

CCF

计算机视觉 专委会简报

COMPUTER VISION NEWSLETTER

2019/06 期
总第 22 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 烨	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
视界无限	视界无限系列研讨会	06
专委新闻	RACV 2019 圆满召开	08
	CCF-CV 专委工作会议暨换届选举顺利召开	10
	CCF-CV 学术新锐奖获奖名单	12
	PRCV 2019 圆满闭幕	13

科技前沿

专题综述	视频行为识别相关研究进展	17
热点追击	人体、人脸、人手表面及运动重建	21
	可微分网络架构搜索的稳定性研究	23

委员风采

委员访谈	南京理工大学潘金山教授访谈	25
委员好消息		27

资源平台

开源代码	图像重光照	30
数据集	微表情	34
招聘信息		37
征文通知		44

CCF-CV 走进高校系列报告会

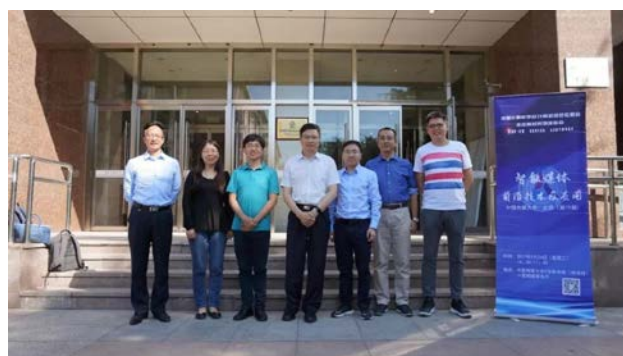
第 78 期 吉林大学



2019 年 9 月 10 日上午，由中国计算机学会计算机视觉专委会（CCF-CV）主办，吉林大学计算机学科（计算机科学与技术学院、软件学院、计算机科学技术研究所、国家级计算机实验教学示范中心、符号计算与知识工程教育部重点实验室、海战场攻防对抗仿真技术教育部重点实验室）、CCF 长春、YOCSEF 长春、CCF 吉大和吉林省计算机学会联合承办的第 78 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”在吉林大学计算机楼 B108 报告厅举行。吉林大学和省内相关高校和企业的 150 多名师生、学者和业内人士参加了报告会。本期报告会邀请了厦门大学纪荣嵘教授、大连理工大学王栋副教授、北京大学刘家瑛副教授和南京理工大学李泽超教授等四位专家学者做特邀报告，吉林大学计算机学院副院长李文辉教授和车翔玖教授担任本次报告会的执行主席。

报告会现场气氛热烈，同学们纷纷针对专家报告提出自己感兴趣的问题，专家给了详细解答，报告会在热烈的气氛中圆满结束。

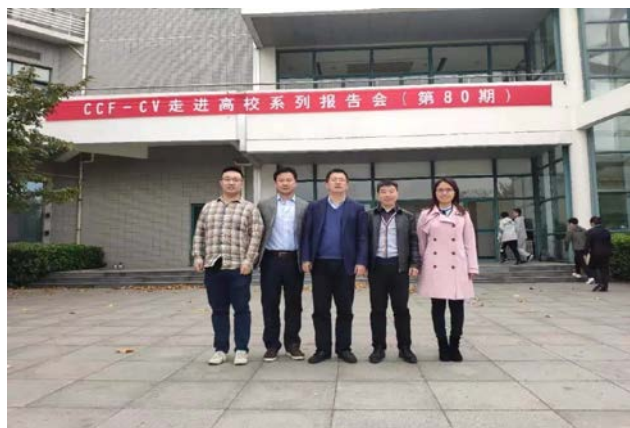
第 79 期 中国传媒大学



2019 年 9 月 24 日下午，由中国计算机学会计算机视觉专委会（CCF-CV）主办，中国传媒大学媒介音视频教育部重点实验室、数据科学与智能媒体学院联合承办的第 79 期 CCF-CV 走进高校系列报告会——“智能媒体前沿技术及应用”在中国传媒大学明德楼一层明德报告厅成功举行。本期报告会邀请了中国科学院自动化所陶建华研究员、张兆翔研究员、西安电子科技大学董伟生教授、南开大学杨巨峰副教授四位专家学者做特邀报告。中国传媒大学媒介音视频教育部重点实验室副主任叶龙副研究员和数据科学与智能媒体学院副院长张远教授担任本次报告会的执行主席。

本期报告会持续了三个多小时，吸引了二百多名师生前来聆听，专家学者们相继介绍了自己的最新研究成果与心得。在提问环节中，在座的学生和老师纷纷提出自己科研方面的问题，专家们也对这些提问做出了详尽的回答。最后，现场听众对四位特邀讲者表示由衷的欢迎与感谢，报告会在雷鸣般的掌声中圆满结束。

第 80 期 中国石油大学



2019 年 10 月 25 日上午,由中国计算机学会计算机视觉专委会 (CCF-CV) 主办,中国石油大学 (华东) 计算机科学与技术学院承办的第 80 期 CCF-CV 走进高校系列报告会——“深度学习前沿与多媒体智能”在中国石油大学 (华东) 逸夫楼一楼报告厅成功举行。本期报告会邀请了北京大学彭宇新教授、西安电子科技大学邓成教授、中科院计算所阚美娜副研究员、上海交通大学张拳石副教授四位专家学者做特邀报告。中国石油大学 (华东) 大数据智能处理创新团队负责人张卫山教授和计算机科学与技术学院院长宫法明教授担任本次报告会的执行主席。

本期 CCF-CV 走进高校系列报告会持续了四个多小时,报告内容在现场引起强烈反响,学术交流气氛活跃,激发全场对学术的创新热情,四位计算机视觉领域的专家学者相继介绍了计算机视觉领域的最新研究成果,具有十分重要的指导意义。专家们以自身的学识和修养深深影响了现场师生,拓宽了听众们的思路 and 眼界,现场通过面对面激烈的探讨交流使大家获益匪浅。最后,参加报告会的现场听众对四位特邀讲者不辞辛苦远道而来,表示由衷的欢迎与感谢,报告会在热烈的掌声中圆满结束。

第 81 期 电子科技大学



2019 年 10 月 26 日,由中国计算机学会计算机视觉专委会主办、电子科技大学生命科学与技术学院承办的第 81 期 CCF-CV 走进高校系列报告会活动——“智能医学影像处理、分析与应用”报告会在电子科技大学 (沙河校区) 学术交流中心 104 报告厅成功举行。本期报告会邀请了四川大学华西医院龚启勇教授,清华大学廖洪恩教授,杭州电子科技大学厉力华教授,西北工业大学夏勇教授四位专家学者做特邀报告,执行主席由电子科技大学生命科学与技术学院李永杰教授担任。来自四川大学、电子科技大学、西南科技大学、西南交通大学、西南民族大学、成都中医药大学、成都理工大学、成都医学院、中国民航飞行学院等高校的 150 多名师生聆听了四位专家在该领域的前沿技术成果和最新动态介绍。

此次持续了四个多小时的 CCF-CV 走进高校系列第八十一期报告内容精彩纷呈,让所有参会人员享受了一场学术盛宴。四位专家同与会师生进行深入互动,对师生在科学研究中遇到的问题进行了耐心细致的回答,所展示的研究成果和科研方法让与会师生受益匪浅。

(责任编辑:马占宇)

CCF-CV 视界无限系列研讨会

第4期 大连理工大学



2019年10月12日，由中国计算机学会计算机视觉专委会主办、大连理工大学电子信息与电气工程学部承办的第4期CCF-CV“视界无限”系列活动——“视觉目标跟踪(VOT)前沿进展与未来趋势”研讨会在大连理工大学创新创业学院成功举行。本期研讨会由计算机视觉专委会常务委员、大连理工大学创新创业学院院长卢湖川教授、大连理工大学王栋副教授、大连理工大学张立和副教授组织和主持。研讨会邀请了安徽大学罗斌教授、哈尔滨工业大学左旺孟教授、清华大学鲁继文副教授、上海交通大学马超博士、微软亚洲研究院彭厚文博士、安徽大学李成龙博士、大连理工大学王栋博士做主题报告及深度研讨。来自全国高校和企业的100余名参会者汇聚大连，就视觉目标跟踪前沿进展及未来趋势进行了热烈的交流和深入细致的探讨。

会议议程依次包括主题报告、海报/系统展示、深度研讨三部分。首先，计算机视觉专委会常务委员、大连理工大学卢湖川教授致辞。卢老

师对与会嘉宾和师生表示欢迎，然后介绍了计算机视觉专委会的组织结构、职能目标、举办的主要活动，以及“视界无限”系列活动的宗旨和目标。卢老师指出，在刚刚结束的RACV 2019会议上，专家们预测了计算机视觉领域未来5-10年的若干研究热点，视频分析和理解尤其是视觉目标跟踪位列其中。大连理工大学科研团队从2008年开始相关领域的研究，与国内其他研究者一起提出了一系列benchmark，取得了多项重要成果。希望通过举办本次活动，深入研讨并继续推动视觉目标跟踪同行之间的交流、促进合作、提升影响力，给国内同行带来启示，并祝愿研讨会圆满成功。

罗斌教授的报告题目是“视觉跟踪的近期研究和进展”。针对视觉跟踪问题的几个特定挑战，分四个方面介绍团队近几年的相关研究成果：(1)针对目标矩形包围盒的背景引入问题，提出图表示学习算法；(2)针对训练数据的正样本弱问题，提出实例对抗生成网络；(3)针对训练数据不足问题，提出多层次挑战感知网络；(4)针对可见光传感器的成像受限问题，引入对光线变化和天气条件不敏感的红外热成像传感器，提出融合可见光和热红外信息的多种模型和方法，有效利用两者的互补优势，实现稳健跟踪。此外，还在报告中指出基于多模态信息融合的视觉跟踪算法是未来趋势。

左旺孟教授的报告题目是“视觉跟踪的相关滤波器方法研究：改进与拓展”，重点从判别项和正则化项两方面针对视觉目标跟踪中的相关

滤波算法进行探索与研究。首先, 通过将 hinge 损失引入相关滤波器模型, 提出了一种支持相关滤波器模型, 并给出了相应的快速求解算法。其次, 在正则化项方面, 提出了一种时域正则化项, 并分析了空域正则化项与 cosine 窗口的关系。最后, 给出了一种相关滤波器与卷积特征的联合学习模型。

鲁继文副教授的报告题目是“深度强化学习与视觉内容理解”, 重点介绍了清华大学自动化系智能视觉实验室近年来提出的面向视觉内容理解的多个深度强化学习方法, 主要包括多智能体深度强化学习、图深度强化学习、和结构化深度强化学习等, 以及它们在视觉跟踪、行为预测、行人再识别、和模型压缩等多个视觉内容理解任务中的应用。

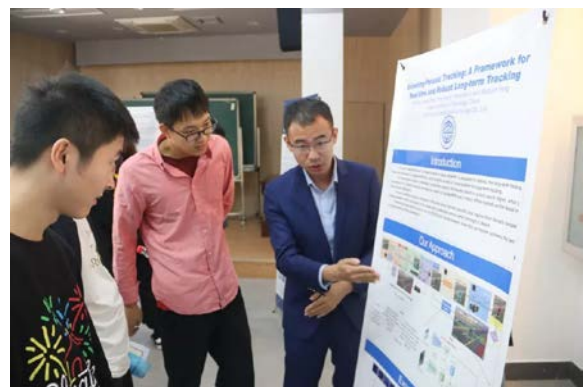
马超博士的报告题目为“基于深度回归模型的视觉目标跟踪方法”, 介绍了如何利用深度回归模型提升视觉目标跟踪算法的精确度和鲁棒性。主要包含三方面内容: (1) 基于多级卷积特征的自适应相关滤波算法; (2) 利用 shrinkage loss 训练深度回归网络并用于跟踪; (3) 基于目标感知回归和非监督回归的视觉跟踪算法。并从自学习和无监督学习角度讨论了监督信息缺乏或丧失情况下如何设计适用于解决跟踪问题的深度回归网络。

彭厚文博士的报告题目为“Deeper and Wider Siamese Networks for Real-time Visual Tracking”, 首先深入分析了为什么之前的跟踪算法不能受益于更深的深度神经网络, 给出跟踪主干网络设计指南。新设计的主干网络能够显著提升算法的精度; 随后, 介绍了团队在 RGB 长时目标跟踪和 RGB-D 长时目标跟踪方面的尝试和初步结果; 最后, 展示了视觉目标跟踪在微软实际项目中的落地应用。

李成龙博士的报告题目为“基于多适配器网络的 RGBT 视觉跟踪”, 重点介绍了如何利用多模态信息处理热红外成像下的视觉目标跟踪问题, 提出包含通用适配器、模态适配器和实例适配器的多适配器网络模型。为了充分利用模态之间的互补优势, 提出模态适配器相互学习模块, 在不同模态之间传播互补特征, 并抑制噪声干扰。

王栋博士的报告题目是“长时视觉跟踪: 全新机遇与挑战”, 重点介绍了长时跟踪问题及其与短时跟踪问题的联系与区别以及新的机遇与挑战, 以及大连理工大学卢湖川教授、王栋副教授团队在 VOT2018 长时比赛冠军算法和 ICCV2019 最新工作为基础, 介绍长时视觉跟踪前沿进展以及深度学习算法在长时目标跟踪问题上的初步探索。

在海报展示环节, 共有来自全国高校的 10 篇论文海报, 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“VOT 的局限及与 MOT 的区别”、“深度特征与时序建模”、“VOT 发展前景”、“CV 领域发展前景”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, 第 4 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)

RACV2019

计算机视觉前沿进展研讨会圆满召开



2019年9月27日-28日，中国计算机学会计算机视觉专委会（CCF-CV）年度学术研讨会RACV2019（Recent Advances on Computer Vision）在北京怀柔圆满召开。与过去几年的RACV论坛不同，按照今年4月份常委会的讨论决议，本次会议在组织形式与研讨内容上均进行了全新改版，定位为国内计算机视觉领域的小规模精品研讨会，通过定向邀请方式汇集领域专家深度研讨计算机视觉领域中的若干核心问题并形成进展报告。根据常委委员前期的讨论票选，本次会议设置了5项研讨主题，近40名嘉宾参加了讨论交流。每项主题首先由特邀嘉宾进行时长1小时的主题发言，之后所有与会嘉宾进行时长1小时的自由讨论。

27日晚上首先进行了“主题一：未来5-10年CV发展趋势”的研讨。该主题由林宙辰、刘日升、阚美娜3位委员负责组织，邀请了查红彬、陈熙霖、卢湖川、刘烨斌、章国锋共5位嘉宾进行主题发言。过去近十年来，深度学习的引入为计算机视觉领域的发展带来了蓬勃生机，放眼未来，计算机视觉悬而未决的基本问题与核心任务还有哪些？几位嘉宾围绕三维视觉、视觉识别、光度视觉、视觉中多模态融合等前沿方向的未来

发展趋势进行了精彩的观点分享。

28日上午进行了“主题二：视觉的层次化处理过时了吗”与“主题三：三维视觉：现状与未来”的研讨。其中，主题二由查红彬、鲁继文、马占宇3位委员负责组织，邀请了胡占义、贾云得、邬霞、郑伟诗、张拳石共5位嘉宾进行主题发言。40多年前，马尔提出的视觉计算理论框架中，层次化处理与表征是其基本思想；如今新的理论方法层出不穷，很多任务得以重新定义与求解，经典的层次化处理理论对于当下的研究有哪些指导意义？嘉宾们围绕马尔视觉框架基本理论、视觉层次化处理的共性机理、视觉层次化与处理模块化、深度学习的层次化处理等议题展开了深入探讨。



主题三由刘青山、章国锋、郭裕兰3位委员负责组织，邀请了查红彬、贾云得、吴毅红、杨睿刚共4位嘉宾进行主题发言。对三维物理世界的建模与理解一直是计算机视觉研究的核心目标，近年来的热点与趋势有哪些？嘉宾们围绕三维视觉中的数据表达形式统一化、三维视觉中的数据标注、几何重建与语义识别、三维视觉未来热点与趋势等议题展开了深入探讨。

28 日下午进行了“主题四：主动视觉：感知、行为与动机”与“主题五：CV 中的可解释性：定义和实现”的研讨。其中，主题四由赖剑煌、谢晓华、王瑞平 3 位委员负责组织，邀请了王亦洲、陈熙霖、徐凯、谢晓华、李家康共 5 位嘉宾进行主题发言。与传统计算机视觉研究的侧重点有所不同，主动视觉历经过去 20 多年发展，虽未产生显著的焦点性成果，却始终吸引着研究者的目光，也被认为是下一代潜在取得突破性进展的技术。嘉宾们针对主动视觉的定义与范畴、主动视觉与传统视觉的 pipeline 之区分、自主学习中的机器意识与计算等议题进行了精彩深入的主题发言分享。



主题五由陈熙霖、纪荣嵘、苏航 3 位委员负责组织，邀请了王井东、林惊、苏航、何旭明、王瑞平共 5 位嘉宾进行主题发言。深度学习与大数据时代，随着各种深度网络模型突破一项又一项计算机视觉任务的性能上限，人们对于复杂模型背后的运行机理与模型结果的可解释性提出了越来越多的期望，从黑盒到灰盒、再到白盒，AI 模型可解释性这一前沿方向的发展路径会是如何？嘉宾们围绕可解释人工智能的核心目标与挑战、可解释性的可能技术途径与度量方式、视觉任务可解释性的自身特点与要求等议题各抒己见，阐述了精彩观点。

研讨会的每一项主题进行过程中，特邀嘉宾主题发言与之后的自由讨论都相得益彰。主题发言视角广阔、见解深入，为主题推进提纲挈领；

自由讨论热情激烈、聚焦热点，嘉宾们不乏“火药味儿”十足的唇枪舌战。有道是真理越辩越明，通过一天多的深入研讨，观点分歧逐渐消解，认知共识则逐渐收敛。伴着雁栖湖畔的落日余辉，研讨会圆满落幕，参会嘉宾们纷纷表示本次会议组织形式新颖、内容丰富，会场氛围轻松而热烈。按照计划，组委会后续将整理相关主题的发言与讨论文稿，形成观点性文档进行发布。



本次研讨会由专委常委山世光与鲁继文两位老师负责组织，专委秘书处成员协助会务组织的相关事宜。



（责任编辑：王瑞平）

CCF-CV

专委工作会议暨换届选举顺利召开



中国计算机学会计算机视觉专委会（CCF-CV）2019 年度工作会议于 2019 年 11 月 9 日在西安曲江国际会议中心成功举办，来自全国高校、科研院所、企业的现任委员 300 余人参加了工作会议。会议由专委会秘书长、中科院自动化所王亮研究员主持。

会议首先由专委会副主任、中科院计算所陈熙霖研究员致辞。陈熙霖研究员代表专委会主任谭铁牛院士感谢委员们从祖国各地到场参会，他回顾了本届专委会四年来的发展历程和主要成绩，积极肯定了专委会在过去一年中在学术交流及社会服务等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，打造精品学术交流活动，为推进计算机视觉学术研究与产业发展继续发挥积极引领作用。

之后，中国计算机学会理事、专委工委委员、西北工业大学於志文教授代表学会致辞。於志文教授对专委工作会议的召开表示祝贺，对专委开展的各项活动给予了积极评价，对专委换届选举工作的重要事项提出了指导意见。他最后预祝

2019 年计算机视觉专委会全体委员会议及换届选举圆满成功。

随后，专委会秘书长王亮研究员向与会委员做年度工作报告。报告介绍了专委会组织结构及过去 5 年获得的计算机学会评优奖励，通报了 4 月份专委会常委会议和 5 月份秘书处工作会议提出和实施的若干新举措，全面回顾了专委会过去一年的各项学术交流活动，指出了工作中存在的问题和拟采取的改进方案，最后介绍了下一年度专委会工作计划。



接下来，进入激动人心的 CCF-CV 颁奖环节。颁奖仪式由专委会副主任、专委提名与奖励工作组组长、北京大学查红彬教授主持，查红彬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，也有为专委会发展尽心尽力、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会大家庭的繁荣兴盛。

本年度 CCF-CV 终身学术贡献奖授予清华大学徐光祐教授，由专委会副主任陈熙霖研究员颁奖。徐光祐教授现场发表了质朴感人的获奖感言，

回顾了早年我国计算机视觉方向的发展历程和个人在其中所做的工作，感谢专委会的认可和颁奖，勉励青年研究者继续做出更大成绩。徐光祐教授的发言赢得了代表们的热烈掌声，他的鼓励与期许极大振奋了委员们的科研工作热情。

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杰出成就的科研人员。本年度该奖项授予专委会副主任陈熙霖研究员，以表彰他“在自然场景理解、物体识别特别是人脸识别、手语识别等方面做出的系统的创造性贡献”。专委会副主任、提名与奖励工作组组长查红彬教授为陈熙霖研究员颁发了奖励证书。陈熙霖研究员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。



本年度继续颁发年度服务贡献奖，该奖项首次由中科视拓冠名，授予中科院计算所王瑞平研究员、西北工业大学魏巍副教授、南开大学杨巨峰教授等3位委员，由专委会副主任赖剑煌教授、王涛研究员和中科视拓副总裁何思谊女士为获奖人颁奖，感谢他们在专委会活动、会议组织、秘书处服务等方面作出的重要贡献。

本年度颁发的另一重要奖项是 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。本年度获得该奖项的3位学术新星是：上海交通大学的方浩树、清华大学的饶永铭和易冉，由专委会学术新锐评审小组组长贾云得教授、秘书长王亮研究员以及中科视拓副总裁何思谊女士为获奖人颁奖。

根据会议日程，接下来进行了专委会换届选

举，选举工作组组长陈熙霖研究员主持了选举过程。本年度共有2位委员申请竞选主任、4位委员申请竞选副主任、2位委员申请秘书长、23位委员申请常务委员。按照流程，候选人一一上台进行了申请陈述，并回答了现场委员的质询。经过参会的专委会现任委员投票表决，北京大学查红彬教授当选新一届专委会主任，中科院自动化所王亮研究员、上海科技大学虞晶怡教授、南京信息工程大学刘青山教授当选副主任，北京邮电大学马占宇副教授当选秘书长，华中科技大学白翔教授、南开大学程明明教授、东南大学耿新教授、厦门大学纪荣嵘教授、中山大学赖剑煌教授、北京大学林宙辰教授、大连理工大学卢湖川教授、清华大学鲁继文副教授、中科院计算所山世光研究员、微软亚洲研究院王井东研究员、爱奇艺王涛研究员等当选为新任常务委员。



在委员建言献策环节中，与会代表进行了热烈的讨论，委员们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，引起了大家的共鸣。委员们纷纷表态将继续积极参与专委会的各项工作，争取通过专委会这个平台为计算机视觉领域的发展做出更大贡献！最后，专委会2019年度工作会议圆满结束！

（责任编辑：王瑞平）

CCF-CV 学术新锐奖获奖名单

根据《CCF-CV 奖励条例》和《CCF-CV 学术新锐奖评选条例》，经过形式审查、提名与奖励工作组初评和评审委员会专家终评，评选出 3 名 2019 年度 CCF-CV “中科视拓 Seeta 学术新锐奖”获奖者，现将结果予以公告。

2019 年度 CCF-CV “中科视拓 Seeta 学术新锐奖”获奖名单

姓名：方浩树
性别：男
单位：上海交通大学
博士生类型：直博
入学时间：2019 年 9 月
导师：卢策吾
推荐人：乔宇 欧阳万里

姓名：饶永铭
性别：男
单位：清华大学
博士生类型：直博

入学时间：2018 年 9 月
导师：鲁继文
推荐人：王瑞平 赫然

姓名：易冉
性别：女
单位：清华大学
博士生类型：直博
入学时间：2016 年 9 月
导师：刘永进
推荐人：查红彬 毋立芳

附：评审委员会名单
贾云得（主任） 王涛 刘青山 山世光 白翔

中国计算机学会计算机视觉专委会
提名与奖励工作组
2019 年 10 月 28 日

（责任编辑：毋立芳）

PRCV2019 圆满闭幕



第二届中国模式识别与计算机视觉大会 (PRCV2019) 于 2019 年 11 月 8 日至 11 日在西安曲江国际会议中心成功举办, 1800 余名专家学者齐聚一堂, 共同研讨相关领域最新理论和技术成果。

本届大会由中国计算机学会 (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, 西北工业大学、西安交通大学、西安电子科技大学和陕西师范大学等单位共同承办, 西咸新区沣东新城管理委员会协办。中国科学院院士、中科院自动化所谭铁牛, 中国工程院院士、西安交通大学郑南宁, 中科院计算所陈熙霖教授, 西北工业大学张艳宁教授共同担任大会主席。大会邀请到计算机视觉领域国内外的多位著名学者做特邀报告, 除了口头论文与墙报论文展示, 还设有企业宣讲和展示环节, 丰富的会议内容、国际水平的学术报告为中国计算机视觉和模式识别发展提供了一场精彩的学术盛宴。

开幕式



PRCV2019 开幕式由大会主席、西北工业大学张艳宁教授主持。张教授对大家的参会表示热烈欢迎! 对全体组委会的辛勤付出表示衷心感谢!

大会主席谭铁牛院士致辞, 对于 PRCV2019 人数创新高表示高兴、振奋, 也显示了我国模式识别和计算机视觉领域的研究盛况。紧接着, 谭院士回顾了 PRCV 会议的由来, 并向为会议成功举办提供支持的各方致谢。最后勉励大家抓住人工智能的机遇多出高水平的研究成果, 继续努力把 PRCV 打造成我国计算机视觉和模式识别领域的高水平交流平台。

大会主席郑南宁院士代表四个学会致辞, 郑院士表示很高兴看到 PRCV 会议的盛况, 说明我国在人工智能、模式识别与计算机视觉领域的人才储备丰厚, 鼓励大家做出更多高水平研究。

西北工业大学党委副书记陈建有代表学校对大家的到来表示欢迎! 并对学校学科建设、人才培养、科学研究等相关情况进行介绍。

大会程序委员会主席、北京大学林宙辰教授对会议程序安排进行了说明, 本届大会的特点是内容丰富多彩, 包括 4 个主旨报告、6 个特邀报告、6 个专题论坛、4 个讲习班、18 篇口头论文、147 篇墙报展示、7 个专题竞赛、40 个企业应用展示等。大会共接受投稿 412 篇, 接收 165 篇, 其中 18 篇 oral 论文, 147 篇 poster 论文。设置最佳论文奖和提名奖各 1 项、最佳学生论文奖和提名奖各 1 项、最佳 oral 论文奖 1 项、最佳 poster 论文奖 1 项。会议论文集由 Springer 出版。

大会主题报告

本届大会安排了四场主旨报告, 报告人均均为计算机视觉领域在国际上享有盛誉的学者。这四场主持报告分别是: 鲁汶大学 (KU Leuven) Tinne Tuytelaars 教授做主题报告《Keep on Learning》, 西安电子科技大学 IEEE Fellow 焦李成教授做主题报告《后深度学习的挑战与思考》, 多伦多大学 (University of Toronto) Kyros Kutulakors 教授做主题报告《Transport-Aware Cameras》, 首尔大学 (Seoul National

University) Kyoung Mu Lee 教授做主题报告《3D Human Pose Estimation: Advances and Challenges》。这些报告思想深邃、内容丰富，既包含对领域历史的全局回顾，又立足当下研究热点与难点问题，为听众带来了许多前瞻性、引领性的叙述观点和思考，激发了大家的浓厚兴趣和强烈反响。



大会特邀报告

本届大会安排了六场特邀报告，报告人均为计算机视觉领域的著名专家学者。这六场特邀报告分别是：香港中文大学终身教授、腾讯杰出科学家贾佳亚教授做主题报告《图像处理进展》，中山大学郑伟诗教授做主题报告《行人重识别：是否已经被解决了？》，南京信息工程大学刘光灿教授做主题报告《基于优化的模式分析》，中国科学院自动化所孙哲南教授做主题报告《虹膜识别研究进展》，西蒙弗雷泽大学（Simon Fraser University）谭平教授做主题报告《A Neural Network for Detailed Human Depth Estimation from a Single Image》，北京理工大学黄华教授做主题报告《流形上的视频运动分析及应用》。这些报告立足当下研究热点与难点问题，引起了参会者的广泛探讨与深入思考。

博士生论坛

博士生论坛首先 3 位 CCF-CV 学术新锐奖获得者上海交通大学方浩树、清华大学饶永铭、清华大学易再做 15 分钟的口头报告，然后由 3 位学术新锐奖候选人上海科技大学陈安沛、浙江工业大学刘儒瑜、西北工业大学谢雨桐做 10 分钟的口头报告。接着，32 名博士生带着自己的论文与 23 位导师进行一对一交流。最后是 panel 环节，查红彬、王亮、山世光、胡润波、鲁继文几



位导师针对高质量博士培养、以及现场博士生提出的问题给出自己的观点。



专题论坛

专题论坛版块分为 5 场，其主题涵盖了计算机视觉当前的多个热门研究方向，包括“类脑视觉”论坛，“认知共融产业机器人”论坛，“三维视觉前沿”论坛，“全民看脸时代的机遇和挑战”，“生物特征识别”论坛以及“遥感大数据和人工智能”论坛。每个论坛的组织形式各具特色，既包括领域专家的学术报告分享，又包括气氛活跃的 panel 观点交锋。



讲习班

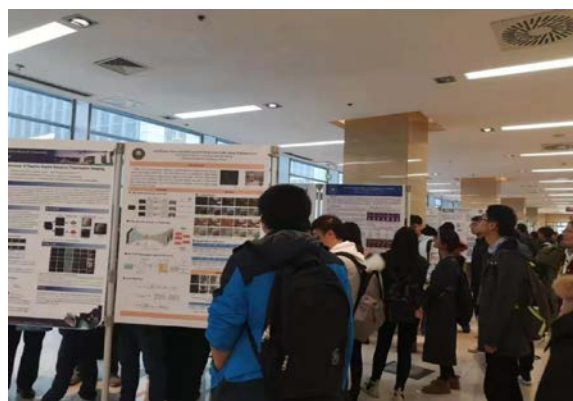
本届 PRCV2019 开设的讲习班板块，包括 4 个热门专题的全景系统性介绍，分别是：“鲁棒和可解释人工智能”，“大规模深度人脸生成与分析”，“迁移学习方法及其在开放环境下的视觉应用”和“视觉检测与跟踪”。4 个讲习班安视频内容智能制作技术挑战赛”，吸引了来自国

排在主会开始前一天的 8 日下午并列进行，每场报告均吸引了众多师生参加，报告之后的问答讨论异常活跃，提前预热了大会火爆的学术交流气氛。



论文宣讲与展示

在 11 月 9-10 日，大会分 6 个主题进行了 Oral 论文的口头宣讲，并在 Poster 展示阶段进行了详细的论文展示交流。青年学子们在这里报告自己的最新工作，交流自己的研究心得，会场内外处处闪耀着激情思想碰撞的火花。



专题竞赛

延续往届大会传统，结合产业应用实际需求，大会组织了“车纹识别技术挑战赛”、“同方威视箱包再识别技术挑战赛”、“鞋印图像数据年龄挖掘挑战赛”、“面向自动阅卷的 OCR 技术挑战赛”、“高速低功耗视觉理解挑战赛”、“大田作物病害图像识别技术挑战赛”、“第一届短内高校与研究院所的多支队伍注册参赛。

经过激烈角逐，北京格灵深瞳信息技术有限公司代表队，厦门大学和华侨大学代表队（并列），大连理工大学代表队，浙江大学代表队与华南理工大学联想研究院代表队，北京理工大学代表队，四川大学代表队，上海交通大学代表队与中山大学代表队分别获得这七项竞赛的第一名。10日下午安排了竞赛专题的交流，各优胜参赛队伍分享了参赛算法与比赛心得，之后进行了颁奖环节，相信这种工业界出题、学术界作答的竞赛模式必将大大促进产学研更加紧密高效的结合。

企业应用展示

大会的成功召开离不开工业界的大力支持。本届大会共收到 20 家企业的赞助，其中金牌赞助商 7 家、银牌赞助商 2 家和铜牌赞助商 11 家。各赞助企业代表介绍了自家企业的前沿技术和主流产品，给会场带来一阵高科技应用的兴趣和热潮。大会期间，赞助企业在预定的区域进行了现场展示，技术负责人为观众实际演示和解说，让观众有了一次别开生面的亲身体验。

晚宴及颁奖

大会次日安排了会议晚宴，晚宴上公布了由大会评奖委员会评选出的 PRCV2019 大会最佳论文奖、最佳论文提名奖、最佳学生论文奖和最佳学生论文提名奖各 1 名，由企业代表赞助奖品并进行颁奖。同时颁发了最佳口头报告奖、最佳海报奖、博士生论坛最佳海报奖、最佳审稿人奖以及最佳服务贡献奖。此外，举行了赞助商颁奖仪式，向提供赞助的广大赞助商致以谢意。

PRCV2020 将在古都南京举行。南京理工大学肖亮教授介绍了 PRCV2020 的策划及准备情况，并播放 PRCV2020 宣传片。

感谢赞助企业

本次大会（PRCV2019）得到深圳市腾讯计算机系统有限公司、北京旷视科技有限公司、精锐视觉智能科技（深圳）有限公司、博世（中国）投资有限公司、北京奇艺世纪科技有限公司、同

方威视技术股份有限公司、北京嘀嘀无限科技发展有限公司、北京澎思科技有限公司、杭州海康威视数字技术股份有限公司、慧眼自动化科技（广州）有限公司、广东康云科技有限公司、北京航天宏图信息技术有限公司、联泰集群（北京）科技有限责任公司、睿视智觉（厦门）科技有限公司、赛灵思电子科技（上海）有限公司、叠境数字科技（上海）有限公司、北京海天瑞声科技股份有限公司、银河水滴科技（北京）有限公司、东南数字紧急发展研究院、联想研究院、Springer 出版社的赞助，对他们的大力支持表示衷心感谢！



感谢幕后的组织团队

为期三天半的 PRCV2019 落幕，会议的成功举办得益于大会组委会的通力合作，特别是张艳宁教授带领的本地团队包括西北工业大学魏巍副教授、王鹏教授、孙瑾秋副教授、赵歆波教授、杨涛教授、张世周博士、徐迪博士等老师以及志愿者学生为大会的全力付出，使得大会能够在周到体贴的安排下圆满顺利地完



（责任编辑：毋立芳）

视频行为识别相关研究进展

北京理工大学 高广宇 刘子铭 吴桐

一、引言

行为识别 (Action Recognition) 是计算机视觉领域的一个重要的研究热点。近年来, 随着大量的视频行为识别数据集被公开 (如图 1), 行为识别相关研究越来越热。文献[1]综述了人体行为识别公开数据集的发展与前瞻, 文献[2]对视频行为识别进行了综述。然而, 由于光照不确定、气候条件、压缩失真、摄像距离等因素产生的视频图像数据难以满足视觉识别和理解要求。视频行为识别也明显受到上述因素影响, 如光照条件、角度变化、复杂背景和同类行为之间的差异, 在很长时间没有重大突破。近年来, 随着深度学习技术的出现和硬件计算能力的提升, 在静态图像和视频数据上的行为识别模型性能得到了显著的提升。

在早期研究工作中, iDT (improved Dense Trajectories) [3] 是性能最好的算法之一, 很多

当时最先进的模型都是基于此算法进行改进的。近几年, 随着深度学习技术的广泛应用, 越来越多的基于深度学习的行为识别方法被提出来并成为主流方法。其中, 双流架构 (Two Stream) [4], 三维卷积网络 (C3D) [5] 以及基于循环神经网络的模型 (RNN based Model) [6] 成为了目前实现行为识别的最主要也是较为成功的三类方法。双流模型基于 2D 卷积网络, 输入包括常见的 RGB 图像和光流图像 (光流是一种描述视频中运动信息的模态, 由视频的相邻帧计算得到, 因此双流网络训练之前需要预先得到的光流图片)。双流网络的计算量和参数量都是较小的, 其较好的性能得益于光流特征的引入。3D 卷积模型是 2D 卷积的扩展, 虽然 3D 卷积网络可以实现视频行为识别的最佳结果, 但是其三维卷积操作也带来参数量的显著增长, 这对计算资源提出了挑战。同时, 也有一些工作对三维卷积操作进行了改进以实现更少的参数量更高的性能。最后, 也有大量的工作基于循环神经网络 (RNN) 变体长短时记忆 (Long Short-Term Memory, LSTM) 网络 [7] 来设计行为识别模型。基于 LSTM 的模型利用了视频数据天然具有的时序特性, 同样可以实现部分数据集上较好的结果。但是, 由于图像和视频本质上存在的复杂多样的视觉变化, 行为识别问题仍然没有被很好地解决。

事实上, 视频行为识别需要考虑的一个重点问题是上下文依赖 (Context Dependencies) 建模。传统二维卷积操作可以用来建模空间上下文信息, 三维卷积操作可以捕捉时空上下文关系, 同时, 循环神经网络被广泛用于建模序列数据的时序上下文信息。但是, 这些操作都仅仅能够建模有限距离内的上下文信息, 而长距离上下文信息的建模才是影响行为识别模型的性能和效果的瓶颈所在。对于长距离上下文的建模也已经有



图 1 来自 UCF101 数据集的视频动作识别示意图[14]

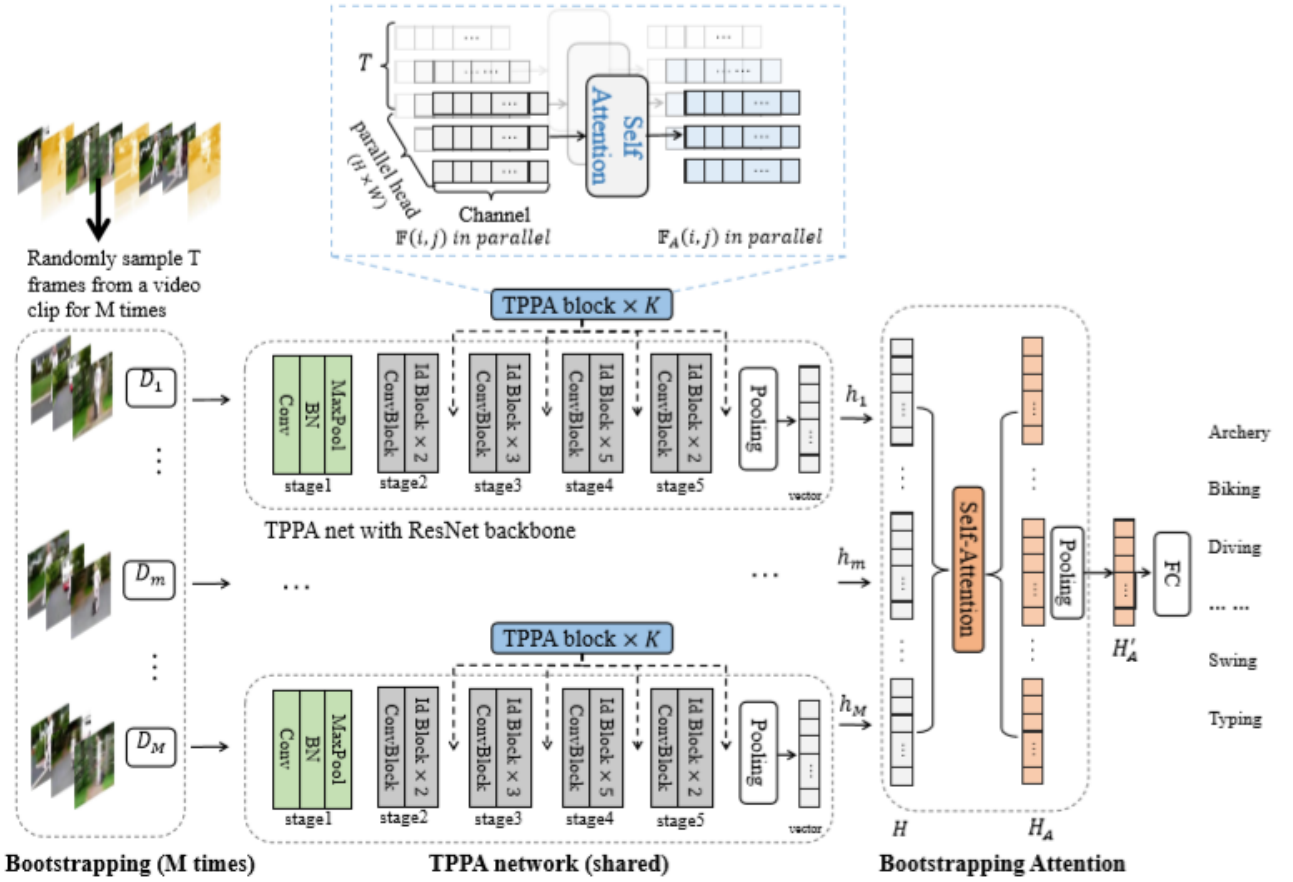


图 2 随机帧采样的集成注意力机制总体架构图

一些工作，例如，非局部网络（Non-local network）证明了在计算机视觉任务（目标检测，视频分类等）上建模长时序信息可以提升模型性能[8]。

此外，视频是高度冗余的数据形式，相邻帧之间可能非常相似。对于大多基于视频数据的行为识别模型（比如 C3D[5]，LSTM[7]），当这些模型在短时序上建模时，容易收缩为在单张图像上的操作。因此，探索长时序建模是另外一个基于视频的行为识别需要重点考虑的理由。

同时，行为通常可以看作是由同一个场景中多个局部特征之间的交互关系决定的，关系网络（Relation Network）证明了同一张图像的不同局部之间有紧密的关系[9]。而且，也有工作提出基于人体不同部位（局部）推理行为结果[10]；基于时空图模型的方法同样证明了局部特征对行为识别的作用[11]。因此，从局部特征角度建模上下文信息是行为识别研究的未来方向之一。

二、基于长时序行为建模的行为识别模型

下面介绍我们提出的一种新的基于长时序行为建模的行为识别模型，如图 2 所示。该工作通过两个途径来建模长时序上下文：(1) 基于自注意力机制（self-attention）[12] 的注意力模块。(2) 基于随机帧采样的自助法注意力（random frames based bootstrapping attention）行为识别框架[13]。

自注意力机制能够建模序列上任意位置与其他位置关系。简单的自注意力机制[12]可以由以下公式表示：

$$X' = \text{softmax}\left(\frac{X \cdot X}{\sqrt{d_k}}\right) \cdot X$$

其中， $X = \{x_i\}, i = 1, \dots, N$ ，是一个包含 N 个向量的矩阵； X' 是注意力操作的输出，形状大小与输入 X 一致； d_k 是其中向量的长度

虽然已经有工作将自注意力机制应用于计算机视觉任务，但是过去的工作仅仅通过注意力

机制建模 t 时刻图像所有像素与 $(t + 1)$ 时刻图像所有像素之间的关系，也就是将整个图像拉伸为一个向量。这种简单的方式不仅带来巨大的计算量而限制了应用范围，同时破坏了图像本身的空间特征。我们的工作充分考虑行为识别中对局部特征的利用，提出了基于时序像素的并行多头注意力模块（temporal pixel based parallel-head attention, TPPA），如图 3 所示。首先，我们定义了时序像素（temporal pixel）概念（所有时刻上在同一空间位置的像素集合， $T \times C$ 的矩阵），并以此为基本单位进行自注意力的计算。然后，图像的每个空间位置都进行同样的自注意力计算，所有空间位置上是并行的计算。这种在时序像素上操作的自注意力机制可以减少需要学习的参数量，让上下文的建模更容易优化。该模块建模长时序上下文是基于空间上局部特征进行，图像的空间特征得以保持。同时，整个操作被封装为一个模块，可以与任何流行的卷积神经网络结合使用。

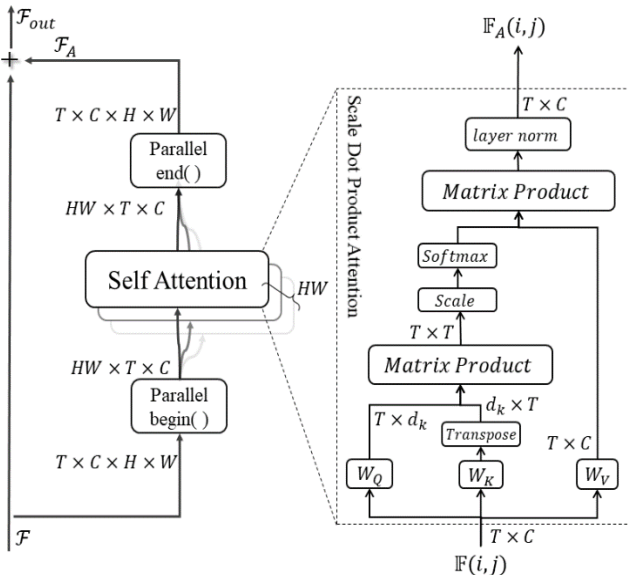


图 3 TPPA 模块的结构

对于大多数的行为类别，与行为识别相关的关键特征往往只是其中的几个关键帧，也就是说，我们可以根据关键帧实现行为识别。为了在一个冗余视频的长时序上下文中捕获这样的关键帧，

我们引入了统计学中的自助法（bootstrapping）来提取关键帧。因为关键帧无法被人为先验得到，采用自助法思想，多次采样又放回的方式，每次采样的帧中可能是噪音也可能是有用的关键帧。多组采样结果通过主干网络得到多组特征，再同样采用自注意力的方式对多组特征重新加权，通过这样的方式增强关键帧的特征，弱化噪音的特征。这种架构被称为基于随机帧采样的自助法注意力（Random Frames based Bootstrapping Attention），整体框架见图 2。这个框架在常见行为识别数据集上都取得了最佳的结果，并且和 TPPA 模块一样，RFBA 架构可以与主流的 3D 卷积网络或者双流网络结合使用。

最后，图 4 展示了我们的模型在最常见的行为识别数据集 UCF101[14]上的实验结果。可以看出，我们所提出的基于长时序动作建模的识别模型最终取得了更好的行为识别结果。

Model	modality	acc@1
Two-stream[24]	RGB	83.6
Two-stream[24]	RGB + flow	91.2
LSTM[5]	RGB	81.0
3D-fused[2]	RGB	83.2
3D-fused[2]	RGB + flow	89.3
I3D[2]	RGB	84.5
TPPA net[ours]	RGB	84.8
RFBA net[ours]	RGB	91.7

图 4 UCF101 数据集上的结果对比

三、总结

在这篇文章中，我们介绍了关于行为识别研究的最新进展，并且介绍了我们所提出的最新行为识别模型——基于随机帧采样的自助法注意力框架（RFBA）。在我们的方法中，我们提出了一个新的注意力模块，基于时序像素的并行多头注意力模块（TPPA），这个新的模块可以更有效地建模长时序上下文信息。

（责任编辑：任桐炜）

参考文献:

- [1] 朱红蕾, 朱昶胜, 徐志刚. 人体行为识别数据集研究进展[J]. 自动化学报, 2018, 44(06):20-46.
- [2] 罗会兰, 王婵娟, 卢飞. 视频行为识别综述[J]. 通信学报, 2018, v.39; No.372(06):173-184.
- [3] Wang, Heng, et al. "Dense trajectories and motion boundary descriptors for action recognition." *International journal of computer vision* 103.1 (2013): 60-79.
- [4] Simonyan, Karen, and Andrew Zisserman. "Two-stream convolutional networks for action recognition in videos." *Advances in neural information processing systems*. 2014.
- [5] Tran, Du, et al. "Learning spatiotemporal features with 3d convolutional networks." *Proceedings of the IEEE international conference on computer vision*. 2015.
- [6] Du, Wenbin, Yali Wang, and Yu Qiao. "Rpan: An end-to-end recurrent pose-attention network for action recognition in videos." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.
- [7] Donahue, Jeffrey, et al. "Long-term recurrent convolutional networks for visual recognition and description." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
- [8] Wang, Xiaolong, et al. "Non-local neural networks." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
- [9] Hu, Han, et al. "Relation networks for object detection." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2018.
- [10] Zhao, Zhichen, Huimin Ma, and Shaodi You. "Single image action recognition using semantic body part actions." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.
- [11] Wang, Xiaolong, and Abhinav Gupta. "Videos as space-time region graphs." *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018.
- [12] Vaswani, Ashish, et al. "Attention is all you need." *Advances in neural information processing systems*. 2017.
- [13] Ziming Liu, Guangyu Gao, A. K. Qin, Tong Wu, and Chi Harold Liu. 2019. Action Recognition with Bootstrapping based Long-range Temporal Context Attention. In Proceedings of the 27th ACM International Conference on Multimedia (MM '19), October 21–25, 2019, Nice, France.
- [14] Khurram Soomro, Amir Roshan Zamir and Mubarak Shah, UCF101: A Dataset of 101 Human Action Classes From Videos in The Wild., CRCV-TR-12-01, November, 2012.



高广宇

北京理工大学副教授, 数据科学与知识工程研究所副所长, 硕士生导师。主要研究方向为深度学习, 计算机视觉和多媒体。Email: guangyugao@bit.edu.cn



刘子铭

北京理工大学硕士研究生, 软件工程专业。主要研究方向为计算机视觉。
Email: liuziming.email@gmail.com



吴桐

北京理工大学硕士研究生, 计算机技术专业。主要研究方向为计算机视觉。
Email: 3220190896@bit.edu.cn

人体、人脸、人手表面及运动重建

清华大学 软件学院 徐枫 张浩

人体及人体局部区域（人脸和人手）的三维表面和运动重建一直以来就是计算机视觉(CV)和计算机图形学(CG)领域的热点研究问题，在教育、电影、游戏、动画等领域有着非常广泛和深远的应用。最近几年，随着 VR/AR，人工智能和 5G 技术的蓬勃发展，人们更期望借助于技术的发展实现远距离全三维通信、虚拟试衣、实时在线三维直播、甚至实现智能机器人对人体各种运动的理解。因此，人体及其局部区域表面及运动重建技术越发得到了学术界和工业界的重视。

1. 人体重建

人体重建可以使用包含深度信息的相机对人体进行拍摄，并通过拍摄得到的图像序列重建出人体的三维几何模型与运动。根据应用的需求也可以重建出人体的表面颜色反射属性或者周围场景的几何模型。

我们的方法[1]可以使用 RGB-D 相机，即颜色-深度相机实现对人体等动态物体的实时重建，得到物体的几何模型、非刚性运动、表面反射率以及环境光照。在该方法中，对于每一帧的输入，首先使用迭代最近点匹配（Iterative Closest Point, ICP）算法求得相机相对于物体的运动；之后对物体的非刚性运动以及环境光照进行联合求解；最后更新物体的几何模型和表面反射率。物体的非刚性运动表达为表面上稀疏的节点带动空间中的体素进行运动，通过线性展开将对运动的求解化为线性问题，然后使用高斯牛顿法进行求解以达到实时重建的效果。表面反射则使用球谐函数表示的环境光与朗伯表面来进行表达。

我们也能够使用深度相机对动态的物体与其周围的静态场景进行实时重建[2]，可以用于重建室内的动态的人与静态场景。该方法将重建

场景和采集的数据分为动态和静态两部分，两部分各自进行运动求解和模型的更新。在该方法中，动态部分的非刚性运动同样使用表面上稀疏的节点进行表达。此外，为了降低重建所占用的空间开销，该方法还使用了哈希体素来管理空间中体素的分配，并进一步使用非刚性运动节点来控制动态物体的体素分配。

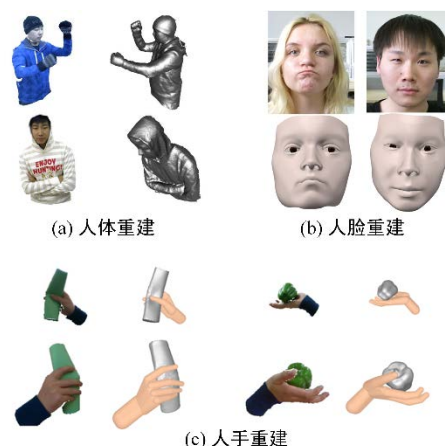


图 1 人体及局部区域重建结果

2. 人脸重建

脸部是人最重要的器官之一，它可以表达人类的身份特征和丰富的情感变化。而在五官之中，眼睛是心灵之窗，它反映了人的关注点、心理活动变化等等。因此人脸的三维重建，尤其是眼部区域的三维重建在计算机图形学和计算机视觉领域都有着重要的意义和广泛的应用。

我们首先利用一套多线性人脸模型和人脸二维特征点进行 RGB 图像中的人脸重建，即将人脸模型上预定义的特征点投影到图像中与二维特征点取得一致。然而，人脸模型对于眼部区域的表达能力十分有限，其中不包含三维眼球模型，且眼皮的形状也与真实图像中相差较大。因此第一步[3]，我们在人脸模型中眼眶的相应位置上

增加了眼球模型来重建眼球的转动。我们根据图像中眼部区域的颜色分布估计眼球模型中虹膜与巩膜的颜色，而后利用基于三维模型表面泰勒展开的光度误差优化方法，通过减小三维眼球投影与二维图像差异的形式重建眼球运动。第二步[4]，我们设计了两组分别表示眼皮的形状特征和运动变化的线性眼皮模型来进行眼皮的三维重建。在图像中，我们利用深度学习的方法提取眼部区域四条主要边界作为二维眼皮特征，包括双眼皮、上眼皮、下眼皮和卧蚕。之后我们通过减小三维眼皮特征点与二维特征投影误差的形式，求解最优的眼皮模型的权重系数，并按其进行线性叠加得到最终的重建结果。最后我们将三维眼球与眼皮的重建结果融合到人脸的重建结果中，使得整体效果更加真实生动。

3. 人手重建

手是人体的重要组成部分，是人与环境进行交互最主要的执行工具。对手和物体的交互过程进行重建对于人体行为的精细化重建，以及机器理解人与环境的交互行为具有重大的意义。我们使用深度相机对交互过程进行拍摄，并通过拍摄得到的图像序列重建出人手的运动，物体的三维几何模型和运动（包含刚性和非刚性运动）[5]。

手与物体交互重建存在许多挑战。这些挑战包括：（1）遮挡，手的自遮挡和手与物体之间的互遮挡；（2）数据特征少，难以通过简单的颜色或深度阈值将手与物体数据分割；（3）手的特征少、数据少、运动复杂，难以进行精准的运动重建；（4）物体的非刚性运动维度高，且没有几何模型。为了解决上述问题，我们首先使用两台相对放置的深度相机对交互过程进行采集，降低遮挡对重建过程的影响；然后训练了一个深度神经网络来对手和物体的深度数据进行识别和分割，为手部运动重建提供指导；其次，借助手的 SphereMesh 模型进行手部重建，综合考虑三维点云拟合误差，二维轮廓匹配误差和手运动的先验信息等，可较好的重建手的运动；最后，借助 DynamicFusion 进行物体表面和非刚性运动重建，针对被手遮挡的物体区域，我们又引入了交互项，从而获得了较为准确的物体几何和非刚性运动。最终实现了手与物体交互过程的重建。

以上系列工作已发表于国际会议 SIGGRAPH 2019, SIGGRAPH 2017, SIGGRAPH Asia 2017, TVCG 2018, TVCG 2016。

（责任编辑：邓成）

参考文献

- [1] Guo K, Xu F, Yu T, et al. Real-Time Geometry, Albedo, and Motion Reconstruction Using a Single RGB-D Camera[J]. ACM Transactions on Graphics, 2017, 36(3):1-13.
- [2] Zhang H, Xu F. MixedFusion: Real-Time Reconstruction of an Indoor Scene with Dynamic Objects[J]. IEEE Transactions on Visualization and Computer Graphics, 2018:1-1.
- [3] Wen Q, Xu F, Yong J H. Real-time 3d eye performance reconstruction for rgb-d cameras[J]. IEEE transactions on visualization and computer graphics, 2016, 23(12): 2586-2598.
- [4] Wen Q, Xu F, Lu M, et al. Real-time 3d eyelids tracking from semantic edges[J]. ACM Transactions on Graphics (TOG), 2017, 36(6): 193.
- [5] Zhang, H., Bo, Z. H., Yong, J. H., & Xu, F. (2019). InteractionFusion: real-time reconstruction of hand poses and deformable objects in hand-object interactions. ACM Transactions on Graphics (TOG), 2019, 38(4), 48.



徐枫

清华大学副教授，主要研究方向为人体及场景三维重建。
Email: feng-xu@tsinghua.edu.cn



张浩

清华大学在读博士生，主要研究方向为人体及人手三维运动重建。Email: zhanghao16@mails.tsinghua.edu.cn

可微分网络架构搜索的稳定性研究

华为 诺亚方舟实验室 谢凌曦

目前，网络架构搜索已经成为自动化机器学习 (AutoML) 领域的重要研究方向。常见的网络架构搜索算法分为两大类，即启发式网络架构搜索和可微分网络架构搜索。相比于启发式搜索算法对计算资源的极度依赖，可微分搜索算法具有搜索速度快、能够快速部署到较大的搜索空间上的优势，因而得到了学界的广泛关注。然而，可微分搜索算法的稳定性一直饱受质疑，这也在一定程度上限制了它们在实际问题中的应用。

在最近的研究中，我们提出了一种假设，将网络架构搜索的不稳定性归结为搜索过程中优化的超网络和搜索结束后保留的子网络之间的结构差异——亦即，最优的超网络并不一定对应于最优的子网络，但是现有的算法大多忽略了这一点。为此，我们提出了一系列的改进方法，以消除这类结构差异带来的负面效果。不失一般性，我们的工作基于 DARTS，一种通用的可微分搜索算法。

在第一个研究课题中，我们注意到了搜索过程中超网络的深度和搜索结束后保留的子网络的深度存在较大差异。也就是说，当前的搜索算法假设不同深度的情况下最优网络具有相同的局部结构，而这显然是不合理的。为此，我们提出了一种渐进加深的搜索算法，使得搜索阶段临近结束时的超网络逼近最终使用的子网络的深度。同时，为了使得超网络搜索适应不同的深度，我们还提出了几种实用的训练策略以增加搜索过程的稳定性。我们的方法在单张 GPU 上只需要 7 个小时就能够完成搜索，并且在 CIFAR 和 ImageNet 数据集上都取得了稳定的提升。这一工作[1]被 ICCV 2019 接收为口头报告论文。

在后续的研究课题中，我们又注意到了超网

络和子网络的拟合性质存在较大差异。超网络中的每个连接都包含有多种不同的基本操作，但是子网络的每个连接只能从这些基本操作中选择一个。这就决定了超网络能够以不同的方式拟合训练数据集，但是这些方式不一定都适用于子网络，因而子网络的最终性能存在较大差异。为此，我们以通道采样的方式来限制超网络的能力，使得过拟合现象得到很大程度的缓解。同时，这种采样方法还降低了搜索过程的计算开销，并且通过增加每个批次内的样本数，提升了搜索的稳定性。我们在单张 GPU 上的运行时间能够进一步下降到 2 个小时以内，并且在 CIFAR 和 ImageNet 上取得很好的效果。这一工作[2]已经提交到 ICLR 2020。

我们最新的一项研究表明，除了上述差异，现有的可微分网络架构搜索还面临一个更加严重的问题，即由分步梯度优化带来的数学近似。我们发现这种近似方式将会带来极大的误差，然而现有的方法都无法直接消除这一误差，只能以早停的方式来减轻误差的伤害。由于搜索过程尚未收敛便要结束，因此随机初始化就会很大程度上影响最后的结果，从而带来搜索的不稳定性质。针对这一问题，我们提出了一种修正误差的方案，以一种巧妙的方式提升了梯度近似的精度。在这一修正的搜索算法下，我们能够确保搜索算法在收敛后仍然取得令人满意的性能。这一工作已经提交到 ICLR 2020 [3]。

总结来说，可微分网络架构搜索是一个迷人而充满前景的研究方向。然而，这套方法存在的问题还有很多，自动化机器学习距离真正的一统天下还有很长的路要走。

(责任编辑：苏航)

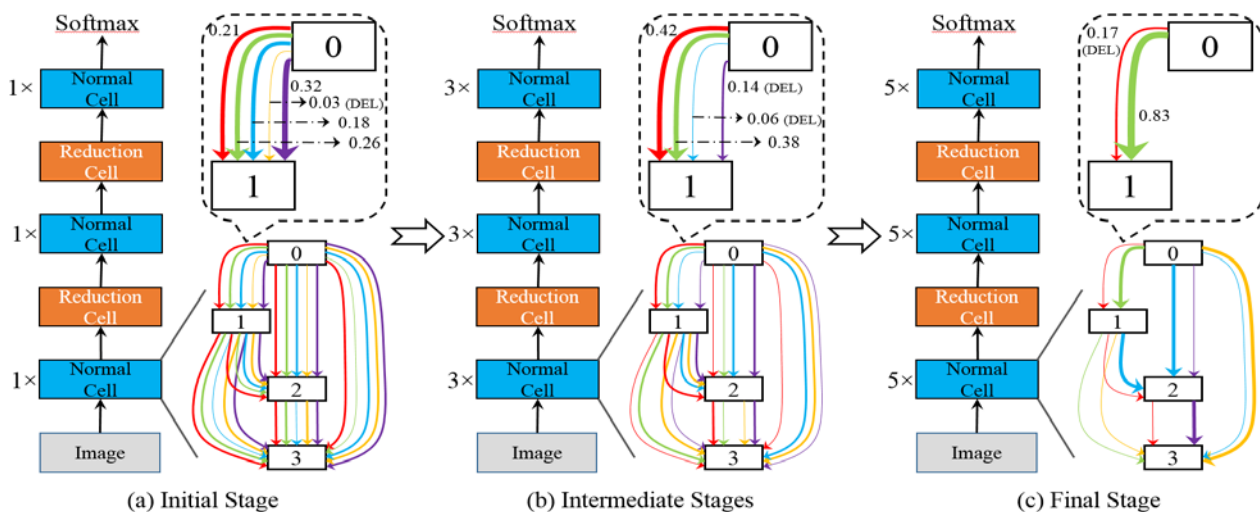


图 1 可微分网络架构搜索总体架构图

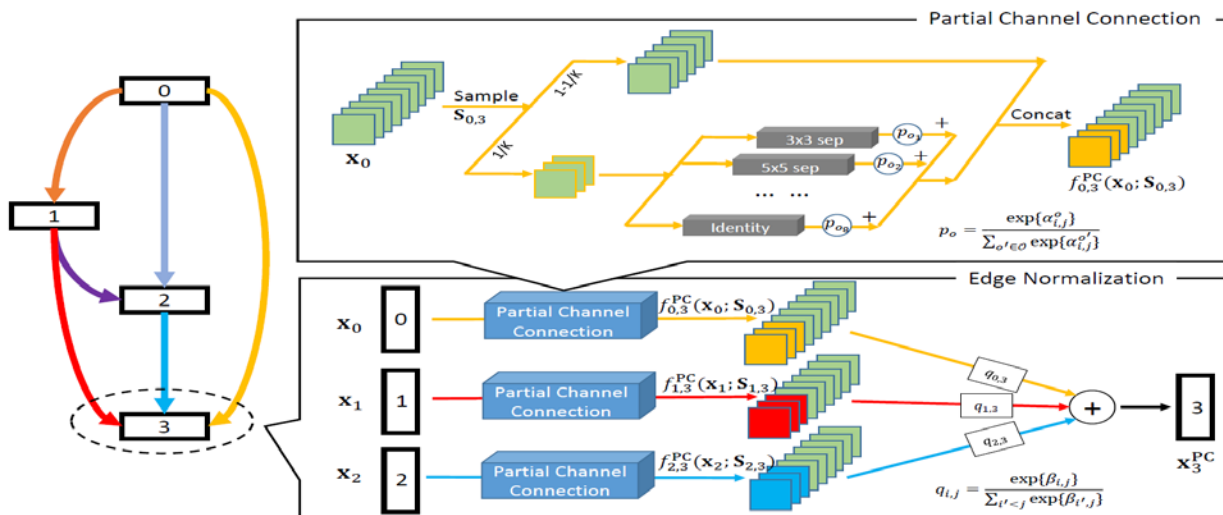


图 2 部分通道连接的可微分网络架构搜索

参考文献

- [1] X. Chen et al., Progressive Differentiable Architecture Search: Bridging the Depth Gap between Search and Evaluation, ICCV, 2019.
- [2] Y. Xu et al., PC-DARTS: Partial Channel Connections for Memory-Efficient Differentiable Architecture Search, arXiv preprint: 1907.05737, 2019.
- [3] K. Bi and C. Hu et al., Stabilizing DARTS with Amended Gradient Estimation on Architectural Parameters, arXiv preprint: 1910.11831, 2019.



谢凌曦

华为诺亚方舟实验室高级研究员，主要研究方向是计算机视觉。

Email: 198808xc@gmail.com

南京理工大学潘金山教授访谈

2019年9月20日，专委秘书处采访了南京理工大学博士生导师、2019年国家优秀青年科学基金获得者潘金山教授。下面是采访实录。

潘老师，您好！首先，请您介绍一下您的研究经历，分享一下您的研究经验。

首先，非常感谢专委会为我们提供这样一个交流平台。我本科的专业是信息与计算科学，硕博的专业是计算数学。2017年博士毕业于大连理工大学数学科学学院。曾在哈佛大学、加州大学 Merced 分校访问过。目前主要从事图像去模糊及相关复原和增强问题的研究。我从本科到博士读的都是数学相关专业，第一次接触计算机视觉相关问题，是从我研究生阶段师从大连理工大学苏志勋老师开始的。那时候第一次感受到数学的魅力不仅在于漂亮的公式和定理推导，许多曾经学过的优化方法和理论也可以用来帮助解决计算机视觉的相关问题。我一直认为打好数学的基础对之后从事计算机视觉相关问题的研究是十分重要的，所以我从本科到博士阶段的相关数学专业课程也为我之后的研究打下了一定的基础。在攻读博士期间，我通过国家留学基金委提供的联合培养方案到加州大学 Merced 分校和哈佛大学进行了访问学习。也是这段学习经历让我更深入快速地接触到了计算机视觉方向的前沿，与许多教授前辈们也产生了很多想法上的火花。所以多交流讨论对于我们做科学研究来说也是十分重要的。

您在图像去模糊、图像复原和增强方面有很多的研究成果，能否跟大家谈谈该领域的一些研究现状和未来发展趋势呢？

图像去模糊及相关复原和增强问题是一个典型的逆问题，其研究目标是从给定的低质量的

图像中有效地计算出高质量的图像。它一方面有助于解决低质量图像难于分析的问题，另一方面使得复原的图像以较好的视觉效果呈现出来，提高终端用户体验或者有利于后期编辑。目前大部分工作基于已有的物理退化模型，从给定的退化图像中挖掘有效的先验知识来估计退化图像到清晰图像的映射。因此，物理退化模型以及先验知识是这一问题的核心。传统的方法大多通过先验建模的方式来解决这一问题。

近年来，深度学习在解决高层视觉问题中取得了很大的成功。受此启发，一些学者开始尝试用深度神经网络解决图像复原及相关底层视觉问题。这类方法通常以要处理的图像/视频等数据作为网络的输入，在不经人为干预的条件下，直接从观测图像/视频中估计出潜在清晰图像/视频。深度学习给解决图像复原和增强提供了一种有效的方法，并且在某些问题（比如超分辨率）上面取得了显著的进展。这也是目前研究相对较多的一类方法。

关于未来的发展趋势，在这里我结合自己的一些研究，说一下自己的看法。考虑到实际退化过程比较复杂，如何从物理成像过程的角度理解退化机理并给出有效的物理成像模型是目前解决相关问题的一个趋势。其次，由于传统的先验建模的方法大多依赖于人工设计的特征，而传统的深度学习的方法大多基于“黑箱”的工作方式，所带来的一个核心的问题就是对于埋藏在深度学习背后的运行机理没有研究清楚。因此，如何将两者有效地结合起来，充分利用两类方法的优点是当前研究的一个趋势。再次，虽然高质量的图像是图像复原和增强所追求的目标，但是，一方面，现有的评价高质量图像的客观指标和主观指标存在着鸿沟；另一方面，复原出的高质量图

像能否真正有助于后续智能化分析问题（如人脸识别、车牌检测与识别等）还有待进一步研究。

您近期在 VALSE 上组织了一期去雨与雾专题“雨雾云收望远山”，作为主办 AC，能跟我们介绍一下您当时的想法么？您对这期的专题有何评价呢？

这一期主要由刘家瑛老师、赵洋老师和我来主办。雾霾、雨等因素是造成户外图像退化的主要因素，在这些环境下拍摄的图像往往质量较低，从而严重影响了后续相关智能化分析任务。近年来关于图像去雾、去雨相关工作被大量提出，相关研究受到了广泛的关注。这个专题的目的是更好地使相关老师和同学了解问题研究现状和未来发展趋势。

我们邀请了相关领域的专家，包括张赫、任文琦、孟德宇、左旺孟等老师。这几位老师对图像去雾、去雨等相关复原问题进行了深入探讨。通过这一专题，专家们对近几年图像去雨和去雾的发展趋势给出了细致的总结和建议，为未来的研究提供了宝贵的建议。

您 2016 年底曾在商汤科技工作过一段时间，能介绍下您与商汤的合作情况及您的主要工作内容么？您是如何看待或评价此段经历的？

商汤科技是 AI 领域的领军企业，在计算机视觉等相关问题上有深厚的技术积累。我在商汤科技的那段时间主要从事图像去模糊和超分辨率问题的研究。通过这段经历，进一步了解到自己所从事的研究和实际需求之间的关联性。这一经历给我最深的启发就是以往自己都是针对一些 well-defined 的问题进行研究，而实际问题往往不是 well-defined，如何从实际问题中凝练出研究问题，从而给出有效地解决方案是很关键的。

您于 2015 年期间在哈佛大学做过一段时间的访问学者，您能介绍下您在这期间的研究经历吗？可否和大家分享一下您对哈佛大学研究氛

围的分析和评价？

我在哈佛大学的视觉计算组跟 Hanspeter Pfister 教授和 Deqing Sun 博士从事图像/视频去模糊问题的研究。

我所在的研究组主要从事计算机图形学、计算机视觉等相关问题的研究。组内报告和交流比较多。每周都会有相应的组会，在组会上每个人都会分享一些自己所看的论文以及研究成果。除此之外，经常会有相关领域的知名学者来组里进行报告交流。通过聆听不同领域的报告以及与这些学者的交流，可以有效地拓宽自己的研究视野，同时也会在交流中产生一些新的想法。

能介绍一下您现在的研究团队，以及您对团队的管理和规划么？

南京理工大学智能媒体分析实验室 (IMAG) 是唐金辉教授建立的一个科研团队。主要成员包括多媒体分析与理解、计算机视觉等方向的老师和同学。我们的研究团队有浓厚的学术氛围，每年都有研究成果在相关领域顶级期刊和顶级会议发表。关于团队的管理和规划，我目前也正在探索和学习中。

您在 TPAMI、CVPR 等国际知名期刊/会议上发表了很多高水平的论文，能为研究生们如何冲刺高水平论文提供一些建议么？

高水平的论文离不开论文本身的质量以及论文的写作。论文的质量离不开知识的积累。关于我自己的科研经历，在这里主要感谢大连理工大学苏志勋教授、北京大学林宙辰教授以及加州大学 Merced 分校的 Ming-Hsuan Yang 教授的指导。他们做出了很多高水平的研究工作并且也将高水平论文的经验和建议放在他们的主页以及相关论文写作的课程上。具体可以参考 <http://faculty.ucmerced.edu/mhyang/course/eecs286-2016/index.htm>。



潘金山

南京理工大学计算机学院教授、博士生导师。2017年毕业于大连理工大学数学科学学院，获博士学位。曾在哈佛大学和加州大学 Merced 分校访问研究两年。主要从事图像去模糊、图

像复原等相关底层视觉问题的研究。近年来，在 CVPR、

ICCV、ECCV 等国际会议以及 IEEE TPAMI、IJCV 等重要国际期刊上发表论文 40 余篇，其中 CCF 推荐 A 类论文 30 篇。根据 Microsoft Academic 统计的近 5 年最具影响力的全球学者排位，位列图像去模糊领域第 2 位。获得了中国人工智能学会优秀博士学位论文奖以及辽宁省优秀博士学位论文奖。2019 年获得国家优秀青年科学基金资助。

（责任编辑：余烨 黄岩 张汗灵）

委员好消息

✪ 2019 年 7 月 11 日，CCF-CV 专委会委员、南京大学任桐炜副教授指导的团队在 ICME 2019 Short Video Understanding 挑战赛中获得第 2 名。

✪ 2019 年 8 月 2 日，CCF-CV 专委会委员、南京大学任桐炜副教授指导的团队在 SHREC 2019 挑战赛 Monocular Image-based 3D Model Retrieval 任务中获得优胜奖。

✪ 2019 年 8 月 2 日，国家自然科学基金委公布了 2019 年度国家杰出青年科学基金建议资助项目申请人名单，共 300 人获得资助，CCF-CV 专委会 3 位委员获得资助：中国科学院西安光学精密机械研究所卢孝强研究员、北京大学彭宇新教授、南京理工大学唐金辉教授。

✪ 2019 年 8 月 16 日，2019 年国家自然科学基金创新研究群体项目评审结果揭晓，34 家单位获得 45 项创新群体研究项目，CCF-CV 专委会委员、东南大学孙长银教授主持的“自主无人系统协同控制理论及应用”项目入选并获得单项目 1050 万资助。

✪ 2019 年 9 月 29 日，中国计算机学会公布了 2019 年度“CCF 科学技术奖”评选结果，共 18 个项目获奖。CCF-CV 专委会委员、大连理工大学卢湖川教授等人完成的项目“复杂场景下视觉目标跟踪理论及方法”获 CCF 科学技术奖自然

科学二等奖，该项目提出了基于稀疏表示的视觉目标跟踪算法，改善了目标跟踪的鲁棒性，并提出目标的超像素结构表达、连续异常噪声模型及判决性局部度量模型等，为目标外观结构建模提供了新方法。“CCF 科学技术奖”授予在计算机科学、技术或工程领域具有重要发现、发明、原始创新，在相关领域有一定国际影响的优秀成果。

✪ 2019 年 10 月 10 日，86 位在各自领域做出一定成就或对 CCF 服务方面有突出表现的专业会员晋升为杰出会员，其中包括 CCF-CV 专委会 6 位委员：北京大学彭宇新教授、中国科学院自动化研究所王亮研究员、同济大学王瀚涛教授、中国科学院心理研究所王甦菁副研究员、华南理工大学许勇教授、中国科学院自动化研究所张兆翔研究员。

✪ 2019 年 10 月 19-20 日，CCF 第十二次会员代表大会在苏州举行，CCF-CV 专委会 15 位委员当选 CCF 理事：合肥寰景信息技术有限公司陈拥权董事长、中国海洋大学董军宇教授、中兴通讯中央研究院董振江院长、北京理工大学黄华教授、中国科学院大学黄庆明教授、西安电子科技大学苗启广教授、太原理工大学强彦教授、字节跳动王长虎博士、爱奇艺公司王涛高级研究员、北京航空航天大学王蕴红教授、哈尔滨工业大学邬向前教授、华南理工大学许勇教授、北京交通

大学于剑教授、华南理工大学余志文教授、西北工业大学张艳宁教授。

✪ 2019 年 10 月 21 日,2019 年度吴文俊人工智能科学技术奖名单公布,共 81 个成果获奖,其中 CCF-CV 专委会 7 位委员的 6 项成果获奖:东南大学孙长银教授等完成的“智能无人系统的控制理论与方法”获吴文俊人工智能自然科学一等奖,重庆大学张磊研究员等完成的“机器嗅觉感知理论与智能计算方法”获吴文俊人工智能自然科学二等奖,中国科学院自动化研究所张兆翔研究员和中国科学院信息工程研究所葛仕明研究员等合作完成的“基于图像视频分析的智能监测技术与应用”获吴文俊人工智能技术发明二等奖,中国科学院深圳先进技术研究院乔宇研究员等完成的“视频序列的深度表征与理解技术及应用”获吴文俊人工智能科技进步二等奖,深圳大学沈琳琳教授等完成的“图像特征提取与增强方法及应用”获吴文俊人工智能自然科学三等奖,复旦大学张文强教授等完成的“光伏电站一站式智能运维关键技术与应用”获吴文俊人工智能科技进步三等奖。

✪ 2019 年 10 月 23 日,CCF 公布了 2019 年度第二批新晋升高级会员名单,47 位在各自领域做出一定成就或对 CCF 服务方面有突出表现的专业会员晋升为高级会员,其中包括 CCF-CV 专委会 4 位委员:中国海洋大学董军宇教授、广州智慧城市发展研究院胡建国教授、国防科技大学刘丽副教授、上海科技大学虞晶怡教授。

✪ 2019 年 10 月 24 日,CCF-CV 专委会委员、南京大学任桐炜副教授指导的团队在 ACMMM 2019 Relation Understanding in Videos 挑战赛 Video Relation Detection 任务中喜获第 1 名,在 ACMMM 2019 Content-based Video Relevance Prediction 挑战赛 Movie 任务中获得第 2 名。

✪ 2019 年 10 月 27 日,CCF-CV 专委会委员、北京大学林宙辰教授 ZERO 实验室组成的队伍取得了 COCO Panoptic Challenge 的创新奖,今年的 COCO 竞赛只颁发了最高分奖项和创新奖项。

✪ 2019 年 10 月 28 日,CCF-CV 专委会委员、

中科院自动化所雷震研究员指导的团队喜获 Light Weight Face Recognition Challenge 2019 挑战赛 DeepGlint-Large Track 任务第 1 名。

✪ 2019 年 10 月 28 日,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、中科院自动化所黄岩博士指导的团队获得 WIDER Face and Person Challenge 2019 挑战赛 Person Search by Language 任务第 1 名。

✪ 2019 年 10 月 28 日,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、中科院自动化所黄岩博士与微软合作指导的团队摘取 VOT 2019 挑战赛 Visual Object Tracking RGBD 任务桂冠;CCF-CV 专委会常务委员、华中科技大学白翔教授与百度联合指导的团队摘得 VOT 2019 挑战赛 VOT 2019 Short Term Challenge 任务桂冠;CCF-CV 专委会委员、大连理工大学卢湖川教授指导的团队摘得 VOT 2019 挑战赛 VOT 2019 Long Term Challenge 任务桂冠。

✪ 2019 年 10 月 29 日,CCF-CV 专委会公布了 2019 年度 CCF-CV“中科视拓 Seeta 学术新锐奖”获奖者名单,共评出 3 位获奖者,其中 CCF-CV 专委会委员、清华大学鲁继文副教授指导的博士生饶永铭获奖。

✪ 2019 年 10 月 29 日,CCF-CV 专委会委员、中山大学林惊教授、李冠彬副教授指导的论文 *Larger Norm More Transferable: An Adaptive Feature Norm Approach for Unsupervised Domain Adaptation* (<https://arxiv.org/abs/1811.0745>) 获 ICCV 2019 最佳论文提名,这是入选论文中唯一一篇来自亚洲的论文。该文提出了一种基于特征范数适配的无监督领域自适应方法,该方法将目标域特征范数逐渐增大,结合源域的分类损失,实现了从源域到目标域的无监督迁移。该方法中的特征范数适配操作可以避免进行显式的领域对齐,针对源域和目标域中的不相交类别具有更强的鲁棒性,相比一类启发式加权算法有了显著提升。此外,该方法在大规模和领域间差异较显著的迁移任务中也获得了较好的性能。

✪ 2019 年 10 月 31 日,北京科委公布了科技领

军人才、科技新星计划等入选名单，共 190 人，其中 CCF-CV 专委会 5 位委员入选。北京邮电大学马占宇副教授获得北京市科技新星计划交叉学科合作课题资助，中科院计算所阚美娜副研究员、北京大学连宙辉副教授、北京航空航天大学刘祥龙副教授、中科院自动化所张俊格助理研究员入选北京市科技新星计划。

✿ 2019 年 11 月 9 日，CCF-CV 专委会工作会议在西安召开，会上，CCF-CV 专委会副主任、中科院计算所陈熙霖研究员获杰出成就奖，CCF-CV 专委会 3 位委员：中科院计算所王瑞平研究员、西北工业大学魏巍副教授、南开大学杨巨峰副教授获 CCF-CV 中科视拓 Seeta 服务贡献奖。

✿ 2019 年 11 月 9 日，CCF-CV 专委会在西安举行换届选举，北京大学查红彬教授当选专委会主任，南京信息工程大学刘青山教授、中科院自动化所王亮研究员、上海科技大学虞晶怡教授当选专委会副主任，北京邮电大学马占宇副教授当选专委会秘书长。

✿ 2019 年 11 月 10 日，PRCV2019 举办颁奖典礼。CCF-CV 专委会委员、北京理工大学黄华教

授等的论文 An Effective Network with ConvLSTM for Low-Light Image Enhancement 获得最佳论文奖；CCF-CV 专委会委员、上海科技大学虞晶怡教授等的论文 Deep Eyes: Binocular Depth-from-Focus on Focal Stack Pairs 获得最佳论文提名奖；CCF-CV 专委会委员、北京航空航天大学王蕴红教授指导的论文 A Temporal Attentive Approach for Video-Based Pedestrian Attribute Recognition 获得最佳学生论文奖；CCF-CV 专委会委员、中科院西安光机所卢孝强研究员指导的论文 Deep Voice-Visual Cross-Modal Retrieval with Deep Feature Similarity Learning 获得最佳学生论文提名奖；CCF-CV 专委会副主任、北京大学查红彬教授指导的论文 Visual Odometry with Deep Bidirectional Recurrent Neural Networks 获口头报告最佳展示奖；CCF-CV 专委会委员、西北工业大学夏勇教授指导的论文 Knowledge-based Collaborative Deep Learning for Benign-Malignant Lung Nodule Classification on Chest CT 获得博士生论坛海报最佳展示奖。

（责任编辑：刘海波）

图像重光照开源代码

北京电子科技学院 宁宁 金鑫, 北京航空航天大学 吴洪宇

图像重光照指的是直接改变图像场景中物体的光照效果或者估计场景光照, 用光照维度对图像进行增强, 并保留图像物体视觉特征的技术。

图像的重光照技术相对于传统的几何建模方法, 无需美工对真实物体的精细建模和光照设计就能真实有效地模拟出物体在新的光照条件下的成像, 大大降低了场景构建的复杂度和渲染时间, 并且具有更高的场景真实感。

利用较少规模的图像数据生成逼真的光照效果, 主要包含两个方面: 一个方面是光照效果生成, 主要是直接改变图像中物体光照效果; 另外一方面是虚实光照融合, 从图像中估计场景光照条件, 然后将三维模型根据图像场景的视角和光照条件进行绘制并融入图像场景, 使得三维模型的光照效果与场景光照保持一致。

本文着重从光照效果生成和虚实光照融合两个方面展开介绍几个图像重光照的开源代码。

1. Post-production Facial Performance Relighting using Reflectance Transfer

论文: Peers P, Tamura N, Matusik W, et al. Post-production facial performance relighting using reflectance transfer[J]. ACM Transactions on Graphics, 2007: 26(3): 521-529.

工作: 该论文介绍的是人脸图像光照迁移。该论文使用商图从图像中提取参考人脸的光照信息, 获得参考人脸的商图后, 使用图像变形方法将参考人脸商图与输入人脸对齐。最后将变形后的商图与输入图像逐像素相乘, 实现光照效果迁移。图1为基于商图的人脸光照迁移结果, 图2为



图1 基于商图的人脸光照迁移结果

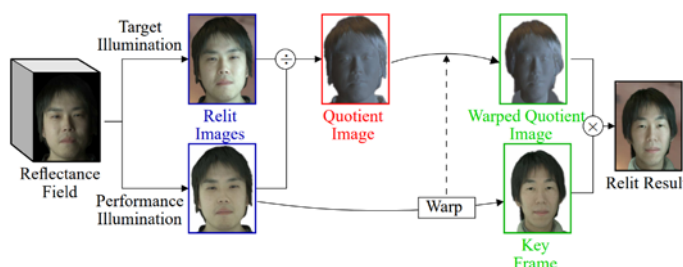


图2 人脸光照迁移算法示意图

论文: http://gl.ict.usc.edu/Research/PFPR/SIGGRAPH2007_PFPR_low.pdf

代码: <https://github.com/AjayNandoriya/FacialRelightingSig07>

2. Face illumination transfer through edge-preserving filters

论文: Chen X, Chen M, Jin X, et al. Face illumination transfer through edge-preserving filters[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2011: 281-287.

工作: 该论文提出基于滤波器的方法, 将单幅参考图像的光照迁移到光照均匀的输入图像中, 该方法首先使用图像滤波器将图像低频部分和高频信号进行分离, 低频部分看作是光照信息, 高频信息看作是人脸细节信息。并将输入人脸图像的低频信息替换成参考人脸的高频信息, 结果如图3所示。图4为基于图像滤波器的人脸光照迁移工作流程。



图3 基于图像滤波器的人脸光照迁移结果

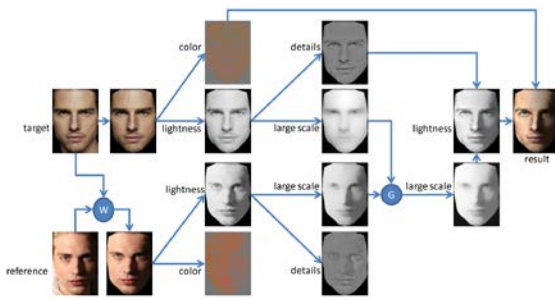


图4 基于图像滤波器的人脸光照迁移工作流程

论文: <http://jinxin.me/downloads/papers/003-CVPR2011/CVPR2011-CameryReady.pdf>

代码: <https://github.com/AjayNandoriya/EPRESERVE>

3. Style transfer for headshot portraits

论文: Shih Y C, Paris S, Barnes C, et al. Style transfer for headshot portraits[J]. ACM Transactions on Graphics, 2014: 33(4): 1-14.

工作: 该论文的主要创新点是提出了基于拉普拉斯金字塔的图像光照风格转换, 利用图像融合算法, 实现风格转换。图5为光照迁移效果, 图6为算法流程图。

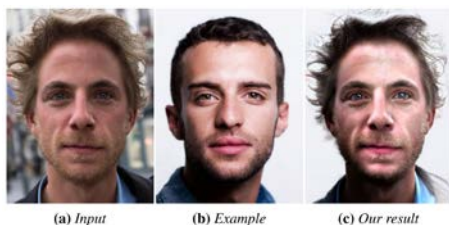


图5 光照迁移效果

论文: http://people.csail.mit.edu/yichangshih/portraitweb/2014_portrait.pdf

代码: <https://github.com/taivop/shih-styletransfer>

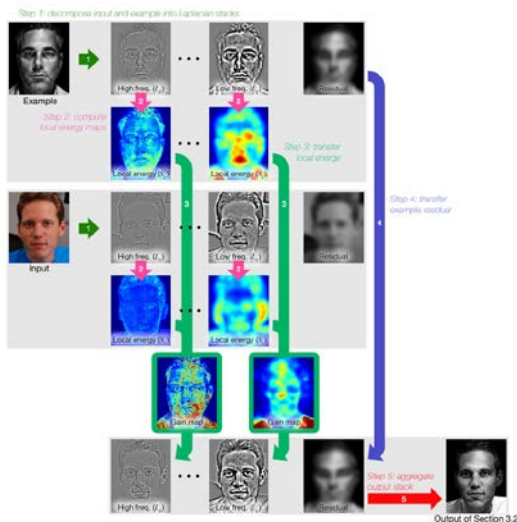


图6 算法流程图

4. Learning Shape, Reflectance and Illuminance of Faces in the Wild

论文: Sengupta S, Kanazawa A, Carlos D, et al. Learning shape, reflectance and illuminance of faces in the wild[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 6296-6305.

工作: 该论文通过输入一张在真实世界中拍摄的环境不受限的RGB人脸图像, 在朗伯假设下, 将图像分解为形状, 反射和光照三方面的信息。首先输入一张图像, 通过卷积神经网络提取出图像的feature map。然后法向特征与反照率特征分别经过卷积后生成最终的Normal和Albedo。这里Normal和光照经过渲染后生成Shading图像, 这个图像就是生成的三维形状的白模在光照下的情况。最终通过形状, 光照, 反照率生成重建的图像。图7为光影迁移结果图, 图8为SFSNet网络结构图。



图7 光影迁移结果

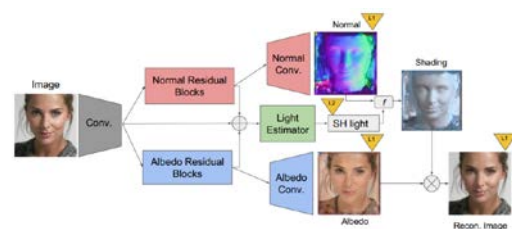


图8 SFSNet 网络结构

论文: <http://people.eecs.berkeley.edu/~kanazawa/papers/sfsnet.pdf>

代码: <https://github.com/bhushan23/SfsNet-PyTorch>

5. Portrait Lighting Transfer Using a Mass Transport Approach

论文: Shu Z, Hadap S, Shechtman E, et al. Portrait lighting transfer using a mass transport approach[J]. ACM Transactions on Graphics, 2017: 37(1): 1-15.

工作：该论文提出了一种基于质量传递（mass transport）人脸光影迁移方法，如图9所示。该方法根据图像重建人脸三维信息，然后在人脸三维信息的约束下将参考人脸的光影信息通过质量传递方法迁移到输入人脸。



图 9 基于质量传递的光影迁移结果

论文： http://www3.cs.stonybrook.edu/~cvl/content/papers/2017/shu_tog2017.pdf

代码： <https://github.com/AjayNandoriya/PortraitLightingTransferMTP>

6. Image Based Relighting Using Neural Networks

论文：Ren P, Dong Y, Lin S, et al. Image based relighting using neural networks[J]. ACM Transactions on Graphics, 2015: 34(4): 1-12.

工作：该论文提出的基于神经网络的图像光影生成方法，降低了光影生成所需要的图像数量。该方法使用光线传输矩阵构建图像每个像素重光照模型。该方法提出的神经网络，仅需要在使用日常的荧光灯或LED灯下采集的图像，便能够重建场景高精度光线传输矩阵，图10为基于神经网络的图像光影生成结果。



图 10 基于神经网络的图像光影生成结果

论文： <https://dl.acm.org/citation.cfm?id=2766899>

代码： <https://github.com/damienfir/image-based-relighting>

7. Deep image-based relighting from optimal sparse samples

论文：Xu Z, Sunkavalli K, Hadap S, et al. Deep image-based relighting from optimal sparse samples[J]. ACM Transactions on Graphics, 2018: 37(4): 1-13.

工作：该论文提出了基于稀疏样本的图像对

象光影生成方法进一步降低光影生成所需要的图像数量，该方法在卷积神经网络基础上，设计了重光照网络（relight-net）建立任意方向光照下的图像与5种光照方向下图像的映射关系，能够对物体表面的直接光照、间接光照以及每个点的光照遮挡关系进行建模。图11为基于稀疏样本的图像对象光影生成结果，图12为网络结构。

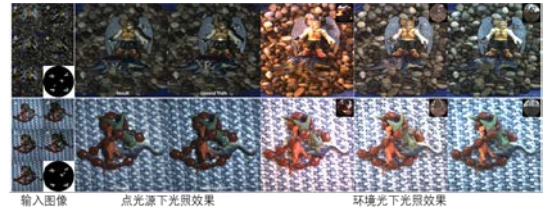


图 11 基于稀疏样本的图像对象光影生成结果

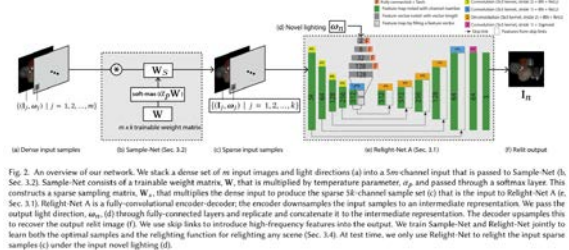


图 12 网络结构

论文： http://cseweb.ucsd.edu/~zex014/papers/2018_relighting/2018_relighting_paper.pdf

代码： <https://github.com/zexiangxu/Deep-Relighting>

8. Rendering Synthetic Objects into Legacy Photographs

论文：Karsch K, Hedau V, Forsyth D, et al. Rendering synthetic objects into legacy photographs[J]. ACM Transactions on Graphics, 2011: 30(6): 15701-15712.

工作：该论文使用面光源对场景光源进行建模，并将虚拟物体产生的遮挡阴影和间接光照与真实图像融合，结果如图13所示。为了能够获得影视级的虚实光影融合效果，首先手工在图像上标记出图像中的主要平面和光源的二维信息，然后利用迭代优化方法恢复光源和主要平面的三维信息，以及场景反射率和相机参数，最后根据上述信息在图像中绘制虚拟物体。



图 13 基于交互的虚实光照融合结果

论文 http://vision.csee.wvu.edu/~doretto/courses/2017-fall-cp/reading/RenderingSyntheticObjectsintoLegacyPhotographs_SIGGRAPH2011.pdf

代码: <https://github.com/kevinkarsch/rsolp>

9. Automatic Scene Inference for 3D Object Compositing

论文: Karsch K, Sunkavalli K, Hadap S, et al. Automatic scene inference for 3d object compositing[J]. ACM Transactions on Graphics, 2014: 33(3): 1-15.

工作: 该论文提出的方法能够自动推算场景几何与光源信息, 使用深度传播, 根据图像相似性将RGBD数据集中的深度信息自动迁移到输入图像; 构建了标注光源信息的图像数据集, 训练光源分类器, 从而从图像中自动标记光源位置; 针对图像中不可见光源信息, 根据图像相似度, 利用光源位置已知的图像数据集作为先验, 自动估计输入图像中不可见光源的三维信息。图14为基于场景自动理解的虚实光照融合结果。



图 14 基于场景自动理解的虚实光照融合结果

论文: http://kevinkarsch.com/publications/tog14-auto_illum.pdf

代码: [https://github.com/kevinkarsch/auto-scene-infer-](https://github.com/kevinkarsch/auto-scene-infer-ence-data)

[ence-data](https://github.com/kevinkarsch/auto-scene-infer-ence-data)

10. Deep Outdoor Illumination Estimation

论文: Hold-Geoffroy Y, Sunkavalli K, Hadap S, et al. Deep outdoor illumination estimation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2373-2382.

工作: 该论文提出了基于深度学习的室外场景光照估计方法, 输入单幅图像, 通过神经网络直接生成室外场景的天空和太阳参数。该方法使用了27万余组训练数据训练卷积神经网络, 每组数据包含1张室外图像和图像拍摄时间的天空和太阳参数。通过天空和太阳参数可直接生成高动态范围环境贴图, 用于虚拟物体的绘制。虚实光照融合结果如图15所示。



图 15 基于深度学习的室外环境光照估计结果

论文: <https://arxiv.org/pdf/1611.06403.pdf>

代码: <https://github.com/PeterZhouSZ/dashcam-illumination-estimation>

(责任编辑: 蹇木伟, 沈沛意)



宁宁

硕士研究生, 北京电子科技学院, 研究方向为计算机视觉。

Email: 2870735138@qq.com



金鑫

北京电子科技学院从事教学与科研工作, 担任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: jinxin.me



吴洪宇

北京航空航天大学虚拟现实技术与系统国家重点实验室从事科研工作。研究方向为虚拟/增强现实、计算机图像处理、物理仿真。

Email: whyvrlab@buaa.edu.cn

微表情数据集

山东大学 黄天欢 贾晔焱 陈雷

微表情 (Micro-expression), 也称面孔微反应, 是一种持续时间极其短暂的面部动作, 它能够揭示一个人试图隐藏的真实情感。此外, 微表情是人们在受到有效刺激的0.5秒内, 不由自主表现出的不受思维控制的瞬时真实反映, 因此微表情在测谎检测、临床诊断、国家安全等方面都有着巨大的潜在应用。鉴于微表情的真实性和潜在应用, 其吸引了国内外学术界和媒体界的高度关注。然而, 虽然多年来一般的面部表情识别 (FER) 在计算机视觉领域得到了深入的研究, 但微表情的自动分析研究却很少。微表情与正常的面部表情相比, 其最显著的特征就是持续时间短, 一般在1/5秒-1/25秒内, 这使得很难从实际对话中找到它们。因此迄今为止, 微表情自动分析研究的最大障碍就是缺乏合适的数据集。

微表情自动分析算法的综合研究和深刻分析与微表情图像数据集密不可分。近年来, 随着计算机、摄像机等硬件设备的进步以及微表情自动分析研究的需要, 微表情图像数据集不论在数量还是质量上都提到了更多的改善和提升。以下本文将介绍目前国际上专业的微表情图像数据集, 包括: SMIC、CASME、CASMEII、CAS (ME)²、SAMM、MEVIEW以及M2MEW。

1. SMIC 数据集

SMIC 数据集 (Spontaneous Micro-expression Database) 由芬兰奥卢大学的研究者提出, 共包含来自16名参与者的164个微表情片段。这些微表情片段是在惩罚犯罪者的审讯室中录制的, 被试者被要求观看能够引起情绪波动的视频, 并尽力不流露出内在情绪; 而记录者则被要求在不观看视频的情况下猜测被试者的情绪。如果被试者的情绪被记录者发现, 则被试者需要填一份冗长的问卷作为惩罚, 以此促使被试者尽力抑制表情, 确保微表情的可靠性。

SMIC数据集分为HS, VIS, NIR三个子数据集, 其中HS由每秒100帧的高速摄像机拍摄, 主要用于研究微表情变化速度快的特征; VIS由每秒25帧的正常视觉相机拍摄; 而NIR子数据集由每秒25帧的近红外相机拍摄, 主要用于解决光照对微表情识别的影响; VIS和NIR子数据集的增加丰富了SMIC数据集的多样性。其中, HS子数据集记录了16名参与者所有的164个微表情片段, 包括51个正类 (高兴) 片段, 70个负类 (悲伤、恐惧、厌恶) 片段以及43个惊讶的微表情片段。VIS子数据集和NIR子数据集分别记录了其中8名参与者的71个微表情片段, 各包括28个正类片段, 23个负类片段以及20个惊讶的片段。

SMIC数据集含有足够多的样本并包含2个子任务, 检测 (微表情/非微表情) 和识别 (分类), 因此SMIC数据集可用于微表情自动分析的全面研究。

论文: <https://ieeexplore.ieee.org/abstract/document/7492264/>

2. CASME & CASME II & CAS (ME)²数据集

CASME、CASME II 及CAS (ME)²数据集均由中国科学院心理研究所傅小兰团队所建立。数据采集时, 均要求被试者在保持面无表情条件下观看不同情绪视频。这种方式能够诱发自然的微表情, 并在一定程度上克服了前期一些微表情数据库中微表情非自然的问题。在采集过程中, 要求被试者保持面无表情的同时保持身体不动, 所以这些微表情比较“纯净”, 夹杂较少的非情绪动作。在微表情样本标注方面, 傅小兰团队综合考虑了面部动作单元 (AU)、视频材料的特点以及被试者的主观报告, 这使得样本的标注更为客观、准确。

CASME 数据集 (Chinese Academy of Sciences Micro-expression Database) 包含了由60帧/秒摄像机拍摄的195个自然微表情样本,

并分为A、B两个部分。其中A部分的分辨率为 640×480 像素，拍摄环境为室内，并使用两盏LED灯从面部两斜角照明。B部分的分辨率为 1280×720 像素，在自然光下拍摄。CASME数据集包含情绪类型有：高兴、悲伤、厌恶、惊讶、恐惧、抑制、紧张。由于有些情绪类型的微表情难以在实验室诱发，所以CASME数据集中微表情的情绪类别分布不均。图1为CASME数据集下部分样例。



图1 CASME 数据集微表情样例

论文：<https://ieeexplore.ieee.org/abstract/document/6553799>

CASME II 是在CASME数据集的基础上发展而来，包含247个微表情样本。与采样率为100帧/秒的SMIC数据集和60帧/秒的CASME数据集相比，CASME II的采样率为200帧/秒，并且有着比SMIC、CASME数据集都大的图像尺寸，这使得CASME II能够提供更详细的面部肌肉运动信息，有利于微表情特征的提取和后续的分类识别。相比于CASME数据集，CASME II引入了样本标注，并充分综合考虑主客观因素，除了基于人脸编码系统（FACS）的基本判断，还添加了受试者的主观回忆来辅助判断样本标签。除此之外，该数据库还对微表情的起始帧，高潮帧，结束帧都做了详细的标注，方便测试使用。考虑到CASME中分类过细且类别分布不均，使得某些情绪类型的样本过少的情况，傅小兰团队在CASME II数据集中将这些微表情分为高兴、悲伤、厌恶、惊讶及其它。图2为CASME II数据集下部分样例。

论文：<https://ieeexplore.ieee.org/abstract/document/7820164>

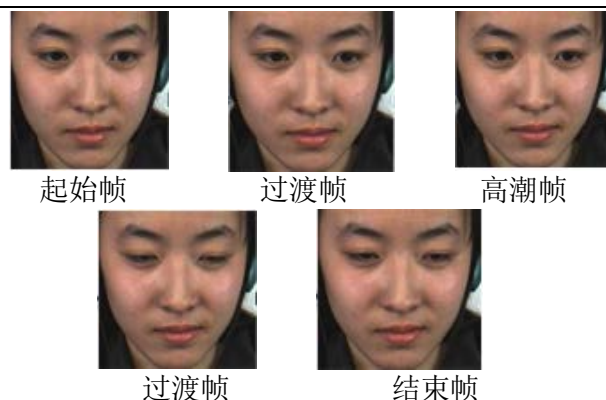


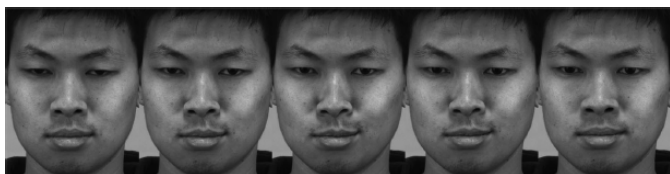
图2 CASME II 数据集微表情样例

CAS(ME)² 数据集 (Chinese Academy of Sciences Macro- and Micro-expression Database) 同样包含A、B两部分。A部分包含87段原始的人脸表情视频片段和未经任何预处理的图像序列，可用于宏表情和微表情的自动检测算法的分析研究；B部分包含357个表情样本，其中宏表情样本300个，微表情样本57个。A、B两部分的数据样本都是在相同的实验条件下的相同个体中收集的，在情绪标定上，CAS(ME)²采用正类、负类、惊讶、其它的分类方式。

3. SAMM 数据集

SAMM 数据集 (Self-Assessment Manikins Micro-expression Database) 由曼彻斯特城市大学的研究者所提出，包含来自32个参与者（16个男人，16个女人）、13个种族的159个自发微表情。这些微表情均在良好的实验室环境中（如：采用光扩散器等一系列灯光设置避免摄像机闪光，让参与者独自呆在实验室的某个空间，使得他们不受其他人的影响）使用200帧/秒的高速摄像机所录制，分辨率为 2040×1088 像素。SAMM数据集利用人脸编码系统（FACS）对微表情样本的起始帧、高潮帧和结束帧进行编码，并根据面部动作单元（AU）进行标注，有蔑视、厌恶、害怕、生气、悲伤、高兴、惊讶这7类微表情。

与SMIC、CASME、CASME II 及CAS(ME)² 相比，SAMM数据集不仅有着最高的分辨率，且受试者的人口结构最为多样化，这能更好的表示人口分布，使得获取的微表情更符合真实世界中各式各样的情感反应。图3为SAMM数据集部分样例。



起始帧 过渡帧 高潮帧 过渡帧 结束帧
图3 SAMM 数据集微表情样例

论文: <https://ieeexplore.ieee.org/abstract/document/7492264>

4. MEVIEW数据集

上述五个数据集均在实验室环境下录制,虽然这些微表情都是真实的,但都是人为诱导的。研究人员要求受试者在观看情感视频时保持尽可能多的中性表情,并记录下他们的正面头部姿势。此外,为了模拟压力因素,采用奖励或惩罚的手段,比如在受试者明显无法抑制情绪的情况下,要求其填一张无聊的表格作为惩罚。这些数据集主要是为微表情的识别而设计的。

由捷克技术大学的研究者所提出的MEVIEW数据集(Micro-Expression Videos in the Wild Database)是一个更为真实的非受控微表情数据集,其内容主要来自互联网上由每秒25帧的正常视觉相机录制的扑克游戏视频和电视采访视频。MEVIEW数据集包含有来自16个参与者的40个视频,包括高兴、惊讶、愤怒、恐惧和蔑视这5中情

绪类型。

论文: <http://cmp.felk.cvut.cz/~cechi/ME/>

5. M2MEW数据集

M2MEW数据集(Macro-to-Micro Expression Warehouse Database)由山东大学賁晔等人所提出,共包括来自36个参与者的300个微表情和900个宏表情,是目前最大的微表情数据库。M2MEW数据库采用诱发微表情的方式,在安静的室内,光线接近自然光的环境中建立。采集过程中,要求受试者在实验室通过观看具有感情色彩的片段自然产生微表情,并利用人脸编码系统以及受试者主观回答共同确定最终的感情分类,其中包括高兴、惊讶、愤怒、厌恶、恐惧和悲伤六种基本类型。图4为M2MEW数据集下部分样例。



图4 M2MEW数据集的一个样本序列

详见: <http://www.dpailab.com/database.html>

申请该数据库可联系: benxanye@gmail.com

(责任编辑: 金鑫, 樊鑫)



黄天欢

硕士研究生, 山东大学信息科学与工程学院。研究方向为机器学习、模式识别。

Email: sduhth@163.com



賁晔

副教授、博导, 山东大学信息科学与工程学院。研究方向为模式识别、图像处理、情感计算。

Email: benxanye@gmail.com



陈雷

副教授, 山东大学信息科学与工程学院。研究方向为图像处理、计算机视觉以及视觉质量评估。

Email: leiguanglg@hotmail.com

招聘信息

一、上海科技大学信息学院

1、上海科技大学&联影智能联合实验室博士后

1) 职位描述:

- 从事视觉理解与医疗影像课题研究;
- 撰写并发表顶级期刊/会议研究论文;
- 协助 PI 组建联合实验室和培养研究生;
- 协助经费申报及管理。

2) 任职要求:

- 拥有或即将获得计算机视觉、机器学习、人工智能或相关专业博士学位;拥有场景理解、目标检测与分割、深度网络模型等相关经验者优先;
- 在本领域知名会议或期刊发表过高质量研究论文;
- 熟悉软硬件平台的设计、调试和测试;
- 具有良好的英文口头表达和写作能力;
- 良好的协调沟通能力与团队合作精神。

2、何旭明教授研究组 PLUS Lab 科研助理

1) 职位描述:

- 参与从事课题研究,有能力和志愿者鼓励开展独立研究;
- 协助搭建和管理科研数据和软件平台;
- 协助对实验室的日常管理(包括电脑硬件设备的购置和管理、实验室网站的更新);
- 协助撰写经费申请书等科研文档。

2) 任职要求:

- 已获得或即将获得计算机科学与工程、应用数学、自动化或统计学等领域的硕士学位;其他专业的学位视情况也可考虑;
- 具有良好的沟通协调能力和团队合作精神;
- 具有良好的中英文书面文字表达能力,文字功底要强;
- 责任心强,诚恳、细心、有耐心。

3、申请方式

- 请应聘者通过人才招聘系统(<http://jobs.shanghaitech.edu.cn/>)提交应聘申请。

- 应聘流程为:注册、填写并提交基本信息、应聘选择岗位;应聘者请同时将简历等相关材料发送至: hexm@shanghaitech.edu.cn。

二、联通在线信息科技有限公司

1、信息安全工程师

1) 工作职责:

- 负责平台的信息安全维护;
- 负责业务开发和上线前的信息安全审核;
- 负责网络信息安全评估和双新评估;
- 负责与省份公司的信息安全沟通工作。

2) 任职资格:

- 2020 届全日制普通高等院校研究生及以上学历;
- 计算机专业;
- 熟悉熟悉 linux 操作系统,mysql 数据库;
- 对信息安全工作有一定了解;
- 有过信息安全实习经验者优先。

2、技术研发工程师:

1) 工作职责:

- 配合业务人员完成软件系统的需求调研和分析;
- 负责与项目相关的技术方案、文档的编写;
- 根据项目具体要求,承担开发任务,按计划完成任务目标;
- 独立完成单元及模块测试,并保证代码质量。

2) 任职资格:

- 2020 届全日制普通高等院校研究生及以上学历;
- 计算机科学与技术、软件工程等相关专业;
- 熟悉 J2EE 技术架构体系;
- 熟悉 SSH 等主流开发框架;
- 熟悉 vue, react, Dojo, JQuery 等前端开发

框架和类库；

- 熟练使用 MySQL, Oracle, DB2 等主流数据库；
- 熟练使用 JUNIT 测试框架。

3、创新技术支撑：

1) 工作职责：

- 负责平台的需求分析、开发管理和平台迭代；
- 负责平台的项目立项、采购、验收等项目管理的工作；
- 负责平台的信息安全工作，包括制定安全策略、组织落实各项安全检查及整改；
- 负责平台的机房及硬件管理，包括硬件资源的需求响应、资源整合迁移等。

2) 任职资格：

- 2020 届全日制普通高等院校硕士及以上学历；
- 计算机相关专业；
- 精通计算机软件开发（JAVA、PHP）、数据库（Oracle）、主流操作系统、网络基础知识；
- 熟悉电信运营商能力开放业务平台的体系架构，包括语音、短信、取号、位置、充值、计费能力体系中的一种或以上；
- 熟悉电信运营商 O 域、B 域的基础架构和业务流程；
- 掌握项目管理的基础理论知识，获得过相关机构颁发的证书优先；
- 英语六级；
- 高度的责任心、自我驱动意识；
- 结果导向，善于利用各种工具、方法独立解决工作中遇到的问题；
- 善于学习、接收最新的互联网前沿思想和知识，自我提升意识强；
- 高度的团队协作意识；
- 有实习生优先。

4、平台开发工程师：

1) 工作职责：

- 协助项目经理，整理产品需求，设计功能细节并撰写相关文档；
- 负责平台整体架构设计与研发，串联前端、底层需求；

- 协调厂家，完成项目进度把控；
- 完成业务数据整理、归纳、统计工作，完成相关材料的编写工作；
- 负责完成领导交办的任务。

2) 任职资格：

- 本科及以上学历；
- 计算机相关专业；
- 熟练使用 Go、Python、C、Java 其中一种语言；
- 拥有良好的编码习惯，理解 io、多线程、高并发、集合等基础框架，对 JVM 原理有一定的了解；
- 对开源技术有浓厚的兴趣，至少了解 Hadoop、Storm、Hive、Kafka、RabbitMQ、Zookeeper 中的一种；
- 对 Nginx、Tomcat、Zookeeper、RabbitMQ 等常用应用服务器中间件有一定的了解；
- 熟悉常用的算法和数据结构，了解常见网络协议（TCP/IP，http 等）；
- 具有扎实的计算机科学功底，较好的编码习惯，扎实的编程基础和数据结构算法基础，对通信互联网领域有较强的兴趣；
- 要求工作认真、细致、有耐心，善于团队合作，有出色的沟通及协调能力，能承受工作压力。

5、软件架构工程师：

1) 工作职责：

- 根据系统需求，分析、设计与实现系统软硬件整体架构，保障系统架构的合理性、可扩展性及经济性；
- 参与底层和重要模块的详细设计，控制质量；负责核心技术问题的攻关，系统优化；指导解决项目开发过程中的技术难题；带领开发团队持续改进系统架构、功能模块、开发需求、提高用户体验；控制架构质量；
- 指导和培养开发团队能力，负责系统开发过程的技术评审工作，检查工程师的代码开发工作，不偏离最初的设计；
- 制定相关技术开发规范，参与制定技术标准，

编写相应的技术文档，并对通用技术进行整理，提高技术复用；

- 能制定合理的开发工作计划，并有效地按照计划完成工作；
- 积极了解业界发展、相关新技术及趋势，促进技术进步和创新。

2) 任职资格：

- 2020 届全日制普通高等院校本科及以上学历，北京大学等国内双一流、欧美名校海归优先，硕博优先；
- 计算机相关专业；
- 熟悉多线程及网络编程；熟悉大容量、高并发服务器的设计及优化技术；了解前端和大数据相关技术；
- 精通 Java 语言，了解 Java 主流工具及高并发、高性能开源工具（netty、disruptor、logback、log4jguava、Apache Commons Pool 等），并熟悉至少一种其它开发语言；
- 熟悉主流开源框架，如 Spring、SpringMVC、MyBatis、Activity、SpringBoot 等；熟悉 Spring 原理；
- 熟练使用 SpringCloud 搭建微服务系统，有过高并发项目经验者优先；
- 熟悉 Java 主流的序列化工具和 RPC 工具（jackson、gson、fastjson、protobuf、Thrift、dubbo 等）；
- 熟悉 Mysql/PostgreSQL 工作原理及程序上如何高性能使用，了解主流开源 NoSQL 系统（Memcached、Redis、MongoDB 等）使用和原理，有 redis 或其他缓存系统大规模使用经验；
- 熟悉 Servlet 容器（Tomcat、Jetty、Resin 等）原理和使用，了解 LVS、Haproxy、Nginx 的配置和调优；
- 熟悉 MQ 原理和使用（RabbitMQ、ActiveMQ、Kafka 等）；
- 熟悉持续集成和自动化测试工具，如 Maven、Hudson、Jenkins；
- 注重利用合理的设计模式和模块化的组织方

式对代码进行架构。

- 喜欢阅读英文技术文档和书籍，关注新技术发展动态，乐于分享技术；
- 英语六级以上；
- 具有良好的抗压能力，学习能力强，善于分析解决问题，有创业精神；
- 有全球 500 强总部、知名互联网或 IT 科技公司实习经验优先，有高并发、高负载、分布式系统项目经验优先，熟悉人工智能和机器学习技术优先，全栈工程师优先。

6、智能网络：

1) 工作职责：

- 负责核心网新技术跟踪研究；
- 负责语音增值业务方案制定；
- 负责语音增值业务测试验证；
- 负责与周边系统对接。

2) 任职资格：

- 2020 届全日制普通高等院校研究生及以上学历；
- 计算机相关专业；
- 熟练使用 JAVA 语言，熟悉 J2EE 架构；
- 熟练掌握 junit、TestNG、ant、maven 等工具使用；
- 英语六级以上；
- 具有良好的抗压能力；
- 有实习生经验优先。

7、视觉感知算法(3D 方向)：

1) 工作职责：

- 负责机器人场景中图像相关机器学习算法研究与应用（车道线定位、检测分割、图像检索、特征匹配、Stereo 等）。

2) 任职资格：

- 具有计算机视觉、图形学、增强/虚拟现实、机器人等相关专业在读本科或以上学历；
- 熟悉 2D 点云或 3D 点云上常用检测、分割算法；了解深度估计、视觉定位、特征匹配等相关算法；
- 熟悉机器学习/深度学习基础知识，了解立体视觉常用知识和算法；

- 熟悉经典 CV 中的常见算法，熟悉 OpenCV；了解图像处理、光流跟踪、图像匹配、描述子等基础知识；
- Python/C++ 编程基础扎实，熟悉 Tensorflow/Pytorch/MXNet 等常用框架。
- 加分项（非必要）：
- 具备 ACM 等大赛奖牌者优先；
- 具有无人驾驶、机器人等场景项目经验者优先；
- 在 CVPR、ECCV、ICCV 或 ICRA、IROS 等相关会议取得过优秀成绩或发表论文或参加竞赛者优先；
- 了解 PointNet、PointCNN 等常见深度学习方法优先；
- 了解 3D 视觉/SLAM 等算法者优先。

投递方式：

官网投递简历：<https://zgl2020.zhaopin.com/>

三、浪潮 2020 校招

1、软件研发工程师：

1) 职位描述

- 负责软件系统功能模块的开发、测试等过程；
- 根据需求设计软件系统技术方案, 对于产品的功能能够提出自己的技术意见和想法；
- 开发过程涉及关键技术、设计总结分享；
- 工作过程产生的资料及时归档。

2) 任职要求

- 具有计算机、软件、数学等相关专业背景，熟悉 C/C++、Java、SQL 等中的任何一种或者多种语言；
- 熟悉各种 Web 开发技术，包括 HTML/JAVASCRIPT/CSS/Ajax 等，并有相关的项目开发经验或成果；
- 追求完美，对于创新及解决具有挑战性的问题充满激情；
- 有责任心、有耐心，具有较强学习能力、总结能力、团队合作能力优先。

2、软件研发工程师（Java）

1) 职位描述：

- 负责浪潮自研发大数据产品、云平台产品、数据库产品测试开发工作，包括测试案例设计、测试脚本开发、测试执行、测试结果报告、产品质量改进等；
- 负责开发测试工具、研究新测试技术方向，提升测试能力及效率；
- 负责软件系统的接口测试、自动化测试、性能测试、安全测试、移动应用测试等工作；
- 负责 DevOps 研发测试平台运维和过程改进工作。

2) 任职要求：

- 计算机科学与技术、信息安全、软件工程、网络、通信、统计学、等相关专业本科及以上学历；
- 至少掌握 C、C++、Python、JAVA 等其中一种语言，掌握面向对象编程思想和设计方法，至少能够熟练使用一种软件开发工具；
- 有一定的学习能力、抗压能力和沟通协调能力。

3、软件研发测试工程师

1) 职位描述

- 负责浪潮自研发大数据产品、云平台产品、数据库产品测试开发工作，包括测试案例设计、测试脚本开发、测试执行、测试结果报告、产品质量改进等；
- 负责软件系统的接口测试、自动化测试、性能测试、安全测试、移动应用测试等工作；
- 负责 DevOps 研发测试平台运维和过程改进工作。

2) 任职要求

- 计算机科学与技术、信息安全、软件工程、网络、通信、统计学、等相关专业本科及以上学历；
- 至少掌握 C、C++、Python、JAVA 等其中一种语言，掌握面向对象编程思想和设计方法，至少能够熟练使用一种软件开发工具；

4、AI 算法工程师

1) 职位描述

- 攻坚相关 AI 关键技术，形成高精度 AI 算法

和模型产品；

- 研究典型场景的 AI 应用算法与模型实现，完成工程项目，形成 AI 应用解决方案；
- AI 算法相关工具的开发；
- 技术交流与培训支持。

2) 任职要求

- 硕士及以上学历，数学、人工智能、计算机视觉、数据挖掘等相关专业；
- 具备一定 AI 算法开发经验，熟悉数理统计、机器学习具体实施方法；
- 能够使用时间序列、模式识别等统计模型且能有效完成报告呈现和数据解读；
- 掌握 R、Python 等编程语言、SASSPSSRMatlab 软件、Hadoop、Spark、Storm 相关技术者优先考虑；
- 有海量数据处理和挖掘经验者优先考虑。

5、Android 开发

1) 职位描述

- 参与爱城市网 Android 端的研发；
- 参与爱城市网 Android 端架构的改良、迭代；
- 参与爱城市网 Android 端的衍生版本研发；
- 参与爱城市网 Android 端的性能优化；
- 参与爱城市网 Android 端的各版本的迭代和维护。

2) 任职要求

- 较强的书面及口头表达能力；
- 具有较强的自主学习能力，乐于接受挑战；
- 有责任心、具有团队合作精神，抗压能力较强。

6、IOS 开发工程师

1) 职位描述

- 根据产品设计完成 IOS 端产品开发、测试、维护；
- 不断迭代 IOS 端功能，解决产品使用过程中的缺陷；
- 参与后端接口定义和项目个性化需求实现；
- 配合研发经理完成相关任务目标。

2) 任职要求

- 熟悉 OC 与 Swift 语法和使用规范；

- 熟悉混编开发，能实现原生与 JS 重度交互；
- 熟悉 IOS 开发规范以及 AppStore 上架流程和审核规则；
- 具有智能设备链接开发以及即时通讯开发经验；
- 计算机相关专业。

7、FPGA 开发工程师

1) 职位描述

- 关注并学习前沿深度学习 CNN、Transformer 等算法、裁剪和量化等压缩算法、加速硬件设计架构等；
- 针对 FPGA 设计高效的深度学习计算架构，并进行优化；
- 编写架构设计的相关文档。

2) 任职要求

- 硕士及以上学历，计算机、软件、数学、机器视觉、电子、微电子等相关专业；
- 具有 TensorFlow、Pytorch 等深度学习开源架构使用经验，熟悉 CNN、Transformer 等网络的典型结构；
- 熟练阅读英文文献，具有一定英文沟通能力；
- 工作态度积极，沟通能力强，学习能力强，具有追求极致的精神。

8、数据研发工程师

1) 职位描述

- 负责数据日常处理工作；
- 数据治理工具研发需求分析；
- 根据院内数据情况使用 SQL 进行数据映射、标注、码值转换；
- 为产品组提供完整、准确的数据信息。

2) 任职要求

- 计算机、信息化相关专业；
- 熟悉 JAVA、python、SQL、shell 等开发语言；
- 具有较强的学习能力、抗压能力和沟通协作能力。

9、芯片研发工程师（软件）

1) 职位描述

- 根据项目需求完成参与芯片方案制定；
- 负责模块级的 verilog RTL 实现；

- 编写验证平台，完成模块级的仿真验证；
- 负责部分设计、开发、测试等相关文档的编写，整理和归档并提供相应的技术支持。

2) 任职要求

- 计算机、电子类、电气类、自动化、通信相关专业，有创新精神；
- 熟悉 FPGA、数字电路、C/C++语言；
- 有一定的大数据、机器学习理论基础者优先；
- 具有较强的学习能力、抗压能力和沟通协作能力。

10、国家重点实验室研究员

1) 职位描述

- 前沿技术调研与项目规划；
- 基于硬件平台的深度学习加速方案设计；
- 面向分布式集群的深度学习框架性能优化技术研究；
- 面向特定应用的处理器、编译技术及系统结构研究；
- 原型系统研发及知识成果转化；
- 国家及省级课题、技术项目的申报与方案设计。

2) 任职要求

- 计算机、电子类、电气类、自动化、通信相关专业，有创新精神；
- 熟悉 C/C++/Python 等语言；
- 熟悉 CPU、GPU、FPGA 架构及计算原理、有一定的计算机系统结构、Tensorflow/Pytorch/Caffe 源码分析经验者优先；
- 具有较强的学习能力、抗压能力和沟通协作能力。

11、虚拟化研发工程师

1) 职位描述

- 负责云计算产品及云计算平台关键模块开发；
- 负责云计算主流技术的调研、技术验证及设计开发；
- 支撑公司云计算产品相关项目的开发及技术支持。

2) 任职要求

- 本科及以上学历，计算机、电子、通信等相关专业；
- 熟悉计算机体系结构、操作系统、计算机网络、数据库等基本知识；
- 至少熟悉一种虚拟化平台；
- 熟悉分布式存储、云安全、桌面协议等技术的优先。

投递方式：

浪潮校招官网：<http://career.inspur.com/campus2020/index.html>

四、货拉拉 2020 校园招聘

1、算法工程师：

1) 职位描述：

- 2020 届毕业生，应用数学/数理统计/机器学习/AI 等相关专业，具有一定的数据结构及算法基础，熟悉机器学习/数据挖掘/自然语言处理等相关算法；
- 熟悉 Python、Java、C++等其中的一种或几种，有 Hadoop/Spark 等大数据开发经验。

2) 任职要求：

- 爱钻研新技术，谦卑不浮躁；
- 良好的逻辑综合分析能力，以及强烈的解决问题的意愿。

2、PHP 开发工程师：

1) 职位描述：

- 2020 届毕业生，计算机相关专业，本科及以上学历；
- 熟练掌握 PHP 语言，熟悉常见的模板引擎或开源框架；
- 会面向对象程序设计（OOP）与基本的设计模式；
- 了解 linux 及基本的操作命令。

2) 任职要求：

- 扎实的计算机基础知识，深入理解数据结构，算法，操作系统等知识；
- 良好的沟通能力，能清晰、准确的在团队成员中传达自己的想法。

3、测试工程师：

1) 职位描述:

- 2020 届毕业生, 计算机相关专业, 本科及以上学历;
- 制定测试计划, 编写测试用例, 部署测试环境, 执行测试, 定位及分析 BUG 并根据 BUG 生命周期;

2) 任职要求:

- 逻辑能力好, 爱钻研新技术, 谦卑不浮躁;
- 计算机基础扎实, 良好的编程技术。

投递方式:

货拉拉校招官网: <https://huolala.zhiye.com/>

五、深交所&深证通 IT 人才招聘

1、金融科技研究类:

1) 职位描述:

- 负责或参与行业金融科技创新应用前瞻性、关键性等技术研究;
- 负责或参与金融科技应用课题立项与预研;
- 负责或参与金融科技项目的系统规划与研发;
- 负责或参与证券期货行业金融科技研究发展中心相关工作。

2) 任职要求:

- 爱钻研新技术, 谦卑不浮躁;
- 良好的逻辑综合分析能力, 以及强烈的解决问题的意愿。

2、系统开发、测试、运维类:

1) 职位描述:

- 负责或参与证券交易、市场监察、业务管理、大数据应用、互联网应用、移动应用、金融云平台等核心应用系统的研发工作;
- 负责或参与软件质量管理、应用项目测试和测试工具开发等工作;
- 负责或参与日常应用系统运维和监控平台、自动化运维平台的建设维护等工作。

2) 任职要求:

- 计算机、数学、统计等相关专业本科及以上学历;
- 具有四级以上英语水平, 具有较好的文字表

达能力, 具有较强的学习能力和分析解决问题的能力;

- 至少具备一项如下技能:
- 熟练掌握 JAVA、C/C++、Python、Shell 等主流开发语言, 有相应项目开发经验或成果, 有多线程编程、网络编程或分布式应用开发经验;
- 熟练掌握 HTML、CSS、Java、script 等 web 开发技术, 有相应项目开发经验或成果, 具有 Ajax 或移动开发经验;
- 熟练掌握 SQL 编程技术, 熟悉 Hadoop 生态圈技术, 有相应项目开发经验或成果, 有数据分析、数据挖掘开发经验;
- 熟练掌握机器学习、深度学习算法和工具, 有相应项目开发经验或成果, 具备自然语言/语音/图形图像处理经验;
- 熟悉软件开发测试流程, 掌握常用测试用例设计方法, 具备设计和开发测试和自动化测试框架能力;

3、信息安全类:

1) 职位描述:

- 负责安全产品的开发和优化工作;
- 负责安全相关数据的分析和建模;
- 负责建设安全开发生命周期体系和推进相关规范。

2) 任职要求:

- 计算机或相关专业本科及以上学历, 信息安全专业优先;
- 具备编程能力, 熟练掌握 Python 或 Java 语言;
- 获得过省级及以上 CTF 比赛奖项、全国大学生信息安全竞赛等奖项优先;
- 有自己的开源项目者优先;
- 具备逆向、PWN 能力者优先。

投递方式:

招聘官网: http://www.sccc.com/join_us/school_recruit/

(责任编辑: 李策, 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 Computer Vision and Image Understanding 特刊及 Engineering Applications of Artificial Intelligence 特刊。

3 会议简介 (ICMVA)

ICMVA 2020 (International Conference on Machine Vision and Applications) 将于 2020 年 2 月 26 日至 2 月 28 日在缅甸仰光举办，论文截止时间为 2019 年 11 月 10 日。会议的目的是为学术界和工业界的研究人员和从业人员提供一个平台，以满足和分享该领域的尖端发展。征稿范围涉及算法、生物医学、应用工厂机器人、智能运输系统、人机交互、多媒体等方向的研究进展。

(责任编辑：贾同)

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
ICMVA 2020	2020.02.26-28	缅甸 Yangon	2019.11.10	http://www.icmva.org
ICPRAM 2020	2020.02.22-24	马其他 Valletta	2019.11.15	http://www.icpram.org
AAAI-MAKE 2020	2020.03.23-25	美国 California	2019.11.17	https://www.aaai-make.info
GCAI 2020	2020.04.06-09	中国 Hangzhou	2019.11.30	http://gcai-2020.info/
DATA 2020	2020.03.29-31	美国 Boston	2019.11.29	https://www.elsevier.com/events/conferences/AI-and-Big-Data-in-Cancer/

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
Computer Vision and Image Understanding	Special Issue on “Adversarial Deep Learning in Biometrics & Forensics”	2019.11.30
Engineering Applications of Artificial Intelligence	Special Issue on Pushing Artificial Intelligence to Edge: Emerging Trends, Issues and Challenges	2019.11.15