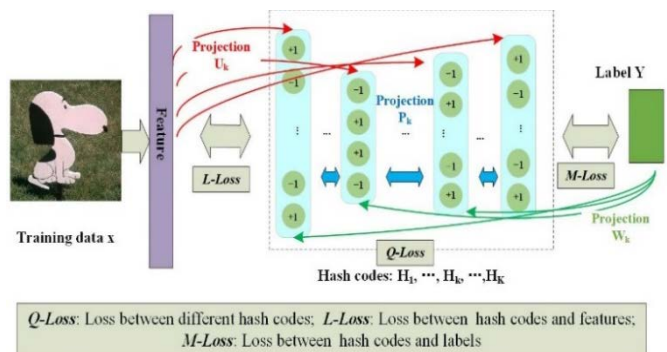


图1 多长度哈希联合学习框架图

进作用。

基于以上动机，我们提出了一种多长度哈希联合学习方法。如图1所示，该方法把不同长度哈希放在一个模型中，通过联合优化和学习，同时得到不同长度的哈希码表示。该方法的模型共分为三个部分：第一部分 (Q -loss) 用于刻画多长度哈希之间的关联关系；第二部分 (L -loss) 用于表达数据原始特征和哈希表示之间的信息损失；第三部分 (M -loss) 的作用是利用样本的标记信息。以上三部分共同构成模型的目标函数，该目标函数可以通过交替优化的方式求解。实验结果表明，该方法因有效的利用了同一样本不同长度哈希码的关联关系，比直接学习同样长度的哈希码，在性能上有较为明显的提升。

以上部分工作已发表于国际会议 AAAI2019。

(责任编辑：任桐炜)



聂秀山

山东财经大学计算机科学与技术学院教授、博士生导师，主要研究方向为计算机视觉、机器学习、多媒体检索等。

Email: niexsh@sdufe.edu.cn



刘兴波

山东大学计算机科学与技术学院博士生，主要研究方向为哈希学习、多媒体信息处理等。

Email: geshiming@iie.ac.cn

南京大学吴建鑫教授访谈

2019年5月5日,专委秘书处采访了南京大学优秀青年学者吴建鑫教授。下面是采访实录。

您在计算机视觉与模式识别等领域做出了很多高水平研究工作,获得了国家自然科学基金优秀青年基金支持,入选了中组部青年海外高层次人才引进计划(青年千人计划),是青年科技工作者学习的楷模。您能否向我们介绍一下您现在主要的研究方向是什么?您是如何选择现在的研究方向的?您对于如何选择研究方向有什么建议吗?

现在最感兴趣的研究方向,我造了一个词组,叫“资源受限的深度学习”。深度学习模型,比如卷积神经网络 CNN,在视觉领域很多问题效果特别好,但是运算量巨大、模型本身参数超多、运行时占用显存和能源也很多,这是深度学习对计算资源的需求非常大。同时,深度神经网络的一个著名特点就是:要学到好的模型,需要海量有准确标注的训练数据,这是深度学习对数据资源的需求非常大。那么,怎么在计算和/或数据资源受限的情况下有效的完成模型的训练(学习)和应用(部署),就是我所感兴趣的问题和研究方向。

事实上,从2000年左右开始从事与研究相关的工作时起,我的每项研究内容多多少少都和计算、数据这两种资源的绝对数量受限或相对数量受限有关。现在回头来看的话,发现还是挺有连续性的。

应该说,这种研究方向上的连续性并不是我主动追求的结果。在确定了自己的研究兴趣和品味之后,如果能在研究工作中有一定的自由度的话,每位研究人员只要能够坚持这种兴趣与品味,即便大领域有着潮起潮落,也能找到其中适合自己的定位与小方向。

您在国际顶级期刊和会议上发表了多篇论

文并被多次引用,相关成果也很多应用在解决实际问题中。在您取得的这些成果中,有哪些工作让您印象特别深刻?有没有哪些不为人知的艰辛或者有趣经历?

我自己印象比较深刻的一个例子是 ThiNet,主要是提出了一种神经网络以滤波器为单位的压缩方法,实验效果很好,也方便在实际应用中使用。ThiNet的关键是利用网络后面一层的输出变化指导前一层的压缩,并通过定义一个清晰的优化问题来求解。

其实在提出 ThiNet 之前,我们还尝试了一些其他方法。但是却发现这些方法,包括已经发表的很多方法,他们的实际效果甚至比不上随机选择。我是在一次吃饭过程中想到了 ThiNet 的主要思想的,随即当天晚上就在纸上推导出优化问题的形式化定义。第二天到办公室扫描后发给罗建豪同学(论文的第一作者)。他对算法和实现做了优化并得到很好的效果。

有趣的是,在后来的一些审稿经历中,不止一次遇到有投稿声称他们首次提出了某种算法(或准则、方法),其实就是 ThiNet 方法,甚至这些投稿还引用了 ThiNet 论文!

高效的科研团队对于深入开展课题研究有着至关重要的作用。您作为南京大学机器学习与数据挖掘研究所的教授,您能否简要介绍一下您所在的科研团队?以及是如何高效管理科研团队的?

机器学习与数据挖掘研究所(LAMDA)是周志华教授建立的一个科研团队,主要包括机器学习、数据挖掘及相关不同应用方向上的教师和同学们。我们团队的学术氛围很浓厚,也已经有很好的成果产出。至于科研团队的管理,我自己也还

是在探索和学习过程中。

作为博士生导师，您在招收研究生时最注重学生的哪些方面？能否跟大家分享一下您在指导学生方面的心得体会？在您的实验室是否有某种奖罚制度用以激励学生？

我在面试研究生的时候主要考察以下几个方面：第一，身心健康、并且知道自己未来的目标是什么；第二，能和导师、同学和睦相处；第三，有一定的数学和计算机基础，有机器学习、计算机视觉相关基础是加分项。

在指导学生时主要感受是三点：第一要有比较广阔的视野，不能因为现在深度学习很热就一头扎进去、别的什么都不管；第二是编程能力，即使将来不做研究也可以从事其他工作；第三是（特别是对博士生来说）培养研究品味，了解什么样的研究是有价值的。

对学生没有惩罚措施（国家和学校也没给我们这个权力），但是研究出有趣的东西或者发表好的论文会有一些小的奖励。

科研基金是开展科研工作的前提，您认为书写基金申请书时应注意哪些问题？能否分享一下您个人申请科研基金的成功经验或者失败原因？

我写基金本子的经验不是太多，审本子的经验倒是略有一些。我觉得从科学或者技术的角度能“自圆其说”很重要，也就是写申请书的时候在科学问题、技术路线这方面自己确实形成了一个整体思路，至少能够说服自己。胡乱拼凑的材料确实很容易被评审专家看出来。从我自己的经验来看，是否有这么一个能自圆其说的整体思路对于申请获批与否很关键。另外，申请材料的格式统一规范、文字简洁通顺、图表美观易懂，这些要求看似简单，其实要花大量精力才能做到，这些方面对于申请科研基金也是很重要的。

您是多个期刊和会议的编委或审稿人，您能

否结合自己的审稿经历，分享一下学术论文写作投稿过程中需要注意哪些方面的问题？或者有哪些建议？

写论文我觉得和写项目申请书很像（当然具体要求不同），所以前面的一些经验这里也适用。对于论文投稿，我建议作者在写完之后可以从审稿人的角度来考虑一下：实验自己觉得够了，但是审稿人会不会觉得不足？有些结论自己觉得理论或实验结果可以支撑，但是审稿人会不会觉得 claim 太强了？自己的方法当然自己觉得好，但是审稿人会不会发现一些其他的缺点？比如说，请领域相近但是没有参与论文研究的同行读一遍就有可能发现很多问题和可以改进的地方。

您认为计算机视觉与模式识别领域未来比较有潜力的研究方向主要有哪些？您认为未来哪些技术的在社会上的应用需求会比较大？

计算机视觉和模式识别涉及的范围很广，值得研究的方向非常多，对其中之一我非常感兴趣：如何打破现有方法的一些限制，使得深度学习能够适用于更多的场景？这些限制包括资源（例如计算和数据资源）方面的限制，也包括深度学习目前暂时还不能处理的一些场景与数据。

在应用需求方面，对高度非线性、动态变化的数据进行有效处理可能是很有社会价值的技术，例如辅助驾驶（及自动驾驶）、传统产业的智能化等方面。

您对现在深度学习技术发展的看法是什么？您认为优势和劣势都有哪些？未来的发展趋势是什么？

我觉得挺乐观的，因为深度学习现在有不少直接影响我们日常生活的落地应用，所以就算对他的投入降低了，大概也是“降温”而已，不太可能会出现“寒冬”。

深度学习的优势和劣势都很明显，对一些应用效果特别好是优势，需要大量资源、应用范围

比较窄、理论基础不足是劣势。

趋势的话不太好评论，我自己倒是愿意往两个相反的方向努力：促进深度学习实用化（“落地”）和构建深度学习理论基础。

您可否可以大家分享一下您的业余生活？是否经常进行体育锻炼？以及您是如何协调工作和家庭的呢？

我大概是比较“宅”的一个人，平常有空的时候喜欢看书。不过家里的书橱里面基本上没有本专业的资料，大多是一些历史、科普、小说之类的，比如阿加莎克里斯蒂的全集有八十几本。

偶尔也会和家人一起去看看电影、出去旅游一下。工作和家庭都需要大量的时间，但是总时间是有限的，所以这两者肯定是冲突的。家人们已经给了我很多支持，我也尽量多花些时间在家里，所以有时候对一些国内外学术活动的邀请只能婉拒了。记得有次在接受了成都的一次活动邀请之后，主办方还挺惊讶的，大概我“宅”的特性已经“恶名远扬”了吧。

关于体育锻炼，我当然也希望能够经常锻炼，但是确实比较忙、人也懒，所以只能（间隔周期

比较长）每次绕着南大仙林校区慢跑一圈。如果有时候特别忙、以至于人很疲惫的话，对我来说出去跑一圈是特别有效的恢复方法。



吴建鑫

南京大学教授，博士生导师，曾获国家优秀青年科学基金资助，入选青年海外高层次人才引进计划（青年千人计划）。现任《自动化学报》编委，《Pattern Recognition》期刊编委（Associate Editor），曾任

重要国际会议 ICCV、CVPR、AAAI、IJCAI 等在内的国际会议领域主席或高级程序委员。主要从事计算机视觉与机器学习领域的研究工作。在重要国际期刊如《IEEE Trans. Pattern Analysis and Machine Intelligence》、《International Journal of Computer Vision》、《Artificial Intelligence》等以及重要国际会议如 ICCV、CVPR、ICML、NIPS、IJCAI、AAAI 等发表论文 90 余篇；据 Google Scholar 统计，发表的论文被 40 余个国家和地区的学者引用 8200 余次，H-index 为 38。

（责任编辑：黄岩 余烨）

委员好消息

✪ 2019年4月11-14日，VALSE 2019（第九届视觉与学习青年学者研讨会）在合肥召开。本届 VALSE 大会首次设立了“VALSE 年度杰出学生论文奖”，共有 3 篇论文获得“VALSE 2018 年度杰出学生论文奖”，6 篇论文获得“VALSE 2018 年度杰出学生论文提名奖”。CCF-CV 专委会委员指导的 6 篇论文获奖，分别是：大连理工大学卢湖川教授指导的论文 *Deep Mutual Learning*（第一作者张莹，发表于 CVPR 2018）和清华大学苏航助理研究员等指导的论文 *Boosting Adversarial Attacks with Momentum*（第一作者董胤蓬，发表于 CVPR 2018）获得“VALSE 2018 年度杰出学生论文奖”；中山大学郑伟诗教授和赖剑煌教授指导的论文 *Person Re-Identification by Camera Correlation Aware Feature*

Augmentation（第一作者陈颖聪，发表于 IEEE T PAMI 2018）、华中科技大学白翔教授指导的论文 *TextBoxes++: A Single-Shot Oriented Scene Text Detector*（第一作者廖明辉，发表于 IEEE T IP 2018）、西安交通大学孟德宇教授等指导的论文 *Kronecker-Basis-Representation Based Tensor Sparsity and Its Applications to Tensor Recovery*（第一作者谢琦，发表于 IEEE T PAMI 2018）和中科院自动化所雷震副研究员等指导的论文 *Single-Shot Refinement Neural Network for Object Detection*（第一作者张士峰，发表于 IEEE CVPR 2018）获得“VALSE 2018 年度杰出学生论文提名奖”。

（责任编辑：刘海波）

从集成聚类到大规模聚类：相关算法及开源代码

华南农业大学 黄栋 中山大学 赖剑煌

聚类 (clustering) 作为无监督学习的有力工具, 已在模式识别与计算机视觉的许多研究主题中得到广泛应用, 其中包括 (但不限于) 图像分割、显著性检测、图像检索、行人重识别等。

在聚类的相关研究中, 集成聚类 (ensemble clustering) 是近几年的热点研究方向之一, 其目标在于融合多个基聚类结果以构建一个更优、更鲁棒聚类。集成聚类的基本流程如图 1 所示, 其过程可分为两个主要阶段, 一是基聚类生成, 涉及如何构建一个兼具多样性与可靠性的基聚类集合, 二是一致性函数, 涉及如何对多个基聚类的信息进行融合以得到集成聚类结果。

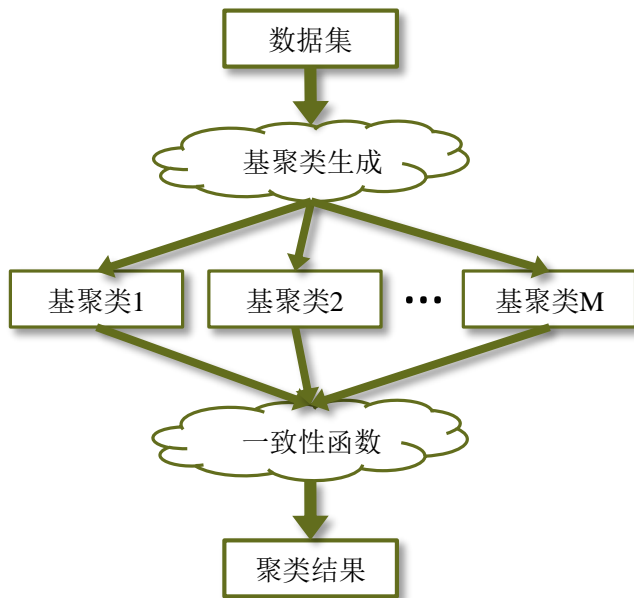


图 1 集成聚类基本流程

近年来, 研究组针对集成聚类的相关难点问题展开研究, 取得了若干研究进展; 特别地, 近期研究工作进一步由常规集成聚类扩展至有限计算资源下的大规模集成聚类, 可在单机环境下实现数千万规模非线性可分数据集的快速、鲁棒聚类划分。接下来, 本文将具体介绍研究组在近

几年集成聚类与大规模聚类研究中的相关研究进展及其开源代码, 涉及基于稀疏图与概率轨迹分析的集成聚类算法、基于因子图模型的集成聚类算法、基于局部熵加权的集成聚类算法、基于簇相似度传播的集成聚类算法、以及有限计算资源下的大规模谱聚类与集成聚类算法。

1. 基于稀疏图与概率轨迹分析的集成聚类算法

论文: Huang D, Lai J H, Wang C-D. Robust Ensemble Clustering Using Probability Trajectories[J]. IEEE Transactions on Knowledge and Data Engineering, 2016, 28(5): 1312-1326.

工作: 数据点与数据点的关系 (以下简称点对点关系) 是集成聚类的基本关联信息。对于基聚类集合中的点对点关系, 过往的集成聚类方法主要有两个方面的局限性, 一是往往忽略了欠稳定点对点问题而容易对集成效果造成不良影响, 二是大多无法根据全局结构信息 (高阶信息) 对局部边进行优化。对此, 该论文提出一种基于稀疏图构造与概率轨迹分析的集成聚类方案 (如图 2 所示)。首先, 该论文采用微簇 (microcluster) 的数据表达形式, 并将微簇作为基本处理单元, 以大幅降低问题规模; 进而, 采用稳定点对选择策略对欠稳定边进行移除, 以构造一个包含少量可靠边的稀疏图。在构建稀疏图的基础上, 该论文提出建立一个新的随机游走转移概率矩阵, 并依此在稀疏图中进行随机游走过程。通过对随机游走概率轨迹的分析, 从包含少量可靠边的稀疏图中得到一个新的、稠密的相似性矩阵, 进而提出了两个新的集成聚类算法, 分别是: 概率轨迹集聚 (probability trajectory accumulation, PTA) 算法和基于概率轨迹的图分割 (probability trajectory based graph partitioning, PTGP) 算法。

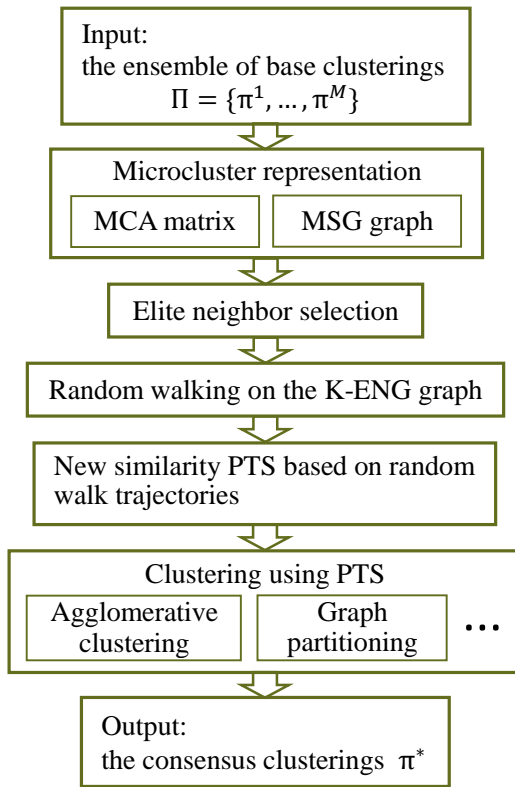


图 2 基于稀疏图与概率轨迹的集成聚类示意图

代码:

<https://www.researchgate.net/publication/284259332>

2. 基于因子图模型的集成聚类算法

论文: Huang D, Lai J H, and Wang C-D. Ensemble Clustering Using Factor Graph[J]. Pattern Recognition, 2016, 50: 131-142.

工作: 该论文提出一种基于因子图模型的集成聚类 (ensemble clustering using factor graph, ECFG) 算法。具体地, 该论文对微簇数据表达形式进行扩展, 提出超点 (super-object) 的概念。相较于微簇, 超点的表达形式具有对基聚类规模的自适应性, 可更为高效地应对大规模数据集。在超点表达与概率建模的基础上, 该论文将集成聚类问题建模为一个最大似然估计问题, 并进而将其转换为一个二值线性规划 (binary linear programming, BLP) 问题。二值线性规划问题是一个 NP 难问题。为近似求解此问题, 该论文提出一种基于因子图模型和最大积置信度传播算法的优化方法, 以得到最终的集成聚类结果。

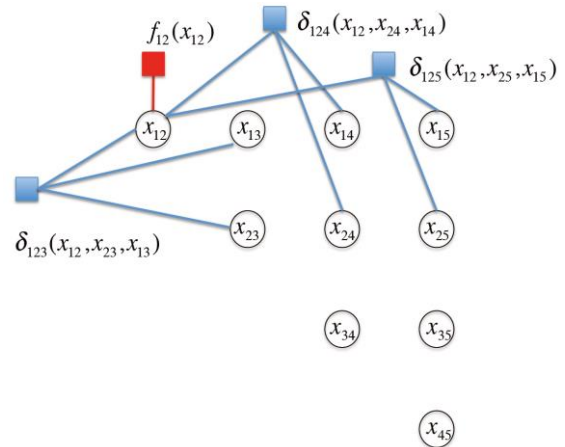


图 3 因子图模型示意图

代码:

<https://www.researchgate.net/publication/282076585>

3. 基于局部熵加权的集成聚类算法

论文: Huang D, Wang C-D, Lai J H. Locally Weighted Ensemble Clustering[J]. IEEE Transactions on Cybernetics, 2018, 48(5): 1460-1473.

工作: 针对各个基聚类的可靠度差异, 现有的集成聚类方法往往采用全局加权策略, 将每个基聚类作为一个整体进行评估与加权, 难以有效应对单个基聚类内部的可靠度差异问题。对此, 该论文提出将每个簇视作基聚类中的一个局部区域, 根据多个基聚类的类标分布情况, 采用熵 (entropy) 理论对基聚类中各个簇的局部不确定性进行评估。在局部不确定性评估的基础上, 提出一种集成驱动的簇可靠度指标 (ensemble-driven cluster index, ECI), 进而将之应用于局部加权机制。最后, 该论文提出两个基于局部加权机制的集成聚类新算法, 分别是: 局部加权线索集聚 (locally weighted evidence accumulation, LWEA) 算法和局部加权图分割 (locally weighted graph partitioning, LWGP) 算法。

代码:

<https://www.researchgate.net/publication/316681928>

4. 基于簇相似度传播的集成聚类算法

论文: Huang D, Wang C-D, Peng H X, et al. Enhanced Ensemble Clustering via Fast Propagation of Cluster-wise Similarities[J]. arXiv preprint arXiv:1810.12544, 2018.

工作: 针对集成聚类的高效运算问题以及多尺度关联问题, 该论文提出一种基于快速簇相似度传播的集成聚类算法。具体地, 该论文以多个基聚类的簇集作为节点集, 根据簇与簇之间的 Jaccard 系数构建加权边, 以得到一个初始簇相似性图。进而, 该论文在所构建的初始簇相似性图上进行随机游走过程, 以得到一个新的、融合多尺度簇集关联信息的簇相似性矩阵。针对传统共联矩阵 (co-association matrix) 仅可反映数据点之间直接关联信息的局限性, 该论文提出将所得到的簇相似性矩阵进一步由簇集合映射到数据点集合, 以构建一个改进型共联矩阵 (enhanced co-association matrix, ECA matrix)。最后, 该论文基于层次聚类和元聚类, 提出两个新的一致性函数, 以构建最终集成聚类结果。

代码:

<https://www.researchgate.net/publication/328581758>

5. 有限资源下的大规模谱聚类与集成聚类算法

论文: Huang D, Wang C-D, Wu J-S, *et al.* Ultra-Scalable Spectral Clustering and Ensemble Clustering[J]. arXiv preprint arXiv:1903.01057, 2019.

工作: 该论文提出两种大规模聚类算法, 分别称为超大规模谱聚类 (ultra-scalable spectral clustering, U-SPEC) 算法和超大规模

集成聚类 (ultra-scalable ensemble clustering, U-SENC) 算法。在 U-SPEC 算法中, 该论文提出一种混合代表点选取 (hybrid representative selection) 策略以及一种快速 K 近邻近似计算 (fast K -nearest representative approximation) 方法, 以构建数据点集合与代表点集合之间的稀疏二部图结构, 进而以迁移分割 (transfer cut) 算法对该二部图结构进行快速分割以得到谱聚类结果。在 U-SENC 算法中, 该论文进一步将多个具有差异性的 U-SPEC 聚类算子在统一的集成聚类框架下进行融合, 以期在显著提升 U-SPEC 算法鲁棒性的同时保持其在大规模聚类过程中的高效性。该论文实验在单机 (i5-6600 CPU + 64 GB 内存) 环境下进行, 涉及 10 个具有代表性的大规模数据集。实验结果表明, 所提出的 U-SPEC 与 U-SENC 算法可在有限计算资源 (普通单机) 下实现几千万规模非线性可分数据集的快速、鲁棒聚类; 时间与空间复杂度均接近于线性。

代码与数据:

<https://www.researchgate.net/publication/330760669>

(责任编辑: 金鑫, 蹇木伟)



黄栋

副教授, 华南农业大学数学与信息学院, 研究方向为模式识别与数据分析, 主要研究内容包括集成聚类、多视图聚类、流数据聚类、大规模聚类等。

Email: huangdong06@163.com



赖剑煌

教授, 中山大学数据科学与计算机学院、广东省信息安全重点实验室主任、中国图象图形学学会副理事长、中国计算机学会计算机视觉专委会副主任。研究方向为模式识别与计算机视觉。

Email: stsljh@mail.sysu.edu.cn

交通场景动态理解数据集

长安大学 严定鑫 房建武 西安交通大学 薛建儒

交通场景动态理解是计算机视觉技术应用于智能交通领域的典范方向之一。相比于以往的静态三维/二维目标检测、语义/实例分割等任务,交通场景动态理解面向于感知并推演驾驶场景或交通监控场景中的交通参与者运动状态变化,是更深层次的场景理解任务。作为交通场景动态理解的主要内容,二维/三维目标跟踪、行为及事件分析、意图及注意预测为更好的交通控制与管理、驾驶规划与决策起着重要作用。随着大规模数据集的不断涌现,且由于交通场景具有丰富的上下文知识,目前的方法主要包含深度学习、在线学习及无监督学习三大方面。

本文针对二维/三维目标跟踪、行为及事件识别、意图及注意预测等任务中常用的数据集进行总结。依据数据的表示方式,可将现有数据集分为跟踪、行为及事件识别、意图及注意预测三大类。另外介绍部分已声明但未公开数据集。

一、跟踪数据集

1. KITTI-3D 目标跟踪数据集

下载地址: http://www.cvlibs.net/datasets/kitti/eval_tracking.php

KITTI-3D 目标跟踪数据集是由德国卡尔斯鲁厄理工学院和丰田美国技术研究院联合创办的 KITTI 数据中的跟踪任务集。其包含 21 个训练序列和 29 个测试序列。主要针对行人和车辆目标,人工标记了三维点云中行人或车辆目标的时序位置三维框。由于行人目标三维点云的稀疏性,很难保证得到一个十分合理的标注,为此,其额外标注了三维点云中目标的左右边界。然后将三维点云投影到 RGB 图像中得到三维目标框的立体显示和对应。



图 1 KITTI-3D 目标跟踪数据集样本

2. UA-DETRAC 数据集

下载地址: <http://detrac-db.rit.albany.edu>

UA-DETRAC 是由纽约州立大学奥尔巴尼分校、中国科学院大学、加州大学圣地亚哥分校、中国科学院自动化研究所等七家单位合作发布,一个集多目标检测与跟踪任务为一体的大规模数据集。该数据集在中国北京市和天津市的道路过街天桥上进行拍摄,并手动标注了包含于 14 万帧图像中的 8250 个车辆,121 万个二维目标对象框。视频以每秒 25 帧的速度录制,分辨率为 960×540 像素。该数据集将车辆分为 4 类(汽车、公共汽车、面包车和其他)并将其分为三个尺度:小型车、中型车和大型车。

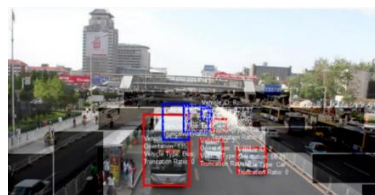


图 2 UA-DETRAC 数据集样本

3. CityFlow 数据集

下载地址: <https://www.aicitychallenge.org/2019-data-sets>

CityFlow 是英伟达推出的首个跨摄像头汽车跟踪数据集,也是目前世界上第一个支持跨摄像头车辆跟踪及重识别的大型数据集。该数据集包含了 3.25 小时的视频,来自横跨美国一座中型城市的 10 个十字路口中的 40 台摄像机,其中两架精密同步摄像机之间的距离为 2.5 公里。数据集涵盖了不同的位置类型,包括十字路口、道路干道和高速公路。该数据集包含超过 20 万个目标框,并且涵盖了多种场景、视角、车辆模型和城市车流状况。并提供了原始视频,相机分布和校正信息来辅助时空分析。可用于测试各种各样的跨摄像头多目标跟踪、单摄像头多目标跟踪、目标检测和重识别的算法。



图3 CityFlow 数据集样本

4. H3D-HRI-US 数据集

下载地址：<https://usa.hondari.com/hdd/introduction/h3d>

H3D-HRI-US 数据集由本田研究所于 2019 年 3 月发布的无人驾驶方向三维多目标跟踪数据集。本数据集是使用 3D 64 线激光扫描仪收集的大型全环绕 3D 多目标检测和跟踪数据集。其包含 160 个拥挤且高度互动的交通场景，在 27,721 帧中共有 100 万个标记实例。



图4 H3D-HRI-US 数据集样本

二、行为及事件识别

1. Daimler 数据集

下载地址：<http://isla.science.uva.nl>

Daimler 数据集是荷兰阿姆斯特丹大学提出的不仅为行人检测任务提供位置信息，而且还提供行人意图信息的数据集。它使用安装在汽车挡风玻璃后面的立体摄像系统（基线 22cm, 16fps, 1176*640 像素）记录图像序列。该数据集包括了四种典型的行人运动类型：在干道或十字路口穿越、驻停、站立片刻后行走。其由 68 个序列组成，其中 12485 幅图像包含（单个）行人。



图5 Daimler 数据集样本

2 Pedestrian direction recognition 数据集

下载地址：<https://www.rovit.ua.es/dataset/pedirecong>

Pedestrian direction recognition 数据集是由阿利坎特大学机器人三维视觉研究组提出的，可用于进行行人运动方向识别的研究。该数据集采用固定摄像机（640×480 像素）获取的

视频集合，包含不同的运动类型，比如向左、向右或向前方移动。进行手工标注及训练测试集划分。其中有 7416 张图像用于训练，1752 张用于验证，1632 张用于测试。图像按右、左、前三个不同的朝向类别进行分类：右向分配 2907 幅图像，左向分配 3099 幅图像，前向分配 1410 幅图像。

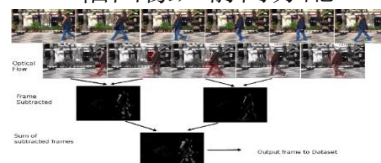


图6 Pedestrian direction recognition 数据集样本

3. nuScenes 数据集

下载地址：<https://www.nuscenes.org>

nuScenes 数据集是自动驾驶公司 nuTonomy 建立的大规模自动驾驶数据集，是迄今为止公布的最大的多模态 3D 自动驾驶数据集。该数据集由 140 万张图像、39 万帧激光雷达扫描和 140 万的 3D 人工标注边界框组成，该数据集包括了 1000 个场景，每个场景长度为 20 秒，在每一个场景中，有 40 个关键帧，并手工标注了每一帧的车辆行为（如移动，停止）和行人行为（如摔倒，站立，行走）。标注了目标大小、类别、可见及遮挡程度等。可用于实现更加智慧的自动驾驶车辆的视觉分析和多传感融合。



图7 nuScenes 数据集样本

4. Street Accidents (SA)数据集

下载地址：<http://aliensunmin.github.io/project/dashcam>

Street Accidents (SA)数据集是由台湾清华大学创建的针对于行车记录仪中事故预测的视频数据集。此数据集包含 1750 个 100 帧的事故数据序列片段，其中有 620 个确定包含事故的正样本和 1130 不包含事故的负样本。在此数据集中，做了视频帧的事故区间限定（最后 10 帧），且标注了时间轴事故发生区间。可以为辅助/自主车辆驾驶提供事故预测基准。



图 8 SA 数据集样本

5. BLVD 数据集

下载地址: <https://github.com/VCCIV/BLVD>

BLVD 是由西安交通大学人工智能与机器人研究所构建的全球首个五维度无人驾驶场景理解数据集。与以往的静态检测、语义或者实例分割任务的数据集不同, BLVD 旨在为动态 4D(3D+时间)跟踪、以驾驶任务为引导的 5D(4D+交互)交互识别和意图预测任务提供一个统一的验证平台。数据来源于中国智能车辆测试中心(IVPCC)所在的江苏省常熟市。该数据集包含 654 个高分辨率的视频片段, 共约 12 万帧, 视频中的 RGB 数据和 3D 点云帧速率为 10 帧/秒。该数据集对行人、车辆及骑行者(包含自行车、摩托车)目标进行了全标注, 共生成 249129 条 3D 框、总长度为 214922 个点的 4902 个独立可跟踪个体、6004 个可用于 5D 交互行为识别有效片段、4900 个可用于 5D 意图预测的独立个体。根据参与者密度(低和高)和光照条件(白天和晚上), 将他们分为四种场景。

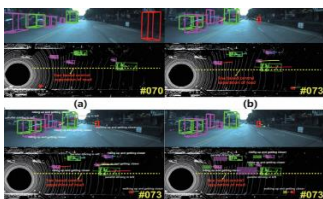


图 9 BLVD 数据集样本

三、注意力预测数据集

1. DR(eye)VE 数据集

下载地址: <http://imagelab.ing.unimore.it/dreeye>

DR(eye)VE 是由意大利 ImageLab 构建的第一个用于驾驶员注意力预测的公开数据集, 可用于研究驾驶员意图, 行人意图等。它收集了 74 个视频序列(每个视频 5 分钟), 共 55.5 万帧。每个序列记录了 8 名驾驶员在实际驾驶过程中注意力数据, 以及驾驶员的位置, GPS 信息、车速和行驶路线。该数据集涵盖广泛交通场景, 视频记录了城市的不同区域(市中心、乡村、高速公路),

呈现了从无交通到非常混乱的交通状况。并根据完全不同的大气条件(晴天、多云和雨天)和一天中的不同时间(白天和晚上)分别进行实验。



图 10 DR(eye)VE 数据集样本

2. JAAD 数据集

下载地址: https://data.nvision2.eecs.yorku.ca/JAAD_dataset

JAAD 数据集是由加拿大约克大学提供, 用于研究自动驾驶环境下的联合注意力, 特别关注行人和驾驶员在交叉口的行为及其影响因素。JAAD 数据集共有 346 个高分辨率的视频片段(大部分为 5-10 秒), 它们是从北美和东欧几个地点拍摄的约 240 小时的驾驶视频中提取出来的, 代表了城市中各种天气条件下驾驶的典型场景。该数据集包含 88000 帧, 2793 个单独的行人并标记了超过 390000 个边界框, 为每个边界框提供了遮挡标记(约 55000(13%)的边界框标记为部分遮挡, 约 49000(12%)标记为重度遮挡)此外, 它为每个行人提供属性列表(如年龄、性别、运动方向等), 并为每个视频提供可见的交通场景元素列表(如停车标志、交通信号等)。



图 11 JAAD 数据集样本

四、已声明但未公开的数据集

1. DADA-2000 数据集

DADA-2000 数据集是由长安大学交通视觉安全(Lotvs)实验室和西安交通大学人工智能与机器人研究所联合提供的面向驾驶视野内行车事故的大型行车驾驶注意力数据集。此数据集为目前首个结合事故检测、预测、驾驶注意预测、结合注意力进行事故分析的大规模数据集。该数据集共 2000 个驾驶事故序列, 共计 658, 476 帧, 根据自车参与与否及事故发生时的参与者类型

分为 54 类。涵盖不同天气，光照，道路环境。记录了每一帧的注意力数据（包括注视图谱与扫视路径），并且手工标注了每个序列的真实异常坐标。详细的介绍参考论文：“DADA-2000: Can Driving Accident Be Predicted by Driver Attention? Analyzed by A Benchmark”，投稿于 ITSC2019。

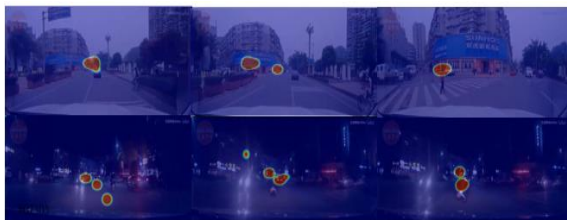


图 12 DADA-2000 数据集

2. BDD-A 数据集

BDD-A 数据集是从加州大学伯克利分校 Deep Drive 实验室在 BDD100K 数据集中选择的带有刹车行为的驾驶视频。该数据集将每个视频剪辑为刹车时间之前 6.5 秒和之后 3.5 秒。共收集 1232 个视频（约为 3.5 小时）。在实验室环境下招募志愿者观看驾驶视频并采用眼动仪收集其注意力数据，相比于 DR(eye)VE 数据集，其更加关注于收集危险情形的驾驶注意力。相关的论文可参考论文：“Predicting Driver Attention in Critical Situations”，发表于 ACCV2018。



图 13 BDD-A 数据集样本

3. A3D 数据集

A3D 数据集是由密西根大学机器人研究所和印第安纳大学提供的，用于视频事故检测。包含 1500 个事故序列（每个序列一个视频），共有 208166 帧。每个视频包含不同时间位置的交通异常事件，并为每个视频标记了异常的开始和结束时间（开始时间定义为不可避免的事故点，结束时间定义为所有参与者恢复正常移动状态或完全停止的时间）。每个交通异常都用一个二进制值标记，指示是否涉及自我车辆。在数据集中，约 60% 的事故涉及自我车辆，而其他事故则是通过从第三人称的角度观察到的。具体可参考论文：“Unsupervised Traffic Accident Detection in First-Person Videos”，投稿于 IROS2019。

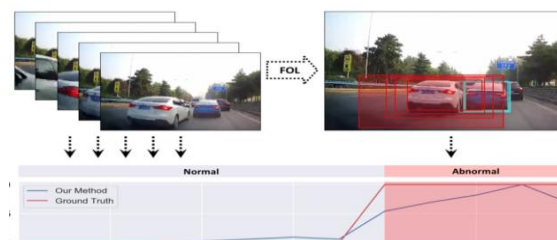


图 14 A3D 数据集样本

（责任编辑：李策，沈沛意）



严定鑫

硕士研究生，长安大学电子与控制工程学院，主要研究方向为驾驶场景注意力检测、异常事件分析。



房建武

副教授，长安大学电子与控制工程学院，现为西安交通大学人工智能与机器人研究所博士后，主要研究向为计算机视觉、交通场景视觉安全。



薛建儒

教授，西安交通大学人工智能与机器人研究所。主要研究方向为无人驾驶中的计算机视觉、视觉导航及场景理解。

招聘信息

一、哈尔滨工业大学（威海）计算机学院招聘人工智能方向教师数名

哈尔滨工业大学（威海）是哈尔滨工业大学“一校三区”办学格局中的重要组成部分，并坚持与校本部统一规格、统一标准、统一要求、统一质量、统一品牌。近年来学校年科研经费连续过亿元，获国家技术发明二等奖 1 项，国家科技进步一等奖 2 项、二等奖 1 项。刘永坦院士获 2018 年度国家最高科技奖。威海校区可充分共享校本部的数字图书馆、研究平台等学术资源。

哈尔滨工业大学（威海）计算机科学与技术学院成立于 2004 年，计算机专业创建于 1989 年，是哈工大计算机学科的重要组成部分，与校本部共享计算机科学与技术国家一级重点学科和软件工程、网络空间安全国家一级学科，具有硕士、博士学位授予权，并共享双一流建设学科资源。计算机应用技术学科被列为山东省“十一五”规划重点建设学科。学院现有专任教师 53 人，其中教授 9 人，博士生导师 6 人，副教授 19 人，山东省泰山学者青年专家、泰山产业领军人才、山东半岛国家自主创新示范区“蓝色汇智双百人才”各 1 人，山东省教学名师 1 人，哈工大教学带头人 1 人。学院还拥有双基地教师 6 人，其中中国工程院院士 1 人、中国计算机学会会士 1 人、黑龙江省教学名师 1 人。

1. **岗位职责：**从事教学和科研工作。

2. 应聘要求：

- 即将毕业的应届博士或博士后；
- 研究方向为人工智能，包括计算机视觉、自然语言处理、语音处理、机器学习理论等；
- 具有一年及以上的海外留学经历；
- 以第一作者发表过 CCF A 类会议或期刊 2 篇及以上；
- 身心健康、热爱教学和科研。

3. **待遇：**因人而异。

4. 联系方式：

请感兴趣的人才直接与张老师联系：
s.zhang@hit.edu.cn。

二、腾讯优图（深圳）招聘

1. 岗位名称：计算机视觉研究员（校招/社招）

1) 岗位要求：

- 熟练掌握计算机视觉和深度学习的基本方法；
- 优秀的分析问题和解决问题的能力，对解决具有挑战性的问题充满激情；
- 较强的算法实现能力，熟练掌握 C/C++ 编程，熟悉深度学习平台如 caffe, tensorflow 等；
- 有较强的代码能力优先，获得过 ACM 或其他商业代码竞赛的荣誉，如 ACM 区预赛金牌、NOI 银牌以上、百度之星决赛等；或代码开源在 github 上并有较大影响；
- 有较强的学术比赛经验或者在重要数据集的 Leaderboard 上排名靠前者优先。

2) 岗位职责：

- 负责计算机视觉和深度学习基本算法的开发与性能提升，涉及的主要问题包括但不限于物体检测、分类、语义分割和图像处理等；推动计算机视觉算法和深度学习在视频理解，工业 AI 以及自动驾驶等领域的性能优化和落地；
- 提出和实现最前沿的算法，保持算法在工业界和学术界的领先。

2. 岗位名称：AI 应用研发工程师（校招/社招）

1) 岗位要求：

- 3 年以上 Linux 后台服务器开发经验，或者 Android/iOS 客户端开发经验；
- 具备良好的编程能力，熟练掌握 C/C++，熟悉 Python，熟悉网络编程或数据库，有

扎实的数据结构和算法基础；

- 对后台或客户端性能优化、架构设计有一定的理解，熟练使用 GDB/LLDB 调试工具；
- 有高性能大容量服务器系统设计经验者优先；
- 熟悉现代 CPU/GPU/DSP 体系结构与编程、有高性能计算经验者优先；
- 有模式识别相关应用开发经验者优先。

2) 岗位职责：

- 负责视频、图像、自动驾驶等场景 A 系统后端架构设计；
- 系统开发、技术攻坚和性能优化；
- 基础架构组件开发和优化。

3、岗位名称：AI 产品经理(校招/社招)

1) 岗位要求：

- 计算机，自动化，数学等相关专业硕士及以上学历，名校机器学习、人工智能、计算机视觉等专业博士优先；
- 熟悉计算机视觉基础算法，对落地场景有深刻理解；
- 有激情好奇心，对人工智能有浓厚兴趣。

2) 岗位职责：

- 负责计算机视觉相关的技术研究和产品策划工作，包括但不限于图像视频理解，工业 AI，内容管理，图像处理产品化工作；
- 跟踪前沿的技术以及产品，结合实际业务场景，推动公司内外协同，制定产品与解决方案。

4、岗位名称：AI 项目经理(社招)

1) 岗位要求：

- 三年以上互联网，通信等相关领域的项目管理经验，对 AI 有较深刻理解；
- 有人工智能研发经验以及对人工智能算法特别是计算机视觉有深刻理解者优先；
- 丰富的沟通、交流和组织能力，出色的团队合作精神；
- 工作认真负责，能承受短期项目压力。

2) 岗位职责：

- 负责项目的协调与组织，发现并跟踪项目问题，管理项目风险；
- 负责组织项目会议，制定流程规范，确保项目有序推进及快速迭代；
- 协助项目负责人负责项目相关的内外部沟通协调工作。

5、岗位名称：AI 运营开发(社招)

1) 岗位要求：

- 责任感强、有较强的逻辑思维能力、沟通能力和抗压能力；
- 计算机或相关专业本科及以上学历，2 年以上相关工作经验；
- 熟悉 C/C++ 开发，有一定的架构能力和良好代码规范；
- 熟悉 Linux/unix 系统与开发环境。

3) 岗位职责：

- 负责将腾讯优图的算法落地成服务；
- 负责服务平台开发、性能优化等工作；
- 建设优图服务及其他产品的可运营体系；
- 负责 GPU 集群的日常运维工作。

6、投递方式

- 1) 找到腾讯优图小伙伴进行内推。
- 2) 直接将简历发送至 youtuxlab@tencent.com。(邮件主题格式：姓名+应聘岗位+预计能入职时间)

三、日本 National Institute of Informatics (NII) 招收实习生

日本国家信息研究所(NII)的 Shin'ichi Satoh 实验室有招收实习生。该工作侧重于基于内容的(content-based)图像和视频分析，如实例搜索、医学图像分析、视频到文本、活动检测等。申请截止日期：2019 年 5 月 31 日。更多“NII 国际实习计划”的相关信息，请访问：<https://www.nii.ac.jp/en/about/international/mouresearch/internship2019-1>，实验室的相关信息，请访问 <http://www.satoh-lab.nii.ac.jp>。

1、招聘要求：

- 就读于 NII 的 MOU 学院的硕士/博士，特

殊情况下招收其他院校的博士；

- 具有较强的自我激励能力和发表过论文的学生优先；
- 实习时间不少于3个月，6个月为佳。

2、相关职责及费用：

- 实习生应根据自己的兴趣，从事科研工作或参加比赛；
- 地点：东京；
- 实习学生在NII期间，NII承担每天5700日元(约17万日元/月)的住宿费。

3、联系方式：

如需要更多的信息或投递简历，请联系 Dr. Zheng Wang (wangz@nii.ac.jp) 和 Prof. Shin'ichi Satoh (satoh@nii.ac.jp)。

四、腾讯 AI Lab 招收计算机视觉实习生

腾讯 AI Lab 长期招聘计算机视觉/图形学方向优秀硕士或博士的算法实习生，工作地点深圳，能实习6个月以上，工作任务主要是做研究项目。

1、应聘要求：

- 对 SLAM 及基于 RGBD 输入的三维重建算法有深入研究，对图形学有了解、有 unity3d、OpenGL 或 OpenGL ES 相关开发经验者优先；
- 对 3D Face、3D Body and Pose 等方面有深入研究，对于 3D 模型的变形与动画、逼真的阴影/光照渲染、非刚性配准等方面相关研究经验者优先；
- 熟练掌握 tensorflow 等深度学习框架，对物体检测(object detection)、动作识别(action recognition)、视觉问答(VQA)、视频人/物交互(Human-Object Interaction)等方向有深入研究者优先。

2、联系方式：

有意者请发送简历到 linchaobao@tencent.com。

五、PayPal 补招实习名额

PayPal 补招暑期实习生了！坐标上海浦东，955 的不养老外企，实习还有转正机会。正式员

工薪酬待遇好，发美股股票。

1、Software Engineer intern 软件开发工程师：

1) 工作内容：

- 负责系统模块的设计，开发（包括前端和后台），测试，上线以及维护的全栈式开发；
- 使用 Java 等高级语言编写具有处理高并发，高可用性，易维护，并具有高拓展性的代码；
- 积极且迅速应对产品在线上运行中可能遇到的问题。

2) 任职资格：

- 计算机等相关专业；
- 拥有扎实的常用设计模型，算法和数据结构的功底；C++，Java；
- UNIX 或 Linux 平台知识和经验，会使用 Python Perl 或 Shell 编程；
- 良好的沟通能力，中英文读写俱流利；
- 熟悉使用 SQL 或 NoSQL 数据库优先；
- 懂得分布式系统优先。

2、数据科学家 Data Scientist Intern：

1) 工作内容：

- 对 PayPal 网上支付的数据进行风险定量分析，运用统计学、数据挖掘、机器学习算法建立反欺诈模型，减少欺诈损失；
- 持续改进与提高风险模型的流程、算法和工具，完善 PayPal 风险管理模型解决方案。

3) 任职资格：

- 计算机，工程，数学，统计等相关专业；
- 出色逻辑思维能力，了解数据分析挖掘或者建模相关知识，具备运用 SAS，R 或 Python 进行数据分析经验者优先考虑；
- 熟练运用一种及以上编程语言，熟悉数据库操作及数据库编程；
- 良好的沟通能力，中英文读写流利。

3、投递方式：

简历投递至：cncampus@paypal.com，邮件命名方式：职位名称-姓名-学校-电话。

六、阿里巴巴优酷招聘计算机视觉实习生

1、岗位职责：

- CV 和深度学习相关研究。

2、应聘要求：

- 计算机、通信、自动化或相关专业硕士以上学历，图形图像处理相关专业背景；
- 熟悉常用的图像处理算法和深度学习算法；
- 动手能力强，熟悉常用 DL 框架；
- 有顶会期刊论文发表者优先。

3、联系方式：

简历请发送到：baichao.wbc@alibaba-inc.com，工作地点：望京阿里中心 B 座。

七、中国电子北京圣非凡公司招聘：

中国电子信息产业集团公司（CEC）是中央直接管理的国有独资特大型集团公司，是中国最大的国有 IT 企业。北京圣非凡电子系统技术开发有限公司是国资委所属中国电子信息产业集团公司（CEC）旗下重要高科技企业，公司以科技创新为核心，产品线包含多型无线专用通信设备与系统、有线通信设备与系统、自动化控制设备与系统。全部产品均具备自主知识产权，技术水平达到国内领先水平。公司具有从事专用通信装备业务所需要的全部资源，且均较完善，其中具有一级保密单位、武器装备承制单位、军工许可证等资质。单位具有国内一流的试验保障设备、仪器仪表和实验设施。在技术能力方面，具有国内一流的数字信号处理技术、无线通信算法技术、嵌入式控制处理技术和完整的通信装备整机研发生产能力，在专用通信业务领域形成了独特的技术优势，以产品的高技术和高附加值获得快速良好的发展。

在未来的日子里，热忱欢迎广大优秀人才加入我们的团队，共创美好明天！

1、硬件设计师(功放设计方向)：

- 掌握射频电路基础、功放设计、微波电路基础；

- 有实际项目开发经验者优先考虑；
- 研究生及以上。

2、硬件设计师(数字信号处理及高速电路设计方向)：

- 熟悉 DSP 软、硬件架构、接口设计有实际项目开发经验者优先考虑；
- 独立完成 DSP 相关的软硬件开发及电路设计；
- 研究生及以上。

3、硬件设计师（FPGA 设计师）：

- 掌握 VHDL/Verlog 语言；
- 独立完成 FPGA 硬件电路的架构设计及调试；
- 独立完成 FPGA 软件架构、系统设计、变成、仿真和调试工作；
- 有实际项目开发经验者优先考虑；
- 研究生及以上。

4、硬件设计师(射频电路设计)：

- 射频电路基础、微波和天线有实际项目开发经验者优先考虑；
- 有实际项目开发经验者优先考虑；
- 研究生及以上。

5、软件设计师（C++方向）：

- 熟练掌握 C、C++ 语言；
- 熟练掌握 Microsoft Visual Studio、QT 等主流开发工具；
- 熟悉 Oracle 数据库或 MySQL、MS SQLServer 数据库；
- 熟练掌握 Office 办公类软件；
- 能够按照标准规范（如 GJB）熟练编写软件开发各类文档；
- 优先考虑有 QT 图形化桌面应用开发经验者，以及有基于 GIS 的指挥调度决策系统或无线通信调度系统开发经验者；
- 有实际项目开发经验者优先考虑；
- 研究生及以上。

6、软件设计师（Java 方向）：

- 熟悉 Java、JSP/Servlet、JavaScript、HTML5 编程，熟练使用 JQuery、Ajax 等开源框架一种以上；

- 熟练使用 spring、Hibernate、struts2 等开发框架；
- 熟悉 XML 解析，熟悉 Web Service 开发；
- 熟悉数据库建模，熟练掌握 Oracle、SqlServer、Mysql 等数据库设计和 SQL 语言使用；
- 熟悉 tomcat、weblogic 等服务器配置与应用；
- 优先考虑有数据监控项目开发经验者；
- 有实际项目开发经验者优先考虑；
- 研究生及以上。

7、基本待遇：

- 正式在编职工；
- 外地生源（本科以上）按照在京中央企业落户管理规定进行落户；
- 技术骨干按照在京中央企业落户管理规定进行落户；
- 异地分居夫妻按照在京中央企业落户管理规定进行落户；
- 提供五险二金（医疗、养老、工伤、生育、失业保险和住房公积金、企业年金）；
- 享受国家规定的带薪年假、探亲假、婚丧假、产假等；
- 享受国家规定的劳动保护、防暑降温等政

策；

- 享受过节费、伙食补贴和交通补贴等公司福利；
- 发展空间；
- 优厚的弹性制薪酬政策，引领行业整体薪酬水平；
- 人性化的工作氛围，提供员工稳定的职业发展环境；
- 扁平化和平台化的组织结构给予员工充分授权和自主创新机会，有利于优秀员工脱颖而出；
- 网状多通道职业发展路径和定制化培训开发计划，加速人才孵化。

8、联系方式：

- 电话：010-66608126；
- 电子邮箱：sffhrwork@163.com；
- 公司网址：<http://www.cec-sff.com>；
- 熟悉数据库建模，熟练掌握 Oracle、SqlServer、Mysql 等数据库设计和 SQL 语言使用；
- 公司地址：北京市昌平区未来科技城中国电子信息安全技术研发基地 B 栋。

（责任编辑：贾同，刘丽）

征文通知

1. 会议征文

计算机视觉领域相关国内外会议的征文通知如表 1 所示。同时，可继续关注每个会议举办的 workshop 或 special session。

2. 期刊征文

计算机视觉领域近期相关期刊专刊的征文通知如表 2 所示，包括 Artificial Intelligence in Medicine, Image and Vision Computing 特刊等。

3. 会议简介 (ICMLA)

ICMLA (International Conference On

Machine Learning And Applications) 是计算机视觉领域最权威的学术会议之一，为全球模式识别领域的研究人员及参与者提供一个交流平台，其内容涉及近期在计算机视觉、机器学习、模式识别方面的最新进展，会议论文集代表了上述领域最新的发展方向和水平。

ICMLA-2019 将于 2019 年 12 月 16 日至 12 月 19 日在美国佛罗里达州举办，论文截止时间为 2019 年 7 月 5 日。征稿范围涉及机器学习在医学、生物学、工业、制造、安全、教育、虚拟环境、游戏等方向的研究进展。

(责任编辑：樊鑫)

表 1 计算机视觉领域相关国内外会议

会议名称	会议时间	会议地点	截稿日期	会议网站
ICMLA 2019	2019. 12. 16-19	美国 Florida	2019. 09. 07	https://www.icmlaconference.org/icmla19
ISoCC 2019	2019. 08. 26-28	加拿大 Vancouver	2019. 05. 30	http://www.isocc.net
ICMIM 2019	2019. 08. 19-22	韩国 Incheon	2019. 05. 15	http://www.icmim.org
NWCOM 2019	2019. 08. 31- 09. 01	阿拉伯联合酋 长国 Dubai	2019. 05. 19	https://csity2019.org/nwcom/index.html
ICIST 2019	2019. 10. 10-12	立陶宛 Vilnius	2019. 05. 25	https://icist.ktu.edu

表 2 计算机视觉领域相关国内外期刊专刊

期刊名称	专刊题目	截稿日期
Artificial Intelligence in Medicine	Trusted virtual Environments for Neuro-Informatics	2019. 09. 30
Image and Vision Computing	Special Issue on Novel Insights on Ocular Biometrics In Image and Vision Computing	2019. 12. 31