

# CCF

# 计算机视觉 专委会简报

## COMPUTER VISION NEWSLETTER

2017/06 期  
总第 10 期

### 专委动态

走进高校系列活动

### 科技前沿

顶会参会感悟

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](mailto: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 湖南大学

#### 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 上海科学技术大学

#### 科技前沿

主编 申抒含 中国科学院自动化研究所  
编委 邓 成 西安电子科技大学  
卢国梁 山东大学  
任传贤 中山大学  
苏 航 清华大学  
王金甲 燕山大学  
杨巨峰 南开大学

#### 委员风采

主编 余 焯 合肥工业大学  
编委 韩爱丽 山东大学  
刘海波 哈尔滨工程大学  
余志文 华南理工大学

#### 资源平台

主编 苗启广 西安电子科技大学  
编委 樊 鑫 大连理工大学  
贾 同 东北大学  
蹇木伟 中国海洋大学  
李 策 兰州理工大学  
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	CCF-CV 走进高校系列报告会	04
走进企业	CCF-CV 委员工作会议	05
信息速递	CCF-CV 简报编委会会议	06
	CCF-CV 标准工作组会议	07
	CCCV2017 圆满闭幕	08

---

#### 科技前沿

专题综述	目标跟踪算法综述	11
热点追击	基于跨域学习的行人再标识	15
	细粒度级别图像检索	16
	基于图像的大场景三维重建	17

---

#### 委员风采

委员访谈	华南理工大学余志文教授访谈	18
委员好消息		21

---

#### 资源平台

数据集	DUTS 数据集	23
开源代码	目标检测及其开源项目	24
招聘信息		26
征文通知		27

# CCF-CV 走进高校系列报告会

## 第 40 期 南昌航空大学

时间：2017 年 9 月 24 日

由中国计算机学会计算机视觉 (CCF-CV) 专委会主办, 南昌航空大学江西省图像处理与模式识别重点实验室和软件学院承办的第 40 期 CCF-CV 走进高校系列报告会在南昌航空大学隆重举行。



本次会议邀请了中国科学院自动化研究所研究员**王亮**博士、厦门大学教授**纪荣嵘**博士、西安电子科技大学教授**董伟生**博士、复旦大学副教授**金城**博士四位专家作了特邀报告。南昌航空大学副校长**黎明**教授、南昌大学**闵卫东**教授、南昌航空大学信息工程学院院长**熊邦书**教授等领导嘉宾出席了本期报告会。报告会由江西省图像处理与模式识别重点实验室主任、南昌航空大学软件学院副院长**储瑒**教授主持, **黎明**副校长致欢迎辞。来自江西省多所院校的近 300 名师生参加了报告会, 聆听了计算机视觉领域的前沿技术成果和最新动态介绍。

本期报告会是 CCF-CV 走进高校系列活动首次在江西省举行, 参会嘉宾和师生总数近 300 人。参会人员除了承办方南昌

航空大学和本地高校相关领域的师生外, 还包括来自景德镇陶瓷大学等江西省其它地市兄弟院校的师生和企业嘉宾。因此, 本次报告会是本系列活动举行以来规模较大、范围较广的一期。

本期报告会为江西省计算机视觉领域的研究人员提供了一次全面、深入地向各位知名专家、学者学习的珍贵机会, 也为全省相关领域的师生提供了良好的交流平台, 促进大家在各自科研方向上形成更多创新思路, 有力推动江西省计算机视觉领域的科学发展和技术进步。



## 第 41 期 武汉大学

时间：2017 年 10 月 21 日



由中国计算机学会计算机视觉 (CCF-CV) 专委会主办、武汉大学遥感信息工程学院承办的第 41 期 CCF-CV 走进高校系列报告会在武汉大学隆重举行。本次会议由武

汉大学遥感信息工程学院**陈震中**教授和阿里巴巴人工智能实验室**王刚**教授担任执行主席, 邀请了西北工业大学教授**何明一**博士、上海交通大学教授**熊红凯**博士、厦门大学教授**纪荣嵘**博士、中山大学教授**郑伟诗**博士等四位专家作了特邀报告。

报告会由武汉大学**陈震中**教授主持, 来自武汉大学、武汉理工大学、华中科技大学、华中师范大学、中国地质大学、中南财经政法大学、湖北工业大学、桂林电子科技大学、武汉卓目科技有限公司、武汉高德红外股份有限公司、北京数字政通科技股份有限公司等全国各地多所高校及企业的 300 余人参加了此次报告会, 共同聆听几位专家对计算机视觉领域前沿技术成果和研究动态的精彩介绍。



此次报告会持续了近三个小时, 报告现场座无虚席, 许多听众更是全程站在场边聆听报告。四位讲者与台下听众互动热烈, 耐心为大家答疑解惑, 以其活跃的学术思维和严谨的科研态度为大家带来了一场学术盛宴。最终, 参加报告会的现场听众对四位讲者表示了由衷的欢迎与感谢, 报告会在雷鸣般的掌声中圆满结束。

(责任编辑: 马占宇)

# CCF-CV委员工作会议



中国计算机学会计算机视觉专委会 (CCF-CV) 2017 年度工作会议于 2017 年 10 月 11 日在万丽天津宾馆举办,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近 300 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副主任、北京大学查红彬教授致辞。查教授感谢新老委员们从祖国各地到场参会,他积极肯定了专委会在过去一年中各方面取得的丰富进展,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成果,希望专委会各项活动进一步提升品牌质量,打造学术交流精品活动。



之后,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,特别肯定了“走进高校”这一品牌活动的良好组织形式。他最后预祝 2017 年中国计算机视觉大会 (CCCV2017) 圆满成功。

接下来,专委会秘书长王亮研究员代表秘书处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评优结果,通报了 6 月份专委年度常委会议作出的若干新举措,详细介绍了 CCCV2017 的筹备情况,全面回顾了专委会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。

根据会议日程,工作会议还进行了新增委员的选举工作。经过专委会常务委员的投票表决,共有 97 名候选人当选为新任委员。



委员建言献策环节,北京大学林宙辰教授、华南理工大学金连文教授、北京科技大学殷绪成教授、大连理工大学孔祥维教授等委员代表就专委会的工作开展情况、目前存在问题、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。

会议的最后一项议程是 PRCV2019 的

申办宣讲,分别由南京理工大学杨健教授领衔的团队和西北工业大学张艳宁教授领衔的团队进行了申办陈述。两个团队的申办报告各具特色,从六朝古都的南京到十三朝古都的西安、从会场选址到赞助预算、从论文投稿到嘉宾报告,参会委员们关注的各个方面均有介绍。经过委员们的提问和投票环节后,最终确定了会议承办方的归属——西北工业大学团队,大家期待 2019 年金秋再次相聚于古城西安!



最后,专委会副主任查红彬教授做总结发言。他首先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 2017 年度工作会议圆满结束!

(责任编辑:王瑞平)

# CCF-CV简报编委会会议

湖南大学 李实英

CCF-CV 专委会简报编委会第一次工作会议于 2017 年 10 月 13 日下午在万丽天津宾馆顺利召开。会议由执行主编**李实英**主持，专委副主任**王涛**和主编**王亮**亲临指导，版块主编申抒含、毋立芳，以及韩爱丽、蹇木伟、李策、刘海波、卢国梁、任传贤、沈沛意、苏航和王瑞平等编委参加。



在会议中编委们分享简报获得好评的

欣慰，互相介绍自己积累素材、选题和编辑的经验，并就如何改进简报排版设计、丰富内容、提高文章质量等方面展开了热烈的讨论。**王亮**主编表扬了科技前沿和委员风采两个版块能够提前进行素材准备，每期保质、保量而且按时完成编辑和校稿，并建议将简报的委员好消息和资源平台部分与专委网站进行资源共享。**王涛**副主任介绍了《中国计算机学会通讯》的编委任期及分工，建议将简报中的好文章转投到学会通讯发表，从而更好地宣传 CCF-CV 专委会和委员的工作。

会议决定，简报将保持双月刊的发布频率以保证简报质量；在征集专委委员意见的基础上，适时调整版块内容，吸收新成

员，调整简报的排版设计。编委们积极献计献策，力求办成一份排版专业、内容充实丰富的专委简报，更好地切合简报读者兴趣和计算机视觉研究发展动向。



本次编委会工作会议的召开得到了计算机视觉专委会和 CCCV2017 大会组委会的大力支持。

(责任编辑：毋立芳)

# CCF-CV青年工作组会议

复旦大学 金城

中国计算机学会计算机视觉专委会 (CCF-CV) 青年工作组第一次工作会议于 2017 年 10 月 12 日在天津宾馆圆满召开，会议由组长华中科技大学**白翔**主持，组员王亮、苗启广、邓成、李策、郑伟诗、刘光灿、丛杨、方昱春、姬艳丽、张姗姗、金城等人参加。

工作组成员在会议中展开了热烈的讨论，对今后一年中的重点任务进行了规划，特别是就如何组织面向国内青年学者的特色活动进行了深入细致的研讨。会议提出

将加强具有研究潜力的青年研究人员之间及其与资深学者之间的合作交流，为其打造展现研究成果的优质平台，从而进一步鼓励创新与合作，为国内计算机视觉领域的可持续发展提供重要支撑。



会议还就配合明年专委会重要学术会议 PRCV 2018 开展的活动提出了初步的方案，也讨论了与国内其他重要会议合作举办活动的可能性，力争为计算机视觉领域的广大研究生和青年学者提供更为广阔的活动舞台。

本次会议的召开得到了计算机视觉专委会和 CCCV 2017 大会组委会的大力支持。

(责任编辑：毋立芳)

# CCF-CV 标准工作组会议

2017年10月13日，中国计算机视觉标准工作组启动会议在天津宾馆成功召开。全国信标委图形图像分委会王聪秘书长、计算机视觉标准组组长陈熙霖教授，副组长王涛博士、王亮研究员，以及来自中科院计算所、自动化所、信息工程研究所、上海科技大学、北京工业大学、西安电子科技大学、湖南大学、湘潭大学、海康威视、商汤科技、格灵深瞳、凯泽科技、爱奇艺、今日头条、微软亚洲研究院、英特尔等20余名产学研专家代表参加了本次会议，会议由工作组秘书毋立芳教授主持。



首先陈熙霖组长介绍了计算机视觉标准制定的意义、工作组业务范围、人员组织和计算机视觉标准体系架构。随着AI深度学习技术的快速发展，计算机视觉及相关产业迎来了蓬勃发展的良好机遇。计算机视觉主要研究视觉信息的感知与智能分析，技术涵盖范围广，应用领域庞大。在交通监控、安全认证、智能驾驶、机器人、媒体、医疗等产业迫切需要计算机视觉标准的制定，以促进产品的规范化发展。视觉标准体系架构包括基础通用（术语、体系框架）、关键技术、应用三个板块，每个板块包括相关多个标准。

全国信标委图形图像分委会王聪秘书长系统介绍了我国信息标准化的背景、管理体制，并对标准的撰写要求进行了详细说明。随着“十三五”技术标准体系的建设，信息标准已经上升到国家战略。制定国家信息标准责任重大，希望以此为基础在

ISO/IEC 国际视觉标准制定中发挥积极作用。



最后，毋立芳教授介绍了计算机视觉术语标准立项进展、任务和挑战。各位专家针对视觉标准术语分类方法和视觉标准应用需求等问题展开深入交流，并对标准组进一步工作计划提出很多启发性的建议。会议气氛热烈，各学术和企业专家代表共同表示努力支持完成中国首个计算机视觉标准的制定，促进计算机视觉产业发展并成功走向国际化。

（责任编辑：毋立芳）

## 通知 专委简报编委会 2018 招新

专委简报是计算机视觉专委委员的重要信息通道，主要内容是发布专委活动和会议信息、介绍本领域科研前沿和热点、宣传和分享委员研究成果和经验、介绍开放的数据库等资源平台，现为双月刊，自2015年11月创刊以来已经发行10期，每期阅读量达几千人次。

为了进一步提升专委简报的信息平

台作用、增加简报编委会的活力，经专委常务委员会同意并简报编委会议讨论决定，专委简报将征集若干名新的志愿者委员，具体要求如下：

1. 对专委工作有热情、有热心；
2. 对简报工作认真负责，能坚持一年或以上。

作为专委简报编委会成员将计入委

员年度贡献。感兴趣加入编委会的委员，请简单介绍自己的工作经历，并参考往期简报内容列出两个感兴趣的版块选项，直接发邮件至 ccfcvn@outlook.com。报名截止到11月25日，专委秘书处讨论后决定人选。

（责任编辑：李实英）

# CCCV2017

## 中国计算机视觉大会圆满闭幕

中国科学院计算技术研究所 **王瑞平** 湖南大学 **李实英**



2017 中国计算机视觉大会 (CCCV) 于 2017 年 10 月 11 日至 14 日在万丽天津宾馆成功举办, 约 1200 名参会者汇聚于此, 共同分享最新的研究成果、创新思想和产业应用。



CCCV 2017 是继 CCCV2015 在西安举办之后, 由中国计算机学会 (CCF) 主办、计算机视觉专委会 (CCF-CV) 和中国民航大学承办、天津大学和南开大学联合协办的计算机视觉领域大会。中科院**谭铁牛**院士、北京大学**查红彬**教授和中国民航大学**杨金锋**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作特邀报告, 展示环节形式多样、内容丰富, 为从事计算机视觉研究的学生、教师和企业研究人员提供了一场精彩的学术盛宴。本届大会的另一大特色是阵容强大的**国际指导委员会**, 其成员均为当前

活跃于国际计算机视觉前沿的资深学者, 具有丰富的国际大会组织经验, 他们的加入极大提升了本届大会的国际化水平。



### 开幕式



CCCV2017 开幕式由大会主席**杨金锋**教授主持。中国民航大学**吕宗平**副校长致欢迎词, 介绍了中国民航大学的发展历史和办学理念, 以及计算机科学与技术学院的特色研究方向与主要研究成果, 并提出将举办更多这样人气旺盛的学术会议。天津科协**卢双盈**副主席热情欢迎参会专家和学者, 并预祝大会成功举办。计算机视觉专委会副主任**查红彬**教授致辞, 介绍了计算机视觉在人工智能研究领域的重要地位和中国计算机视觉研究盛况, 并感谢组委会、赞助企业、志愿者和参会人员为大会成功举办做出的贡献。大会程序主席、南开大学**程明明**教授介绍了会议日程

安排, 本届大会的特点是内容丰富多样, 包括特邀报告、专题论坛、口头论文、墙报展示、专题竞赛、企业应用展示、讲习班等七个版块。



### 特邀报告

本届大会安排了五场特邀报告, 五位特邀讲者均为计算机视觉领域在国际上享有盛誉的学者。这五场特邀报告分别来自: 微软研究院首席研究员、IJCV 前主编 **Katsushi Ikeuchi** 教授做主题报告:



《AI, Robotics and Computer Vision: Retrospective and Perspective Overview》, 苏黎世联邦理工学院 (ETH Zurich) 教授、IEEE 成员 **Marc Pollefeys** 教授做主题报告: 《Semantic 3D Scene Understanding》, 南加州大学教授、IEEE/IAPR/AAAI 成员 **Gérard Medioni** 教授做主题报告: 《A Decade of Research on Faces in the Wild》,



中科院计算所研究员、IEEE/IAPR 成员、CCF 会士**陈熙霖**研究员做主题报告：《从物体识别到场景理解—计算机视觉的未来十年机遇和挑战》，多伦多大学教授、IEEE TPAMI 现任主编 **Sven Dickinson** 教授做主题报告：《The Role of Symmetry in Human and Computer Vision》。这些特邀报告思想深邃、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、引领性的学术观点和思考，激发了大家的浓厚兴趣和强烈反响。



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：北京大学**林宙辰**教授主持的“深度视频分析”论坛、图像视频大数据产业技术创新战略联盟副秘书长**黄晨**博士主



持的“计算机视觉的应用与未来”论坛、北京大学**刘家瑛**副教授主持的“医学影像分析”论坛、哈尔滨工业大学**左旺孟**教授主持的“大规模视频场景下的手势识别”论坛、以及电子科技大学**李永杰**教授主持的“从生物视觉到计算机视觉”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。



论文宣讲与展示



CCCV2017 大会共收到有效投稿 465 篇，录用 187 篇，其中 Oral 论文 32 篇、Poster 论文 155 篇。在 12-14 日上午，共分 7 个主题进行了 Oral 论文的口头宣讲，在 12、13 日的下午进行了 Poster Spotlight 快放，并在之后的 Poster 展示时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，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“基于 FPGA 的视觉计算”、“心血管 OCT 易损斑块识别技术”、“遥感图像目标提取技术”等三个专



题竞赛，吸引了来自国内高校与研究院所的共计 183 支队伍注册参赛。经过激烈角逐，上海交通大学代表队、哈尔滨工业大学个体化心血管代表队、哈尔滨工业大学深圳研究院 HGC 代表队分别获得这三项竞赛的第一名。14 日上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。



## 企业展示与宣讲

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 21 家企业的赞助，其中钻石赞助商 4 家、铂金赞助商 3 家、金牌赞助商 4 家、银牌赞助商 3 家和铜牌赞助商 7 家。大会在 12、13 日上午安排了两场工业界前沿技术宣讲，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

大会首日安排了会议晚宴，晚宴上公布了由大会评奖委员会评选出的 CCCV2017 大会**最佳学生论文提名奖、最佳学生论文奖、最佳论文提名奖和最佳论文奖各 1 名**，由企业代表赞助奖品并进行颁奖。此外，举行了由计算机视觉专委会评定的终身学术贡献奖和服务贡献奖颁奖仪式。2017 年度的**终身学术贡献奖**颁发给为中国计算机视觉领域的发展做出突出贡献的中科院自动化所**马颂德**研究员，专委会副主任**查红彬**教授为其颁发证书和奖品。**马颂德**研究员获奖后发表了声情并茂、热情洋溢的获奖感言。2017 年度的专委**服务贡献奖**颁发给了中国民航大学**杨金锋**教授、南开大学**程明明**教授和湖南大学**李实英**副教授，专委会副主任**王涛**博士为三人颁发证书和奖品。



## 讲习班

本届 CCCV2017 大会首次增设了讲习班版块，包括 4 个热门专题的全景系统性介绍，分别是：行人再辨识研究进展（讲者：**廖胜才**，**杨阳**，**石海林**）、计算机视觉中的基于学习的偏微分方程（讲者：**刘日升**，**方聪**）、基于图像的大规模场景三维重建（讲者：**申抒含**，**崔海楠**）、距离度量学习及其在计算机视觉中的应用（讲者：**鲁继文**，**王瑞平**）。4 个讲习班安排在主会开始前一天的 11 日下午两两并列进行，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。

## 大会期间举办的若干工作会议

伴随本届大会的进行，本着高效、精简的原则，CCF-CV 专委中的若干工作小组在同期举办了相应的工作会议，并且达成了各自的预期目标，达到了各自的预期成果。

**工作会议之一：CCF-CV 专委会 2017 年度工作会议。**

**工作会议之二：CCF-CV 专委会青年工作组第一次工作会议。**

**工作会议之三：CCF-CV 简报编委会第一次工作会议。**

**工作会议之四：全国信标委图形图像分委会计算机视觉标准工作组启动会议。**

## 感谢幕后的组织团队

为期 4 天的 CCCV2017 大会圆满落幕，会议的成功举办得益于大会组委会的通力合作，特别是中国民航大学**杨金锋**教授、**贾桂敏**教授，南开大学**程明明**教授、**杨巨峰**教授，天津大学**胡清华**教授、**朱鹏飞**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顺利完成。



## 延伸阅读：计算机视觉专业委员会

中国计算机学会计算机视觉专委会 (CCFCV) 成立于 2013 年 10 月 26 日，是直属于中国计算机学会的计算机视觉领域的专业分支机构。计算机视觉专委会的目标是就计算机视觉学科的专业内容更好地开展学术/技术交流、发展战略研究，促进国内学者间的了解与合作，推动国内计算机视觉学科发展，提升我国计算机视觉研究在国际领域的影响力。

计算机视觉专委会自成立以来，开展了 CCF-CV 走进高校系列报告、CCF-CV 走进企业系列交流活动、专委简报、专委微信公众号以及专委网站建设 (ccfcv.ccf.org.cn) 等工作，更好地服务于专委委员，推动计算机视觉领域国内外同行以及学术界和工业界之间的紧密联系和发展。

(责任编辑：虞晶怡)

# 目标跟踪算法综述

大连理工大学 卢湖川

## 一、引言

目标跟踪是计算机视觉领域的一个重要问题，在运动分析、视频压缩、行为识别、视频监控、智能交通、机器人导航等领域有着广泛的应用。如图 1 所示，目标跟踪的主要任务是给定目标物体在第一帧视频图像中的位置，通过外观模型和运动模型估计目标在接下来的视频图像中的状态。目标跟踪主要可以分为 5 部分，分别是运动模型、特征提取、外观模型、目标定位和模型更新。运动模型可以依据上一帧目标的位置来预测当前帧目标可能出现的区域，现在大部分算法采用的是粒子滤波或相关滤波的方法来建模目标运动。随后，提取粒子图像块特征，利用外观模型来验证运动模型预测的区域是被跟踪目标的可能性，据此进行目标定位。由于跟踪物体先验信息的缺乏，需要在跟踪过程中实时进行模型更新，使得跟踪器能够适应目标外观和环境的变化。尽管在线目标跟踪的研究在过去几十年里有很大进展，但是由被跟踪目标外观及周围环境变化带来的困难使得设计一个鲁棒的在线跟踪算法仍然是一个富有挑战性的课题。本文将最近几年本领域相关算法进行综述。

## 二、目标跟踪研究现状

### 1. 基于相关滤波的目标跟踪算法

在相关滤波目标跟踪算法出现之前，大部分目标跟踪算法采用粒子滤波框架来进行目标跟踪，粒子数量往往成为限制算

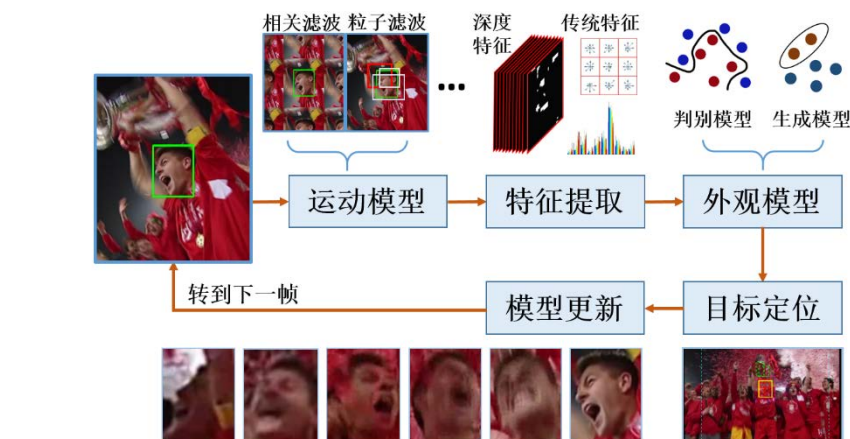


图 1 目标跟踪算法流程图

法速度的一个重要原因。相关滤波提出了一种新颖的循环采样方法，并利用循环样本构建循环矩阵。利用循环矩阵时域频域转换的特殊性质，将运算转换到频域内进行计算，大大加快了分类器的训练。同时，在目标检测阶段，分类器可以同时得到所有循环样本得分组成的响应图像，从而根据最大值位置进行目标定位。相关滤波用于目标跟踪最早是在 MOSSE [1] 算法中提出的。发展至今，很多基于相关滤波的改进工作在目标跟踪领域已取得很多可喜的成果。

### 1.1. 特征部分改进

MOSSE[1] 算法及在此基础上引入循环矩阵快速计算的 CSK[2] 算法均采用简单灰度特征，这种特征很容易受到外界环境的干扰，导致跟踪不准确。为了提升算法性能，CN [3] 算法对特征部分进行了优化，提出 CN (Color Name) 空间，该空间通道数为 11 (包括黑、蓝、棕、灰、绿、橙、粉、紫、红、白和黄)，颜色空

间的引入大大提升了算法的精度。

与此类似，KCF [4] 算法采用方向梯度直方图 (HOG) 特征与相关滤波算法结合，同时提出一种将多通道特征融入相关滤波的方法。这种特征可以提取物体的边缘信息，对于光照和颜色变化等比较鲁棒。

方向梯度直方图 (HOG) 特征对于运动模糊、光照变化及颜色变化等鲁棒性良好，但对于形变的鲁棒性较差；颜色特征对于形变鲁棒性较好，但对于光照变化不够鲁棒。STAPLE [5] 算法将两种特征进行有效地结合，使用方向直方图特征得到相关滤波的响应图，使用颜色直方图得到统计得分，两者融合得到最后的响应图像并估计目标位置，提高了跟踪算法的准确度，但也使得计算稍微复杂了一些。

深度特征能够有效地建模物体语义信息，对于物体外观及周围环境变化具有很好的适应能力。Ma 等人提出 HCFT[6] 算

法，将深度特征与相关滤波算法相结合，取得了很好的效果。算法指出深度神经网络不同层的特征具有不同的特点，浅层特征包含更多的位置信息，但语义信息不明显；深层特征包含更多的语义信息，抗干扰能力较强，但位置信息弱化。如图 2，算法利用 VGG 网络的三层输出特征分别训练三个相关滤波器，将得到的响应图按权重相加进行最终定位。

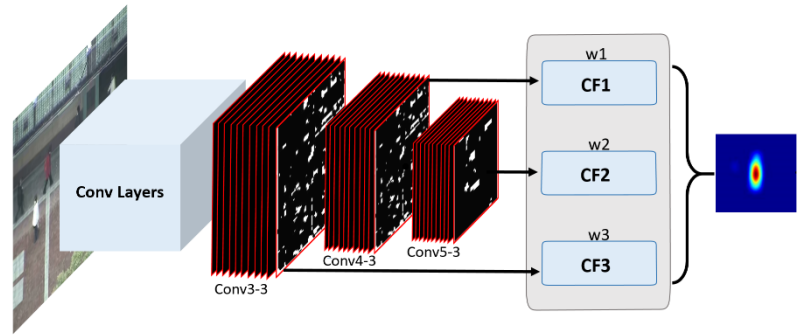


图 2 HCFT 算法流程图

## 1.2. 引入尺度估计

尺度变化是目标跟踪中一个比较常见的问题。SAMF[7]和 DSST[8]算法在 KCF[4]的基础上引入了尺度估计。前者使用 7 个比较粗的尺度，使用平移滤波器在多尺度图像块上进行检测，选取响应值最大处所对应的平移位置和目标尺度；后者分别训练平移滤波器和尺度滤波器，使用 33 个比较精细的尺度，先用平移滤波器进行位置估计，然后在该位置处使用尺度滤波器进行尺度估计。这两种尺度估计的算法也是后来算法中经常使用的两种方法。

## 1.3. 基于分块算法

此外，跟踪过程中的物体通常用一个矩形框选取，由于跟踪的物体一般不为矩形，目标图像块不可避免地引入背景信息，导致跟踪不准确。为解决这一问题，可以采用分割算法或关键点算法来表示目标。基于分割的物体表示虽然可以很好地表示目标形状，但计算量过大，会影响跟踪算法的速度；而基于关键点的目标表示难以获取跟踪目标的整体特征，也不是一个很好的方法。为解决这一问题，Li[9]等人提出选取有效局部图像块来表示物体结构，利用每个局部图像块的响应图像来计算其置信度，这些响应图像按一定方法组合之后可以大致表示出图中物体的形状。算法采用霍夫投票的方法融合多个响应图像，估计目标位置和尺度。

## 1.4. 边界效应改进

边界效应也是影响滤波器性能提升的一个主要问题。在训练阶段，由于密集样本是经过中心图像块循环移位得到的，只有中心样本是准确的，其他的样本都会存在位移动边界，导致训练的分类器在物体快速移动时不能准确地跟踪。大部分算法的解决方案是在图像上加上余弦窗，弱化图像边界对结果的影响，这样只要保证移位后图像中心部分是合理的就可以。虽然增加了合理样本的数量，但仍不能保证所有训练样本的有效性；另外，加入余弦窗也会使得跟踪器屏蔽了背景信息，只接受部分有效信息，降低了分类器的判别能力。

为了克服边界效应，SRDCF[10]采用大的检测区域，并且在滤波器系数上加入权重约束，越靠近边缘权重越大，越靠近中心权重越小，这样可以使得滤波器系数主要集中在中心区域。由于滤波器在整个搜索区域内移动来检测图像块的相关性，因此在克服边界效应的同时不会忽略边缘物体的检测。

## 2. 基于深度学习的目标跟踪算法

### 2.1. 基于判别模型的算法

基于判别模型的深度学习算法的出发点是训练分类器用来区分前景和背景。分类器性能的好坏决定了跟踪器的性能。深度学习最早应用于目标跟踪领域的 DLT[12]算法就是基于判别模型实现的。

该方法先使用 4 个堆叠的栈式降噪自编码器在大规模自然图像数据集上进行无监督训练获得物体表征能力，然后在解码器后面加入分类层用于目标和背景图像块。由于训练数据不足，网络并没有获得很好的表征能力，算法精度不高，但是它提出的“离线预训练+在线微调”的方法为深度学习在目标跟踪中的应用提供了一个可行的方向，之后很多算法都采用这种方法来进行目标跟踪。

近几年来，随着大规模图像分类比赛的流行，很多典型卷积网络被应用到图像处理领域，也出现了很多基于这些网络的目标跟踪算法。2015 年，王立君等人提出的 FCNT[13]算法就利用了 VGG-16 网络。算法提出深度神经网络不同层的特征具有不同的特点，浅层特征含有较多位置信息，深层特征含有更多语义信息，而且深度特征存在大量冗余。因此，算法针对 Conv4-3 和 Conv5-3 两层输出的特征图谱，训练特征选择网络分别提取有效的特征，然后将选好的特征输送到各自的定位网络中得到热力图，综合两个热力图得到最终的热力图用于目标定位。算法利用不同层特征相互补充，达到有效抑制跟踪器漂移，同时对目标形变更加鲁棒的效果。

为了扩展 CNN 在目标跟踪领域的的能力，需要大量的训练数据，但这在目标跟踪中是很难做到的。MDNet[14]算法提出了一种解决该问题的思路。算法采用

VGG-M 作为网络初始化模型，后接多个全连接层用作分类器。训练时，每一个跟踪视频对应一个全连接层，学习普遍的特征表示用来跟踪。跟踪时，去掉训练时的全连接层，使用第一帧样本初始化一个全连接层，新的全连接层在跟踪的过程中继续微调，来适应新的目标变化。这种方法使得特征更适合于目标跟踪，效果大大提升。由此可以看出，通过视频训练的网络更适合目标跟踪这一任务。

### 2.1. 基于生成模型的算法

基于生成模型的深度目标跟踪算法主要通过神经网络来学习模板与候选样本之间的相似程度。除了与传统的相关滤波算法结合之外，大部分算法都采用粒子滤波框架。基于粒子滤波框架的深度学习算法，由于需要提取多个图像块的深度特征，算法速度难以达到实时性要求。

为提升算法速度，David Held 等人提出 GOTURN[15]算法，如图 3 所示，将上一帧的目标和当前帧的搜索区域同时经过 CNN 的卷积层，级联特征输出通过全连接层，回归当前帧目标的位置，由于没有模板更新，算法速度可达 100 多帧/秒。与此类似，YCNN[16]算法构建孪生网络输出目标概率图，得到目标状态。这种基于模板匹配的目标跟踪算法，用一条支路保存模板信息，为目标跟踪提供先验信息，取代了全连接层在线更新，算法速度一般较快。

## 三、目标跟踪的最新研究进展

### 1) 强化学习成功应用

近几年，强化学习方法受到广泛关注，作为一种介于半监督与无监督之间的训练方法，非常适用于目标跟踪这个缺乏训练样本的领域。ADNet[17]算法是强化

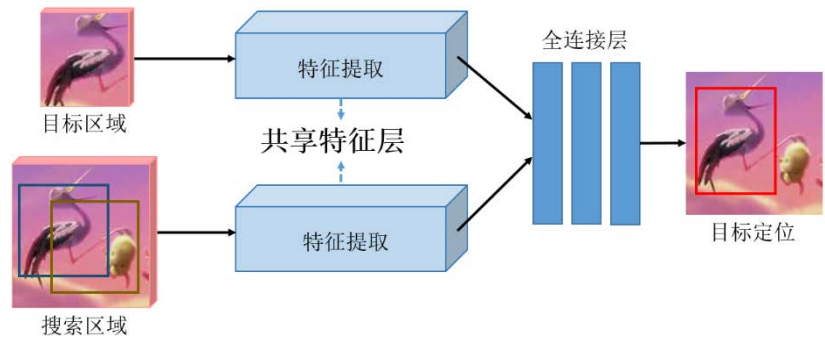


图 3 GOTURN 算法流程图

学习在目标跟踪领域的一个成功应用。算法通过强化学习得到一个智能体来预测目标框的移动方位及尺度变化，在当前帧中，以前一帧的目标位置为初始点，经过多次方位估计、位移和尺度变化，得到最后的目标位置。与此不同，算法[18]利用策略学习得到智能体用于模板选择。无监督或弱监督学习是目标跟踪领域的一个新兴方向，具有巨大的潜力。

### 2) 速度方面有效改进

虽然深度特征具有传统特征无法比拟的抗干扰能力，但是一般提取速度较慢，而且特征中存在大量冗余。当算法精度达到一定标准之后，很多方法开始着力解决算法速度问题。孪生网络[19]是其中的一个典型例子，采用两路神经网络分别输入目标模板和搜索图像块，用来进行模板匹配或候选样本分类。其中一路神经网络对于模板信息的保存可以提供跟踪物体先验信息，取代网络在线更新，大大节省了速度。另外，对深度特征进行降维或自适应选择也是加速算法的有效途径。

### 3) 传统跟踪算法的网络化

深度学习方法具有优秀的特征建模能力，相关滤波算法利用循环矩阵的性质进行计算，速度很快。很多算法着眼于将相关滤波的思路应用到网络中去。CFNet[20]算法将滤波器系数转换为神经网络的一层，推导出前向与后向传播的公式，实现

了网络的端到端训练，算法速度可达每秒几十帧，满足了实时性的要求。传统算法发展至今，有很多可以借鉴的巧妙方法，如果能将其应用到神经网络中去，将对算法效果有较大的提升。

## 四、未来发展方向

目标跟踪的未来发展方向可总结为下面三个方面：

1) **无监督或弱监督方向**。限制深度目标跟踪算法发展的一个主要原因是训练样本的缺乏，没有像图像分类任务那样的大规模训练样本，而且在线跟踪时只能给定第一帧的信息，难以训练一个适合当前跟踪物体的网络模型。近几年来，无监督或弱监督方法受到人们的广泛关注。也有一些算法开始尝试将强化学习应用到目标跟踪领域。对抗网络可以生成迷惑机器的负样本增强分类器判别能力。这些无监督和弱监督的方法可以有效地解决目标跟踪领域样本不足的问题。

2) **速度改进方向**。速度是评价在线目标跟踪算法的一个重要指标。由于深度神经网络复杂的计算及模型更新时繁琐的系数，现存大部分深度目标跟踪算法速度都比较慢。很多深度目标跟踪算法采用小型神经网络（如 VGG-M）来提特征。另外，跟踪中只给定第一帧目标位置，缺少跟踪物体先验信息，这就要求模型实时更

新来确保跟踪精度，而这在深度目标跟踪算法中往往非常耗时。一些算法采用孪生网络结构来保存先验信息，代替模型在线更新，使得算法速度得以提高。深度特征的高维度也会影响跟踪算法的速度，如果能够提出有效的特征压缩方法，不管对算法速度还是精度都会有所帮助。只有高速且有效地算法才具有实际的应用价值。

**3)有效网络结构探索。**影响深度神经网络效果的两个主要因素是网络结构和训练数据。现存大部分深度目标跟踪算法均采用 CNN 结构，虽然 CNN 结构具有很好的特征提取能力，但是难以建模视频帧中的时间连续性信息。一些算法采用 RNN 结构构建目标跟踪模型，但效果并不是很突出，仍需探索发展。还有一些新型的网络架构（如 ResNet 和 DenseNet 等），在图像分类领域取得很好的效果，这些网络架构是否能在目标跟踪领域成功应用也是令人期待的方向。

### 参考文献：

- [1] Henriques J F, Rui C, Martins P, et al. Exploiting the Circulant Structure of Tracking-by-Detection with Kernels, In Proc. of the ECCV,2012:702-715.
- [2] Rui C, Martins P, Batista J. Exploiting the circulant structure of tracking-by-detection with kernels, In Proc. of the ECCV, 2012:702-715.
- [3] Danelljan M, Khan F S, Felsberg M, et al. Adaptive Color Attributes for Real-Time Visual Tracking, In Proc. of the CVPR, 2014:1090-1097.
- [4] Henriques J F, Rui C, Martins P, et al. High-Speed Tracking with Kernelized Correlation Filters. TPAMI, 2014, 37(3):583-596.
- [5] Bertinetto L, Valmadre J, Golodetz S, et al. Staple: Complementary Learners for Real-Time Tracking. 2015,38(2):1401-1409.
- [6] C. Ma, J. Huang, X. Yang, M. Yang, Hierarchical convolutional features for visual tracking, In Proc. of the ICCV, 2015:3074-3082.
- [7] Li Y, Zhu J. A Scale Adaptive Kernel Correlation Filter Tracker with Feature Integration, In Proc. of the ECCVW, 2014: 254-265.
- [8] Danelljan M, Häger G, Khan F S. Accurate scale estimation for robust visual tracking, BMVC, 2014.
- [9] Li Y, Zhu J, Hoi S C H. Reliable Patch Trackers: Robust visual tracking by exploiting reliable patches, In Proc. of the CVPR, 2015:353-361.
- [10] M. Danelljan, G.Häger, F. S. Khan, M. Felsberg, Learning spatially regularized correlation filters for visual tracking, In Proc. of the ICCV, 2015:4310-4318..
- [11] Mueller M, Smith N, Ghanem B. Context-Aware Correlation Filter Tracking. In Proc. of the CVPR, 2017.
- [12] N.Wang, D. Yeung, Learning a deep compact image representation for visual tracking, in Proc. of the Advances in Neural Information Processing Systems, 2013:809-817.
- [13] L. Wang, W. Ouyang, X. Wang, H. Lu, Visual tracking with fully convolutional networks, In Proc. of the ICCV, 2015: 3119-3127.
- [14] H. Nam, B. Han, Learning multi-domain convolutional neural networks for visual tracking, In Proc. of the CVPR, 2016: 4293-4302.
- [15] D. Held, S. Thrun, S. Savarese, Learning to track at 100 FPS with deep regression networks, in Proc. of the ECCV, 2016:749-765.
- [16] J. Redmon, S. K. Divvala, R. B. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, In Proc. of the CVPR, 2016:779-788.
- [17] S. Yun, J. Choi, Y. Yoo, K. Yun, J. Y. Choi, Action-decision networks for visual tracking with deep reinforcement learning, In Proc. of the CVPR, 2017.
- [18] Choi J, Kwon J, Lee K M. Visual Tracking by Reinforced Decision Making. arXiv abs/1702.06291.
- [19] L. Bertinetto, J. Valmadre, J. F. Henriques, A. Vedaldi, P. H. S. Torr, Fully-convolutional siamese networks for object tracking, in Proc. of the ECCV, 2016:850-865
- [20] J. Valmadre, L. Bertinetto, J. F. Henriques, A. Vedaldi, P. H. S. Torr, End-to-end representation learning for correlation filter based tracking, In Proc. of the CVPR, 2017..

(责任编辑：王金甲)



卢湖川，CCF 计算机视觉专委会常务委员、IEEE 和 CCF 高级会员、大连理工大学电子信息与电气工程学部副部长、教授。研究领域为计算机视觉和模式识别，目前担任 IEEE Transaction on Cybernetics 编委。近五年，发表 CCF A 类论文 40 余

篇，Google Scholar 总引用近 7000 次。曾获教育部自然科学二等奖及多项国际学术论文奖。

Email: lhchuan@dlut.edu.cn

# 基于跨域学习的行人再标识

中山大学 谢晓华 赖剑煌 郑伟诗

行人再标识 (Person Re-Identification) 亦称行人再识别或者重识别, 是近年来计算机视觉领域的研究热点问题之一。其主要任务是将监控摄像头网络中不同摄像头所捕捉到的行人进行匹配关联。运用行人再标识技术, 可以获取摄像头网络下行人的移动轨迹, 给公安监控、客流管控等带来较大的便利。

在实际应用中, 监控视频网络中不同摄像头的视域一般不存在交叠, 且不同视域对应的场景、视角、光照条件均有所不同。加之行人本身存在较大的非刚性形变, 这些因素均给行人再标识带来了较大的挑战。从技术上而言, 两个不同视域下的行人匹配往往成为一个跨域识别问题 (Cross-Domain Recognition)。在这种情况下, 传统采用同一字典特征表达、共同子空间映射、对称度量学习的方法一般难以获得令人满意的识别效果。

针对行人再识别中的跨域识别问题, 中山大学研究组开展了深入的摸索, 提出了一系列具有影响力的理论分析和新方法: 1) 在特征提取阶段利用摄像头之间的关联信息, 提出了视角相关的行人深度特征提取方法; 2) 首次提出非对称距离学习的概念, 并提出跨摄像头鉴别成分分析, 有效解

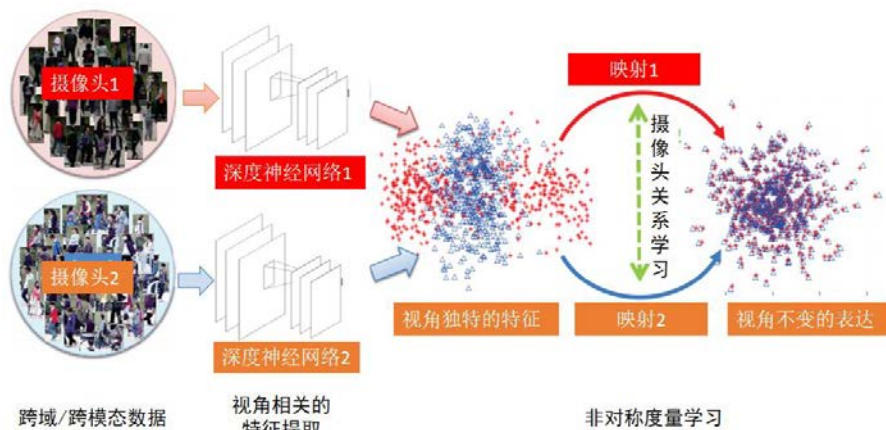


图1 基于视角相关特征提取、非对称度量学习的行人再标识框架

决了跨域特征之间的匹配问题; 3) 提出基于镜像表示的跨域行人匹配框架, 对不同摄像头学习出不同的投影矩阵, 表达出不同摄像头的相关性; 4) 提出联合深度学习的排序框架, 并将之应用在行人再标识问题上, 新提出的学习排序算法与深度卷积神经网络形成无缝连接。这些研究成果构建出一个有效解决行人再识别中跨域匹配问题的新框架 (图1)。此外, 我们还提出利用上下文信息提高行人再识别的准确度, 其中采用时间、空间约束以及群组等线索来辅助行人识别。

在实际监控应用中, 夜晚一般利用深度相机或者红外相机进行数据采集。因此需要解决“可见光-红外”或者“可见光-深度”

跨模态图像之间的匹配问题。对此, 我们对跨域行人再标识框架进行相应拓展, 发展出若干跨模态行人再标识方法。这些方法进而与哈希编码结合, 通过联合学习获得快速的行人再标识算法。

以上工作获得了包括国家自然科学基金委-广东省联合基金中心项目(重大)在内的多个科研项目支持。基于这些工作已在包括IEEE TPAMI、IEEE TIP、PR、IEEE TCSVT、ICCV、CVPR、IJCAI在内的高水平国际期刊和会议发表相关论文十多篇, 并获得CCCV 2017唯一的最佳学生论文奖, 获得全国研究生“智慧城市”竞赛技术一等奖。

(责任编辑: 苏航)



谢晓华, CCF 计算机视觉专委会委员, 中山大学特聘研究员。主要研究领域为模式识别、计算机视觉。

Email: xiexiaoh6@mail.sysu.edu.cn



赖剑煌, CCF 计算机视觉专委会副主任, 中山大学教授。主要研究领域为模式识别、计算机视觉。

Email: stsljh@mail.sysu.edu.cn



郑伟诗, CCF 计算机视觉专委会委员, 中山大学教授。主要研究领域为模式识别、计算机视觉。

Email: zhwshi@mail.sysu.edu.cn

# 细粒度级别图像检索

南京大学 吴建鑫 魏秀参

图像检索(Content-based image retrieval)是近年来计算机视觉和多媒体领域的研究热点之一。传统图像检索任务是给定目标图像,在图像数据库中检索类似复制的图像(near-duplicated images),并返回相似度最大者(如图1所示)。可以发现,传统图像检索图像是在不同光照不同时间下同一地点的图像,这类图像不会有形态、颜色、甚至是背景的差异。而“细粒度级别”的图像检索是我们近期提出的一个图像检索领域的新问题。



图1 传统图像检索任务。左侧单列为目标图像,右侧为返回的检索结果

如图2,细粒度级别图像检索需将同为“绿头鸭”的图像从众多不同鸟类的图像中返回;同样,需要将“劳斯莱斯幻影”从众多不同品牌不同车型的图像中检索出来。细粒度图像检索的难点,一是图像标记的粒度非常细微(如鸟的不同子类);二



图2 细粒度级别图像检索任务

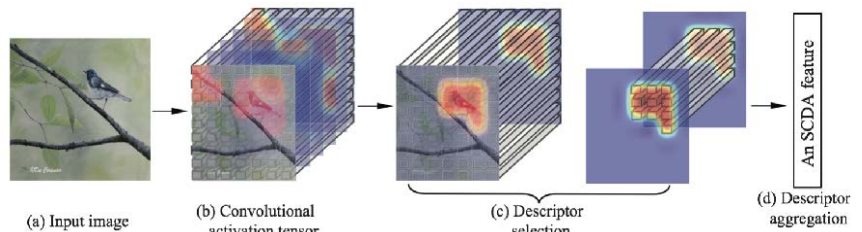


图3 针对细粒度级别图像检索的选择性卷积特征描述子融合方法

是对细粒度图像而言,哪怕是属于同一子类的图像本身也具有形态、姿势、颜色、背景等巨大差异。可以说,细粒度级别图像检索是图像检索和细粒度级别图像分析领域的一项具有新鲜生命力的研究课题。

针对上述挑战,我们提出了首个基于深度学习的细粒度级别图像检索方法“选择性卷积特征描述子融合”(SCDA)。同其他深度学习框架下的图像检索工作一样,在SCDA中,细粒度级别图像作为输入送入基于ImageNet预训练的卷积神经网络中得到卷积特征/全连接特征。区别于传统图像检索的深度学习方法,针对细粒度图像检索问题,我们不仅发现卷积特征优于全连接层特征,同时提出要对卷积描述子进行选择以提高细粒度图像检索效果——保留对应主要物体的特征描述子,而去除背景或噪声对应的特征描述子。不过由于图像检索任务没有提供图像标记信息,这就要求算法在无监督条件下依然可以完成物体的定位,从而根据定位结果进行卷积

特征描述子的选择。

具体而言,SCDA将最后一层卷积特征张量( $H \times W \times D$ )在深度方向做加和,之后可得 $H \times W \times 1$ 的矩阵,我们称其为“汇合图”。之后,计算其中所有 $H \times W$ 元素的均值,此均值 $m$ 便是该图物体定位的关键:汇合图中大于 $m$ 的元素位置的卷积特征需保留;小于的则丢弃。而这些被保留下来的位置,也就对应了应保留卷积特征描述子的位置。对保留下来的深度特征,分别做以平均和最大池化操作,之后级联组成最终的图像表示。

实验结果表明,在细粒度级别图像检索中,SCDA获得了最好结果;同时SCDA在传统图像检索任务中,也可取得同目前传统图像检索任务最好方法相差无几(甚至优于)的结果。该研究的相关成果于2017年发表于国际期刊IEEE Trans. on Image Processing.

(责任编辑:杨巨峰)



吴建鑫, CCF 计算机视觉专委会委员, 南京大学教授。主要研究领域为计算机视觉和机器学习。

Email: wujx2001@nju.edu.cn

魏秀参, 南京大学博士生。主要研究领域为计算机视觉和机器学习。

Email: weixs@lamda.nju.edu.cn

# 基于图像的大场景三维重建：混合式从运动恢复结构

中国科学院自动化研究所 崔海楠 申抒含

从运动恢复结构是指通过摄像机运动恢复场景的三维结构：通过输入多幅二维图像，输出场景稀疏点云与对应摄像机姿态，因此通常也被称为稀疏重建。

根据摄像机姿态的估计模式不同，传统方法大致可以分为两类：增量式稀疏重建和全局式稀疏重建。对于大场景三维重建而言，增量式方法鲁棒，但是由于反复调用捆绑调整，效率较低，同时误差的累积容易造成场景漂移；全局式方法只进行一次捆绑调整，效率高但是通常依赖于两视图的外极几何关系，对特征匹配外点不够鲁棒。为此，中国科学院自动化研究所机器视觉组提出了一种混合式的从运动恢复结构算法，充分继承了增量式和全局式两种模式的优点：全局式估计摄像机旋转矩阵，增量式估计摄像机位置。

传统全局式获取摄像机旋转矩阵的方法主要是通过最小化两视图相对旋转矩阵与绝对旋转矩阵之间的误差。该方法只考虑了外极几何图中边上的误差，而没有考虑到图像之间的连接紧密程度，速度虽快但是当场景存在分块性质时求取不够准

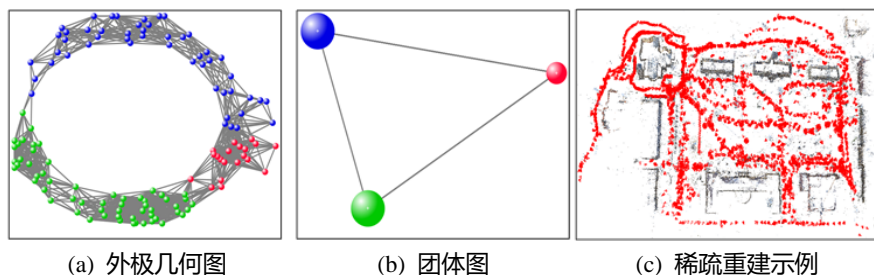


图 1 混合式从运动恢复结构计算框架

确。为此，课题组提出了一种基于社交网络中团体检测的方法，全自动地确定外极几何图中是否存在团体分块性质，以及如何进行团体检测。在团体分块之后，每个团体内部的图像之间的连接紧密程度，通常要高于整个外极几何图中的图像连接紧密程度，因此更适合进行旋转矩阵均匀化。首先对每一团体内部进行局部摄像机旋转矩阵求取，然后再利用团体之间的外极几何边上的相对旋转矩阵，进行一次整体的摄像机旋转矩阵均匀化，使得所有团体内部的旋转矩阵都对齐到同一个全局坐标系中，从而获得所有摄像机的旋转矩阵。

在已知摄像机的旋转矩阵之后，传统方法采用基于相对平移均匀化的方式计算摄像机位置。这种方法依赖于相对平移的

计算精度，而在室外大场景三维重建中，城市建筑的重复纹理频发，特征匹配外点多，使得相对平移计算十分不准确。因此，课题组提出在已知摄像机旋转矩阵的情况下，增量式求取摄像机的位置，通过RANSAC机制和迭代捆绑调整操作对抗特征匹配外点。由于摄像机旋转矩阵已知，只需要两个点就可以估计摄像机的位置，减少了传统P3P或者EP4P方法在RANSAC时的最小配置，使得摄像机位置的求取更加稳定和鲁棒。此外，在迭代捆绑调整过程中，固定摄像机旋转矩阵而仅优化摄像机位置的操作，不仅可以极大地加速了捆绑调整的速度，同时也有效减少了增量式过程中的误差累积。该研究的相关成果已相继发表于CVPR 2017、3DV 2017。



**崔海楠** 中国科学院自动化研究所助理研究员。主要研究领域为三维计算机视觉、基于图像的大规模场景三维建模。

Email: hncui@nlpr.ia.ac.cn

**申抒含**，CCF 计算机视觉专委会委员，中国科学院自动化研究所副研究员。主要研究领域为三维计算机视觉、场景三维重建、机器人三维环境感知。

Email: shshen@nlpr.ia.ac.cn

# 华南理工大学余志文教授访谈

2017年11月3日晚，专委秘书处在线采访了华南理工大学计算机科学与工程学院副院长、优秀青年学者余志文教授。

余老师好，感谢您百忙之中安排时间接受我们的访谈，让我们有机会分享您研究过程中的经验！

您在短短几年间就在机器学习、模式识别、多媒体方向上做出了令人瞩目的成绩，能否和大家分享一下您在科研道路中获得成功的经验？您有什么成功的秘诀吗？

说成功的话还不敢当，科研的道路是无止境的。科研的路也是没有捷径的，只有多阅读文献才能推动我们前进。但在科研的路上，更重要的还是我们持之以恒的决心，困难总是会有的，看了很多论文但是没 idea、实验做了好几个月但是出不了好结果、论文投了好多次还是被拒等种种挫折相信很多人都会经历，关键在于我们怎样保持乐观的态度，屡败屡战，越挫越勇，通过不断的努力跳出这个瓶颈期。其次是研究的内容最好不要太分散，在前期投入研究的时候，如果这也想学、那也想学，最终的结局只会是各个方向都只懂一些表面的东西，真正要对算法做优化或提出新模型却无能为力。

您发表了多篇高水平论文并被多次引用，今年还荣获国家优秀青年科学基金，在您取得的这些成就中，哪些工作让您印象特别深刻？

应该是我发表第一篇论文的经历。当时刚刚开始在香港城市大学修读博士学位，想把自己的科研成果投到国际知名的学术期刊上。第一篇论文内容是模式识别

相关的，给第一家期刊投过去后，不久就收到了拒绝信；再投第二家，还是被拒绝。屡次被拒，这在求学初期是一件很常见的事。你觉得自己辛苦多时获得的很好的研究，在很多经验丰富的学术前沿专家看来，其实并没有什么价值。但是对于我们这些上课考试一直都能获得好成绩的学生，一下子难以接受这样的打击。还好我当时的导师一直在鼓励我、支持我，让我摆脱了沮丧，最终坚持了下来。后来根据审稿意见对论文进行了大量的修改，终于在第三次投稿时被《IEEE Transactions on Multimedia》采纳。我想对于绝大部分的学生而言，这样的经过诸多波折最后才获得成功的经历，应该是科研道路上最令人难忘的，就像是开启迈向成功的第一步。

您曾主持、参与 20 多个科研项目，发表 100 多篇学术论文，其中包括 27 篇 IEEE/ACM Transactions 期刊长文，您能否分享一下您在撰写论文方面的一些经验？论文写作中需要注意些什么问题呢？

作为一个过来人，我可以跟大家分享一下这方面的经验。对于很多没写过论文的学生来说，写论文的确是一件比较有挑战的事情，一开始我也写的不是很好，一段话即使自己认真琢磨了还是存在语法和逻辑错误，文章总体结构也欠清晰，后来经过导师的多番指点才逐步掌握写作技巧，在一次次的锻炼中理清了论文写作的思路。在论文写作中重要的一点是主线清晰，要以读者的身份问自己：为什么文章要分成这几部分？每个部分有什么关联？当你有了答案之后，文章的结构就会清晰明了，逻辑通顺。其次很重要的一点就是

要充分展现出自己的创新点，让别人知道你论文的贡献之处。在细节上，想要做到句子优美、用词精准，那就需要多阅读相关的文献，了解常用的专业术语以及描述方式。此外，实验要有说服力，怎样多方面展现自己算法的优点，让读者确信你的算法是有效的。做到这些后，审稿人就没什么理由拒绝你的稿件了。

作为 CCF 理事、杰出会员，IEEE、ACM 高级会员和一些刊物的编委，能否结合一些印象比较深的亲身经历，谈谈您在评价一份稿件时看重的内容闪光点？

我觉得最重要的还是创新点，这篇文章所述的工作有没有新意，应该是审稿人最看重的。创新点是这个工作的“不同”，有创新点才意味着这个工作确实是你做的，是一个产出的成果。其次就是这个工作的 motivation，你这个工作是基于什么开展的，可以解决目前哪些不足。从问题出发，找到解决问题的方案，才能体现这个工作的价值。最后就是理论和实验论证，特别是理论论证，严谨的理论论证是非常具有说服力的，而大量的实验论证则说明你对工作进行了验证和修正。

您在集成学习方面颇有成就，能为我们介绍一下这个领域，以及您当初为什么会选择此领域进行深入研究呢？

当初开始从事集成学习是因为看得出集成学习是机器学习里面的一匹黑马，而从目前集成学习在学术界、工业界上的表现也印证了这一点。简单借用“盲人摸象”的典故来描述集成学习：假如把一个大待研究的大数据集比作一匹大象，我们用

于对数据进行挖掘的学习模型就是一个摸象的盲人，一个盲人去摸象是难以感知大象的整体外观的，但如果同时有很多个盲人去摸象，再把感知的信息结合起来，那么，我们就能得出大象的全貌。围绕集成学习这一原理，我们团队提出把多个学习模型有机融合起来的方法，以及各种新的集成学习模型，从海量的数据中快速获取有用的信息。

您担任华南理工大学计算机学院的副院长，在学术研究之外还要进行学院管理，请问您是如何平衡行政与科研两方面的工作的？

我觉得我们作为科研工作者，应该具备一种认知，那就是不管其他工作再怎么忙，研究和教学的工作都不能放下。担任副院长后确实多了很多杂乱的行政工作，需要花费大量的时间，但这并不能成为自己懒惰的借口。作为一名学者，最根本的就是做学术研究，这是我们的使命。以前我比较喜欢和学生一起在实验室学习研究，但是现在因为要处理一些行政工作，不得不长期留在办公室。不过这也只是换了一个地方继续做科研，虽然氛围没有实验室好，但是相关的工作不能落下。只要不是出差，我都会让学生陆续来我办公室一起讨论研究课题，即使讨论过程中偶尔会有其他的学生或者老师带着学院的公务过来打断，但是不能因为怕被打断就不再继续进行科研探讨，这是削足适履的做法。看论文和一些专家的博客分享等也可以在一些碎片化的时间里完成，即使是出差路上的时间也可以利用，总之要时刻保证知识面停留在相关学术发展的最前沿。

能否介绍一下您现在的科研团队？

我每年大概会招 4 个左右的研究生，以及 1 个博士，所以目前团队的规模不算



小。硕士生有十几个，博士有 5 个，平时的培养压力也比较大，需要花不小精力顾及每一位学生的培养。研究方向比较多样化，包括聚类、分类、半监督、集成学习、遗传算法、粒子群算法、自然语言处理、深度学习、强化学习、迁移学习等，一个细致的方向通常我只会安排一到两个学生，通过每周讨论会大家互相之间交流，使得大家了解到机器学习各个方面的技术，成为一个全能型的机器学习科研人员。

研究团队的组建和管理是研究工作顺利开展的基础，能不能分享一下您在学生指导方面的心得？

在管理学生方面，我还需要不断的探索和改进。刚当导师时我才刚从学校里出来，觉得自己还有很多很好的想法没有实现，因此指导学生时总是抓得很紧，好像很迫切地希望把自己当时的想法灌输给他们，让他们来帮我实现，甚至考虑到我当初写论文时吃的暗亏，发表的论文绝大部分内容都由我自己一个人来写或者修改，这其实是不利于学生的科研发展的。几年后慢慢地我开始意识到这样的指导是不正确的，开始转变自己的心态。让学生自己去选择一个感兴趣的方向，然后我再从旁作指导，给他们一些合理的建议，又或者是我指定一个大的方向，让学生自己去探索具体应该做哪一个领域的内容，尽量让他们自己去解决问题，而不是过多干涉学生的想法。而且实践证明这样的学生发展

会更好，能力上能更快的得到提升。

但是还是需要适时地给学生一些压力，我认为人在紧张的状态下更能够发挥自己全部的潜能。有些学生特别是硕士研究生，自律性不够，也没有太高的科研积极性，这些学生往往需要一些外部的压力才能更好地完成科研工作。

在科研成果显著的同时，在课堂教学方面您是否也有些独特的心得？

课堂教学的第一个问题就是教材的选用。有些老师为了减少备课时间，沿用以前的课件及教材。当这个领域的新知识出来后，平均得过三年左右才能出现在他们的课件当中，学生学到的老知识老技术，跟不上技术发展的潮流。作为教师应该不惜一切代价给学生授予最前沿的知识，及时更替教学材料，除了讲授课程内容，还应该根据学科的发展动态、前沿研究，补充相应的知识，将知名学者、著名期刊等全面而系统地介绍给学生，以培养学生的科研兴趣。对于研究生课程的教学则应更加严格，除了要求他们阅读课程相关的前沿技术论文，还应要求他们去分析论文中的不足、总结精辟之处、提出改进方法，有时间条件的应该要求学生撰写课程论文，论文中应有充分的实验，从而全面培养学生的科研能力。

您曾在香港城市大学修读博士学位，能否结合您自身的求学经历，谈谈香港与大陆在科研人才培养方面的异同？

我觉得香港的学术氛围比较好，有了环境的熏陶，学生就会不自觉地静下心来学习，向更高水平的研究者靠拢。我在城大读书的时候，每周导师都会主持两到三次的讨论会，大家轮流汇报一下最近的工作，让每个人都了解团队里的其他成员最近在做什么课题，可以进行相关的交

流学习；另外还需要和大家分享一下最近看的一些论文，自己讲一遍可以梳理内容，对论文的理解更加深刻，同时听的同学也能以一个较快速的方式了解到学术前沿问题。而在国内老师们可能对讨论会就不是那么重视了，通常就是让一两个学生到办公室汇报一下工作，或者是老师去实验室看看学生最近做的一些工作，而不是一个公开的讨论平台。而且在国内做科研，感觉是外部诱惑比较多，大家很难安安静静地一心为科研，总是容易想着说今晚会有什么有趣的活动要参加，或者是周末要去哪里玩。相比之下我在香港学习的时候，大家都是在实验室学习到晚上九点多，一些像我这种英语底子不是很好，需要“勤能补拙”的学生会学习到十一二点，周末除了外出购买一些生活用品外，都会在实验室工作学习，日常除了运动基本就没有什么业余的活动。所以很多老师和学生都羡慕在香港做课题，成果产出率更高，往往能获得不少的成果。

**您成功地把机器学习应用于模式识别、多媒体等应用方向，取得了良好的成果。请问您从事的工作是偏科研还是工程？**

工程很多时候是对现有技术的合理利用，而科研是在技术层次上的创新。要做高水平的学术，我还是把自己定位为科研人员而不是工程人员。我从事的基本都是基础理论研究工作，没有在研发产品方面投入精力。理论研究是整个科研体系的基石，只有把科研的一砖一瓦堆砌起来，才能促进产业界的进步，国家未来的长足发展才有足够的推动力。现在市场上很多先进又好用的软件，像 Siri、天天 P 图，都是通过基础算法的研究而生产出来的。而我们，就是设计这些算法的人。

**您在集成学习、聚类、半监督学习、多目标优化方面已经取得比较丰厚的成果，在**

**未来几年里，您有想过开辟新的研究方向吗？**

那肯定是有这样的想法的。特别是一些学生因为年轻，容易产生一些乍一眼看上去好像有点天马行空的想法，但是如果他愿意用心去做，我也会支持，而且确实有些学生的想法最后也有了一定的成果产出。前段时间我的合作导师 IEEE Fellow 陈俊龙教授 (Philip Chen) 提出了一种叫做 board learning 的算法，是与最近非常热门的 deep learning 相对而提出的，我觉得这个想法就非常新颖，让团队里的每个学生都好好学习这种思想，如果可以，我们也尝试这个方向上能不能取得一些突破。

**能否谈一谈人工智能未来几年您认为比较有前途的研究方向？**

在未来几年里，媒体大数据依然是人工智能的重中之重，所以我们一直把模式识别、多媒体作为研究重点之一。深度学习在近几年比较火，在处理媒体大数据中取得了很好的效果，虽然深度学习不是我们的研究重点，但我们也一直保持着对这个领域的关注。算法的可靠性、准确性与效率性一直是我们的重点，深度学习的准确性是大家认同的，但是它的效率性却一直被诟病。最近我的合作导师陈俊龙教授提出的 board learning 很好地解决了深度学习效率性的问题，通过网络的平铺，进行更丰富的特征提取，仅通过很浅的网络就能达到深度学习所能达到的准确性，这使得 board learning 省去了漫长的训练时间，是一种既高效又准确的算法。随着 board learning 理论的完善，我相信在未来几年里，会有很大的发展。

**您对目前的工作成果是否满意，有没给自己设定一些发展目标呢？**

我觉得还是不要谈目标吧，因为我自己本身是挺享受学术研究这个过程的。对于我带的一些没有太大科研动力的研究生，我可能会给他们定一些小目标来敦促他们，但是像我们这种对科研感兴趣的人，谈目标感觉就太功利了一些，很有可能把兴趣变成了完完全全的任务，能动性反倒是大大地降低了。我觉得华南理工大学是一个很不错的平台，它能帮我申请到科研的经费，在经费上不需要太多的担忧，可以不受外界干扰地，踏踏实实地做科研，做些感兴趣的，有意义的研究。

**很多人眼里，余老师您是一位非常勤奋的研究者，请问在生活中余老师是什么样一个人呢？能否跟大家分享一下您生活的其他方面？**

在生活方面我是一个比较 easy going 的人，尤其是在跟学生相处方面。平时在办公室、讲台上可能看到我是一个比较认真严肃的人，但是在工作之余我从来都是放下教师架子，跟学生打成一片。以前行政事务没那么繁忙的时候，经常会跟学生一起去打羽毛球娱乐一下，周末也会跟家人一起去户外活动，增进感情。我觉得在生活中张弛有度对做好科研大有裨益。

**您一般每周工作多长时间？您是如何平衡科研与日常生活的？**

在时间管理上我也没什么心得，反正就是有空余时间就工作，我很享受工作的过程。最近这两年和香港城市大学合作申请了一个项目，经常需要出差到香港交流，花在路上的时间特别长，有些还难以利用，而且有时还需要利用自己周末甚至是寒假的时间。所以其实我也没有很好地区分学校的科研与家庭的日常生活，只能说该做哪一件事的时候就要专注地做好。在学校就要认真地和学生一起学习讨论，出差参加研讨会和各种会议时也要抓紧机会多接

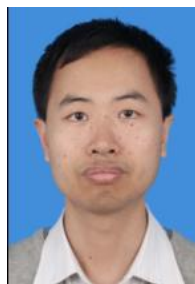
收一些新的知识和想法，提高科研效率。在保证科研成果产出的情况下尽量腾出一些时间来陪伴家人，和家人外出旅游，郊外踏青，碰上节日假期回老家陪陪父母。

对于现在的硕士生来说，毕业之后是继续走学术道路好还是离校找工作好呢？现在有很多研究生觉得很迷茫，您对在读的研究生有哪些建议？

这个问题主要看学生自己的意向。有的同学动手能力比较强，倾向于工程实现，那么毕业之后去企业工作对自己会比较有帮助；而有的同学则对科研比较感兴趣，看文献时有独到的理解，这些同学更适合走学术的道路。我们作为导师应该尊重学生的意愿，让学生得到最大程度发展。但对于那些没有明显倾向的同学，我个人建议，如果自身对科研是有点兴趣的，或者自己在这方面有一定潜力的，都应该尝试一下继续走学术的道路，给自己定高一点的目标，做高水平的学术，为工业界贡献更多高效准确的算法，推动社会科学的发展。

您在机器学习方面取得了很多成果，能为刚入门的年轻学者提供一点建议吗？

我觉得开展科研工作的初期，最重要的是给自己选定一个合适的方向。因为在刚开始做科研的时候，往往会碰上很多失败，只有自己选定了一个方向，才能很好地坚持下去。刚开始入门时要多读一些相关的论文，了解经典的算法以加深对问题的理解，同时了解这个学科方向的发展前沿工作。科研方向做得比较窄，保持对一个方向深入的思考，才比较容易做出成果来。当你对这个问题有了更深刻的认识，才可能得到一些创新的点子，才能做出有价值的成果。



余志文，华南理工大学计算机学院教授、博士生导师、副院长、国家优秀青年科学基金获得者，CCF 理事、CCF 杰出会员、ACM

广州分会副主席、CCF

YOCSEF 广州主席、IEEE、ACM 高级会员，CCF 人工智能与模式识别专委会委员、CCF 计算机视觉专委会委员、CAAI 机器学习专委会委员。于 2008 年在香港城市大学计算机系获博士学位，之后在香港理工大学电子计算系做过博士后研究，2011 年入选教育部新世纪优秀人才支持计划，2013 年获得广东省自然科学基金资助，2017 年获国家优秀青年科学基金资助。主要从事机器学习、数据挖掘、模式识别和人工智能方面的研究，发表学术论文近 100 多篇，其中 27 篇 IEEE/ACM Transactions 系列论文，如：TKDE, TEC, TCYB, TMM, TCSVT, TCBB, TNB 等，1 篇论文获国际会议最佳论文奖，担任 CCF 会刊《计算机应用》编委。主持和参与的项目 20 多项，其中主持国家、省部级项目 11 项。更多信息见个人主页：

[www.hgml.cn/yuzhiwen](http://www.hgml.cn/yuzhiwen)

(责任编辑：韩爱丽 余焯)

## 委员好消息

- 2017 年 9 月 18 日，CCF-CV 专委会委员、中科院自动化所副研究员、银河水滴科技创始人兼首席执行官黄永祯博士入选 2017 年《财富》中国 40 位 40 岁以下的商界精英。黄永祯毕业于中科院自动化所模式识别国家重点实验室，师从著名计算机视觉专家谭铁牛院士。其自主研发产品已经被广泛用于安防反恐、机器人、智能家居、智能商业等领域。

- 2017 年 9 月 21 日，北京市科委发布了 2018 年首都科技领军人才培养工程和北京市科技新星计划入选人员名单。CCF-CV

专委会秘书长、中科院自动化所王亮研究员入选首都科技领军人才培养工程，CCF-CV 专委会委员、北京航空航天大学李甲副教授入选北京市科技新星计划。2018 年度首都科技领军人才培养工程入选人员共 30 人，科技新星计划共 120 人入选。

- 2017 年 10 月 1-4 日，在美国丹佛举办的 IJCB 2017 上，CCF-CV 专委会委员、清华大学冯建江副教授、鲁继文副教授和清华大学深圳研究生院郭振华副研究员指导的论文 Fingerprint Pose Estimation Based on Faster R-CNN 获最佳学生论文提名。

- 2017 年 10 月 12 日，CCCV2017 最佳论文获奖名单出炉。CCF-CV 专委会副主任、北京大学查红彬教授指导的论文 Structure-aware SLAM with Planes in Man-Made Environment 获最佳论文，CCF-CV 专委会委员、中山大学朱俊勇副研究员、赖剑煌教授（专委会副主任）和谢晓华研究员指导的论文 Person Re-identification on Heterogeneous Camera Network 获最佳学生论文。CCF-CV 专委会委员、南京信息工程大学刘青山教授（专委会常委）和袁晓彤教授指导的论文

Hyperspectral Image Classification Using Spectral-Spatial LSTMs 获最佳学生论文提名。会议共收到有效投稿 465 篇, 录用 187 篇, 评出最佳论文、最佳学生论文、最佳论文提名和最佳学生论文提名各 1 篇。

● 2017 年 10 月 12 日, 2017 年度 CCF-CV 颁奖典礼在天津举行, CCF-CV 专委会委员、中国民航大学杨金锋教授、南开大学程明明教授和 CCF-CV 专委会副秘书长、湖南大学李实英副教授荣获本年度 CCF-CV 服务贡献奖。

● 2017 年 10 月 14 日, CCCV 2017 竞赛结果揭晓。在心血管 OCT 易损斑块识别技术挑战赛 (中科院西安光机所、北京健康促进会心血管影像中心承办) 中, CCF-CV 专委会委员、东北大学贾同教授, 清华大学深圳研究生院郭振华副研究员指导的队伍分别获得三等奖。在遥感图像目标提取技术挑战赛 (北京航天宏图承办) 中, CCF-CV 专委会委员、哈尔滨工业大学深圳研究生院徐勇教授指导的 HGC 队获得冠军, CCF-CV 专委会委员、中科院自动化研究所樊彬副研究员指导的 blueskyforever 队获得季军。

● 2017 年 10 月 23 日, 在 ICCV 2017 的 HANDS Workshop 上, CCF-CV 专委会委员、华南理工大学金连文教授等指导的论文 YOLSE: Egocentric Fingertip Detection from Single RGB Images 获得最佳论文奖。

● 2017 年 10 月 23-27 日, 在美国加州山景城举办的 ACM MM 2017 大会上, CCF-CV 专委会委员、中国人民大学金琴副教授在 MSR Video to Language Challenge (MSR-VTT) 竞赛中卫冕冠军, 金琴团队是在去年举行的第一届 MSR-VTT 夺得冠军后, 今年再度取得了主观人工评测和客观性能评测的双冠军。团队发表的论文 "Knowing yourself:

improving video caption via in-depth recap" 也获得了挑战赛最佳论文奖。此外, 在大会第一天的 AVEC 2017 (Audio/Visual Emotion Challenge and Workshop) 中也夺得了 Affect Sub-Challenge 的冠军。

● 2017 年 10 月 26 日, CCF-CV 专委会委员、中科院自动化所副研究员黄永祜博士作为创始人兼首席执行官的银河水滴科技入选清科 2017 年度中国最具投资价值企业 50 强 (2017 Venture 50-新芽榜 50 强)。银河水滴科技是专注于机器视觉和人工智能的高科技创新型公司, 开发的基于步态的身份识别系统突破传统生物特征识别近距离、需配合的局限, 将成就身份识别领域的终极形态。

● 2017 年 10 月 27 日, CNCC 2017 大会举行颁奖, CCF-CV 专委会副主任、爱奇艺首席科学家王涛领导的爱奇艺技术团队研发的 "HCDN——大规模混合架构在线视频数据分发技术"、"视频情感识别及其大规模应用" 分别斩获本届大会 CCF 技术发明二等奖、CCF 科技进步优秀奖。

● 2017 年 10 月 28-29 日, 在 CCBR 2017 上, CCF-CV 专委会委员、中科院自动化所雷震副研究员指导的论文 Detecting Face with Densely Connected Face Proposal Network 获得最佳学生论文一等奖。CCF-CV 专委会委员、北京航空航天大学黄迪副教授指导的论文 2D-3D Heterogeneous Face Recognition Based on Deep Canonical Correlation Analysis 获得最佳学生论文二等奖。

● 2017 年 10 月 29 日, ICCV 2017 公布了 Joint COCO and Places Recognition Challenge 竞赛排名情况, CCF-CV 专委会委员、中科院自动化所刘静研究员带领的团队在 Places 场景解析挑战任务中, 击败谷歌、今日头条、亚马逊、牛津大学、清

华大学等国内外著名 AI 实验室以及知名科研院所, 获得最终冠军。

● 2017 年 10 月 29 日, CCF-CV 专委会委员、旷视科技首席科学家、研究院院长孙剑带领的旷视科技研究院团队获得了 2017 COCO & Places 图像识别大赛三项冠军 (COCO 物体检测和人体关键点, Places 物体分割), 成为了第一个在 COCO 竞赛中获得冠军的中国企业。2017 年 10 月 31 日, 旷视科技完成 4.6 亿美元 C 轮融资, 此融资金额刷新了人工智能领域融资记录。旷视科技成立于 2011 年, 主要从事计算机视觉相关技术研发与应用落地, 最早以 "Face++" 的名字为外界熟知。

● 2017 年 11 月 2 日获悉, CCF-CV 专委会秘书长、中科院自动化所王亮研究员在研究生培养工作中成绩卓越 (所培养的研究生黄岩博士荣获中国科学院院长特别奖), 获得了中国科学院优秀导师奖。

● 2017 年 11 月 2 日获悉, CCF-CV 专委会委员、中科院自动化所助理研究员黄岩博士荣获中国科学院院长特别奖, 中科院共计 40 位优秀青年研究人员获此殊荣。黄岩曾获得 CVPR-Workshop 最佳论文奖、ICPR 最佳学生论文奖以及 RACV 最佳墙报奖等学术奖项; 曾获得博士生国家奖学金、攀登一等奖学金、百度奖学金、北京市优秀毕业生等荣誉奖励。

● 2017 年 11 月 3 日获悉, 人力资源社会保障部公布了 2017 年国家百千万人才工程入选人员名单。共 411 人入选, 授予 "有突出贡献中青年专家" 荣誉称号, CCF-CV 专委会委员、中科院自动化所副所长刘成林研究员、CCF-CV 专委会常务委员、中科院计算所山世光研究员入选。

(责任编辑: 刘海波 余志文)

# 数据集：DUTS 数据集介绍

山东财经大学 蹇木伟 国防科技大学 刘丽

显著性检测起步阶段的传统算法大多数是利用启发式的先验知识(如背景先验、中心先验、对比度先验等),通过无监督的计算模型来进行的。这些传统方法在早期公布的数据集(如 MSRA1000)上仅采用颜色纹理等低级特征就已经取得了很好的效果。随着显著性检测的不断发展,待检测图像的场景变得越来越复杂,低级特征和无监督的方法遇到了瓶颈。

随着数据量的增多和硬件设备的进步,近几年基于深度神经网络(Deep Neural Networks)的方法在众多计算机视觉领域取得了巨大的突破,而现有的数据集也变得不再适合于训练复杂的深度神经网络模型。这些数据集本意是用来评估算法的优劣,但由于图像样本不充足,其上的评估也很难做到公正合理。同时面对各种各样的数据集,如何选择数据集作为训练深度神经网络的训练集也成为不容忽视的问题。

其次,不同的方法采用不同的数据集进行训练会导致在评估阶段的不一致性和不公平性。受上述问题的启发,因此创建 DUTS 数据集(<http://saliencydetection.net/duts/>)。该数据集提供了统一的训练集和测试集,旨在推动显著性检测的发展趋向于规范化,提高评估不同算法时的公平性。

## 1 DUTS 数据集

DUTS 数据集含有 10553 张训练图像以及 5019 张测试图像,这些图像均来自于 ImageNet 和 SUN 这两个主流数据集。DUTS 图像中既含有简单场景也包括复杂场景的图像,如图 1 所示。

### 1.1 训练集

DUTS 训练集图像来自于 ImageNet DET 数据集的训练集和验证集。首先对所有图像计算显著图,根据计算得到的显著图与真值目标框的重合度进行粗略的筛选,再结合人工筛选,删除其中包含过多目标(5 个及以上)以及前景过于简单(背景颜色单一,前景突出)的图像,最终保留 10553 张图像进行真值标注。



图 1 DUTS 中的图片

### 1.2 测试集

DUTS 测试集图像来自于 ImageNet DET 数据集的测试集以及 SUN 数据集。ImageNetDET 测试集图像的选取同上述方法。SUN 数据集由于没有真值目标框,仅采用人工的方法筛选,同时删除与 DUT-OMRON 重复的图片,最终共保留 5019 张图像。所有图像最终都归一化为  $400 \times X$  或  $X \times 400$  的尺寸( $X \leq 400$ )。

## 2 Ground-truth

DUTS 数据集目前仅含有像素级别的前景标记的真值,共有 50 名人员参与并进行了三轮真值的标注。

首先将训练集和测试集图像放到一起,打乱顺序并均匀的分给所有人员。所有人员采用统一的标注工具,对分到的图像进行第一轮标记,用矩形框标注出图像中的显著物体。接下来收集所有图像,再次打乱顺序并均分给所有人员,进行第二轮像素级的前景标记。标记者根据第一轮

标注出的目标框,把框内的物体标注成前景(白色),框外的物体标注成背景(黑色)。最后收集所有图片并再次打乱顺序均分给所有人进行第三轮的校验工作,检查前景标记的好坏。

## 3 评价

据我们所知,在显著性检测领域,DUTS 数据集是目前最大的,包含有明确的训练集、测试集的显著性检测数据集。为了今后研究的公平性,DUTS 的训练集将用来进行统一的深度神经网络的训练,同时该数据集的测试集和其他公开的数据集一起组成测试集进行公平的模型评估。图 2 对 DUTS 测试集进行评估,对 15 种最先进算法的 P-R 曲线进行了比较。

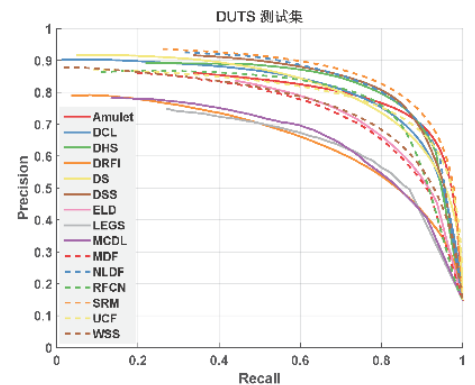


图 2 15 种方法的 P-R 曲线比较图

[1]J. Deng, W. Dong, etc. Imagenet: A large-scale hierarchical image dataset. CVPR, 2009

[2]J. Xiao, J. Hays, etc. SUN database: Large-scale scene recognition from abbey to zoo. CVPR, 2010

[3]L. Wang, H. Lu, etc. Learning to Detect Salient Objects with Image-level Supervision. In CVPR, 2017

(责任编辑:樊鑫)

# 开源代码：目标检测及其开源项目

大连理工大学 李豪杰 马新柱 樊鑫

## 1 目标检测介绍

目标检测 ( Object Detection ) 是计算机视觉中的主要研究分支之一。近年来，随着深度学习的快速发展，目标检测领域也取得了非常显著的成绩。

下面对该领域内一些主要的方法及其开源项目作简要介绍。

## 2 目标检测及其开源项目

### 2.1 RCNN 系列

随着深度学习在图片分类任务上的巨大成功，开始有一些学者尝试用深度学习来解决目标检测问题。Ross B. Girshick 首先提出了 RCNN 方法，成功的将目标检测转化为了分类问题，并应用于深度学习架构，取得了在当时极为突出的检测结果。并且在随后的研究中，又相继提出了 Fast RCNN 和 Faster RCNN，大幅度的提高了处理速度。

#### 2.1.1 RCNN

RCNN 将目标检测转化为了分类问题。它首先利用一些其他算法 ( 如 selective search ) 得到有可能是某个物体的区域 ( 称为 proposal )。然后将这些 proposals 送入网络，检测其类别，从而实现了物体检测的目的。

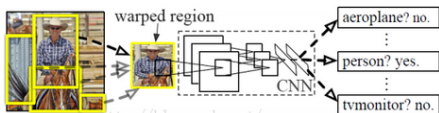


图 1 RCNN 架构示意图

- 论文链接：

<https://arxiv.org/abs/1311.2524>

- 开源项目链接：

<https://github.com/rbgirshick/rcnn>

- 开源项目简介：

项目使用的深度学习框架为 CAFFE ( Convolutional Architecture for Fast Feature Embedding )。项目为论文作者发布，完成了 RCNN 的主要功能，提供了 MATLAB 接口和预训练模型。目前已有其他框架 ( 如 Tensor Flow, PyTorch 等 ) 的改写版本。

#### 2.1.2 Fast RCNN

Fast RCNN 为 RCNN 的改进版本。因为 RCNN 中的 proposals 存在大量的重叠区域，这就导致了重复计算。Fast RCNN 针对这一点做出改进，使得计算量得到共享，从而加速了整个计算过程。

- 论文链接：

<https://arxiv.org/abs/1504.08083>

- 开源项目链接：

<https://github.com/rbgirshick/fast-rcnn>

- 开源项目简介：项目使用 CAFFE 框架编写，使用了 VGG16 网络架构。训练阶段比 RCNN 快 9 倍，测试阶段比 RCNN 快 200 倍，测试精度也比 RCNN 略有提高。项目提供了 Python 接口和作者的预训练模型。

#### 2.2 Faster RCNN

Faster RCNN 为 Fast RCNN 的改良版。在 Fast RCNN 仍然需要一些传统方法 ( 如 selective search ) 来提供 proposals，这一机制使得 proposals 的提取成了整个过程

的速度瓶颈。Faster RCNN 针对这一点做出改进：摒弃了传统方法，设计了 RPN ( region proposal networks )，利用分类部分的共享卷积权重预测 proposal，大大的提高了 proposals 的生成速度。

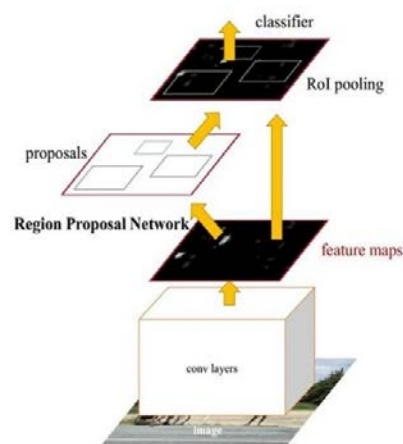


图 2 Faster RCNN 架构示意图

- 论文链接：

<https://arxiv.org/abs/1506.01497>

- 开源项目链接：

<https://github.com/ShaoqingRen/faster-rcnn>

- 开源项目简介：项目使用 CAFFE 架构编写。同时支持多种特征提取网络 ( ZF, VGG 等 )。速度方面有了明显提升，在 TITAN X 显卡上的测试结果为 59ms/151ms ( ZF/ VGG ) 一帧。在优化速度同时，还对内存占用做了优化。相对的，Fast RCNN 内存占用为 3G/8G。项目提供了 MATLAB 接口和 ZF/VGG 网络的预训练模型。

### 2.3 YOLO 系列

YOLO ( you only look once ) 是一种全新的目标检测架构，它将整张图片作为

输入,通过划分不同的子区域,然后给出每个区域的边框预测和概率,并依据概率大小对所有边框分配权重。最后设置阈值,输出得分(概率值)超过阈值的检测结果。

### 2.3.1 YOLO

YOLO 旨在提升目标检测的速度,设计了新的架构,采取分块的理念来取代 RCNN 中的 proposals。在速度方面有了重大突破,能够实时处理数据。

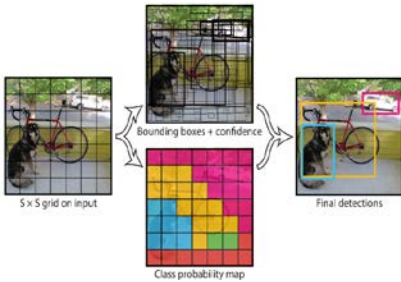


图 3 YOLO 检测过程示意图

- 论文链接：  
<https://arxiv.org/abs/1506.02640>
- 开源项目链接：  
<https://pjreddie.com/darknet/yolo/>
- 开源项目简介

项目团队为 yolo 编写了专用深度学习架构:darknet。该架构小巧灵活,没用多余模块,极度契合 yolo。项目提供了多种 yolo 架构变种 如 yolo yolo-small, yolo-tiny。其中最小的 yolo-tiny 在 TITAN X 显卡上能够达到 140 帧/秒。代码采用 C/CUDA 编写,使用 opencv 对图片进行输入/输出操作。提供了在 VOC 数据集上的预训练模型。

### 2.3.2 YOLO9000

YOLO9000 为 YOLO 的改进版。它融合了其他的先进思想,采取了多种优化策略,使得检测速度和精度都有了进一步的提高。同时提出了 Word Tree 的概念,能够实时检测 9000 类物体(一般方法为 20-80 类)。

- 论文链接：  
<https://arxiv.org/abs/1612.08242>
- 开源项目链接：  
<https://pjreddie.com/darknet/yolo/>
- 开源项目简介:项目采用 darknet 编写,给出了 yolov2(架构与 yolo9000 相同,训练方式不同,只支持相对较少类别的检测)与 yolo9000 的网络模型,以及其变种。项目提供了在 VOC 数据集上的预训练模型。

### 2.4 SSD

SSD (single shot multibox detector) 理念与 yolo 相似,同样不采用 proposals,通过子区域块来预测物体。而且 SSD 能够根据待检测目标尺度大小调整用于预测的特征图尺度,从而得到更精确的结果。

- 论文链接：  
<https://arxiv.org/abs/1512.02325>
- 开源项目链接：  
<https://github.com/weiliu89/caffe/tree/ssd>
- 开源项目简介:项目采用 CAFFE 架构编写,提供了 SSD300 和 SSD512 两个变种以及其对应的预训练模型。

以 SSD512 为例:其平均精度能够达到 79.8 (yolo 为 63.4),速度能够达到 19 帧/秒 (yolo 为 45 帧/秒)。项目提供了 python 接口。注:上述数据指的是初始版本的 yolo。

### 2.5 R-FCN

R-FCN (object detection via region-based fully convolutional networks) 使用全卷积网络改进 Faster RCNN。同时又进一步加大了能够共享的权重占整个网络中所占用的比例。使得精度和速度都有了提高。

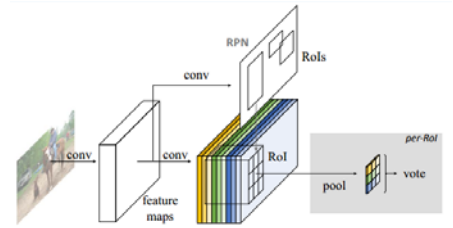
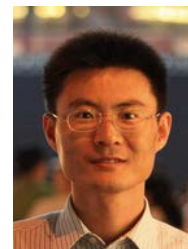


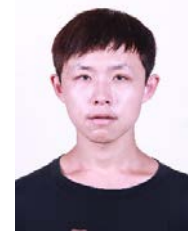
图 4 R-FCN 架构示意图

- 论文链接：  
<https://arxiv.org/abs/1605.06409>
- 开源项目链接：  
<https://github.com/msracver/Deformable-ConvNets>
- 开源项目简介:项目采用 MXnet 框架编写,支持多 GPU 训练。给出了基于多种不同数据集(voc,coco,cityscapes等)的预训练结果。项目提供了 python 接口。需要 opencv 作为输出输出支持。

### 作者简介:



李豪杰,大连理工大学国际信息与软件学院,教授,博士生导师,研究方向为多媒体信息检索,计算机视觉。。



马新柱,硕士在读,大连理工大学,研究方向为基于深度学习的目标检测与识别。



樊鑫,教授,博士生导师,主要研究方向为计算几何,图像处理与机器视觉。。

(责任编辑:蹇木伟)

# 招聘信息

## 一、腾讯 2018 博士生招募

### 1、岗位名称：机器学习方向

岗位部门：技术研究员

需求人数：不定

岗位要求：

(1) 2017、2018、2019 博士或博士后学位获得者；

(2) 计算机、应用数学、模式识别、人工智能、自动化控制、量子计算、神经科学等专业；

(3) 基础学科能力扎实、有丰富的项目经验，在 AI 领域顶级会议或期刊上发表过学术论文优先。

### 2、联系方式

请应聘者以 E-mail 形式发送简历，请在邮件主题上注明：姓名+应聘岗位名称。

简历投递截止时间：招满即止。

简历投递入口：

[Join.qq.com](http://Join.qq.com)

## 二、北京大学信息工程学院招聘信息

1、岗位名称：北京大学深圳研究生院信息工程学院招聘博士后

需求人数：2 人

岗位要求：

(1) 在相关领域有扎实的学术成果和研究经验；

(2) 数据挖掘与人工智能。基于大数据的深度学习。图像/视频处理、图像/视频分析研究

(3) 具有熟练采用英语写作及交流的能力；

(4) 基于大数据的深度学习；

### 2、待遇：

(1) 博士后在站期间待遇每年 30 万(含深圳市政府对博士后给予每年 12 万元的生活补助)；

(2) 实验室提供一流的办公条件，良好国际合作空间；

### 4、联系方式

请应聘者以 E-mail 形式发送简历，包括个人简历(含学历、工作经历、主要研究成果索引)；博士研究生毕业证书和博士学位证书复印件或应届博士毕业的证明材料；博士论文、文章代表作。

简历投递截止时间：招满即止。

联系人：邱老师

简历接收邮箱：

[gli@pkusz.edu.cn](mailto:gli@pkusz.edu.cn)

## 三、清华大学深圳研究生院智能计算实验室招聘信息

### 1、岗位名称：博士后

需求人数：1 人

岗位要求：

(1) 具有独立开展科研工作的能力，责任心强，有良好的团队合作精神，在研究方向相关内容上具有实际工作经验；

(2) 专业方向：数据挖掘及应用、人工智能、电子商务、视频处理、高性能计算等；

(3) 具有信息领域等相关专业的博士学位，具有扎实的理论基础和实践经验；具有良好的文献读写与交流能力；

### 2、待遇：

(1) 参照《清华大学深圳研究生院博士后管理办法》及深圳市相关条例(1.目前考

核合格者经深圳市人事局审核通过，除了正常工资外，还可享受深圳市政府每年 12 万元全额生活补贴 年收入 17 万左右；两年日常科研补助 1.5 万元；

(2) 出站后自主择业留深工作的，可申请 30 万元的留深科研补助；符合相关人才条件的可以申请认定为深圳市高层次专业领军人才或孔雀人才，并申请享受相关政策 and 待遇；；

(3) 除以上基本工资外，奖金等福利待遇。

### 3、联系方式

请应聘者以 E-mail 形式发送简历，包括年龄 35 周岁以下，近 2 年获得或即将获得计算机、电子、自动化等相关专业学科博士学位。本人简历(从大学至今，时间不间断，附近照)。充分反映本人学术及教学水平的有关材料。

简历投递截止时间：招满即止。

联系人：李老师

电话：0755-26036769

简历接收邮箱：

[li.xiu@sz.tsinghua.edu.cn](mailto:li.xiu@sz.tsinghua.edu.cn)

## 四、中国科学院深圳先进技术研究院劳特伯影像中心招聘信息

### 院劳特伯影像中心招聘信息

#### 1、岗位名称：博士后

需求人数：多名

岗位要求：

(1) 计算机专业，生物医学工程专业，电子信息工程专业等有图像或信号处理背景的博士毕业生；

(2) 博士毕业 3 年内，年龄不超过 35 周岁；

(3) 有发表过 1 篇以上 SCI 论文经验；

(4) 愿意从事医学信号处理，未来从事

神经成像、放射组学、快速成像、机器学习某一方面相关研究方向；

## 2、待遇：

(1) 深圳先进院在职博士后，即有技术职称，并且每年有资格参加员工晋升；深圳先进技术研究院基本薪资 + 深圳市政府对博士后的补贴 = 每年大于 22 万；

(2) 进站报到后可落户深圳，其配偶及其未成年的子女可办理户口随迁；还有国

家级，省级等各级补助。博士后出站待遇；

(3) 能够提前晋升副研；留深工作三年后可成功申请深圳市级领军人才，享受约 160 万个人购房补贴；

(4) 出站留深博士后可有 30 以上万元的科研资助；

(5) 协助申请各类人才项目；

(6) 五险一金：全额缴纳五险一金；我单位有独立的档案保管权和人事代理权，可无条件接受户口、档案；附加其他 3H 等优厚条件；

## 3、联系方式

请应聘者以 E-mail 形式发送简历，简历投递截至时间：招满即止。

联系人：李老師

电话：0755-26036769

简历接收邮箱：

[ss.wang@siat.ac.cn](mailto:ss.wang@siat.ac.cn)

(责任编辑：李策)

# 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，可继续关注每个国际会议举办的 workshop。

## 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 Neurocomputing 和 Advances in Multimedia。

## 3 会议简介

**计算机视觉和模式识别大会 CVPR** 由 IEEE 计算机协会主办，CVPR 是年度主要的计算机视觉活动，会议的主要内容是计算机视觉与模式识别技术，会议组成包括重要会议(论文)、研讨会和讲习班。CVPR 为学生、学术界和工业界的研究者提供了一个高质量、低成本来展示研究成果，是一个特别有价值的交流机会。

IEEE 计算机视觉和模式识别大会 CVPR 2018 将于 2018 年 6 月 18-23 日

在美国犹他州的盐湖城 (Salt Lake City, Utah, USA) 举行，会议致力于计算机视觉、模式识别、机器学习和人工智能，包括颜色检测、跟踪、运动、物体识别、立体声和物体检测。

在近几年的 CVPR 大会中，深度学习越来越成为计算机视觉领域的主流，每年收录的论文数量一般 300 篇左右。

(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
CVPR 2018	2018.06.18-23	Salt Lake City, Utah, USA	2017.11.15	<a href="http://cvpr2018.thecvf.com">http://cvpr2018.thecvf.com</a>
ICCGV 2018	2018.02.25-27	日本 东京	2017.11.15	<a href="http://www.iccgv.org">http://www.iccgv.org</a>
ICACI 2018	2018.03.29-31	中国 厦门	2017.11.15	<a href="http://www.icaci2018.org">http://www.icaci2018.org</a>
CGIIT 2018	2018.02.23-25	中国 香港	2017.11.25	<a href="http://www.cgiit2018.org">http://www.cgiit2018.org</a>
CSIG 2018	2018.05.11-13	中国 扬州	2018.01.15	<a href="http://www.csig.org.cn">http://www.csig.org.cn</a>
ICDIP 2018	2018.05.11-14	中国 上海	2018.03.10	<a href="http://www.icdip.org">http://www.icdip.org</a>

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
Neurocomputing	Robustness, Security and Regulation Aspects in Current Biometric Systems	2017-11-15
Advances in Multimedia	Learning-Based Multimedia Analyses and Application	2018-03-16

# COMPUTER VISION NEWSLETTER

---

## 计算机视觉

---

专委简报

---

### 诚征封底广告

---

**对象**：计算机视觉和人工智能领域高科技企业、学术团体或个人等

**内容**：产品技术、服务理念、科研成果、学术贡献及其他推介

**费用**：双方协商（联系 Email：[ccfcvn@gmail.com](mailto:ccfcvn@gmail.com)）

**【简介】**专委简报是中国计算机学会计算机视觉专委会（CCF-CV）与委员之间的双向信息通道，每年6期，单月出刊。目前已出四期，每期阅读量几千人次。读者包括活跃于高校、科研院所、高科技企业的计算机视觉和人工智能相关领域的教师、学生和研发人员。CCF-CV于2013年10月成立，在CCF专委评估中连续三年分别获得2014年度“特色活动奖”、2015年度“综合进步奖”、2016年度“优秀专委奖”。

---

---