

CCF

计算机视觉 专委会简报

COMPUTER VISION NEWSLETTER

2019/01 期
总第 17 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
专委新闻	专委会与滴滴出行签订合作协议	06
	计算机视觉标准工作讨论会召开	07
	CCF-CV 专委会工作会议成功举办	07
	PRCV 2018 圆满闭幕	10
	RACV 2018 成功举办	13
会议推介	PRCV 2019 征文通知	14
	PRCV 2019 诚邀申请组织竞赛	15
	PRCV 2019 诚招赞助	16

科技前沿

专题综述	人脸属性转换进展概述	18
热点追击	基于选择性蒸馏的自然环境低分辨率人脸识别	22
	视觉代表模式稀疏选择与发现	23

委员风采

委员访谈	上海交通大学倪冰冰教授访谈	24
委员好消息		28

资源平台

开源代码	图像视频生成技术	29
数据集	三维目标与场景	31
招聘信息		34
征文通知		38

CCF-CV 走进高校系列报告会

第 63 期 湘潭大学

时间：2018 年 11 月 17 日

第六十三期“CCF-CV”走进高校系列报告会活动在湘潭大学计算中心 A 座 217 报告厅成功举行。来自湖南大学、湖南师范大学、湖南工程学院、上饶师范学院、湖南世优电气有限公司等高校和企业的人员和约 150 名师生，聆听了计算机视觉领域的专家学者在该领域的最新前沿技术及应用报告。与会专家学者有中山大学赖剑煌教授，南京信息工程大学刘青山教授，厦门大学纪荣嵘教授，复旦大学姜育刚教授和西北工业大学王鹏教授。会议报告由执行主席湘潭大学信息工程学院欧阳建权教授，张东波教授和肖芬教授主持。

本次报告会，通过本次报告与会人员与特邀专家面对面的沟通、交流和研讨，达到了相互学习与借鉴，促进合作与发展的目的，与会的专家学者对本次报告会的组织工作和与其目标给予了充分的肯定。



第 64 期 昆明理工大学

时间：2018 年 11 月 29 日

由中国计算机学会计算机视觉专委会主办，信息工程与自动化学院承办的第六十四期“CCF-CV 走进高校系列报告会”活动在昆明理工大学成功举行。与会的专家学者有北京大学彭宇新教授，南京信息工程大学刘光灿教授，复旦大学金城教授和北京航空航天大学刘偲副教授。报告会执行主席为昆明理工大学信息工程与自动化学院院长余正涛教授和昆明理工大学信息工程与自动化学院刘骊副教授。



此次活动持续四个多小时。报告内容精彩，现场气氛热烈，交流互动活跃。参会师生均对四位讲者表示由衷的欢迎与感谢，报告会在热烈的掌声中圆满结束。通过本次参会师生与特邀专家面对面的沟通、交流和研讨，达到了相互学习与借鉴的目的，与会专家学者对本次报告会的组织工作给予了高度评价。

第 65 期 郑州航空工业管理学院

时间：2018 年 12 月 13 日

第 65 期 CCF-CV 走进高校系列报告会活动——“计算机视觉前沿技术及应用”报告会在郑州航空工业管理学院图书馆第二报告厅成功举行。本期报告会邀请了中科院深圳先进技术研究院乔宇研究员，郑州大学徐明亮教授，南开大学杨巨峰副教授三位专家学者做特邀报告。来自华北水利水电大学、河南农业大学、中原工学院、郑州工程技术学院、IEEE 郑州大学分会的研究生以及郑州航空工业管理学院的 200 余名师生聆听了三位专家的报告。最后报告会在热烈的掌声中圆满结束。



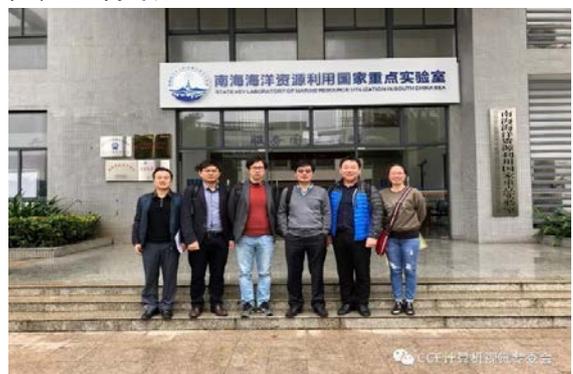
专家们耐心解答师生代表提出的问题。他们所展示的研究成果让学生们享受了一次学术盛宴，启发了参会人员的研究思路。最后报告会在热烈的掌声中圆满结束。

第 67 期 海南大学

时间：2019 年 1 月 5 日

第 67 期 CCF-CV 走进高校系列报告会活动“计算机视觉前沿技术及应用”在海南大学南海海洋资源利用国家重点实验室 809 会议室成功举行。本期报告会邀请了复旦大学张军平教授、江西财经大学方玉明教授、西安交通大学沈超教授、南京理工大学唐金辉教授四位专家做特邀报告。报告由海南大学信息科学技术学院靳婷副教授主持。

四位专家学者的报告深入浅出，报告现场座无虚席。在场的部分师生与专家学者们进行了深入的讨论与交流。



(责任编辑：马占宇)

第 66 期 厦门大学

时间：2018 年 12 月 20 日

第 66 期 CCF-CV 走进高校系列报告会活动——“计算机视觉前沿技术及应用”报告会在厦门大学克立楼报告厅成功举行。本期报告会邀请了浙江大学章国锋教授，西安交通大学孟德宇教授，西安电子科技大学邓成教授，华南理工大学金连文教授，中山大学郑伟诗教授，哈尔滨工业大学左旺孟教授六位专家学者做特邀报告。来自厦门大学、集美大学、华侨大学等高校的 100 余名师生聆听了六位专家的报告。



专委会与滴滴出行签订合作协议

首届中国模式识别与计算机视觉学术会议（PRCV2018）于2018年11月24日在广州白云国际会议中心盛大开幕。同期，滴滴与中国计算机学会计算机视觉专委会（CCF-CV）签订合作伙伴协议。CCF会士、理事、副秘书长，CCF计算机视觉专委会副主任，中科院计算所副所长陈熙霖与滴滴出行副总裁，智能出行部和地图事业部负责人郑小虎（Tiger Qie）分别代表CCF-CV与滴滴出行在协议上签字。



签约仪式伊始，陈熙霖代表中国计算机学会计算机视觉专委会致欢迎词，他十分诙谐幽默地把CCF以及计算机视觉专委会和滴滴的渊源娓娓道来，并指出企业和学会不仅仅要在宏观层面建立广泛合作，例如滴滴和CCF共建大数据联合实验室，更要和各个专委会就专项话题展开深入合作。陈熙霖表示这次签约仪式是合作的美好开始，后续希望双方继续加强互动与交流，推动计算机视觉在工业界的应用和落地。

Tiger 作为滴滴出行的代表向与会老师表示热烈欢迎和美好祝愿。Tiger 介绍到滴滴与计算机视觉专委会渊源已久，表示滴滴一直利用人工

智能技术与结合海量数据优势推动出行场景下视觉应用的落地，以提升用户体验，创造社会价值，建设安全、开放、可持续的移动出行新生态为己任，并期望在国家推动新一代人工智能发展战略下，滴滴与计算机视觉专委会的深入互动与交流可以加速推动AI与产业的发展融合，为高质量发展提供新动能。最后双方代表共同表示，计算机视觉专委会与滴滴出行在推动我国计算机视觉学科发展、加强产业交流与合作的目标上高度统一，建立积极的合作伙伴关系是水到渠成。



CCF-CV专委会副主任、爱奇艺首席科学家王涛博士，CCF-CV专委会秘书长、中科院自动化所王亮研究员，CCF-CV专委会常务委员、南京信息工程大学刘青山教授，CCF-CV专委会常务委员、北京大学林宙辰教授，CCF-CV专委会常务委员、上海科技大学虞晶怡教授，计算机视觉专委会委员代表和滴滴出行增强现实团队总监徐迅博士，地图事业部视觉计算团队负责人许鹏飞博士，AI Labs图像组负责人沈海峰博士等人出席见证。

（责任编辑：马占宇）

计算机视觉标准工作讨论会召开

全国信息技术标准化技术委员会图形图像分委会计算机视觉标准组第三次会议于 2018 年 11 月 25 日在广州白云国际会展中心岭南大会堂贵宾室 A 召开。计算机视觉标准工作组陈熙霖组长、王涛副组长，全国信标委图形图像分委会王聪秘书长，AVS 标准组资深专家、中科院自动化所卢汉清研究员，以及计算机视觉标准组成员出席会议。



首先，陈熙霖组长讲话。计算机视觉术语标准已正式立项，该项工作与同类型国际标准同步进行，鼓励大家抓住机遇，积极参与到国家标准和国际标准制定工作中。

其次，王聪秘书长介绍了近期我国国际化取得的重要进展和成绩，中国国家电网有限公司董事长舒印标当选为 IEC 第 36 届主席(2020-

2022)，ITU-T 赵厚霖秘书长成功连任（2019-2022）。王聪秘书长也成功当选 ISO/IEC JTC1/SC24/WG7 召集人（2019-2022），为计算机视觉领域国际标准化打下基础。然后，王聪秘书长给大家详细讲解了术语标准定义规范和注意事项，并对已有的术语定义草稿，针对性地提出了修改建议。另外，建议标准工作组相关人员，积极参与 ISO/IEC JTC1/SC24/WG7 关于计算机视觉框架标准制定工作。

最后，大家讨论了计算机视觉术语标准起草人的权利和义务，共同制定计算机视觉术语标准工作进度，争取在 2019 年 1 月底完成计算机视觉术语标准英文版，并提交相关国际标准组织。



（责任编辑：毋立芳）

CCF-CV

专委会工作会议成功举办

中国计算机学会计算机视觉专委会（CCF-CV）2018 年度工作会议于 2018 年 11 月 24 日在广州白云国际会议中心举办，来自全国高校、

科研院所、企业的现任委员及候选新申请委员 300 多人参加了工作会议。

会议由专委会秘书长、中科院自动化所王亮

研究员主持。会议开场首先由专委会主任、中科院自动化所谭铁牛院士致辞。谭院士感谢新老委员们从全国各地到场参会，他深情回顾了专委会成立五年来的快速成长壮大，积极肯定了专委会在过去一年中学术交流各方面取得的丰硕成果，鼓励委员们在当前人工智能热潮中聚焦学术前沿做出高质量研究工作，希望专委会大家庭一起精诚合作，各项活动进一步提升品牌质量，打造学术交流精品活动，为推进视觉学术研究与产业发展继续发挥积极引领作用。



之后，中国计算机学会常务理事、中国科技大学吴枫教授代表学会致辞。吴枫教授对专委会工作会议的召开表示祝贺，对专委会开展的各项活动给予了积极评价，特别肯定了专委会在成立短短5年时间内的迅速发展和建立的一系列学术交流品牌活动。他最后预祝2018年中国模式识别与计算机视觉大会（PRCV2018）圆满成功。

接下来，中国计算机学会专业委员会工作委员会代表、中科院计算所蒋树强研究员代表专委工委致辞。在致辞中，蒋树强研究员热情洋溢地回顾了过去几年连续参加CCF-CV专委工作会议的体会，积极肯定了专委历经几年快速成长所取得的各项丰硕成果，并祝专委会在各位主任和常委们的带领下齐心协力、越办越好。

随后，由专委会秘书长王亮研究员向与会代表做年度工作报告。报告简要介绍了专委组织结构及历年评优结果，通报了5月份专委年度常委会议和6月份秘书处工作会议所作出的若干新举措，全面回顾了专委过去一年的学术交流活动，指出了工作中存在的问题和改进方案，最后介绍了下一年度专委工作计划。

接下来，进入激动人心的CCF-CV颁奖仪式环节。颁奖仪式由专委会副主任、专委提名与奖励工作组组长查红彬教授主持，查教授详细介绍了本年度设立的奖项评选范围与评选规则，包括：终身学术贡献奖、杰出成就奖、服务贡献奖、学术新锐奖。4个奖项的获奖人既有早年投入我国计算机视觉专业奠基性研究的老一代科学家，又有新一代崭露头角、冉冉上升的学术新星，也有为专委发展尽心尽力、无私服务的中生代科研工作者，充分展示了计算机视觉专委大家庭的繁荣兴盛。

本年度CCF-CV终身学术贡献奖授予复旦大学吴立德教授、北京交通大学袁保宗教授两位前辈，由专委会主任谭铁牛院士为两位获奖人颁奖。吴教授现场发表了激情洋溢的感言，回顾了早年我国计算机视觉方向的起步与发展，赢得了代表们的热烈掌声。袁教授为专委发来了感谢信并委托赵耀教授代他现场宣读，信中指出计算机视觉是人工智能发展中最重要部分之一，勉励专委会在国家大力鼓励创新的潮流中做出更大的成绩



贡献。两位前辈的鼓励与期许，极大振奋了专委会代表们的科研工作热情。

本年度首次颁发 CCF-CV 杰出成就奖，以最高标准评选出计算机视觉领域取得了公认的杰出成就的科研人员。该奖项授予专委会主任谭铁牛院士，以表彰他“在生物识别与视频监控领域的创新理论发展、关键技术突破、实际应用推广等方面做出了卓越的贡献”。专委会副主任陈熙霖研究员为谭院士颁发了奖励证书。谭院士发表了热情风趣的获奖感言，衷心感谢专委会同仁的支持与认可。CCF-CV 专委会正是在谭院士的带领下于 5 年前成立，并在此后获得了快速成长，极大地促进了我国计算机视觉领域的学术进步，推动了产学研全链条合作发展。

本年度继续颁发年度服务贡献奖，授予北京工业大学毋立芳教授、上海科技大学虞晶怡教授、中山大学郑伟诗教授等 3 位委员，由专委会副主任赖剑煌教授和 CCF 代表吴枫教授为获奖人颁奖，感谢他们在专委会活动、会议组织、秘书处、工委小组工作等方面作出的重要贡献。

本年度首次颁发另一重要奖项 CCF-CV 学术新锐奖，以鼓励创新性研究，促进青年人才成长，奖励计算机视觉及相关领域在读低年级博士生。首次获得该奖项的 3 位学术新星是：清华大学董胤蓬、华中科技大学廖明辉、中科院自动化所张士峰，由专委会副主任王涛博士和常委贾云得

教授为获奖人颁奖。

根据会议日程，工作会议还进行了新增常委和委员的选举。本年度共收到 13 位委员申请常委，71 位候选人申请成为委员。按照日程，13 位常委候选人一一上台进行了申请陈述，经过参会的专委会现任委员投票表决，共有 4 名候选人当选为新任常务委员，他们是上海科技大学虞晶怡教授、厦门大学纪荣嵘教授、清华大学鲁继文副教授、北京工业大学毋立芳教授。新委员的选举则由现任常务委员通过投票表决，最终共有 38 名候选人当选为新任委员。

委员建言献策环节引起了与会代表的热烈讨论。王蕴红、王涛、陈胜勇、刘家瑛、汪增福、殷绪成、欧阳万里、杨金锋、何晖光、程明明、孟德宇、李家康、强彦等新老委员代表就专委会的工作开展情况、目前存在问题、下一步目标规划、未来发展方向等发表了各自见解与建议，引起了委员们的共鸣。

最后，专委会主任谭铁牛院士做了总结发言。他感谢与会代表的积极参与和对专委会的认可，建议专委会整理与会代表的意见和建议、持续提升各项工作质量，勉励新老委员再接再厉，积极参加专委会的各项工作，争取通过专委会这个平台为学科发展做出更大贡献！会后，全体代表合影留念，专委会 2018 年度工作会议圆满结束！



（责任编辑：王瑞平）

PRCV2018 圆满闭幕

首届中国模式识别与计算机视觉大会（PRCV2018）于2018年11月23日至26日在广州白云国际会议中心成功举办，约1600名参会者汇聚于此，共同分享最新的研究成果、创新思想和产品应用。

大会开幕式



PRCV2018 由中国人工智能学会、中国计算机学会、中国自动化学会和中国图象图形学学会联合主办，中山大学协同中国人工智能学会模式识别专业委员会、中国计算机学会计算机视觉专业委员会、中国自动化学会模式识别与机器智能专业委员会、中国图象图形学学会视觉大数据专业委员会共同承办，广州南沙开发区（自贸区南沙片区）管委会协办，来自全国高校、科研机构以及工业界等多领域专家、学者代表约1600人汇聚一堂。现场出席的嘉宾有大会主席谭铁牛院士，中山大学校长罗俊院士，大会主席、北京大学查红彬教授，程序委员会主席、中国科学院自动化所副所长刘成林研究员，程序委员会主席、中国科学院计算所副所长陈熙霖研究员，程序委员会主席、清华大学周杰教授，中山大学数据科学与

计算机学院院长钱德沛教授，会议组织委员会主席、中国科学院自动化所王亮研究员、中山大学数据科学与计算机学院郑伟诗教授，广州南沙开发区（自贸区南沙片区）管委会潘玉璋副主任、促投局苏敏处长，伊利诺伊大学厄巴纳—香槟分校教授、IEEE/ACM Fellow David Forsyth，加拿大约克大学教授、IEEEPAMI、IJCV、CGF 副主编 Michael S. Brown。开幕式由程序委员会主席、中山大学数据科学与计算机学院赖剑煌教授主持。

大会主题报告

会议邀请到 David Forsyth, Michael S. Brown, Tamara Berg 和 Zhengyou Zhang 四位专家作大会主题报告，为与会者带来了视觉理解、成像模型、语义与图像、智能机器人等领域的权威讲座。来自相关领域的专家学者在此欢聚一堂，共庆中国模式识别和计算机视觉领域的盛事。



大会特邀报告和分组报告

24日下午 PRCV2018 开始进入大会特邀报告和分组报告环节，其中每个环节包括一个特邀报

告和三个分组报告。本环节共涉及八个子领域，分别是 24 日举办的“图像分析”、“学习和应用”、“物体检测”，25 日举办的“分类”、“视频分析”、“行人再识别”，以及 26 日举办的“分割检测”以及“图像处理及生成”。特邀讲者学识渊博，旁征博引，深入浅出；分组报告技术前沿，方法创新，听众热情高涨，如沐春风。



度几何表征学习与语义分析、智能医学影像分析、深度学习的理论理解、第六届中国图像视频大数据产业创新论坛、类脑计算与模式识别、第二届视觉大数据高峰论坛、IEEE Fellow 论坛——如何提高学术影响力、PRCV 青年学者论坛等。专家学者带来了精彩纷呈的前沿成果，与会听众产生了学术与科研的强烈共鸣。



墙报展示及顶刊顶会交流

24 日下午 PRCV2018 大会进行了顶刊顶会交流，吸引了众多参会人员的关注和交流。25 日和 26 日，PRCV2018 大会进行了论文墙报展示环节。参会人员与论文作者进行面对面交流，论文作者如数家珍，听者聚精会神，学术思想碰撞，科研经历共鸣。



大会分论坛

大会期间同时举办了多场前沿技术论坛，包括：第三届计算机视觉研究与应用创新论坛、深

企业宣讲

人工智能领域目前已经取得了很多人振奋的成就。人工智能一个很重要的领域就是模式识别和计算机视觉。24 日上午，在企业宣讲环节里面，我们看到了人工智能的落地情况。这些代表企业有滴滴出行、高新兴科技集团股份有限公司、图普科技公司、Face++ 旷视、商汤科技和英伟达公司。25 日上午，深圳极视角公司、深圳捷尚视觉科技公司、深圳力维智联技术有限公司和图匠数据等赞助企业在会议期间进行了宣讲。

大会讲习班

在 23 日和 26 日的下午，大会安排举行了不同主题的讲习班，其中分别有：智能影像增强计算、全光人脸图像分析、基于高阶统计建模的深层卷积神经网络、在线视觉跟踪基础及前沿、生物视觉机理与计算机视觉应用、基于 FPGA 技术

的从端到云自适应智能计算。



PRCV 之夜

25 日，大会进行了 PRCV2020 承办方竞选，并举办由高新兴、图普科技、旷视、商汤和滴滴等企业全额赞助的 PRCV 之夜活动。经投票，在 PRCV 之夜宣布南京理工大学获得了 PRCV2020 的承办权。



大会致谢

首届中国模式识别与计算机视觉学术会议 (PRCV2018) 得到北京商汤科技开发有限公司、北京旷视科技有限公司、广州图普网络科技有限公司、

公司、北京滴滴无限科技发展有限公司、高新兴科技集团股份有限公司、北京字节跳动科技有限公司、精锐视觉智能科技(深圳)有限公司、慧眼自动化科技(广州)有限公司、北京爱奇艺科技有限公司、北京地平线信息技术有限公司、驭势科技(北京有限公司)、杭州海康威视数字技术股份有限公司、广州像素数据技术股份有限公司、广州视源电子科技股份有限公司、北京新唐思创教授科技有限公司、NVIDIA 技术服务(北京有限公司)、北京联想软件有限公司、赛灵思电子科技有限公司(上海)有限公司、浙江捷尚视觉科技股份有限公司、北京海天瑞声科技股份有限公司、软银机器人贸易(上海)有限公司、苏州超集信息科技有限公司、深圳力维智联技术有限公司、深圳极视角科技有限公司、广州图匠数据科技有限公司、Springer 出版社的赞助,对他们的大力支持表示衷心感谢!



大会同时获得了众多志愿者的支持,对他们为大会的筹备和顺利举办所付出的辛勤工作表示衷心感谢!



(责任编辑: 王瑞平)

RACV2018 成功举办

2018年11月24日，中国计算机学会 CCF-CV 专委会主办的第三届计算机视觉研究与应用创新论坛（RACV2018）在广州白云国际会展中心2号楼深圳厅成功举办，近300名嘉宾现场参加。



CCF-CV 专委会副主任、中山大学赖剑煌教授致辞，并介绍了 RACV 论坛的创办公理念和历史。本论坛由赖剑煌教授，专委会副主任、爱奇艺资深科学家王涛博士和专委会副秘书长、北京邮电大学马占宇副教授主持。

CCF-CV 专委会副主任、北京大学查红彬教授在“基于三维数据流融合的场景重建与传感器定位技术”中介绍了同时定位与地图（SLAM）技术研究发展进程，讲述了他们研究团队利用传感数据流和在线学习方法在三维重建、SLAM 等方面的重要研究工作。

CCF-CV 专委会委员、复旦大学姜育刚教授在“面向视频识别的深度学习方法及应用”报告中介绍了他们实验室近年在视频分析领域的研究成果，包括视频识别相关数据集构建、深度学习方法，以及在高铁部件损失检测及暴恐视频分析等方面的应用研究。

CCF-CV 专委会常务委员、叠境数字科技有限公司创始人、上海科技大学虞晶怡教授在“Learning to Build a New Reality”报告中

介绍了人工智能、计算机视觉和光场成像研究的发展历程，展示了他们团队围绕 A/VR+AI 理念在光场 3D 重建、材质复现方面的研究及应用工作。

CCF-CV 专委会委员、Momenta 研发总监任少卿博士在“无人驾驶的技术与挑战”报告中介绍了公司在环境感知、定位和地图，可行驶区域检测，车辆车型识别，以及行人姿态估计和动作识别方面的一系列具有很高准确率和鲁棒性的软件算法。

CCF-CV 专委会委员、字节跳动人工智能实验室总监王长虎博士在“抖音背后的那些黑科技”报告中介绍了抖音短视频的增长态势及其影响，以及深度学习和计算机视觉技术在抖音中的应用，包括问题视频精确筛选和视频酷炫玩法等方面的技术应用，让现场观众大开眼界。



在王涛博士主持的 Panel 环节，姜育刚教授、虞晶怡教授、任少卿博士和王长虎博士等嘉宾就“未来五年计算机视觉面临的挑战与机遇”议题各抒己见，给出各自对计算机视觉技术中最具挑战性研究问题及其应用领域和产品的思考。现场观众踊跃发言，积极参与互动。本次论坛在近300名参会嘉宾的热烈掌声中圆满结束。

（责任编辑：马占宇）

PRCV2019 征文通知

第二届中国模式识别与计算机视觉大会

PRCV 2019

2019年11月8-11日

曲江国际会议中心, 西安

征文通知

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、西安交通大学以及西安电子科技大学等单位共同承办。

本届会议将主要汇聚国内从事模式识别和计算机视觉理论与应用研究的科研工作者及工业界同僚, 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。现向广大科技工作者公开征集高质量、原创性的优秀学术论文英文。大会录用的稿件将在会上展示, 会议论文集将由 Springer 出版社出版, 并被 EI 和 ISTP 检索。优秀论文将推荐到国内外高质量期刊的特刊。

征文范围 (包括但不限于)

- 模式分类与聚类分析

- 机器学习
- 神经网络与深度学习
- 特征提取与特征选择
- 结构模式识别
- 文档分析
- 手写字符识别
- 人脸识别与姿态识别
- 生物特征识别
- 语音识别与说话人识别
- 信息检索
- 性能评测和基准数据库
- 物理模型驱动视觉分析及三维重构
- 计算机视觉基础理论
- 底层视觉理解、图像处理
- 三维视觉
- 目标检测与识别
- 动作识别
- 计算摄像学、传感与显示技术
- 运动目标检测及跟踪
- 多媒体数据分析
- 生物医学图像分析
- 遥感图像
- 最优化方法
- 机器人视觉
- 模式识别与计算机视觉应用

重要日期

- 投稿截止日期: 2019年4月15日
- 录用通知日期: 2019年6月15日
- 终稿提交日期: 2019年7月15日
- 会议举办日期: 2019年11月8-11日
(责任编辑: 毋立芳)

PRCV2019

诚邀申请组织竞赛

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、西安交通大学以及西安电子科技大学等单位联合承办。

诚挚邀请公司或研究机构在第二届中国模式识别与计算机视觉大会（PRCV2019）组织与承办视觉与学习方面的各种技术竞赛。视觉信息分析与理解在安防监控、考勤安检、遥感测绘、无人系统、网络信息过滤、文档分析、工业检测、医疗诊断等方面有着重要的应用价值。随着计算机视觉与模式识别理论、方法的研究进展，特别是以深度学习为代表的新一代人工智能理论在计算机视觉与模式识别研究中的巨大成功，学界已经提出了大量的、各具特色、面向各种应用的算法，极大地推动了相关技术的发展与应用落地。另一方面，无论是从学术研究的角度还是从产业

应用的角度，都迫切需要对这些算法的性能进行公平合理的考核与评测，评估其性能，发现各算法的优势与不足，并推动进一步的研究与应用。技术竞赛则是实现这一目的的重要手段。在 PRCV 及其前身 RACV 2016、CCCV 2017、PRCV 2018 的竞赛单元中，国内多家公司、研究所承办了 20 余项相关技术竞赛，共吸引了 1000 余支队伍报名参加，成功架起了学术界与工业界之间的沟通桥梁，有力推动了计算机视觉与模式识别学科及相关产业的发展。

有意愿组织或承办竞赛的公司或单位需要于 2019 年 5 月 15 日前将竞赛申请书以电子邮件的形式提交给竞赛单元主席。申请结果将会在 2019 年 5 月 31 日之前通知组织者申请是否被大会采纳。所有被采纳的竞赛将会在 PRCV 2019 网站上公布。大会将会为每个竞赛分配各自的网页链接，用于公布竞赛细则、报名方式和竞赛结果等信息。

竞赛申请书需包括以下信息：

1. 竞赛名称、申请人邮箱
2. 竞赛目的与意义
3. 竞赛组织方
4. 竞赛参与者要求
5. 报名方式
6. 主要时间节点
7. 竞赛数据使用方式
8. 任务设置
9. 结果的评价方式

10. 结果提交方式
11. 奖项设置和奖励方法
12. 专题研讨会组织
13. 知识产权归属
14. 参赛团队注册方式

申请书需要以 word 文档的形式提交给竞赛单元主席。

PRCV 2019 大会还将为竞赛承办方分别配专门的研讨会时间和地点（单个竞赛的研讨会总时

长不超过半小时，细则另行公布），并协助竞赛组织者对参会的竞赛获胜者颁发奖励证书。

如果竞赛组织申请人对上述信息有任何疑问，请及时联系竞赛单元主席。

- 桑农 (nsang@hust.edu.cn, 华中科技大学)
- 左旺孟 (wmzuo@hit.edu.cn, 哈尔滨工业大学)
- 尹翰林 (iverlon1987@126.com, 西北工业大学)

（责任编辑：毋立芳）

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、西安交通大学以及西安电子科技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。

PRCV2019 根据赞助额度设钻石、金牌、银牌和铜牌等赞助级别。赞助级别和享有的赞助权益，分列如下：

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，为了更好地保障赞助权益，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。

备注

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）

人脸属性转换进展概述

西安电子科技大学 邓成 魏坤

一、引言

近年来，大数据和计算机技术飞速发展，一些新型的技术应运而生。其中，人脸属性转换因其广泛的应用性而引发了持续关注。人脸图像属于较为精细的图像，其中包含着多种属性，比如头发颜色、眼镜、年龄、性别等等。如图 1 所示，人脸属性转换，属于图像合成的范畴，其希望在仅输入一张人脸图像的情况下，完成各种属性的转换，这无疑是一个巨大的挑战。面对海量数据中各种复杂的背景、表情和姿势，以及不配对的人脸图像，如何在转变属性的同时生成逼真的人脸图像，是一项非常重要并且具有挑战性的课题。

早期的图像合成 (Image Synthesis) 任务是基于卷积神经网络 (Convolutional Neural Network, CNN) 来完成的，生成的图像虽然有一些效果，但是太不真实。随着生成对抗网络 (Generative Adversarial Networks, GANs) [1] 的提出，越来越多的图像合成任务使用 GANs 来完成。训练过程中，生成器和判别器不

断对抗，使生成器生成的图像更加逼近真实图像数据的分布，从而使生成的图像更加真实，极大地提高了生成图像的视觉质量。短短几年之内，GANs 演化出一系列变种[2][3]，让 GANs 的训练更加稳定且能够应用于各种复杂的任务。与此同时，基于 GANs 的人脸属性转换方法也飞速发展，能够生成更加逼真的人脸，告别了以往方法中生成的鬼脸与伪影。

众所周知，大量且优质的数据是深度学习能够快速发展的前提条件。然而对于人脸属性转换而言，我们很难收集到大量的关于同一个人所有属性的图像，不仅需要大量的时间成本和人力成本，而且还有一些属性的照片收集起来并不现实，比如性别以及年龄等，这在研究初期严重限制了人脸属性转换的发展。随着 Cycle Generative Adversarial Networks (Cycle-GAN) [4] 的提出，为不配对数据的图像合成提供了新的方向，也为人脸属性转换提供了飞速发展的契机。



图 1 人脸属性转换

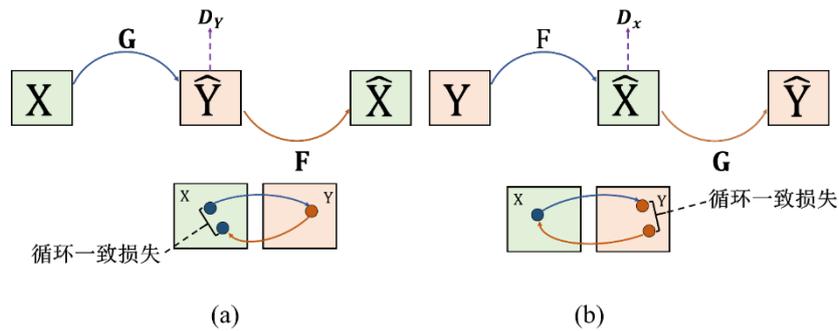


图2 域迁移示意图

对于人脸图像而言，各种属性之间相互纠缠耦合，导致在完成属性转换的同时经常会伴随其他无关属性也随之发生变化，这大大降低了人脸属性转换的有效性。为了解决这一问题，有以下两种途径：（1）通过人脸的解耦特征表示，将需要转换的属性与其他无关转换的属性分离，在转变相关属性的同时能够保证无关属性不变；（2）通过视觉注意力机制。因为与风格转换等任务需要图像全局发生改变不同，人脸图像属性转换通常只需要局部区域发生变化。引入视觉注意力机制能够判断人脸图像中哪些区域是需要改变的哪些是不需要改变的，这也能在一定程度上保证无关属性不发生变化。

二、人脸属性转换现状

1. 不配对人脸属性转换

对于同一个人，我们很难集齐（她）包含所有属性的照片，这不仅需要耗费大量的人力物力，而且有的属性图像是无法获得的，这极大地提高了人脸属性转换的难度。随着Cycle-GAN, Dual GAN [5] 和 Disco GAN [6] 的相继提出，不配对图像翻译得到了有效地解决。

对于人脸属性转换任务而言，输入图像和生成图像分别是两个域，输入图像可以看成是源域，而生成图像属于目标域。成对的图像翻译需要建立源域和目标域之间一一映射关系，而不成对的图像翻译不需要建立训练数据之间一对一的映射，就可以实现源域到目标域的迁

移，这大大降低了人脸属性转换任务对于数据集的要求。如图2所示，这种方法通过对源域图像进行两步变换：首先尝试将其映射到目标域，然后返回源域得到二次生成图像，从而消除了在目标域中图像配对的要求。使用生成器将图像再映射到目标域，并且通过生成器与判别器之间的不断对抗，能有效提高图像质量。

除此之外，还设计了循环一致损失。通过计算输入图像和二次生成图像之间像素级的差值，来迫使生成器生成更拟合于目标域的图像，同时致使生成器具有保持人脸细节特征的能力。

2. 多域属性转换

对于多种人脸属性转换任务来说，Cycle-GAN 并不是一个很好的选择。如果我们需要学习 n 种人脸属性的相互转换，Cycle-GAN 则需要训练 $\frac{n \times (n-1)}{2}$ 次才能完成我们的需求。显而易见，这样只能处理两个域转换的算法并不能我们的需求，既浪费时间且不灵活。在这种情况下，StarGAN [7] 和 Att-GAN [8] 等算法通过结合Cycle-GA和Conditional Generative Adversarial Networks (CGAN) [9] 解决了不配对图像多域属性转换问题。

如图3所示，通过在输入图像上添加目标类标信息来控制目标域，添加方式为将one-hot形式的类标与图像在特征图维度上拼接起来。通过编码器，可以将类标信息编码成和目标域有关的特征来引导图像的转换过程。

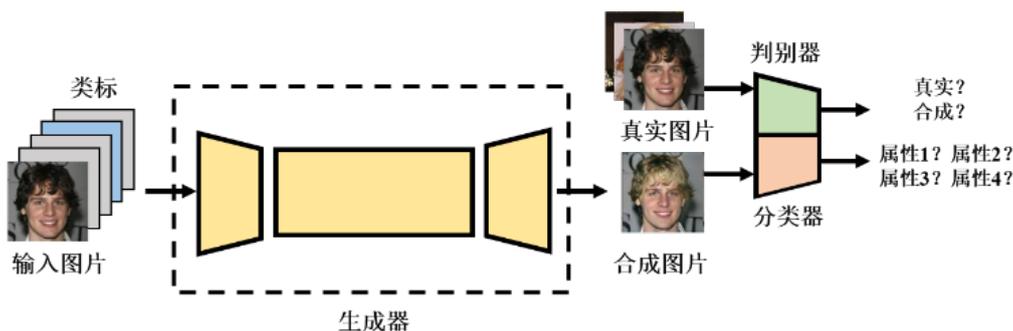


图3 多域属性转换模型

除此之外，在判别器的顶端加了一个多域属性分类器，能够提供属性分类损失，指导生成器生成带有目标属性的人脸图像。在判别器的优化过程中，用真实图像和原始类标信息来训练分类器。在生成器优化的过程中，用合成图像和目标类标信息计算属性分类损失来指导生成器。

为了保证生成器具有保持人脸细节的能力，还会再将合成图像与输入图像的原始类标信息输入到网络中二次生成重建图像，再计算与输入图像的差值来计算重建损失。除此之外，重建损失还能保持人脸的身份信息，避免了属性发生转换导致图像身份信息的丢失。

3. 解耦特征表示

人脸图像包含着多种属性，相互纠缠耦合在一起。在人脸属性转换的时候，经常会导致与转换属性无关的属性发生了变化，这大大影响了属性转换的效果。

为了能够将人脸各种属性解耦，DNA-GAN [10] 和 DDAE [11] 相继被提出。如图4所示，通过简单的几层编码器，将输入的人脸图像解耦成头发颜色、眼睛和年龄等一个个特征，每个特征分别被表示成一个特征维度，每个维度的不同取值表示各个特征的取值。这些算法通过改变目标属性特征的取值且保持其他属性特征取值不变来完成属性转换，有效地保证了无关属性稳定不变且使属性转换具有可控性和可解释性。

另外，这种方法还可以同时改变多种属性的取值，使网络可以只需要一步就完成单张图像多种属性的转换，而无需通过多步完成单张图像多种属性的转换，极大地提高了人脸属性转换的可控性和转换效率。

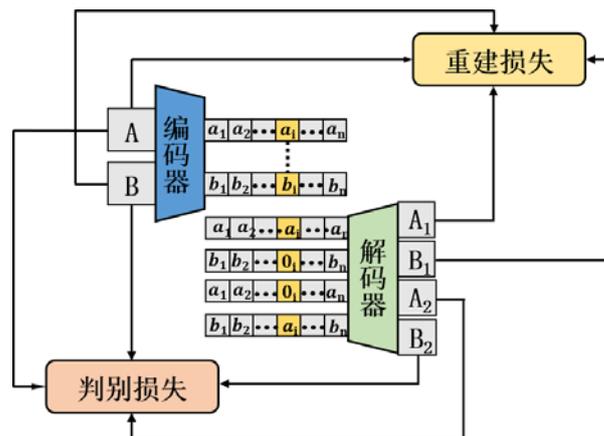


图4 解耦属性转换模型

三、未来发展方向

随着GAN的提出与发展，基于GAN的人脸属性转换成为计算机视觉领域最为活跃的研究方向之一。本文主要关注基于GAN的人脸属性转换及其相关应用，并分别从不对人脸转换，多域属性转换以及解耦特征表示进行深入讨论。

目前人脸属性转换已经取得一定的进展，但是由于模型参数巨大、数据集不完全，人脸属性转换依然面临着许多急需解决的问题。因此，我们认为未来发展方向包括快速的人脸属性转换、视频人脸属性转换以及少量图像生成。

（责任编辑：邓成）

参考文献:

- [1] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]// Proceedings of Advances in Neural Information Processing Systems. 2014: 2672-2680.
- [2] Arjovsky M, Chintala S, Bottou L. Wasserstein gan[J]. arXiv preprint arXiv:1701.07875, 2017.
- [3] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of wasserstein gans[C]// Proceedings of Advances in Neural Information Processing Systems. 2017: 5767-5777.
- [4] Zhu J Y, Park T, Isola P, et al. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks[C]// Proceedings of the IEEE Conference on International Conference on Computer Vision. 2017: 2242-2251.
- [5] Yi Z, Hao (Richard) Zhang, Tan P, et al. DualGAN: Unsupervised Dual Learning for Image-to-Image Translation[C]// Proceedings of the IEEE Conference on International Conference on Computer Vision. 2017: 2868-2876.
- [6] Kim T, Cha M, Kim H, et al. Learning to Discover Cross-Domain Relations with Generative Adversarial networks[C] // Proceedings of International Conference on Machine Learning. 2017: 1857-1865.
- [7] Choi Y, Choi M, Kim M, et al. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation[J]. arXiv preprint arXiv:1711.09020, 2017.
- [8] He Z, Zuo W, Kan M, et al. Arbitrary Facial Attribute Editing: Only Change What You Want[J]. arXiv preprint arXiv:1711.10678, 2017.
- [9] Mirza M, Osindero S. Conditional generative adversarial nets[J]. arXiv preprint arXiv:1411.1784, 2014.
- [10] Xiao T, Hong J, Ma J. DNA-GAN: Learning Disentangled Representations from Multi-Attribute Images[J]. arXiv preprint arXiv:1711.05415, 2017.
- [11] Liu Y, Wei F, Shao J, et al. Exploring disentangled feature representation beyond face identification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 2080-2089.



邓成

西安电子科技大学教授，博士生导师。
主要研究方向为机器学习，计算机视觉
和模式识别。

Email: chdeng@mail.xidian.edu.cn



魏坤

西安电子科技大学博士研究生。主要研究方向为机器学习
和人脸属性转换。Email:

weikunsk@gmail.com

基于选择性蒸馏的自然环境 低分辨率人脸识别

中国科学院信息工程研究所 赵胜伟 葛仕明

深度学习和大数据的出现推动了人脸识别的飞速发展。在知名人脸数据集 LFW (Labeled Faces in the Wild) 上, 现有模型已取得超过 99.5% 的识别精度。然而, 对于诸如视频监控、无人机监控、智能驾驶等自然环境下获取的低分辨率人脸, 现有人脸识别模型效果并不理想。这一方面是因为低分辨率人脸存在严重的信息缺失, 另一方面是因为自然环境下的人脸与模型训练中的人脸存在巨大的分布差异。此外, 这些场景通常需要在计算资源受限(如嵌入式)的情况下达到高速处理的要求, 这给自然环境下的低分辨率人脸识别带来了更大的挑战。

近年来, 许多研究人员尝试使用高分辨率人脸知识辅助构建低分辨率人脸识别模型, 提出的方法大致可以分为两类: 超分辨率方法和嵌入方法。超分辨率方法首先对低分辨率人脸进行超分辨率重建, 再对重建后的高分辨率人脸进行识别。由于引入了额外的计算, 这类方法的处理速度普遍较慢。嵌入方法将高分辨率人脸和低分辨率人脸映射到同一特征子空间, 通过求解一定的约束条件, 将高分辨率人脸中的知识迁移到低分辨率人脸中。这类方法成功的一个关键在于: 如何有效将从高分辨率人脸中获取的最富信息知识迁移到低分辨率人脸识别模型中?

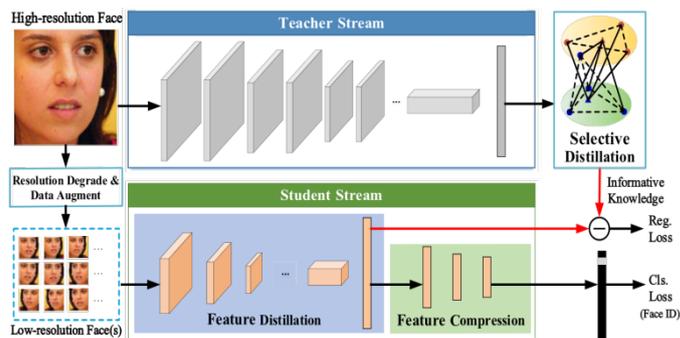


图 1 选择性蒸馏方法框图

针对这一问题, 我们提出了基于选择性蒸馏的低分辨率人脸识别方法。该方法通过图优化从大量高分辨率人脸知识中选择性地蒸馏出最富信息的信息, 再通过教师-学生框架联合多任务训练, 将选择后的知识迁移到低分辨率人脸识别模型中, 从而进一步提升低分辨率人脸识别的精度。如图 1 所示, 该方法一共有两个流, 分别是教师流和学生流。其中教师流以高分辨率人脸为输入, 学生流以低分辨率人脸为输入。学生流一方面从选择后的高分辨率人脸知识中模仿学习, 另一方面从低分辨率人脸中自我学习。此外, 学生流中的特征压缩模块极大地降低了人脸特征的维度, 从而加快了大规模人脸检索的速度。

以上工作已发表于国际期刊 IEEE TIP2019。

(责任编辑: 崔海楠)



赵胜伟

中国科学院信息工程研究所硕士生, 主要研究方向为计算机视觉、深度学习、人工智能安全等。

Email: zhaoshengwei@iie.ac.cn



葛仕明

中国科学院信息工程研究所博士生导师, 主要研究方向为计算机视觉、深度学习、人工智能安全等。

Email: geshiming@iie.ac.cn

视觉代表模式稀疏选择与发现

重庆大学 王洪星

模式与规律发现一直是人类认知物理世界的重要途径。发现视觉数据背后潜在的代表模式便显得异常重要。代表模式承载了原始数据的重要信息，却具有相对更小的体量，便于存储、传输、分析、处理和应用，其对视觉数据摘要与检索等问题具有重要意义。

然而，如何选择视觉样例使其尽可能代表原始数据及其结构信息仍然是一个开放性且极具挑战性的问题。中心点聚类是模式发现的代表方法，找到类簇中心点作为代表模式。此类方法通常需要预设代表模式的个数和聚类中心点。因此，为了生成不同数量的代表模式，需要使用者每一次都重新执行聚类算法。近年来，稀疏字典选择被应用于代表模式的选择与发现。此类方法利用稀疏表示原理从数据中选取一组基，最大化其对全体数据的线性重构能力。这组被选中的基被称为数据字典，对应于数据的代表模式。该类方法的一个显著特点是无需预设代表模式的个数，很容易根据使用需要生成不同数量的代表模式。图1展示了该类方法的主要过程：将待处理的各视觉数据（例如图像）描述为特征向量，在特征空间对其进行稀疏重构，对重构结果进行分析得到代表模式，从而完成对视觉数据的整体概要。

研究显示，稀疏字典选择发现的代表模式在特征空间中处于全体数据的凸包上，揭示了数据的最大外延。这一特性也致使其欠缺对数据局部

特征空间结构的表达能力，从而难以应对具有复杂特征分布的数据。另外，凸包点极易因噪声扰动而变化，影响了代表模式的稳定性。针对上述问题，我们提出了几种代表模式稀疏选择方法。

(1) 设计结构化稀疏选择方法，在稀疏条件下考虑样例差异性与局部敏感性，使选出的代表模式能充分表达数据局部结构及多样性分布。

(2) 量化视觉数据的备选先验信息，将其引入稀疏选择模型，从而保证包含重要视觉信息的数据不被漏选，同时避免噪声数据被误选。

(3) 从多视角描述视觉数据的特征，设计多视角稀疏选择方法，实现视觉模式在不同视角下的信息传递与融合，使选出的视觉模式体现多视角的代表性，同时增强代表模式的稳定性。

(4) 我们利用超球面映射，将全体数据重构简化为数据中心重构，提出超球面稀疏选择方法，实现快速鲁棒多样的代表模式选择与发现。

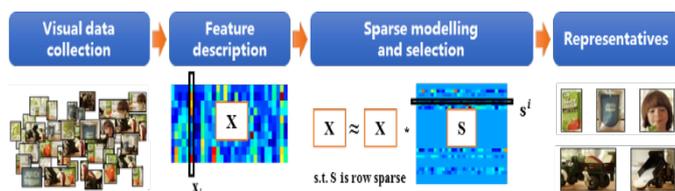


图1 代表模式的稀疏选择过程

该研究相关成果发表在 IEEE CVPR、IEEE TIP、Pattern Recognition、IEEE SPL。

(责任编辑：任传贤)



王洪星

CCF 计算机视觉专委会委员，重庆大学副教授，博士生导师，主要研究方向为计算机视觉与模式识别。

Email: ihxwang@cqu.edu.cn

上海交通大学倪冰冰教授访谈

2019年1月3日，专委秘书处采访了上海交通大学青年千人倪冰冰教授。下面是采访实录。

倪老师，您好！您在计算机视觉、机器学习与多媒体计算领域做出了很多高水平研究工作，获批中组部第十一批青年千人专家，发表了一系列高水平的论文。您能跟我们详细介绍一下您的研究经历，分享一下获得成功的经验，以及您取得这些成就的动力么？

大家好。我是2010年在新加坡国立大学获得博士学位，2015年底回国工作的，目前在上海交通大学电子工程系任教。从博士研究生阶段开始，我的主要研究内容都是围绕智能视频媒体分析这个主题，取得了一些进展。我觉得对于计算机视觉研究来说，想要取得一点成绩，创新性是比较重要的因素。我认为创新有三个层次，第一个层次创新是能够发掘一个前人没有发现的新的有价值的问题，比如过去我们大多数研究者都在图像与视频的识别方面开展工作，而近两年来，图像与视频的生成这个全新的主题变得越来越popular，这个就是好的创新的例子。第二个层次是对已有的问题，提出颠覆性的、非常有效的解决方法，比如深度卷积神经网络的提出一下子把大规模图像识别的问题从较低的准确率提高到超过人的识别水准，对于整个领域起到了极大的推动作用。当然还有第三种创新是对已有的算法、方法做一些增量性的改动，提高性能，这样的创新，对于研究者来说，风险比较小，但是影响力也是有限的。我认为对于一个研究者来说，应该更加关注于前两种的高层次创新。诚然，这需要更大的探索耐心与毅力，但我相信回报也一定是很厚重的。

您“十年来专注智能视频分析研究”，能跟大家分析一下这方面的研究难点所在，以及近几

年在这方面所存在的市场空间么？

我从博士阶段开始，主要关注的研究方向就是视频内容的智能分析。主要是对图像视频中的人、车、物的静态属性（例如人脸、物体检测定位、人脸识别、物体细粒度标签识别、人群密度估计等），以及动态属性（个体行为识别，群体互动行为识别等）进行智能分析。视频理解有很多挑战。首先因为摄像机角度、复杂背景光照、复杂物体运动等原因，我们在识别过程中会遇到目标尺度变化大，目标的外观变化大等问题，因此如何提取、构建具有较强分辨力的、又能兼容目标尺度、外观等变化的统一视觉表征，是本领域的最重要的难点。第二，因为视频是运动变化的，因此如何对视频内的各种目标进行跨时间空间的关联，也是比较关键，在很多大型视频监控问题上，得到同一目标的时空路径，对于可能发生的事件推理，是相当关键的。第三是目前视频理解的困难之一在于存在有标注的数据量还是比较少，而视频内容的变化又是巨大的，使得很多有监督模型的训练比较困难，所以很多视频理解的算法，特别是复杂的行为识别算法，目前工业界还没有广泛的落地应用。最后是计算的问题，因为视频数据体量很大，目前通用的深度学习模型也比较庞大，如何压缩这些模型，使得效率与精度兼备，这也是目前以及将来很长一段时间在视频领域需要解决的问题。

除了这些传统的智能视频分析的研究，目前我的团队更加关注与聚焦于互动创意视频媒体领域。近期我们在智能的视频生成方面做了一系列的工作，比如基于空间约束的人物运动视频生成、人脸的实时3D表情迁移和从sketch到动画的自动生成等。我们认为新媒体的重大需求就是内容的智能产生，以及内容的个性化产生，而目

前人工智能在视频方向的发展，恰恰能够满足这两个方面的需求。我们可以看到未来有更多的媒体内容是计算机程序自动编辑产生的，而且对于不能的观众，它能根据观众的喜好产生不同的内容。我们团队最近的一个非常有趣的工作是根据 NBA 视频自动产生体育解说，目前效果非常好，几乎到了以假乱真的地步，在很多国际主流科技媒体上，也被广泛报道。当然，这些技术的难点在于生成内容的解空间非常巨大，因此如何进行有效的空间约束和空间分解，可能是提高内容生成质量的关键所在，我们也在这个方向进行努力探索。

您多次参加行为识别、检测竞赛，并获得了 2012 年国际模式识别旗舰会议 ICPR 行为识别竞赛第一名、2014 年国际计算机视觉旗舰会议 ECCV 行为识别竞赛第二名、2015 年 THUMOS 2015 行为检测国际竞赛第一名等，能跟大家分享一下您参加这些竞赛的经历和经验么？

我觉得参加竞赛，与学术研究有一定的相通点，也有很大的不同。比如相通点是都是要对于算法进行一定的创新设计，目前深度学习的时代，很多开源的工具、算法，对于竞赛的参赛者来说，起点和基础都是类似的，这个时候，如果能对模型进行一定的创新，更加能够适应场景的问题，最后一定会脱颖而出。异同点在于对于竞赛来说，最后的工程细节也非常重要，比如你用模型融合的方法提高性能，那么如何融合，如何给合适的融合系数，在何阶段进行模型融合，是否应该对于子问题进行建模，数据应该如何前处理以及后处理，这些细节都能够极大地影响到最后的成绩。对于参赛者来说，事先如何规划好团队的实验方案，如何在最短最可控的时间内进行尽可能多的模型以及方法尝试，都是至关重要的。

您于 2010-2015 年在美国伊利诺伊大学香槟分校新加坡高等研究院担任研究科学家，请问能介绍一下您这段时间的工作经历么？您承担的主要工作是什么，有哪些经历是您难忘的？在这

段期间的工作经历中，最大的获益是什么呢？

在高等研究院的 6 年研究经历对我来说是非常重要的。它是我博士毕业后进行独立科研的重要时期。在这期间，我的很多科研能力得到了训练，例如独立进行项目申请书的撰写，独立进行研究生研究课题的制定与指导，独立进行科研项目的进度管理与团队管理。我在高等研究院时期的研究方向主要是跨模态的视频理解，主要是融合可见光、深度、红外等视频模态进行目标识别与行为识别。当时我们研究组的主管教授是来自香槟分校的著名信息论学者 PIERRE MOULIN 教授，几年的共事，他影响我最深的是他对于研究工作的每处细节都是事必躬亲，我们可以时常看到他在办公室废寝忘食地进行公式推导，上课教案准备，以及帮助学生一字一句地进行科研论文的修改（实际上他当时已经是 UIUC 的荣誉教授，是国际信息论的权威）。许多年后我自己在带团队的时候，也努力做到在科研的每个阶段一定要和学生深入互动，不管是 idea 准备、实验设计以及论文撰写，都尽量和学生一起，做到最严谨、最完美的状态。科研无小事，只有对细节的全面把握，才能做出最第一流的研究成果。这几年，当然最难忘的还有第一次指导自己的研究生发表他的第一篇顶会论文，第一次带领团队夺得国际算法大赛冠军。作为导师，为学生的科研事业打开第一扇窗户，那是最有成就感的。

在美国伊利诺伊大学香槟分校新加坡高等研究院担任研究科学家 6 年之后，您为什么选择返回祖国高校工作呢？又是什么机缘让您去了上海交通大学？

回国的原因是几方面的，一方面其实是考虑个人的学术生涯的发展，国内高校在人工智能、计算机视觉方面的发展如火如荼，有一定知名度年轻学者比较容易在国内高校快速建立自己的团队，可以加速产生更多更有影响力的学术成果。这样的大环境，国外是没有的。另外一方面，我本科就是毕业于上海交通大学电子工程系，当时

我们学院在计算机视觉方面，科研实力相对比较薄弱，我们学院的领导很希望加强这方面的学科以及科研建设，对我回国任教提供了强有力的帮助与支持，使我深受感动，所以更加坚定了我回国任教，建立自己的科研团队，为母校做贡献的信心。

能分析一下上海交通大学的研究氛围，和您之前在美国、新加坡等地研究氛围的异同么？

我个人感觉中国、美国、新加坡的研究氛围还是有一定的区别的，或者说各自有各自的偏重与特色。比如我在美国谷歌公司时候，虽然我们公司是公司，应该是以盈利为导向的，但是我们可以很明显感受到，谷歌公司内部是很鼓励天马行空的科学研究的，特别是在算法研究部门，一些很好的、很超前的科研想法，只要是跟产品线有一定的关联度，公司都会非常支持你去创新。比如我当时在谷歌的 YOUTUBE 视频分析部门，提出了几个当时非常新颖的研究 idea，例如通过视频内容推荐相应的背景音乐，通过海量的学习自动产生个性化的、智能合成的视频和音乐等，得到了团队领导的大力支持，有几个项目和团队同事一起努力也最终获得了 YOUTUBE 产品的上线应用，还是非常有成就感的。谷歌也支持工程师们发表高水平学术论文，参加国际顶级会议。

新加坡的研究所或者高校，比较偏重研究和产业结合。比如我们高等研究院，从管理层，到每个课题组的 PI，都非常鼓励我们走出去和企业交流，从中挖掘一些可以产业落地的研究想法。我们当时课题组，与新加坡的各大医院广泛合作，将视频行为理解的技术运用到医院的日常流程管理、老年病人复健训练、以及病人防跌倒的实际项目中。

回国以后我感觉到高校的研究氛围是非常自由和友好的，可以根据自己的科研兴趣，制定相应的科研方向，如果团队足够大的话，也可以同时探索几个方向。当然，很多时候我们也关注

国家的重大需求以及产业的需求，比如我们团队目前主要的研究力量放在互动创意新媒体的方向，和国内好几个新媒体的企业（短视频、直播）保持着非常良好的合作关系。

能介绍一下您现在的研究团队，以及您如何管理这个研究团队的么？

我目前在上交大的团队大概有 40 人的规模，包括博士生、硕士生和本科生。团队规模算是比较大的，我们的日常科研进展管理通常采用科研小组的形式，即每一位资深的博士生负责一个方向，每个方向会有若干硕士生和提前进组的本科实习生，科研进展讨论会也基于小组的形式进行，至少保证每位每周一次与负责博士以及我进行进展讨论。另外我们每周会有一次论文学习分享会，面向所有同学，主要是将近期的顶会顶刊论文进行分享汇报与讨论，在讨论过程中逐渐形成新的观点与创新点。目前我们几个科研小组分别是互动创意新媒体内容生成/智能艺术、深度学习的可解释性、智能医疗影像等。我感觉，对于一个团队的形成与磨合，单单靠这些 routine 式的组织手段通常也是不够的。关键还是要让学生不断地感受到科研给他们带来的成就感，比如能发表顶级的论文，能经常有比赛的获奖，这些成果也会无形中更加坚定同学的科研信念，同时每位同学都觉得科研有奔头，无形中也增加了团队的凝聚力与战斗力。当然我们团队是很友爱的团队，老师学生也经常一起玩，比如一起踢球，一起观星，一起郊游、聚餐，总之活动非常丰富。

从微信上看，您有一个非常可爱的女儿。对于研究者，事业和家庭都要兼顾是有一定难度的，能否跟大家介绍一下您是如何处理事业和家庭之间关系的？

科研工作，虽然工作时间上比较灵活。但是几乎也是等同于每天都是上班状态。特别是有的时候可能会连续几天出差。所以家里还是太太照顾的比较多，作为丈夫和父亲，还是比较愧疚的。

最近女儿在准备幼升小的考试，我也尽量抽出时间对女儿进行一些辅导。其实从与女儿的互动中，能够感受到巨大的快乐，自己在也不断成长。特别希望未来有更多的时间和家人一起共享。

我爱人对我的工作是非常支持的，她本人也



倪冰冰

上海交通大学电子系教授，博士生导师。中组部第十一批青年千人专家。2010-2015 年于美国伊利诺伊大学香槟分校新加坡高等研究院 (University of Illinois at

Urbana-Champaign, Advanced Digital Science Center Singapore)担任研究科学家。2005 年在上海交通大学电子工程系获学士学位；2011 年在新加坡国立大学 (National University of Singapore) 电气与计算机工程系获博士学位。博士期间，先后在微软亚洲研究院和谷歌公司美国总部工作，担任算法研究员。主要研究方向为计算机视觉、机器学习与多媒体计算，专长人脸识别，视频理解，媒体生成，智能医疗，以及智能艺术。

发表论文 100 余篇，其中 IEEE T-PAMI、IJCV 等 SCI 期刊源论文 30 余篇，包括 IEEE/ACM 汇刊论文 20 余篇。CVPR、ICCV、NIPS 等中国计算机学会推荐 A 类会议论

是人工智能的超级粉丝。这里可以透露一下，她目前也在从事人工智能的教育事业（与几位志同道合的朋友创办了 AI300 学院这个在线人工智能教育平台）。目前在人工智能职业教育，人工智能高中教育，以及少儿人工智能启蒙教育方面已经初具规模。

文 40 余篇。所发表论文的 Google Scholar 引用次数为 3000 余次，H-因子 28。获得美国专利授权 2 项。获 ICPR2012-HARL、CVPR2015-THUMOS 行为检测竞赛第一名、ECCV2014-ChaLearn 行为识别竞赛第二名。2011 年度泛太平洋多媒体会议 (PCM2011) 最佳论文奖。2017 阿里巴巴天池国际 AI 医疗竞赛第一名 (1/2887)。2018 年 BOT 国际智能零售算法大赛第一名 (1/480)。十年来专注智能视频分析研究，近两年来进一步聚焦跨时空复杂事件识别、视频精细语义识别、视频预测生成研究方向，在视频行为分析的精细性、跨时空性、可预测性方面，率先提出一系列解决方案，突破了传统算法语义识别粒度粗、特征时空关联度差的瓶颈。倪博士被 IBM 公司 T. J. Watson 研发总部评选为 2010 年全球多媒体与信号处理领域十大新锐之一 (Emerging Leaders in Multimedia and Signal Processing)。担任国际计算机视觉顶级会议 ICCV2019 领域主席。

(责任编辑：余焯 黄岩 张汗灵)

委员好消息

✪ 2018年11月16日，中国自动化学会公布了2018年会士名单，包括了22位自动化、智能化和信息化科学与技术领域的卓越科技工作者。CCF-CV专委会委员、西安交通大学薛建儒教授入选。

✪ 2018年11月21日获悉，CCF-CV专委会秘书长、中科院自动化所王亮研究员和CCF-CV专委会委员、上海交通大学杨小康教授入选 IEEE Fellow，本年度共有40位国内高校学者入选。

✪ 2018年11月23-26日，首届中国模式识别与计算机视觉学术会议（PRCV 2018）于2018年11月23-26日在广州举行，CCF-CV专委会委员、浙江工业大学陈胜勇教授团队的论文 *Oscillation Detection and Parameter-Adaptive Hedge Algorithm for Real-Time Visual Tracking* 获 PRCV2018 最佳论文奖，CCF-CV 专委会委员、南京信息工程大学刘青山教授（常务委员）和袁晓彤教授团队的论文 *Integrating Convolutional Neural Network and Gated Recurrent Unit for Hyperspectral Images Spectral-Spatial Classification* 获最佳学生论文奖。

✪ 2018年12月12日，科技部公示了拟作为2018年创新人才推进计划入选对象名单，共306名中青年科技创新领军人才、50个重点领域创新团队、200名科技创新创业人才和32个创新人才培养示范基地。CCF-CV专委会委员、北京大学彭宇新教授入选中青年科技创新领军人才。

✪ 2018年12月20日，中国电子学会网站公示了2018年度优秀博士学位论文评选结果，本年度从110篇参评论文中共评选出22篇获奖论文。CCF-CV专委会委员、上海交通大学杨小康教授指导马超完成的博士论文《Shallow and Deep

Learning for Robust Online Object Tracking》和 CCF-CV 专委会委员、西安电子科技大学高新波教授指导彭春蕾完成的博士论文《基于概率图模型的异质人脸图像合成与识别》入选。

✪ 2018年12月24日，中国图象图形学学会公示了石青云女科学家奖-青年奖评审结果，5位获奖人员中有4位是CCF-CV专委会委员，她们是中科院计算所阚美娜副研究员、北京大学刘家瑛副教授、国防科技大学刘丽副教授、西北工业大学孙瑾秋副教授。中国图象图形学学会石青云女科学家奖-青年奖设立于2018年，旨在鼓励广大图像图形学科技工作者继承和发扬老一辈的科学精神，表彰在图像图形学研究事业发展中做出突出贡献的女性科技工作者，激发广大女科技工作者的创新创造热情。

2019年1月8日，2018年度国家科学技术奖励大会在京举行。CCF-CV专委会委员、清华大学冯建江副教授和清华大学深圳研究生院郭振华副研究员参与完成的“大人指掌纹高精度识别技术及应用”和 CCF-CV 专委会委员、东南大学郑文明教授参与完成的“心理生理信息感知关键技术及应用”获得国家技术发明二等奖。2018年度国家科学技术奖共评选出278个项目和7名科技专家。其中，国家最高科学技术奖2人；国家自然科学奖38项，其中一等奖1项、二等奖37项；国家技术发明奖67项，其中一等奖4项、二等奖63项；国家科学技术进步奖173项，其中特等奖2项、一等奖23项、二等奖148项；授予5名外籍专家中华人民共和国国际科学技术合作奖。

（责任编辑：刘海波）

图像视频生成技术开源代码

北京电子科技学院 于明学 金鑫

图像视频生成技术，旨在让模型通过训练，能生成新的具有人类感官的连续图像视频。该技术通常依据深度学习中的生成模型，该模型从数据集中学习到图像以及图像与图像之间的特质，如图像中的形状、颜色、纹理和图像之间的动作的连续性等，并应用这些特质以生成新的连续图像。图像视频生成技术一般分为两种，一种是无条件生成，一种是有条件生成。无条件生成是指生成模型的输入只有一个先验的噪音，而有条件的生成是指生成模型的输入带有额外的标签信息，如类别信息等。下图为图像视频生成示意图。



图 1 图像视频生成示意图

本文着重介绍几个图像视频生成技术的开源代码，包括图像视频生成的数据集和算法，算法主要包括 VAE 和 GAN，以及他们的具体应用。

1. 图像生成技术数据集 cityscapes

介绍: cityscapes 数据集关注于城市街道场景的语义理解，它的 label 图带有语义类别信息和实例信息，常作为有条件图像生成模型的数据集。数据集包含 30 个类别，有 5000 张详细标注的精细图片，以及 20000 张粗略标记的粗糙图片。

主页: <https://www.cityscapes-dataset.com>

2. 图像生成技术数据集 CelebA

介绍: CelebA 数据集是一个大型的人脸属性数据集，拥有超过 20 万张名人图像，每张图像都有 40 个属性标注。该数据集包含 10177 个身份标注信息。常作为无条件生成的数据集使用。

主页: <https://www.cityscapes-dataset.com>

3. VAE (变分自动编码器)

论文: Kingma, Diederik P, Welling, *et al.* Auto-Encoding Variational Bayes[C]. In The 2nd International Conference on Learning Representations, 2014.

工作: VAE 也是 AE 的一种，原理是构建一个隐变量 Z 生成目标数据 X 的模型。VAE 假设了 Z

服从某些常见的分布，然后希望训练一个模型 $X=g(Z)$ ，这个模型能够将原来的概率分布映射到训练集的概率分布。VAE 与 AE 的不同之处在于 VAE 的编码结果和额外的噪音作用后，输入到解码器中。下图为 AE 与 VAE 的示意图。

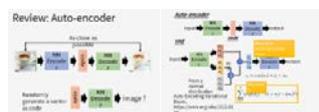


图 2 AE 与 VAE 示意图

代码: <https://github.com/bojone/vae>

4. GAN (生成对抗网络)

论文: Goodfellow I, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial networks[C]. Advances In Neural Information Processing Systems, 2014:2672-2680.

工作: GAN 主要包括生成器和判别器两个部分，生成器主要用来学习真实图像分布，从而让自身生成的图像更加真实，以骗过判别器。判别器则需要对接收的图片进行真假判别。在整个过程中，生成器努力地让生成的图像更加真实，而判别器则努力地去识别出图像的真假，这个过程相当于一个二人博弈，随着时间的推移，生成器和判别器在不断地进行对抗，最终两个网络达到了一个动态均衡：生成器生成的图像接近于真实图像分布，而判别器识别不出真假图像，对于给定图像的预测为真的概率基本接近 0.5（相当于随机猜测类别）。下图为 GAN 原理示意图：

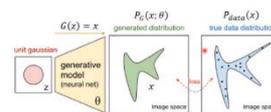


图 3 GAN 原理示意图

4.1. DCGAN(深度卷积生成对抗网络)

论文: Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[C]. In The 2nd International Conference on Learning Representations, 2016.

工作: DCGAN 是 GAN 的一个变体。其生成模型和判别模型都运用了深度卷积神经网络。DCGAN 将 GAN 与卷积网络结合起来，以解决 GAN

训练不稳定的问题。DCGAN 提出了一系列有助于 GAN 稳定训练的方法，如取消全连接层，批量归一化，以及采用 Leaky ReLU 激活函数等。下图为 DCGAN 模型示意图：

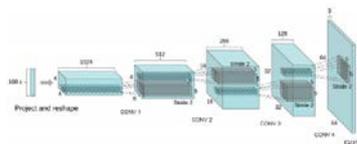


图 4 DCGAN 模型示意图

代码：https://github.com/Newmu/dcgan_code

4.2. Image-to-Image Translation with Conditional Adversarial Networks

论文：Isola P, Zhu J Y, Zhou T, *et al.* Image-to-image translation with conditional adversarial networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017:5967-5976.

工作：该模型是一个具有条件限制的 GAN 网络，通过简单的添加额外输入数据，来构建我们所需要的条件化的生成器和判别器。该模型可以用于学习多模态模型，通过提供了图像标记应用程序的初步示例，演示了该方法如何生成不属于训练标签的描述性标签。同时，该网络生成器还使用了 unet 网络，判别器借鉴了马尔可夫随机场的思想，提出了 PatchGAN，进一步提升了生成精度。下图为模型示意图：

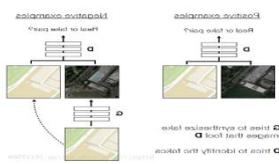


图 5 Image2Image 模型示意图

代码：<https://github.com/williamFalcon/pix2pix-keras>

4.3. High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs

论文：Wang T C, Liu M Y, Zhu J Y, *et al.* High-resolution

image synthesis and semantic manipulation with conditional GANs[J]. IEEE Conference on Computer Vision and Pattern Recognition, 2018:8798-8807.

工作：该论文提出了一种利用条件生成对抗网络（conditional GANs）来合成高分辨率、照片级真实的图像的新方法。采用多个生成器以及判别器，由 label 图生成了更加高清的真实图像 [8]，图像像素可达 2048×1024 。该网络除 label 图信息之外，还添加了 instance 图信息，进一步提升了不同物体轮廓细节信息。下图为模型示意图

代码：<https://github.com/NVIDIA/pix2pixHD>

4.4 Vid2Vid

论文：Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, *et al.* Video-to-Video Synthesis[C], Advances In Neural Information Processing Systems, 2018:1152-1164.

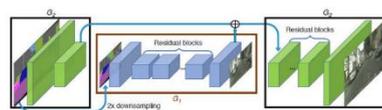


图 6 pix2pixHD 模型示意图

工作：该论文解决了生成的视频很难保证前后帧的一致性，容易出现抖动。Vid2Vid 建立在 pix2pixHD 基础之上，加入时序约束。因此可以实现高分辨率视频生成。同时该模型，生成器和判别器都加入了光流信息，对前景和背景分别建模。下图为 Vid2Vid 示意图：



图 7 Vid2Vid 示意图

代码：<https://github.com/NVIDIA/vid2vid>

（责任编辑：沈沛意）



于明学

北京电子科技学院在读研究生，研究方向为图像视频生成。
Email:349588189@qq.com



金鑫

现于北京电子科技学院从事教学与科研工作，担任北京电子科技学院可视计算与安全实验室负责人。研究方向为计算机视觉、虚拟/增强现实、人工智能安全。
个人主页：www.jinxin.me

三维目标与场景数据集

牛津大学 胡庆拥

国防科技大学 郭裕兰

三维场景理解是计算机视觉领域的重要研究方向之一。作为三维场景理解的主要内容，三维目标检测、三维目标识别和三维语义分割在现实生活中有着重要应用。例如，三维目标检测是无人驾驶技术的一个重要模块，对驾驶行为的决策起着重要的作用。随着深度学习的发展，以及数据库建设的不断完善，深度学习已经成为三维场景理解的主流方法。

本文针对三维目标检测、三维目标识别及三维语义分割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 CAD、RGB-D 及 Point Cloud 数据集三大类。

一、CAD 数据集

1. ShapeNet 数据集

数据集地址：<https://www.shapenet.org>

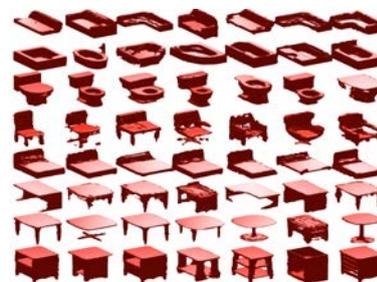
ShapeNet 是由斯坦福大学、普林斯顿大学以及芝加哥丰田技术研究中心于 2015 年发布的大型三维形状数据集。该数据集包含了一系列来自不同语义类别的三维模型，并根据 WordNet 方法进行归类。该数据集包含大约 3135 类目标，总共包含约 220000 个三维模型。其中，ShapeNetCore 和 ShapeNetSem 是 ShapeNet 公开发布的两个子集，其中 ShapeNetCore 包含了 55 类常见目标以及 51300 个不同的三维模型。相比于 ShapeNetCore，ShapeNetSem 的标注更加丰富，尺寸也更小。ShapeNetSem 由大约 270 类目标的 12000 三维模型组成，除了人工确认的类别标签以外，ShapeNetSem 中的模型还标注了现实中的维度。



2. ModelNet 数据集

数据集地址：<http://3dshapenets.cs.princeton.edu>

ModelNet 由普林斯顿大学、港中文以及麻省理工学院于 2015 年发布，是一个大型的三维 CAD 模型的数据集。该数据集包含 660 类目标（如椅子、沙发等）的 15128 个三维 CAD 模型。该数据集常用于三维目标识别及分类的测试与评估。其中，ModelNet10 与 ModelNet40 是 ModelNet 最常用的两个子集，分别包含 10 类和 40 类常见的目标。



二、RGB-D 数据集

3. NYU-D 数据集

数据集地址：http://cs.nyu.edu/~silberman/datasets/nyu_depth_v2.html

NYU-D 数据集是由纽约大学，伊利诺伊大学以及微软公司团队于 2011 年首次发布（2012 年发布 NYU-D v2），旨在通过提供更多的三维数据，进一步推动三维场景（包括但不限于场景分割、识别等任务）的发展。该数据集包含了带有详细标注的从 464 个室内场景中获取的 1449 张 RGB-D 图像，并且图像中的每个物体都标有其具体类别和实例级的语义标签。

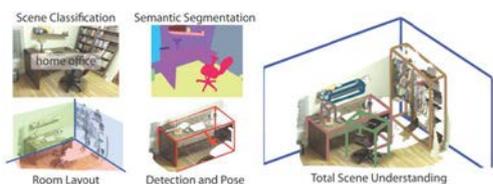


4. SUN RGB-D 数据集

数据集地址：<http://rgbd.cs.princeton.edu>

SUN RGB-D 数据集由普林斯顿大学视觉与机器人团队在 2015 年发布，旨在解决三维场景

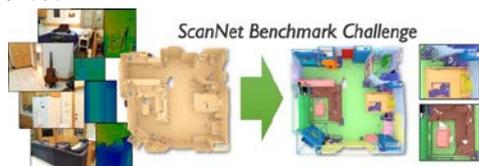
理解问题，可用于场景分类、语义分割、目标检测等六种主流任务的测试与评估。该数据集包含 10335 幅 RGB-D 图像，47 类场景类别以及大约 800 类目标。此外，该数据集中的数据还有大量的标注，包含 146617 个二维多边形以及 58657 个具有精确姿态标注的三维目标框，还包括三维的房间布局以及场景类别。此外，由于该数据集中的数据是通过四种不同的传感器（RealSense、Xtion、Kinect v1、Kinect v2）获取的，该数据集还可用于研究算法在不同传感器数据之间的泛化性能。



5. ScanNet 数据集

数据集地址：<http://www.scan-net.org>

ScanNet 数据集由斯坦福大学、普林斯顿大学以及慕尼黑工业大学团队于 2017 年首次发布（2018 年发布 ScanNet v2），旨在通过提供更多真实的三维数据，进一步推动三维场景理解的发展。该数据集包含了带有大量标注的从 1513 个场景中获取的 250 万幅 RGB-D 图像（从 707 个独立的空间中采集得到）。此外，该数据集还提供了对相机姿态、估计的校准参数、三维表面重建、纹理网格、大量的对象级语义分割和对齐的 CAD 模型的标注。ScanNet 中的数据来自对一系列小型空间（浴室、壁橱、杂物间）和大型（公寓、教室、图书馆）空间的扫描，且包含了各种实例级的语义标签。



6. Matterport3D 数据集

数据集地址：<https://niessner.github.io/Matterport>

Matterport3D 数据集是由普林斯顿大学，斯坦福大学以及德国慕尼黑工业大学团队于 2017 年发布，旨在为室内场景提供多视角、高质量的

RGB-D 图像，可用于语义分割、场景重构和关键点匹配等五种主流任务的测试与评估。该数据集一共包含了 90 个建筑物场景的 194400 张 RGB-D 图像和 10800 张全景图像。与传统的三维数据集不同，该数据还提供了每个视角的深度和颜色 360 度全景图。



三、Point Cloud 数据集

7. Sydney Urban Objects 数据集

数据集地址：<http://www.acfr.usyd.edu.au/papers>

</SydneyUrbanObjectsDataset.shtml>

Sydney Urban Objects 数据集是由悉尼大学团队于 2013 年发布的，旨在进一步提升目前三维方法在视角变化剧烈和遮挡等非理想条件下（即代表实际的城市成像条件）的性能，可用于三维场景匹配和识别等任务。该数据集包含了用 Velodyne HDL-64E 激光雷达扫描的各种常见城市道路物体模型（主要包括车辆、行人、标志和树木这四类物体），这些物体模型主要是在澳大利亚悉尼的中央商业区附近采集到的。

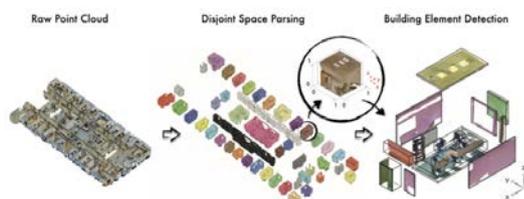


8. 斯坦福大型三维室内空间（S3DIS）数据集

数据集地址：<http://buildingparser.stanford.edu/dataset.html>

斯坦福大型三维室内空间（Stanford Large-scale 3D Indoor Spaces, S3DIS）数据集在 2016 年 CVPR 上首次公开，该数据集可用于进行三维目标语义分割的测试与评估。该数据集主要由使用 Matterport 扫描仪从 5 个大型室内区域（分别为 1900, 450, 1700, 870 和 1100 平方米）获取的点云组成，总点数超过 2 亿。该数

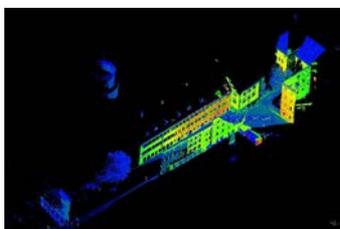
数据集提供了非常精细的 12 种语义类型标签, 包括七种结构性模块(天花板、地板、墙壁、横梁、柱子、窗户、门)以及五种常见的家具(桌子、椅子、沙发、书柜、木板)。与传统的语义分割数据集不同, 该数据集考虑了大型室内空间中可能出现的新的语义标签, 例如房间和门廊等。该数据集适合用于对大场景三维语义分割进行训练与评估。



9. Semantic3D 数据集

数据集地址: <http://www.semantic3d.net>

该数据集由 ETH Zurich 的 IGP 和 CVG 实验室于 2017 年发布, 旨在进一步提升目前三维语义分割域内数据驱动方法的性能。Semantic3D 中的数据由静态地面激光扫描仪获得的 30 次扫描结果组成, 总标注点数超过 40 亿。该数据集由 8 种语义类别的数据组成, 还包含一系列城市以及郊区户外场景, 例如教堂、街道、铁轨、广场、村庄、足球场和城堡等。数据的类别标签分别为人造地形、自然地形、高植株、低植被、建筑物、硬景观 (hard scape), 扫描缺陷 (scanning artefacts) 和汽车。此外, 该数据集还提供了原始数据集 (semantic-8) 以及采样后的数据集 (reduced-8) 的下载链接。



10. Virtual KITTI 数据集

数据集地址: <http://www.europe.naverlabs.com/Research/Computer-Vision/Proxy-Virtual-Worlds>

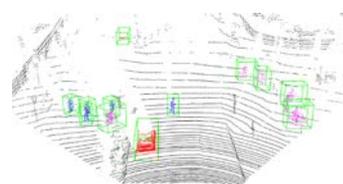
Virtual KITTI 数据集由德国亚琛工业大学的视觉计算课题组于 2017 年发布, 该数据集为了解决当时户外大场景的语义分割数据集不足的问题, 通过模仿真实的 KITTI 数据集而提出。鉴于 KITTI 数据集包含大型户外场景中五个不同的单目视频序列, 且每个视频序列都是由深度和像素级的 13 个语义标签标注, 将 KITTI 数据集中的二维深度图像投影到三维空间便得到 Virtual KITTI 数据集。



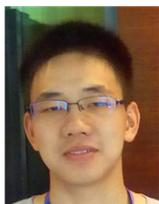
11. KITTI 三维目标检测平台

数据集地址: http://www.cvlibs.net/datasets/kitti/eval_object.php?obj_benchmark=3d

KITTI 三维目标检测平台包括 7481 个训练图像、7518 个测试图像以及相对应的点云数据。实际测试时, 主要测试算法对车辆、行人以及骑行者这三类目标的检测性能, 并根据目标尺寸以及遮挡程度设置简单、中等以及困难等三种不同的难度。同时, 该测试平台还提供了对应数据鸟瞰图 (bird's eye view) 检测性能的评估, 评估准则沿用了二维目标检测中的 PASCAL 准则。



(责任编辑: 李策、贾同)



胡庆拥

博士研究生, 牛津大学计算机系, 主要研究方向为计算机视觉、三维场景理解。硕士毕业于国防科技大学电子科学学院。



郭裕兰

讲师, 国防科技大学电子科学学院, 主要研究方向为三维视觉与机器学习。发表 PAMI/IJCV 等论文 60 余篇, 任 IET Computer Vision 期刊编委, CSIG 三维视觉专委会秘书长。

招聘信息

一、中国科学院心理研究所王甦菁博士招收访问和实习学生

中国科学院心理研究所主要从事计算机视觉在行为学，心理学方面的相关应用研究，根据相关应用，建立计算机模型。重点关注使用微表情，远程心率监测，注视点跟踪等多种计算机视觉技术，结合相关心理学知识来检测人的欺骗行为，特别是谎言检测。另外，实验室还致力于科技助残，把一系列的研究成果实现产品转化。例如智能轮椅，智能代步车等。已经参与和承担 10 多项国家科研项目，在国内外发表论文 50 余篇，获批国家专利 1 项。

王甦菁博士是国内第一个研究微表情检测和识别的科研工作者，采用计算机视觉技术，对微表情检测和识别展开了卓有成效的工作。通过计算机科学与心理学学科交叉研究，在自发微表情诱发、微表情检测和微表情识别等方面取得了系统性的创新成果。具体包括：诱发自发的微表情方法、微表情数据库的建立，微表情稀疏模型和彩色空间模型的建立，基于光流的微表情检测和识别方法的建立。这些工作解决了微表情检测和识别方面的一些理论难题，并推动了在心理学和刑侦学等应用领域的发展。构建了严格标注的三个微表情数据库，被国内外数百研究团队用于研发和测试微表情识别算法。自主研发了自动微表情识别系统，“微表情识别方法”荣获 2018 年第八届吴文俊人工智能科学技术奖一等奖。

本实验室有二台计算服务器，数十块 1080Ti GPU 卡，4 台计算工作站，一台 PTZ 摄像机和三台深度相机，提供了较好的实验平台。

1、岗位名称：嵌入式开发工程师

岗位要求：自动化专业，熟悉 STM32 编程；单片机驱动直流电机，舵机；串口通信等。

岗位职责：进行嵌入式应用开发，主要包括：结合无人车相关技术进行智能小车的研发工作。

2、岗位名称：KINECT 开发工程师

岗位要求：熟悉 KINECT 编程，VS 编程环境，C++，C#。

岗位职责：研发注视点跟踪相关算法；针对已有和研发出的算法进行工程应用落地。

3、岗位名称：深度学习工程师

岗位要求：熟悉 PYTHON，C，C++，MATLAB，LINUX，CAFFE 或 TENSORFLOW。

岗位职责：阅读相关文献，根据文献或合作导师意见实现相关算法，完成论文写作。

4、应聘基本条件：

- 计算机、自动化专业；
- 大三，大四（已经保研）、研一，研二年级的学生；
- 实习期不少于 3 个月，每周工作不少于 4 天。

5、相关待遇：

每月 800 到 4800 不等，视具体的工作能力而定。工作能力突出，报酬视具体能力还可以更高。实习结束后，可以开具相关实习证明。提供在国内参加计算机视觉相关会议的机会。可在实验室指导完成相关毕业设计。如果态度认真，可以保证实习生发表相关论文。

6、联系方式：

有意者请联系 QQ: 50563889，更多关于合作导师王甦菁副研究员的信息见 <http://sujingwang.name>

二、高新兴科技股份有限公司招聘信息

高新兴科技股份有限公司（以下简称“公司”）是业内优秀的智慧城市物联网产品和服务提供商。公司 1997 年成立，2010 年上市（股票代码 300098），长期致力于感知、连接、平台等物联网核心技术的研发和行业应用，以物联网技术为核心，聚焦公共安全、智慧城市，探索车联网等重点行业应用。2018 年着手构建系统性战略，提出“一横四纵”战略布局，即把物联网、大数据及人工智能技术横向贯通到“公共安全、大交通、通信、金融”四大纵向行业板块。

公司业务涵盖智慧城市、平安城市、智能交通、执法监督、无线通信、通信安全、金融安全、铁路安全等，为全国各省市的公安、交警等部门及铁路、银行、运营商等多领域提供产品和服务。

公司是国家计算机系统集成一级企业、国家高新技术企业，注重自主创新，引领前沿技术，参与国家公共信息安全、信息技术软件管理、增强现实系统、视频图像分析、公安信息通信网信息安全、警务云平台存储、汽车电子标识等多项国家和行业标准制定；公司尤为重视产品的安全性，产品已通过国家保密局、中国信息安全测评中心在内的多项权威认证。

集团总部设在广州，并在深圳、西安、杭州、成都、天津、南京等设有子公司或研发中心。集团公司旗下单位包括 10 家分子公司以及集团研发平台、供应链平台、职能平台等 3 大平台。公司官网：<http://www.gosuncn.com>

1、岗位名称：算法研究员

岗位需求：

- 硕士及以上学历，模式识别、计算机、图像处理、信号处理、电子、自动化相关专业；
- 掌握图像处理、模式识别及计算机视觉技术等基本理论；
- 熟练掌握 C/C++ 开发语言，熟悉 OpenCV、Matlab、或者类似图形库；
- 掌握深度学习相关算法等基本理论，并做过大量的算法训练，模型优化与改进等研究；
- 良好的沟通能力与团队合作精神；
- 具备以下开发经验之一者优先：
 - 1) 具有行为分析或视频摘要检索算法开发经验者优先；
 - 2) 具有车牌识别或人脸识别算法开发经验者优先；

岗位职责：负责人脸识别核心算法、车牌识别算法及特征搜索等核心算法研发。

2、公司主要福利待遇：

- 法定福利：
社会保险及住房公积金；
商业意外保险（全体员工）；

法定节假日、带薪年假。

- 交通生活：
提供多选择住宿；
交通、通讯补贴申请；
公司往返市区主要线路的免费班车；
每天免费的工作餐（早、午、晚）及水果。
- 表彰激励：
内部培训师奖励；
骨干持股，长期激励；
伯乐人才推荐奖、文章发表奖；
年终奖、年度评优、季度导师评优、日常奖励。
- 特别礼金：
员工结婚贺礼、生育礼金；
妇女节、端午节、中秋节、春节等传统节日礼品；
- 团队福利：
免费年度体检；
慈恩基金援助；
免费图书借阅；
档案挂靠办理（符合条件）；
户口落户办理（符合条件）；
季度生日会、跑步、爬山、晚会等文娱活动；
免费篮球、羽毛球、台球文娱体育设施；
优秀员工每满一年将有调薪、调岗的机会。

3、联系方式

钟小姐，电话：020-3206-8888-5117，136-3130-5278，邮箱：zhongmeihao@gosuncn.com。邮箱投递简历，邮件主题请写：应聘岗位+姓名。

三、大连理工大学大工-立命馆国际信息与软件学院几何计算与智能媒体技术研究所人才招聘

大连理工大学大工-立命馆国际信息与软件学院几何计算与智能媒体技术研究所，由大连理工大学和北京大学教师组成，目前有教授 7 人（含双聘院士 1 人），副教授 9 人，讲师 5 人。研究所面向国家重大战略需求和大数据智能及应用，从事几何计算和数字媒体智能处理方向的理论和应用研究。近三年来，实验室在 IEEE TPAMI、IEEE TIP、IEEE TVCG、IEEE TNNLS、IEEE TMM、IEEE TCSVT 等知名国际期刊发表 SCI 论文 100 余

篇,在 NIPS、ICCV、CVPR、ECCV、ACM MM、AAAI、IJCAI 等国际顶级计算机视觉会议上发表论文二十余篇。依托泛在网络与服务软件辽宁省重点实验室,团队建有“3D 几何与多视角图像采集”平台和“3D 打印实验室”,已发布大规模 3D 人脸标定数据库,水下图像数据库等。团队目前承担 20 多项国家级研究课题(包括“水下敏捷机器人抓捕性能智能评测模型与融合平台”、“复杂拓扑三维流形的六面体网格生成理论及算法研究”等 3 项国家自然科学基金重点项目),10 余项国防科研项目、国际合作项目和企业委托项目。研究所为教师博士后提供有竞争力薪酬,为研究生提供了丰富的项目研究和实践机会,以及灵活优厚的科研津贴。团队主页见 <http://dutmedia.org>

1、招聘岗位:“星海青千”杰出人才

1) 岗位要求:

原则上要求申报人符合国家“千人计划”青年项目申报条件:

- 属自然科学或工程技术领域,年龄不超过 39 周岁;
- 取得博士学位,并有 3 年以上海外科研工作经历;
- 在海外知名高校、科研机构或知名企业研发机构有正式教学或科研职位经历;
- 是所从事科研领域同龄人中的拔尖人才,有成为该领域学术或技术带头人的发展潜力;
- 一般应未全职在国内工作;已经在国内工作的,回国时间应在一年内;
- 对国外取得博士学位人员,且在读期间已取得突出研究成果或其他有突出成绩的人,可以突破年龄、任职年限等限制,破格申报。

2) 薪资待遇:

- 编制:事业编制;
- 薪酬:人民币 25 万-35 万/年;
- 科研支持经费:人民币 100-150 万;
- 安家:人民币 25-60 万(含大连市、学校支持,不包含学部学院额外安家补贴),提供人才公寓;
- 子女教育:学校为全体引进人才提供子女入

学、入托资源,我校附属幼儿园、小学、中学、高中均为大连市优质教育资源,在大连市基础教育领域具有良好口碑;

- 配偶工作:协助解决配偶工作。

2、招聘岗位:“博新计划”特别资助博士后

1) 岗位要求:

- 毕业于海外高水平大学(前 200)、国内一流大学建设高校、中科院科研院所且业绩突出的博士毕业生;
- 有扎实的理论基础、较强的科研能力和潜力、明确的科研方向和一定的科研成果;
- 机器学习/计算机视觉/多媒体技术/计算机图形学研究方向优先;
- 年龄 35 岁以下;获得博士学位不超过 3 年。

2) 岗位职责:

- 研究与解决计算机视觉与机器学习等领域的科学问题,撰写和发表 CCF-A 类论文;
- 在课题组指导下合作或独立申请国家或省市级基础研究项目;
- 协助课题组指导博士和硕士研究生。

3) 岗位待遇:年人力资源成本不低于 40 万,此外享受学校教职工同等待遇。

3、招聘岗位:科研博士后

1) 岗位要求:

- A 类:招收对象主要为毕业于海外高水平大学(前 200)、国内一流大学建设高校、中科院科研院所且业绩突出的博士毕业生;
- B 类:招收对象毕业于国内高校一流学科、强势学科(排名前 10 名或前 10%)且业绩突出的博士毕业生;
- C 类:招收对象为毕业于除 A、B 两类外的其他高校博士毕业生。
- 有扎实的理论基础、较强的科研能力和潜力、明确的科研方向和一定的科研成果;
- 年龄 35 岁以下,获得博士学位不超过 3 年;
- 机器学习/计算机视觉/多媒体技术/计算机图形学研究方向优先。

上述 3 类博士后分档通过学校组织的评审(函评或会评)确定。

2) 岗位职责:

- 研究与解决计算机视觉与机器学习等领域的科学问题, 撰写和发表 CCF-A 类论文;
- 在课题组指导下合作或独立申请国家或省市级基础研究项目;
- 协助课题组指导博士和硕士研究生。

3) 岗位待遇:

- A 类: 年人力资源成本不低于 30 万;
- B 类: 年人力资源成本不低于 25 万;
- C 类: 年人力资源成本不低于 20 万, 合作导师最低资助额不低于 7 万元。

此外均享受学校教职工同等待遇。

4、招聘岗位: 博士、硕士研究生

1) 招聘条件:

- 乐于沟通、尊重他人, 诚实守信;
- 计算机/软件工程/应用数学相关专业毕业, 机器学习/计算机视觉/多媒体技术/计算机图形等研究方向优先;
- 有浓厚的科研兴趣, 希望从事理论或工程应用研究, 有强烈的发表高水平论文的心愿;
- 大连理工大学本校、985 高校或有研究生院的高校的推荐者优先考虑。

2) 岗位职责:

- 研究与解决计算机视觉与机器学习等领域的科学问题, 撰写和发表英文论文;
- 博士协助课题组指导硕士研究生, 并承担相应的实验室管理工作。

3) 岗位待遇: 在国家 and 学校奖学金的基础上, 给予一定的生活及科研补贴, 具体情况面议。

5、联系方式:

有意者欢迎联系咨询雷娜老师:

nalei@dlut.edu.cn

四、快手 MMU 部门招聘

快手 MMU (Multimedia Understanding) 部

门, 聚焦多媒体内容的理解与应用, 团队分为多个研发 team, 分别聚焦在图像、人脸、语音、音乐、文本、多模态算法和应用等方面, 由算法工程师和应用工程师两部分组成。

1、招聘岗位: 计算机视觉算法负责人

工作职责:

- 负责视觉相关算法的研发, 包括但不限于 image classification、multi-label、captioning 等相关工作;
- 推动海量视频数内容分析服务上线到公司相关业务上, 并对结果负责;
- 搭建梯队合理的视觉算法团队。

任职资格:

- 统招硕士及以上学历, 在计算机视觉 (CVPR, ICCV, ECCV 等) 相关会议发表过论文优先;
- 对视频内容分析方向有浓厚兴趣, 有丰富应用经验;
- 对深度学习基础理论和大规模实践有较深理解。

2、招聘岗位: 多模态算法方向负责人

工作职责:

- 负责综合用户行为、视频内容等多种模态的模型设计和算法研究;
- 将算法应用到不同的业务场景中;
- 搭建梯队合理的视觉算法团队。

任职资格:

- 计算机相关专业硕士研究生及以上学历, 在机器学习等相关国际顶级会议发表过论文优先;
- 精通常用机器学习模型和算法;
- 对前沿技术研究抱有浓厚兴趣, 良好的技术视野和深度。

3、联系方式:

联系邮箱: maxulong@kuaishou.com, 邮件主题: 应聘岗位+姓名+来源

(责任编辑: 樊鑫、蹇木伟)

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 International Journal of Computer Vision, Computers & Electrical Engineering 特刊等。

3 会议简介 (ICCV)

ICCV(International Conference on Computer Vision)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研究水平。ICCV 始于 1987 年，其后两年举办一届。

ICCV-2019 将于 2019 年 10 月 27 日至 11 月 3 日在韩国首尔举办，论文截止时间为 2019 年 3 月 1 日。征稿范围涉及三维计算机视觉、动作识别与跟踪、大数据、医学图像分析、深度学习、图像处理、优化方法、机器人视觉、识别、检测、分割等方向的研究进展。

(责任编辑：金鑫)

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
ICASSP 2019	2019.05.12-17	英国 Brighton	2019.02.18	https://2019.ieeeicassp.org/
ICCV 2019	2019.10.27-11.03	韩国 Seoul	2019.03.01	http://iccv2019.thecvf.com
CAIP 2019	2019.09.02-05	意大利 Salerno	2019.04.01	http://caip2019.unisa.it
PRICAI 2019	2019.08.26-30	斐济 Yanuca	2019.02.28	https://www.pricai.org/2019/
IJCAI 2019	2019.08.10-16	中国 Macao	2019.02.25	http://www.ijcai19.org
ACL 2019	2019.07.28-08.02	意大利 Florence	2019.03.04	http://www.acl2019.org/EN/index.xhtml
SIGKDD 2019	2019.08.03-07	美国 Anchorage	2019.02.03	http://www.kdd.org/kdd2019
MCA 2019	2019.05.27-31	日本 Tokya	2018.12.07	http://www.mva-org.jp/mva2019

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
Computers&Electrical Engineering	Artificial Intelligence and Computer Vision	2019.02.15
International Journal of Computer Vision	Generative Adversarial Networks for Computer Vision	2019.03.31