

CCF

计算机视觉 专委会简报

COMPUTER VISION NEWSLETTER

2018/05 期
总第 15 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	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	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	沈沛意	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	山东财经大学
	金 鑫	北京电子科技学院
	李 策	兰州理工大学
	刘 丽	国防科学技术大学

COMPUTER VISION NEWSLETTER

CONTENTS

目录

专委动态

走进高校	走进高校系列活动	04
走进企业	走进企业系列交流会	05
专委通知	CCF-CV 2018 年度学术新锐奖报名	06
	CCF-CV 2018 年委员增选申请	07
	CCF-CV 2019 年走进高校系列报告会活动申请	08
专委大会	PRCV 2018 注册通知	09
	PRCV 2020 承办单位征集	10

科技前沿

专题综述	图匹配年度进展概述	12
热点追击	视觉盲计算：基于密码学的云环境人脸隐秘识别	18
	显著性引导的视频目标分割	19

委员风采

委员访谈	清华大学冯建江副教授访谈	20
委员好消息		22

资源平台

开源代码	基于深度学习的视觉问答技术	24
数据集	异常行为检测	28
招聘信息		30
征文通知		35

CCF-CV 走进高校系列报告会

第 59 期 燕山大学

时间：2018 年 7 月 20 日

第五十九期 CCF-CV 走进高校系列报告会“计算机视觉前沿技术及应用”在燕山大学世纪楼学术报告厅成功举行。报告会邀请了中国科学院自动化所王亮研究员、西安电子科技大学苗启广教授、中国科学院沈阳自动化所丛杨研究员、南开大学杨巨峰副教授和北京邮电大学马占宇副教授出席并作精彩报告。报告会由燕山大学信息科学与工程学院计算机科学与工程系主任张世辉教授、副主任吴培良副教授担任执行主席。报告会最后，燕山大学信息科学与工程学院副院长练秋生教授作总结发言。



五位专家的报告异彩纷呈，在座师生纷纷向五位专家提问，并就其中的算法设计思想、计算机视觉发展趋势等问题进行深入探讨。此次 CCF-CV 走进高校系列报告会持续了四个小时，报告内容丰富多样，报告会互动热烈，为现场听众奉献了一场计算机视觉领域的学术盛宴。

第 60 期 新疆大学

时间：2018 年 9 月 8 日

第六十期 CCF-CV 走进高校系列报告会在新疆大学圆满结束。本次报告会由新疆大学信息科学与工程学院承办，执行主席为库尔班·吾布力教授，其团队所在实验室是新疆多语种信息技术重点实验室，该实验室依托于新疆大学，其前身是新疆大学多语种信息技术重点实验室。本次报告会的主题为“计算机视觉学科专业领域的学术与技术交流”，报告会由库尔班·吾布力教授进行主持，分别由华南理工大学金连文教授、北京大学施柏鑫研究员、杭州电子科技大学俞俊教授和四川大学赵启军副教授四位学者进行主讲。



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

（责任编辑：马占宇）

CCF-CV 走进企业系列交流会

第 15 期 图森未来

时间：2018 年 8 月 10 日

第十五期 CCF-CV 走进企业系列交流会活动“走进图森未来”，在北京市朝阳区世贸商务楼五层图森未来会议室成功举办。本期活动由 CCF-CV 秘书处召集，图森未来校园关系负责人王金波负责现场组织协调。图森未来对此次交流活动给予了高度重视和大力支持。图森未来联合创始人 COO 郝佳男博士、合伙人首席科学家王乃岩博士联袂参加了本次活动。

首先，公司联合创始人 COO 郝佳男博士分享了公司的成长历程和发展航路，全面介绍了图森未来围绕无人驾驶卡车的技术、平台、应用场景所进行的方向布局和技术研发。报告介绍了成立近 3 年来所取得的突破。报告最后分享了公司的发展愿景——“致力打造 L4 级别无人驾驶重卡解决方案，构建全球最大的公路货运运力平台”，并介绍了多项已成功商用的全流程解决方案。参会人员纷纷对图森未来的蓬勃发展由衷点赞！

接下来，公司合伙人首席科学家王乃岩博士带来“图森未来在 L4 自动驾驶卡车中的最佳实践”的主题报告。王博士的报告充满了技术干货，系统全面地综述了无人车系统的各个模块，包括前端的多种传感器与后台的多种核心算法。报告以大量运行实例生动地展示了计算机视觉前沿技术特别是物体检测与分割、环境感知与时序理解、多传感器融合等在无人驾驶研发中的实际应用，分析了实际场景中的挑战与应对方案，引导

听众真切感受到了图森未来在自动驾驶卡车应用场景中卓越的研发实力和实践优势！



活动间歇，参会师生分别介绍了自己的研究领域，与主办方展开互动。图森未来为参会人员精心准备了茶饮，大家边讨论交流边观摩公司技术演示，充分体会无人驾驶货运领域的快速发展与技术魅力！活动结束后，图森未来为每位参会人员准备了精美的纪念品，大家一同合影留念。

（责任编辑：王瑞平）

CCF-CV 2018 年度学术新锐奖报名



为推动中国计算机视觉领域的科技进步，鼓励创新性研究，促进青年人才成长，由 CCF-CV 青年工作组提议，特设立中国计算机学会计算机视觉专委会 (CCF-CV) 学术新锐奖 (以下简称“学术新锐奖”)。本奖项由 CCF-CV 设立，每年评选一次，由 CCF-CV 评选和颁发，每年最终获奖者不超过 3 人。欢迎计算视觉领域类的优秀学生积极报名，CCF-CV “学术新锐奖”将成为您专业领域启航的强大助推力量之源。

一、如何成为正式候选人

参加学术新锐奖评选应具备如下条件：

1. 候选人在申请时系低年级博士生 (普博生入学后的前两年；或直博生入学后的前三年；或硕博连读自硕士入学开始起前三年。以申请当年 8 月 31 日之前状态为准)

2. 科研工作取得了突出的成绩。如：成果发表在计算机视觉领域顶级期刊或学术会议上，在权威学术竞赛中取得优异成绩，或为重大科研项目的完成与成果转化做出了重要贡献。

3. 候选人需获得导师同意，且由两名 CCF-CV 专委会委员 (不含导师) 推荐，方可成为正式候选人。

二、推荐方法

1. 每位候选人需提交申请书 1 份，以及由两名 CCF-CV 专委会委员填写的《CCF-CV 学术新锐推荐表》及其电子版，提交发送至电子邮箱 ccfcvaward@sina.com

2. 每位 CCF-CV 专委会委员同一年最多推荐两位候选人。

三、评奖办法

在 CCF-CV 常委会领导下，由 CCF-CV 提名与奖励工作组负责评奖工作。

评奖过程分初审和复审两个阶段：

1. 初审阶段由 CCF-CV 提名与奖励工作组负责组织，进行资质审查，并推荐不超过 10 名正式候选人；根据正式候选人情况，由 CCF-CV 提名与奖励工作组提名评审委员会名单 (不超过 7 名成员)；评审委员会名单与正式候选人名单提交 CCF-CV 常委会审议通过，通过后的评审委员会名单与正式候选人名单同时公示，公示期为 7 天。

2. 复审阶段由评审委员会负责组织，根据正式候选人情况，最终推荐不超过 3 名 CCF-CV 学术新锐获奖者，报 CCF-CV 常委会审议通过。

3. CCF-CV 学术新锐终评委员会将邀请当选者在当年 PRCV 或年会上以公开学术报告形式介绍其工作。

4. 在评议过程中, 若发现被评论文有作者抄袭、剽窃等学术道德问题时, CCF-CV 提名与奖励工作组有权取消该申请人的参评资格; 若在颁奖之后发现上述问题 CCF-CV 将取消其获奖资格, 收回证书, 并予公布。

5. CCF-CV 在专委会年度学术活动 PRCV 大会上, 对获奖者及其导师颁发获奖证书, 获奖证书由 CCF-CV 主任签署, 获奖者免 PRCV 会议注册费。

四、评奖时间

1. 报名时间: 2018 年 8 月 15 日-2018 年 9 月 15 日;

2. 初审阶段: 2018 年 9 月 16 日-9 月 30 日;

3. 复审阶段: 2018 年 10 月 31 日前。

备注: 其他未尽事宜, 请关注微信公众号 CCF 计算机视觉专委会了解详情, 可于公众号原文下载 CCF-CV 学术新锐奖评选条例、CCF-CV 学术新锐推荐表。

(责任编辑: 马占宇)

CCF-CV 2018 年委员增选申请

自 2013 年 10 月成立以来, 中国计算机学会计算机视觉专业委员会发展迅速, 举办了很多有影响力的活动, 搭建了全方位、高水平、大规模的计算机视觉领域交流平台。专委会成立四年以来, 已经发展委员 338 人, 连续四年在 CCF 专委评估中分别获得“特色活动奖”、“综合进步奖”、“优秀专委奖”、“年度特别奖”。为了保持专委会的活力、促进国内外视觉领域人员的交流和合作, 专委会现开放 2018 年计算机视觉专委会的委员增选工作。

申请时间: 2018 年 5 月 16 日—2018 年 10 月 31 日

申请流程: 填写申请表(请从专委会网站上下载 ccfcv.ccf.org.cn), 发送给秘书处(ccfcv@139.com), 主题“2018 新委员申请-姓名-单位”。(注: 推荐人必须是现任专委委员, 名单

可以从专委网站查询。电子版申请表中需填写推荐人姓名和意见, 委员增选成功后可以补签签名)

申请资格: 任职国内外学术界或企业界副教授或等同级别以上的人员, 拥有计算机视觉相关领域的高水平研究成果, 是 CCF 会员, 且积极参加计算机学会计算机视觉专委会的各项活动。特别优秀的讲师、企业人士亦可考虑。

申请需知: 每个现任专委委员可以推荐最多 3 名候选人, 主任委员、副主任委员、秘书长、常务委员可以推荐最多 5 名候选人。本次申请结果将在“2018 年中国模式识别与计算机视觉学术会议(<https://prcv-conf.org/2018>)”期间举行的专委工作年会上投票确定(申请者届时必须“注册参会”)

(责任编辑: 毋立芳)

CCF-CV 2019 年走进高校 系列报告会活动申请通知

借 AI 东风，布 CV 之道：为了更好地推动计算机视觉学科专业领域的学术与技术交流，促进国内外学者间的了解与合作，全面推动国内计算机视觉的学科发展，提升我国计算机视觉研究在国际领域的影响力，中国计算机学会计算机视觉专委会在全国范围的高校和科研院所等开展了 CCF-CV 走进高校系列报告会活动。

自 2015 年 11 月开始以来，本活动得到了高校、讲者、听众的大力支持。截至 2018 年 9 月底，CF-CV 走进高校系列报告会活动已成功举办了 60 届，遍及祖国大江南北 30 余座城市，已邀请讲者做分享专题报告 200 余场，活动现场人均听众 200 人次，微信公众号平均阅读 1000 余次，在征得讲者同意的前提下，于爱奇艺视频网站上上传共享报告会的现场盛况视频，并于专委会网站分享讲者的报告 PPT，在国内外计算机视觉领域引起了强烈反响。在这里，向对活动顺利开展提供支持和帮助的承办方以及分享精彩报告的讲者们表示由衷的感谢！

值此金秋时节，我们面向全国的高等院校、科研院所等机构征集 2018 年度活动申请，希望能够进一步贯彻落实《新一代人工智能发展规划》纲要中有关计算机视觉领域发展的精神，继续推动计算机视觉领域的学术交流、思想碰撞，为广大师生提供更多近距离接触行业领域专家、互动沟通的机会。

活动细则、申请表等材料请点击如下链接查询。请感兴趣的申请人下载活动申请表，填写相关信息后发给专委会秘书处。

联系方式：

马占宇：mazhanyu@bupt.edu.cn；

毋立芳：lfwu@bjut.edu.cn



特别提示：

- 1、原则上同一城市举办两次活动的间隔须在三个月以上，同一城市/学校在满足此条件下，可以举办多次活动；
- 2、除去假期等因素影响，每年实际安排活动在 20 期左右，申请者请从速。

附：相关链接

CCF-CV 官方网站

<http://ccfcv.ccf.org.cn/>

历次活动精彩分享，请微信关注 CCF 计算机视觉专委会公众号。

（责任编辑：马占宇 毋立芳）

PRCV 2018 注册通知



中国模式识别与计算机视觉学术会议 (PRCV) 是由中国模式识别学术会议 (CCPR) 和中国计算机视觉大会 (CCCV) 合并而来, 由中国人工智能学会 (CAAI)、中国计算机学会 (CCF)、中国自动化学会 (CAA) 和中国图象图形学学会 (CSIG) 联合主办, 定位为模式识别与计算机视觉领域国内顶级学术盛会。

首届中国模式识别与计算机视觉学术会议 (PRCV2018) 将于 2018 年 11 月 23-26 日在广州举行, 由中山大学承办。本届会议将主要汇聚国内从事模式识别和计算机视觉理论与应用研究的广大科研工作者及工业界同僚, 共同分享我国模式识别与计算机视觉领域的最新理论和技术成果, 为大家提供精彩的学术盛宴。现向广大科技工作者公开征集高质量、原创性的优秀学术论文 (英文)。大会录用的稿件将在会上展示, 会议论文集将由 Springer 出版社以 LNCS 系列出版, 并被 EI 和 ISTP 检索。Springer 将为最佳论文提供 1000 欧元的奖励。优秀的论文经扩展后将推荐到国内外高质量 SCI 杂志发表。

会议邀请到 David Forsyth, Tamara Berg, Zhengyou Zhang 和 Michael S. Brown 四位专家作大会报告。同期将举办顶会与顶刊交流会, 行业专业应用竞赛, 模式识别、计算机视觉与机器学习领域前沿理论与方法讲习班。

注册时间及会议费标准

全额注册:

适用对象: 适用于需注册论文者及教师、研究员、博士后等企事业单位工作人员。

注册时间及收费标准: 2018 年 10 月 1 日及以前: 2800 元; 2018 年 11 月 22 日及以前: 3000 元; 会议现场: 3200 元。(以汇款时间为准)

会员注册:

适用对象: 具有以下会员身份之一: 中国人工智能学会 (CAAI)、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)。

注册时间及收费标准: 2018 年 10 月 1 日及以前: 2200 元; 2018 年 11 月 22 日及以前: 2400 元; 会议现场: 2800 元。(以汇款时间为准)

半注册:

适用对象: 全日制在校学生。

注册时间及收费标准: 2018 年 10 月 1 日及以前: 1300 元; 2018 年 11 月 22 日及以前: 1500 元; 会议现场: 1800 元。(以汇款时间为准)

备注:

- 每篇录取的论文必须有至少一位论文作者在 2018 年 9 月 20 日前向会议全额注册 (非会员注册) 并缴纳会议费。每份全额注册费只能用于两篇论文。未全额注册的论文将不能被发表。

- 全日制在校学生不含在职研究生和博士后; 学生注册时须提供学生证的扫描件, 否则需补齐全额注册费。

- 更多注册及缴费详情请登录 PRCV 官方网站 (<https://prcv-conf.org/2018>) 进行查询。

(责任编辑: 毋立芳)

PRCV2020 征集承办单位



中国模式识别和计算机视觉大会 (PRCV, Chinese Conference on Pattern Recognition and Computer Vision) 是全国模式识别学术会议 (CCPR)、中国计算机视觉大会 (CCCV) 等相关会议合并后联合举办的全国性学术会议。

全国模式识别学术会议 (CCPR) 由中国自动化学会、人工智能学会和模式识别国家重点实验室联合主办, 2007 年、2008 年在北京举办 (中科院自动化所承办), 2009 年在南京举办 (南京理工大学承办), 2010 年在重庆举办 (重庆大学承办), 2012 年在北京举办 (中科院自动化所承办), 2014 年在长沙举办 (湖南大学承办), 2016 年在成都举办 (电子科技大学承办)。中国计算机视觉大会 (CCCV) 由中国计算机学会主办、计算机视觉专委会承办。首届 CCCV 于 2015 年在西安举办 (西安电子科技大学承办); 第二届 CCCV 于 2017 年 10 月 12-14 日在天津举办 (中国民航大学承办, 天津大学和南开大学协办)。

中国人工智能学会模式识别专委会、中国计算机学会计算机视觉专委会、中国自动化学会模式识别与机器智能专委会、中国图象图形学学会视觉大数据专委会四个专委会的委员名单重叠度高, 每年各专委会组织的相关学术年会或研讨会 (如 CCPR、CCCV 等) 内容相近。因此, 为了推进本领域的学科发展, 打造一个高质量的学术交流与合作平台, 经讨论一致同意合并过去各个专委会的类似会议, 联合举办中国模式识别与计算机视觉大会。

2018 年起, 中国模式识别和计算机视觉大会 (PRCV) 由中国人工智能学会、中国计算机学会、中国自动化学会和中国图象图形学学会联合主办, 每年举办一次。2018 年由中山大学 (广州) 承办, 2019 年由西北工业大学 (西安) 承办, 现征集 2020 年承办单位。

一、会议地点

中国模式识别和计算机视觉大会 (PRCV) 将在全国各地城市轮流举办。2020 年的举办地应该是在广州、西安以外的地区。

二、会议时间

会议召开时间宜在气候比较适宜的季节, 一般在秋季, 以 9-11 月为宜。2020 年 PRCV 应该避开领域内主要的相关国际学术会议的时间。

三、关于组委会

承办单位应该有一个实验室或部门在国内模式识别和计算机视觉界有一定学术基础和组织能力。大会主席和程序委员会主席应由本领域的知名学者担任 (根据 PRCV 指导委员会建议, 程序委员会主席每个专委会各派一人担任)。请不要把非本学术领域的领导列为大会主席或程序

委员会主席，担任实际组织工作的组织委员会主席除外。

承办单位在申请时不能确定大会主席和程序委员会主席的，可先空缺，在确定承办会议后，由协调指导委员会推荐指定大会主席和程序委员会主席。

四、会议程序

PRCV 论文集将由 Springer 正式出版。

会议将公开征集论文，录用论文分口头报告 (Oral) 和张贴报告 (Poster)。要求每篇录用论文至少有一名作者注册参加会议并报告论文，否则该论文将从论文集删除。

主会会议日程一般为 3 天。会议上除了特邀报告和论文报告，还可以组织专题论坛、讲习班等形式的交流和讨论。

五、财务预算

会议积极争取企业赞助，同时适当收取注册费，用于支付论文集出版和会议举办过程中的实际费用。对自动化学会、计算机学会、人工智能学会、图象图形学学会会员，以及学生，注册费应至少优惠 20%。

会议结束后，组委会应向第一主办学会和协调指导委员会提交详细的财务收支报告，会议结余资金应向联合主办学会上交。

六、承办申请书

申请书应包含以下内容：

1. 会议时间、地点（高校、酒店或会议中心）
2. 承办单位研究基础、会议设施、接待设施介绍
3. 组委会主要人员及其研究经历、组织经验
4. 程序安排基本想法介绍
5. 财务预算，包括注册费标准、预计收支情况等

请在 2018 年 10 月 15 日前将申请书发给中

国模式识别和计算机视觉大会协调指导委员会联系人王亮。申请人将被邀请在 2018 年 PRCV 上介绍会议申请报告。如有超过一个申请单位，将在会上由参会代表投票决定承办单位。



附：中国模式识别和计算机视觉大会 (PRCV) 协调指导委员会名单：

主席：谭铁牛（中科院自动化所）

委员（按姓氏拼音排序）：

陈熙霖（中科院计算所）

刘成林（中科院自动化所）

权 龙（香港科技大学）

芮 勇（联想集团）

查红彬（北京大学）

郑南宁（西安交通大学）

周 杰（清华大学）

秘书长：

王 亮（中科院自动化所）

联系人：

北京中科院自动化所模式识别国家重点实验室，王亮

邮编：100190

电话：010-82544643

邮箱：prcv@nlpr.ia.ac.cn

（责任编辑：毋立芳）

图匹配年度进展概述

上海交通大学 严骏驰 汪润中 杨小康

一、引言

图匹配 (graph matching) 是计算机视觉、计算机图形学、模式识别等学科的基本问题之一。相对于向量化的表示形式, 图结构包含了节点之间的相互关系, 在刻画结构化的信息时具有显著优势。图匹配是一种常用的图数据分析技术, 它试图求解两个或更多的图结构之间节点对应关系。在计算机视觉领域, 图匹配方法从局部特征出发, 建立了图像特征点之间的一一对应关系, 进而为相关度量模型的构建提供可能。对于图匹配问题, 如何获得图像特征、计算相似度, 以及在求解时权衡效率和精度, 进而实现精确、高效的匹配, 是当前计算机视觉领域一项重要的研究课题。

在待匹配的图之间, 如果只考虑节点到节点的相似度 (一阶相似度), 匹配问题就变成了一个线性指派问题 (LAP, Linear Assignment Problem), 可以使用匈牙利算法 [1] 在多项式时间内求解。为了加入结构性的约束, 图匹配方法通常会同时考虑边对边的相似度 (二阶相似度)。此时, 图匹配问题可以被一般化成 NP-难的二次指派问题 (QAP, Quadratic Assignment Problem), 无法在线性时间内求解。因此, 研究者们提出了许多松弛算法, 试图在求解速度和匹配精度之间取得平衡。

在计算机视觉领域, 典型的图匹配算法一般包含了提取特征、计算相似度、求解 QAP 三个步骤。在提取特征时, 常使用 SIFT [2]、DAISY [3] 等通用的人工描述子。近年来, 也有工作 [4-7] 通过深度学习模型获取用于匹配的图像特征, 取得了令人鼓舞的效果。在计算相似度时, 点集、边集的相似度用相似度矩阵 \mathbf{K} 表示。以二图匹配为例, 假设两图分别有 n_1, n_2 个节点, 则 $\mathbf{K} \in \mathbb{R}^{n_1 n_2 \times n_1 n_2}$ 。其中 \mathbf{K} 的对角线元素代表点对点的一阶相似度, 其余非对角元素代表边对边的二阶相似度。 \mathbf{K} 中

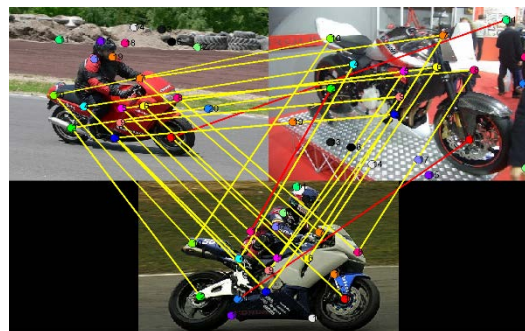


图 1 图匹配示意图 [18]

的元素由前一步提取的图像特征计算得到。点对点的一阶相似度, 往往由图像特征点在特征空间的欧氏距离得到; 而二阶相似度刻画图像中的物体结构, 根据定义不同, 可能与一阶相似度或特征点对的空间位置等信息有关。在求解 QAP 时, 定义排列矩阵 $\mathbf{X} \in \{0,1\}^{n_1 \times n_2}$, 以及 $\mathbf{x} = \text{vec}(\mathbf{X})$ 为矩阵 \mathbf{X} 的向量化的版本, 则图匹配问题的目标函数可以写为

$$\max_{\mathbf{X}} J(\mathbf{X}) = \mathbf{x}^T \mathbf{K} \mathbf{x} \quad (1)$$

$$\text{s. t. } \mathbf{X} \mathbf{1}_{n_1} = \mathbf{1}_{n_1}, \mathbf{X}^T \mathbf{1}_{n_2} \leq \mathbf{1}_{n_2}, \mathbf{X} \in \{0,1\}^{n_1 \times n_2}$$

如前所述, 上式在一般情形下为 NP-难问题, 不存在多项式时间内的全局最优解。因此, 多种近似方法 [8-10, 12, 14] 被提出, 以合理的复杂度求解上述问题, 得到尽量精确的结果。

在近几年的图匹配研究中, 多图匹配是一个主要的研究方向 [5, 7, 9-10, 13, 18-19]。由于特征提取算法不可避免的误差缺陷以及存在随机噪声, 二图匹配算法找到的全局最小值并不一定对应于语义上的正确匹配。相比于二图匹配, 多图匹配融合了多个图像信息, 可以有效地抑制二图匹配中的随机噪声, 获得更准确的匹配结果; 此外, 视频处理、三维重建等场景都需要输入多个图像, 多图匹配算法具有非常广阔的实际应用前景。在二图匹配的基础上, 多图匹配加入了一个重要约

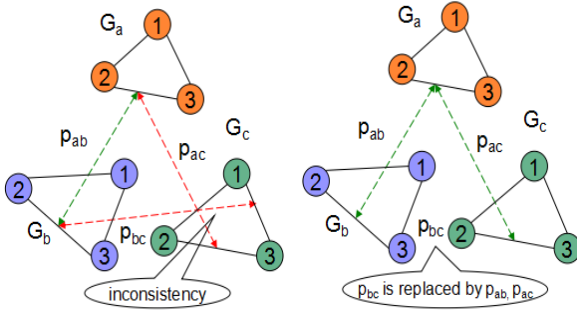


图 2 多图匹配的循环一致性示意图[19]

束：循环一致性 (cycle consistency)，即两图的匹配结果和通过其他任意路径传递后的匹配结果保持一致，如图 2 所示。对于三张图的情况，令 $\mathbf{X}_{12}, \mathbf{X}_{23}, \mathbf{X}_{13}$ 分别为保存它们之间两两匹配结果的排列矩阵，则循环一致性要求

$$\mathbf{X}_{13} = \mathbf{X}_{12}\mathbf{X}_{23}$$

多图匹配算法可分为两类，一类方法在所有图的特征空间直接进行优化；另一类需要通过二图匹配算法，预先生成两两匹配的结果，在循环一致的约束下寻找与两两匹配最相近的多图匹配结果。最新的研究工作基于多图匹配的场景，研究了分布式图匹配[5]，增量图匹配[13]等新问题，以解决已有的多图匹配算法无法处理大规模多图匹配的缺陷。

图匹配求解算法和图上特征的提取，是图匹配研究工作的两大重点和难点。图匹配求解作为一个 NP-难问题，需要在复杂度和精度之间权衡，寻找更接近全局最优的局部最优解。此外，如何在求解过程中抑制噪声、寻找其他形式化图匹配问题的方式，也是图匹配求解的重要研究方向。在当前阶段，图匹配求解算法的效率和精度是制约图匹配问题向大规模、高精度方向发展的瓶颈。另一方面，在计算机视觉领域，当前图匹配算法使用的图像特征基本都是通用特征，同时也有使用在其他问题（例如分类问题[7]）上预训练的卷积神经网络进行特征提取的工作，而图匹配问题与其他问题对特征的抽取，往往存在一定的差异，因此选择通用的人工描述子或者在其他任务上预训练的模型会为图匹配带来偏差。采用当前深度学习领域端到端学习的思路，就能够改善图匹配

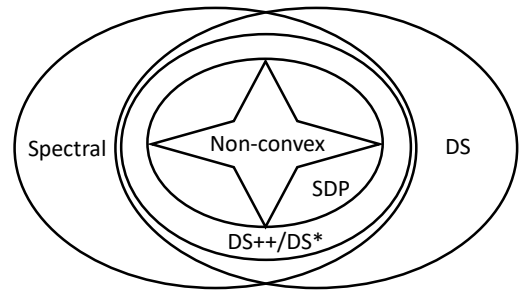


图 3 不同松弛条件的取值集合

特征提取步骤的效果，自动学习最适合图匹配任务的相关特征。

二、图匹配研究现状

1. 图匹配求解算法

图匹配作为一个 NP-难的 QAP 问题，其求解算法需要平衡求解的效率和精度。图匹配求解算法作为图匹配问题的根基，一直是研究的重点。一类最常见的求解方法为连续域优化，即将原目标函数(1)中 \mathbf{X} 的离散约束松弛为连续的。例如，一种常见的双随机 (DS, Double Stochastic) 松弛的目标函数为

$$\max J(\mathbf{X}) = \mathbf{x}^T \mathbf{K} \mathbf{x}$$

$$\text{s. t. } \mathbf{X} \mathbf{1}_n = \mathbf{1}_n, \mathbf{X}^T \mathbf{1}_n = \mathbf{1}_n, \mathbf{X} \in [0,1]^{n \times n} \quad (2)$$

\mathbf{X} 中的每个元素在连续域 $[0,1]$ 上取值。此处，认为两图中的特征点一一对应，且两图中特征点的数目相同 $n_1 = n_2 = n$ 。双随机矩阵定义为每行、每列求和均为 1 的非负矩阵，由公式 (2) 中的前两项约束体现。双随机松弛将图匹配问题的目标函数松弛为连续域的凸函数，可以使用凸优化方法进行求解。除了双随机松弛，常见的松弛技术还包括谱松弛 (Spectral Relaxation)、半正定规划 (Semi-Definitive Programming) 等连续域方法，能够较为便捷地求解图匹配问题。

Dym 等[8]系统地分析了图匹配问题中谱松弛、双随机松弛和半正定规划的性能优劣，在双随机松弛方法的基础上提出了 DS+ 和 DS++ 算法。如图 3 所示，半正定规划是所有现存方法中最接近原始非凸问题的松弛条件，对应着最接近全局最优解的结果；但半正定规划的最大缺点是具有

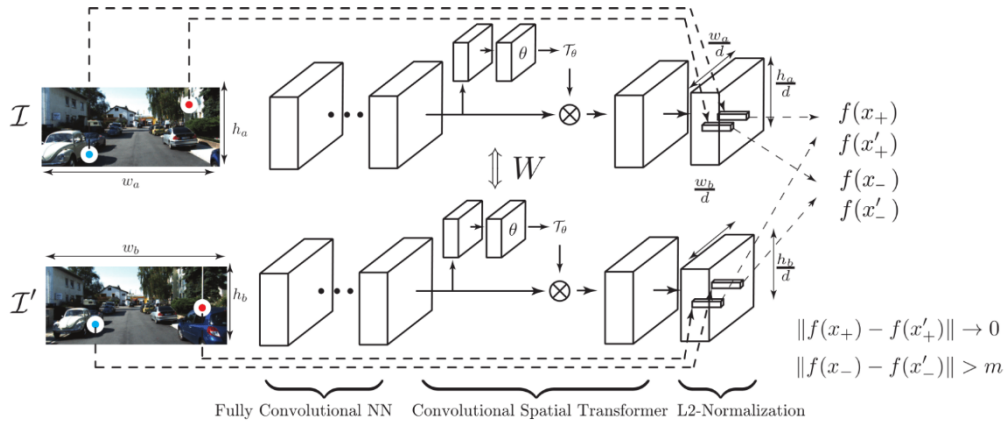


图 4 全局一致性网络 (UCN) 结构图[4]

$O(n^4)$ 规模的参数，这会极大地降低图匹配求解的速度。相比之下，谱松弛和双随机松弛仅有 $O(n^2)$ 规模的参数，求解速度更快但精度更低。为了更好地平衡求解精度和速度，作者提出了具有 $O(n^2)$ 规模的参数，且约束条件更加接近半正定规划的 DS++ 算法。在公式 (2) 的基础上，算法引入了非凸的惩罚项 $-a\|\mathbf{X}\|_F^2$ ，其中 $\|\mathbf{X}\|_F$ 是矩阵 \mathbf{X} 的 Frobenius 范数

$$\min_{\mathbf{X}} E(\mathbf{X}) - a\|\mathbf{X}\|_F^2 + a \cdot n \quad (3)$$

$$\text{s.t. } \mathbf{X}\mathbf{1}_n = \mathbf{1}_n, \mathbf{X}^T\mathbf{1}_n = \mathbf{1}_n, \mathbf{X} \in [0,1]^{n \times n}$$

a 越大，目标函数就越接近离散约束下的目标函数；当 a 足够大时，即可得到离散取值的排列矩阵 \mathbf{X} ；选取合适的 a ，即可获得与原条件最接近的凸约束。需要注意，论文中定义的 $E(\mathbf{X}) = \mathbf{x}^T \mathbf{W} \mathbf{x}$ 衡量的是匹配损失，与公式 (1) 中 $J(\mathbf{X})$ 定义的相似度正好相反，在问题公式化时，需要最小化公式 (3)。另一项工作 [14] 采用了类似的思路，将 DS++ 算法进一步泛化，提出了 DS* 算法，并展示了它在多图匹配等问题上的卓越性能。

Maset 等 [10] 提出了基于谱分解 (Spectral Decomposition) 的 MATCHEIG 算法，能够快速求解多图匹配问题。MATCHEIG 通过特征值分解和重构，将两两匹配的结果转换成满足循环一致性条件的多图匹配结果。相对于其他需要迭代优化的算法，MATCHEIG 算法具有解析解，且算法实现非常容易。Tron 等 [9] 提出了基于密度聚类的多图匹配算法 QuickMatch，具有很多优点。

QuickMatch 直接从特征空间出发，输出离散的排列矩阵，不需要预先计算两两匹配的结果，也不需要后处理步骤；QuickMatch 使用多图匹配的性质自动计算聚类大小 k ，不需要人工定义；算法的准确度和计算效率都很高，能够处理大规模的数据集。Jiang 等 [20] 提出了基于拉格朗日乘子法的二图匹配算法 MPGM，能够直接输出接近离散的稀疏化匹配结果，因此不需要复杂的后处理步骤。Wang 等 [21] 提出了 ATGM 算法，首先将源图变换到由目标图节点在特征空间中围成的连续域，再在该连续域中进行匹配。ATGM 算法将参数规模从 $O(m^2n^2)$ 降低到了 $O(mn)$ (m, n 分别为源图和目标图的节点数)，显著地提高了运算速度，同时还获得了更高的匹配精度。Vestner 等 [12] 在二图匹配问题中引入图核，用图核取代二阶相似度计算中常用的欧氏距离。实验表明，引入图核计算二阶相似度，能够显著地提升算法的计算效率。

2. 使用深度学习提取特征

深度学习模型已经在许多计算机视觉任务上取得了重大突破。作为一种强大的特征提取器，许多工作 [4, 6] 训练深度学习模型以提取物体间的一致性特征，并将其用于匹配算法中。

Wang 等 [7] 使用了在 ImageNet 分类问题上预训练的 AlexNet [16] 作为特征提取器，在多张图片之间寻找共有的语义特征。该工作在使用神经网络提取特征的基础上，进行了一个类似多图匹配的优化流程，通过非监督的形式找出了图集中

共有的语义信息和对应的关键点。然而，该工作直接使用了在分类问题上预训练的神经网络，而分类问题与图匹配问题之间的差异可能会导致通过该方法计算得到的特征对匹配问题而言并非最优。近年来，使用端到端神经网络学习匹配特征的方法日益受到关注。Choy 等[4]提出了全局一致性网络 UCN (Universal Correspondence Network)，使用端到端学习的方法，直接获得一个保留了空间一致性和语义一致性的特征空间，如图 4 所示。在训练时，对于具有对应关系的特征点对，网络最小化它们在特征空间的距离；对于其他负样本，网络使它们在特征空间的距离大于一个固定值 m 。UCN 还在网络中加入了空间变换网络[15]，使具有不同旋转和缩放的物体能够保持相似的特征输出。

值得一提的是，在 CVPR2018 的最佳论文提名工作中，Zanfir 等[6]首次提出了端到端的深度图匹配模型。作者选择了 VGG-16[17]作为主干网络，将特征输入可学习的相似度层，获得点对点、边对边的相似度矩阵 \mathbf{K} 。之后，将相似度矩阵输入固定参数的图匹配层，即可得到图匹配结果。在训练时，输入待匹配的两个图像，同时将图匹配结果作为监督信息，网络即可通过前向和反向传播自动优化参数。模型的图匹配层由传统的图匹配求解器组成，在反向传播时，作者推导了图匹配求解器的梯度公式，从而实现了完整的端到端训练。实验结果显示，完全端到端的深度图匹配算法在计算上是可行的，并且该算法在 Pascal VOC 关键点数据集上取得了目前最高的匹配精度。

3. 围绕多图匹配的前沿研究

多图匹配作为图匹配问题中目前一个重要的发展方向，受到了研究者的普遍重视。对于多图匹配问题，除了开发更加精确、高效的多图匹配算法，还需要考虑如何在有限时间内快速求解大规模图匹配问题。大规模图匹配的需求引出了对增量图匹配[13]和分布式图匹配[5]的研究。此外，如果一幅图像中同时存在多个物体，就需要算法能够同时处理图割和图匹配问题[11]。

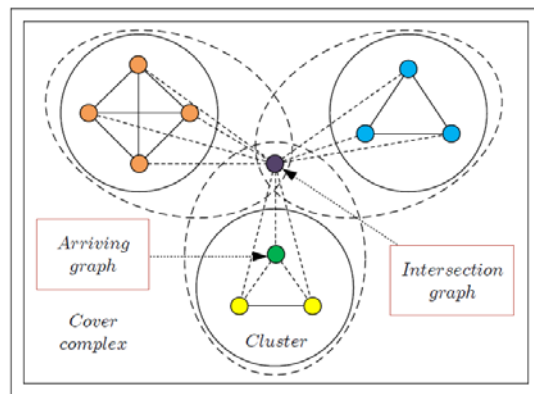


图 5 IMG M 图聚类结构图[13]

值得一提的是，Yu 等[13]首次考虑了增量图匹配的场景，提出了增量图匹配算法 IMG M (Incremental Multiple Graph Matching)。在实际应用中，图片数据经常是随着时间不断加入的，例如由连续帧组成的视频监控数据。现有的多图匹配算法，都聚焦于对已有的全部数据一次性求解，忽略了在线图匹配的应用场景。如果使用现有的多图匹配算法，每次有新的数据到来，都需要对整个数据集重新计算，显然会大大降低算法的效率。因此，增量图匹配具有非常广阔的应用前景。如图 5 所示，IMG M 算法基于随机性和差异性将图聚类，每个聚类内是全连接的拓扑结构；同时，通过一个由算法自动选取的参考图，在多个聚类间通过参考图传递多图匹配的循环一致性信息。实验表明，IMG M 能够在保证一定精度的前提下，提升在线图匹配场景的效率。

Hu 等[5]提出了保证循环一致性的分布式多图匹配框架 DMatch。现有图匹配算法基本都是单机运行算法，由于图匹配问题的复杂性，这些算法几乎无法处理大规模图匹配问题。采用分布式并行计算，能够显著提升大规模图匹配的计算速度，推进大规模多图匹配的实际应用。DMatch 算法使用聚类将整个图集分割成多个子集，子集之间的重叠部分负责传递全局一致性，每个子集内部通过运行现有的多图匹配算法进行优化。在实际运行时，每个子集内的优化过程可以分配到多台分布式节点，或一台计算机的多个核心上。作者在理论上证明了在 DMatch 算法的框架下，子集

内部匹配的一致性能保证全局一致性。

除此之外，Yu 等[11]提出了一个新的场景：如果一幅图像中同时存在多个具有对应关系的物体，如何将它们分开并进行匹配。作者从图的角度出发，同时考虑图割和图匹配问题，构造了一个新的目标函数

$$\begin{aligned} \max_{\mathbf{x}, \mathbf{y}} \quad & \mathbf{x}^T \mathbf{K} \mathbf{x} - \lambda_1 \mathbf{y}^T \mathbf{L}^g \mathbf{y} + \lambda_2 \mathbf{y}^T (\mathbf{I} - \mathbf{X}) \mathbf{y} \quad (4) \\ \text{s.t.} \quad & \sum \mathbf{X}_{ij} = \mathbf{1}, \sum \mathbf{X}_{ij} = \mathbf{1}^T, \mathbf{X}_{ii} = 0, \\ & \mathbf{X} \in \mathbb{S}^+, \|\mathbf{y}\|_2 = 1 \end{aligned}$$

在公式(4)中，第一项衡量图匹配，第二项衡量图割，第三项用于维持图匹配与图割结果的一致性。实验表明，通过对图割和图匹配的联合优化，能够实现同时对同一幅图片中相似物体的分割和匹配。在这项工作中，暂时只考虑了同一幅图片中只存在两个物体的情形。

三、未来发展方向

图匹配问题作为计算机视觉中的基础问题之一，具有一贯的学术价值和广阔的应用前景。本

文介绍了图匹配算法的基本知识，系统梳理了近年来图匹配研究领域在多图匹配、深度图匹配等方向取得的进展，并对其中的重要算法与主要思路展开讨论。我们认为，多图匹配和深度图匹配将成为图匹配领域未来发展的重要推动力。

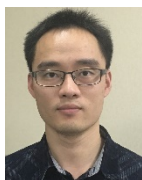
近几年图匹配相关的研究工作普遍基于多图匹配这一大方向，提出了许多新的多图匹配算法，同时也涌现出了分布式图匹配、增量图匹配等新的研究方向。具有 NP-难性质的图匹配的求解一直是制约图匹配问题精度和效率提升的主要障碍，在未来，以多图匹配为中心的理论与应用仍将会是图匹配领域的研究重点之一。特征学习和相似度学习将会是图匹配领域另一个重要的研究方向。如何将深度学习与图匹配尤其是多图匹配更加紧密地结合，如何为了图匹配任务设计更加优秀网络结构，如何在提升模型准确度的同时提高深度图匹配的计算效率，如何提高深度图匹配模型的鲁棒性，都是非常具有潜力和前景的研究方向。

(责任编辑：苏航)

参考文献

- [1] Munkres J. Algorithms for the Assignment and Transportation Problems. *Journal of the Society for Industrial & Applied Mathematics*, 1957, 5(1):32-38.
- [2] Lowe D G. Object Recognition from Local Scale-Invariant Features. *Proceedings of International Conference on Computer Vision*, 1999:1150-1157.
- [3] Tola E, Lepetit V, Fua P. DAISY: an efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2010, 32(5):815-30.
- [4] Choy C B, Gwak J Y, Savarese S, et al. Universal Correspondence Network. *Proceedings of Advances in Neural Information Processing Systems*, 2016: 2414-2422.
- [5] Hu N, Thibert B, Guibas L. Distributable Consistent Multi-Graph Matching. *The IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [6] Zafir A, Sminchisescu C. Deep Learning of Graph Matching. *The IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [7] Wang Q, Zhou X, Daniilidis K. Multi-Image Semantic Matching by Mining Consistent Features. *The IEEE Conference on Computer Vision and Pattern Recognition*, 2018.
- [8] Dym N, Maron H, Lipman Y. DS++: A Flexible, Scalable and Provably Tight Relaxation For Matching Problems. *ACM Transactions on Graphics*, 2017, 36(6): 184-198.

- [9] Tron R, Zhou X, Esteves C, et al. Fast Multi-image Matching via Density-Based Clustering. Proceedings of International Conference on Computer Vision, 2017:4077-4086.
- [10] Maset E, Arrigoni F, Fusiello A. Practical and Efficient Multi-view Matching. Proceedings of International Conference on Computer Vision, 2017:4578-4586.
- [11] Yu T, Yan J, Zhao J, et al. Joint Cuts and Matching of Partitions in One Graph. The IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [12] Vestner M, Kimmel R, Cremers D, et al. Efficient Deformable Shape Correspondence via Kernel Matching. Proceedings of International Conference on 3d Vision, 2017:517-526.
- [13] Yu T, Yan J, Li B, et al. Incremental Multi-graph Matching via Diversity and Randomness based Graph Clustering. European Conference on Computer Vision, 2018.
- [14] Bernard F, Theobalt C, Moeller M. DS*: Tighter Lifting-Free Convex Relaxations for Quadratic Matching Problems. The IEEE Conference on Computer Vision and Pattern Recognition, 2018.
- [15] Jaderberg M, Simonyan K, Zisserman A, et al. Spatial Transformer Networks. Proceedings of Advances in neural information processing systems, 2015:2017-2025.
- [16] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. Proceedings of Advances in Neural Information Processing Systems, 2012:1097-1105.
- [17] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Representation Learning, 2015.
- [18] Yan J, Cho M, Zha H, et al. Multi-Graph Matching via Affinity Optimization with Graduated Consistency Regularization. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 38(6):1228-1242.
- [19] Yan J, Tian Y, Zha H, et al. Joint Optimization for Consistent Multiple Graph Matching. Proceedings of International Conference on Computer Vision, 2013:1649-1656.
- [20] Jiang B, Tang J, Ding C, et al. Graph matching via multiplicative update algorithm. Proceedings of Advances in Neural Information Processing Systems, 2017: 3187-3195.
- [21] Wang F, Xue N, Zhang Y, et al. Adaptively Transforming Graph Matching. European Conference on Computer Vision, 2018.



严骏驰

上海交通大学特别研究员，博导。获 CCF 优秀博士论文，ACM 中国优秀博士论文提名。主要研究方向为计算机视觉与机器学习，特别是结构与时序信息的建模与学习。

Email: yanjunchi@sjtu.edu.cn



汪润中

上海交通大学本科生，主要研究方向为深度学习与图匹配。

Email: roucherwang@sjtu.edu.cn



杨小康

上海交通大学长江学者特聘教授，博导。主要研究方向为图像处理与机器学习。

Email: xkyang@sjtu.edu.cn

视觉盲计算：基于密码学的云环境人脸隐秘识别

北京电子科技学院 金鑫

随着云计算技术的深入应用、网络传输速度的不断加快、图像视频数据的持续增长，人们将越来越多的图像视频数据存储在云端，并且逐渐习惯于依赖云端的强大计算能力和优秀算法处理与分析图像视频。这些智能应用在带来巨大便利的同时，也带来了存储在云端的用户影像数据的隐私泄露问题，一旦云存储服务器被攻破，这些携带大量用户隐私的影像数据将直接泄露。人们逐渐意识到：图像视频的云存储与计算带来的内容隐私泄露风险是网络空间安全的新挑战。

在该背景下，北京电子科技学院可视计算与安全课题组相继提出了基于密码学的云环境人脸盲检测、识别与检索算法，力求在不泄露内容隐私的前提下，在云端隐秘地完成人脸检测、识别与检索。

检测：2017年，本课题组提出图像的随机基图像表示，将待检测图像随机地分解为若干随机基图像，只将随机基图像随机地发送给云端，云端在随机基图像上进行 VJ 算法物体检测的特征计算，将计算结果发回给终端后，终端计算真实的特征响应，再利用百万富翁协议隐秘地与云端阈值比较大小之后确定是否为人脸。该方法同时保护了用户图像的内容隐私以及云服务提供商所训练的算法参数。实验结果表明基于随机基图像的方法将传统的 Blind Vision 方法的计算速度提高了 $10^{\sim}50$ 倍。

识别：2015年，本课题组提出了基于稀疏表示的人脸图像盲识别方法，如图 1 所示，利用稀疏表示将人脸表示成十进制向量，并且提出了十进制向量的欧氏距离隐秘计算方法。在只损失了 4% 识别率的情况下，将 SCIFI 的人脸表示向量长度从 3000bit 降低到 1600bit，并且能够抵御 Luong 等人的攻击方法。随后 2017 年，北京工业大学的毋立芳教授团队提出了基于深度学习的人脸隐秘识别方法，进一步将人脸表示向量长度降低到 256bit。

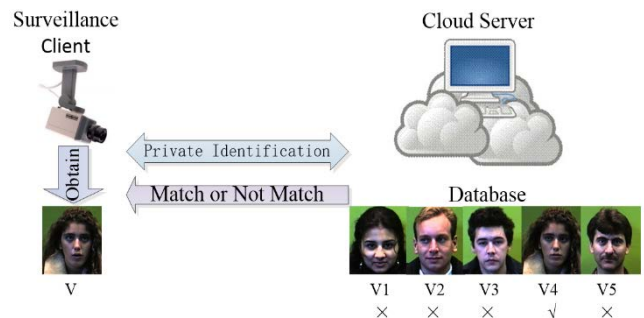


图 1 云环境人脸图像隐秘识别

检索：2017年，本课题组提出了云环境人脸隐秘检索方案，即百度云盘、时光相册等现有云环境明文人脸检索方法的隐秘版本，能够同时保护用户照片的内容隐私与人脸检测服务提供商的模型参数，云端仅提供存储与计算资源。

（责任编辑：任传贤）



金鑫

CCF 计算机视觉专委会委员，现于北京电子科技学院计算机科学与技术系从事教学与科研工作，担任北京电子科技学院可视计算与安全实验室负责人。研究方向为计算美学、图像理解等。

Email: jinxin@besti.edu.cn

显著性引导的视频目标分割

北京理工大学 沈建冰

无监督的视频目标分割旨在自动分割视频中的前景物体，它对于动态场景理解、视频压缩、视频检索等问题具有重要意义。

由于缺乏对目标的标注，传统的无监督视频分割算法主要关注视频过分割或超体素分割，即将视频分割成若干细粒度的、具有内聚性的、时空连续的像素集团。由于该类算法不考虑视频目标的定义和分割，为降低无监督视频分割的歧义性，新算法引入了一定的人类先验，强调物体表观信息和运动信息对于分割视频目标的重要性，提出基于点轨迹的运动分割，尤其是对于视频中的单一主要目标，和目标采样相结合的新算法，达到了较高的分割精度。

但是，我们观察到，由于目标内的运动不一致性和前景背景运动信息的复杂性，单纯基于运动信息的算法容易产生过分割。基于目标采样的算法，在保证一定精度的条件下，过多的牺牲了计算速度。因为对于每一帧都需要进行目标采样，需要耗费大量时间，同时也需要复杂的算法对分割目标进行推断和选择。

基于以上考虑，我们提出一种显著性引导的无监督视频目标分割算法。我们重新强调显著性对于无监督视频分割的重要性，尤其是强调视频显著性物体检测的重要性。近年来，图像显著性物体检测获得了较大进展，主要得益于背景先验的应用，但对于视频显著性检测的研究相对滞后。显著性物体检测强调检测人类视觉系统关注的显著性物体，和视频目标分割具备天然的关联性。

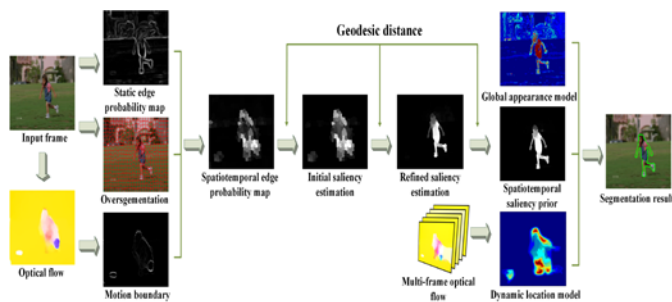


图 1 显著性引导的视频目标分割算

将视频显著性和视频目标分割相结合，具备以下优点：1) 视频显著性检测将物体表观信息、运动信息和近年来证明有效的物体显著性先验假设相统一，为视频目标提供了一个一般的、可靠的、且符合人眼视觉特性的高层描述；2) 省略了耗时的目标采样和复杂的分割目标选取的过程，为视频分割提供了一个统一的且直观的框架，视频显著性检测针对视频目标进行分析，视频分割基于检测结果进行精确分割；3) 填补了视频显著性检测和视频分割间的空白。如图 1 所示，算法分为：显著性检测和视频目标分割。在显著性检测部分，我们利用帧间的连续性和前景/背景的表观差异进行优化。在视频分割部分，我们综合考虑显著性、前景/背景表观模型、多帧的位置累加信息和时空连续性，得到最终分割结果。

基于以上算法，我们在数据集如 SegTrackV1、SegTrackV2 和 FBMS 上取得了较高的分割精度和较快的分割速度。该研究的相关成果于发表于国际期刊和会议 IEEE TPAMI、CVPR 2015。

(责任编辑：王金甲)



沈建冰

CCF 计算机视觉专委会委员，北京理工大学教授，主要研究方向为计算机视觉和机器学习

Email: shenjianbing@bit.edu.cn

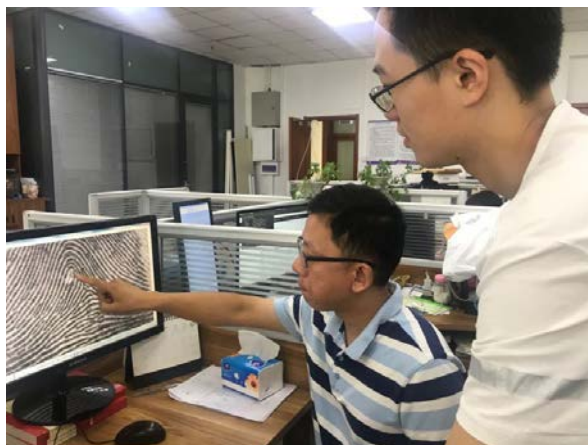
清华大学冯建江副教授访谈

2018年8月26日，专委秘书处采访了清华大学自动化系、杰出青年学者冯建江副教授。下面是采访实录。

冯老师您好，您在图像处理与模式识别等领域做出了很多高水平研究工作，获得了国家自然科学基金优秀青年基金，中国电子学会科技进步一等奖1次、教育部自然科学二等奖2次，是青年科技工作者学习的楷模。您能否向我们介绍一下您现在主要的研究方向是什么？是如何选择现在的研究方向的？

我当前的主要研究领域是指纹识别和医学图像分析。在指纹识别领域，我主要关注犯罪现场指纹的识别问题。现场指纹是嫌疑人不经意间遗留在案发现场的指纹，对于案件侦破非常重要，被称为物证之首。但是现场指纹包含的信息量很少，而有限的信息又埋藏在极强的噪声里，还需要在规模达到亿级的指纹档案库中搜索匹配的档案指纹。现有的现场指纹识别技术虽然已经很有实用价值了，帮助警方破获了大量案件，但是技术性能的提高空间还很大。在医学图像领域，我主要研究基于影像的癌症和心血管病辅助诊断。癌症和心血管病是世界人口的主要死因，早期诊断的意义重大。随着我国居民健康意识和投入的提高，影像设备数量和水平的提高，近年来医学影像数据大幅增长，相比之下影像医生存在严重短缺。这是图像处理科研人员大有可为的一个领域。

您在指纹方面做了大量的工作，能否和大家分享一下您是怎么选择这个研究方向的？您在这个研究方向上前行时，有没有遇到什么挫折，您又是怎么克服的呢？



指纹识别是我本科毕业设计时候的选题。当时实验室有指纹识别和多媒体通信两个方向可选，我觉得指纹识别很直观，就选择了它。之后就没有中断，至今已经研究18年了。研究中最大的挫折是在读博士时候。当时博士论文预定的主题是基于纹线的指纹识别。选这样的题目是源于一个朴素的想法，既然传统的指纹识别算法都是基于细节点，而细节点是纹线上的特殊点，那么我用信息量更高的纹线做识别，识别率肯定更高。我信心满满地开工了，三、四年时间把图像增强、纹线提取、纹线后处理、纹线匹配等各个环节的算法都做出来了，但是最终的识别率始终不如基于细节点的方法。更头疼的是，由于整个方案不是主流的统计模式识别方法，缺少理论框架，很难改进。读了四年博士，如果再换方向，等拿到博士学位，恐怕椅子都要坐穿了。最后只好在完美主义和现实主义中间做了妥协，将基于细节点的方法和纹线的方法进行融合，实现了提高识别率的目标。

您所获得的成果中，哪一项是最令您骄傲的？能和大家分享下您的感受么？



特征提取是现场指纹识别的关键步骤，它需要在强噪声的干扰下，尽可能准确而全面地提取出现场指纹的特征。自动指纹识别技术诞生后的40多年时间里，出现了很多种特征提取算法。这些算法对于用传感器或者油墨捺印得到的指纹图像的处理效果比较好，但是对于现场指纹却始终无法取得满意的结果。因此，在实际应用中，现场指纹的特征一直是由指纹专家人工提取的，费时费力。我研究这个问题一段时间后，有了一种朦胧的意识，觉得算法应该学习指纹专家。但是指纹专家在脑子里是如何提取指纹特征的，没人能说得清楚。大概过了两年多的时间，我才找到工程上可行的方法，将指纹的统计规律融入算法，把现场指纹的特征提取性能提高了一大截。这项工作得到了学术界同行和工业界专家很高的评价，在我国某些警用指纹识别系统中也得到了成功应用。

指纹识别的应用已经很普遍了，似乎指纹识别技术已经进入成熟期了。下一步可能有哪些新的应用？

从上世纪70年代的警用自动指纹识别系统，到本世纪初的指纹考勤机，再到近几年的智能手机、指纹锁，指纹识别技术日益进入我们的日常生活。随着指纹传感器和识别算法的性价比逐步提高，指纹识别技术有可能嵌入到更多的物品中，赋予身份识别功能。

作为博士生导师，请问您是如何选拔研究生，

以及培养研究生的？能否和大家分享一下您的经验？

清华的很多本科生在大二、大三的时候会主动报名参与实验室的科研项目。我会根据学生的兴趣，安排他们参与某研究生的课题。其中有的学生能够一直坚持下来。经过半年到一年的考察，应该能从中发现科研兴趣高、有潜质的学生。在研究生的培养方面，我觉得个性化的培养很重要。不同研究生的基础、个性存在很大差异，选题的难度不同，在不同科研阶段所需要的帮助不一样，毕业条件也无法统一要求。这就需要导师在每个学生上花足够的时间。

您在国际学术刊物及会议上发表论文80余篇，包括模式识别领域顶级期刊T-PAMI上10篇长文。能否分享一下您在发表高水平论文方面的经验？

我们领域的高水平论文通常具有这样的特点，选题重要、方法新颖、性能突出、表达清晰。如何做到前三点，实际上是关于如何做科研。这个题目太大了，我恐怕讲不好。我这里假设选题、方法、性能都很好了，如何把论文写好。下面的三条建议可能对研究生有参考价值。一、多读经典论文。不但平时做研究要读，在写论文时候要再次翻阅，学习他们写作方面的优点；二、多积累素材。在汇报进度、阅读文献、听学术报告等各种活动中，有意识地积累有用的文字表达、图表，这样将来的写作效率高。三、论文完成后，请他人（合作者、导师等）从读者、审稿人的角度，用非常挑剔的眼光对论文挑毛病。

您曾在美国高校进行博士后研究工作，在国外的工作经历，您最大的收获是什么？

最大的收获是发现教授可以很有成就感，而且这种成就是实实在在的，是源于对社会独特的贡献。因为这个发现，我才决定到大学发展。

您理想中的科研环境处于一种什么样的状态，

对科研资金、管理制度有什么样的期待呢？

科研项目可以粗分为自由探索和目标导向。两种项目的经费分配和管理方法应该有所区别。我更喜欢在宽松的环境下做自由探索。但是，如果想做真正落地的大项目，就需要大量的人力、财力和物力，采用专业的工程管理、经费管理手段也是必要的。

科研工作需要投入大量精力，您在科研中取得了这些成绩，肯定要付出大量时间和精力。请问您是如何平衡家庭和事业的？

对于我来说，家庭和事业都重要，很难笼统地说哪个更重要。但是具体到某一件家庭中的事情或者工作中的事情，相对重要性还是可以判断的。

委员好消息

★ 2018年1月5日，教育部公示了2017年度长江学者建议人选名单。根据《“长江学者奖励计划”实施办法》，经高校推荐、通讯评审、会议评审等程序，共产生463名2017年度长江学者特聘教授、讲座教授和青年学者建议人选。5月底以来，各高校陆续公布了最终入选人员名单。CCF-CV专委会委员、复旦大学姜育刚教授和杭州电子科技大学俞俊教授入选青年学者。青年长江学者是长江学者奖励计划的青年学者项目，2015年由中华人民共和国教育部设立，旨在遴选一批在学术上崭露头角、创新能力强、发展潜力大，恪守学术道德和教师职业道德的优秀青年学术带头人，每年遴选200名左右。

★ 2018年6月22日，CCF-CV专委会委员、大连理工大学李培华教授团队在CVPR2018大规模物种识别竞赛iNaturalist Challenge (iNat18)中荣获冠军。基于该团队研发的统计聚合深度卷积网络模型，在识别率上分别高出第2名Deep Learning Analytics公司和第3名Baidu VIS

冯建江

清华大学自动化系副教授，博士生导师。研究方向为图像处理与模式识别。担任国际学术期刊《Image and Vision Computing》编委，国际生物特征识别学术会议



(International Conference on Biometrics, 2014-2016) 领域主席，中国生物特征识别学术会议(2015-2018)程序委员会主席。在国际学术刊物及会议上发表论文80余篇，包括模式识别领域顶级期刊T-PAMI上10篇长文。获中国发明专利10余项、美国发明专利4项，指纹技术授权国内外领军企业。获国际会议最佳论文奖/最佳学生论文奖3次、中国电子学会科技进步一等奖1次、教育部自然科学二等奖2次、国家自然科学基金委优秀青年基金。

(责任编辑：余焯，黄岩，张汗灵)

团队1.1%和1.8%。相对于粗粒度、仅包含1000类、类别样本均衡的ImageNet竞赛，iNat18由于以下特点更具有挑战性：(1)属于精细粒度分类问题——类别之间差异细微；(2)类别数规模更大——类别数是ImageNet竞赛的8倍多，包含8142类的动物、植物和菌类物种；(3)不同类别样本分布极不均衡，呈现长尾分布——每类最多有1000张样本图像、最少却仅有2张。该竞赛由谷歌公司、加州理工和康奈尔大学组织，得到微软公司赞助。

★ 2018年7月25日，CCF-CV专委会委员、上海交通大学杨小康教授等合著的论文Using Free Energy Principle for Blind Image Quality Assessment荣获《IEEE Transactions on Multimedia》的最佳论文奖。该论文将脑科学和神经科学的研究成果引入视觉感知过程的建模中，解决了无参考图像感知质量评价的难题，目前已经被引用近200次，连续3年入选ESI高被引论文，是IEEE多媒体汇刊自2015年来所有

发表的论文中被引用总次数最高的文章。该奖项由 IEEE 电路与系统分会、信号处理分会通信分会和计算机分会联合设立，授予三年内在 IEEE 多媒体汇刊中发表的优秀论文。奖项综合考虑论文的创新性、实用性、时效性和书面表达，每年最多有一篇论文入选（有时空缺）。此次是该奖项设立 16 年来首次由国内单位独立获得。

✪ 2018 年 7 月 25 日，由 CCF-CV 专委会委员、清华大学鲁继文副教授等合著的论文 Rank-Consistency Multi-Label Deep Hashing、由 CCF-CV 专委会委员、清华大学苏航老师等合著的论文 Essay-Anchor Attentive Multi-Modal Bilinear Pooling for Textbook Question Answering、由 CCF-CV 专委会委员、杭州电子科技大学俞俊教授等合著的论文 Deep Point Convolutional Approach for 3D Model Retrieval 以及由 CCF-CV 专委会常务委员、中科院计算所山世光研究员和 CCF-CV 专委会副主任、中科院计算所陈熙霖研究员等合著的论文 Continuity-Discrimination Convolutional Neural Network for Visual Object Tracking 在 ICME 2018 国际会议上荣获铂金最佳论文奖。

✪ 2018 年 7 月 23-27 日，ICME 2018 在美国圣地亚哥举行。CCF-CV 专委会委员、上海交通大学林巍峭副教授获得 ICME 2018 杰出领域主席奖。ICME 2018 共有领域主席 53 人，其中 5 人获此殊荣。

✪ 2018 年 8 月 3 日，科技日报正式公布了《国家自然科学基金委员会关于公布 2018 年度国家杰出青年科学基金建议资助项目申请人名单的通告》，CCF-CV 专委会常务委员、南京信息工程大学刘青山教授入选。和往年一样，今年建议资助的国家杰青共有 200 人，来自 112 家单位的科学家入选。国家杰出青年科学基金项目主要支持在基础研究方面已取得突出成绩的青年学者自主选择研究方向开展创新研究，促进青年科学技术人才的成长，吸引海外人才，培养造就一批进入世界科技前沿的优秀学术带头人。1994 年，正式命名为“国家杰出青年科学基金”。

✪ 2018 年 8 月 5-8 号，ICFHR 2018（第 16 届手写识别前沿国际会议）在美国 Buffalo 召开。CCF-CV 专委会委员、华南理工大学金连文教授带领的团队，在会议举办的 Thai Student Signature and Name Component Recognition and Verification 竞赛中，获签名鉴别任务第一名。

✪ 2018 年 8 月 17 日，国家优青出炉，CCF-CV 专委会 4 位委员获得资助：清华大学鲁继文副教授、浙江大学章国锋副教授、江西财经大学方玉明教授、郑州大学徐明亮副教授。优秀青年科学基金也被称为“小杰青”，是国家杰出青年基金的铺垫性科技支撑基金，设立于 2012 年，本年度共资助 400 项。

✪ 2018 年 8 月 23 日，CCF-CV 专委会 3 位委员，北京航空航天大学王蕴红教授、南京航空航天大学陈松灿教授和微软研究员王井东研究员入选 IAPR Fellow。IAPR Fellow 每两年评选一次，当选率少于会员总数的 0.25%。

✪ 2018 年 8 月 24 日，ICPR 2018（第 24 届国际模式识别大会）论文颁奖，CCF-CV 专委会副主任、北京大学查红彬教授等的论文 Scalable Monocular SLAM by Fusing and Connecting Line Segments with Inverse Depth Filter 以及 CCF-CV 专委会主任、中科院自动化所研究员谭铁牛院士和 CCF-CV 专委会委员、中科院自动化所王伟助理研究员等的论文 DeepFirearm: Learning Discriminative Feature Representation for Fine-Grained Firearm Retrieval 获得 ICPR 2018 最佳科学论文奖。

✪ 2018 年 8 月 24 日，2018 年国家重点研发计划“生物安全关键技术研发”重点专项公示，CCF-CV 专委会委员、上海海事大学周日贵教授负责的项目“特殊生物资源检测与溯源技术研究”进入公示名单。国家重点研发计划是 2014 年由原来的 973 计划、863 计划、国家科技支撑计划、国际科技合作与交流专项、产业技术研究与开发基金和公益性行业科研专项等整合而成，旨在集中力量解决国家重大战略科技问题。

（责任编辑：刘海波）

基于深度学习的视觉问答技术 开源代码

北京电子科技学院 周兴晖 金鑫

视觉问答技术 (Visual Question Answering, VQA) 的主要任务是给定一张图像和一个关于该图像内容的文字问题，视觉问答的任务旨在从若干候选文字或语句回答中选出正确的答案。其本质是分类任务，也有工作是用 RNN 解码来生成文字回答。视觉问答也是计算机视觉和自然语言处理两个领域的交叉任务。



图 1. 图像语义分割示意图

本文着重介绍几个基于深度学习的视觉问答技术数据集和部分开源代码。

1. 视觉问答数据集

1.1 DAQUAR 数据集

介绍: 它是最小的 VQA 数据集之一。它由 6795 训练和 5673 测试的问答对构成。该数据集还可以整理为一个更小的数据集, 仅包含 37 个类别, 称为 daquar1-37。daquar1-37 仅包含 3825 训练和 297 个测试的视觉问答对。

主页: <https://www.mpi-inf.mpg.de/departments/computer-vision-and-multimodal-computing/research/vision-and-language/visual-turing-challenge/>

1.2 COCO-QA 数据集

介绍: COCO-QA 包含 78, 736 个培训和 38, 948

个测试问答对。大多数问题询问图像中的对象 (69.84%), 其他问题涉及颜色 (16.59%), 计数 (7.47%) 和位置 (6.10%)。所有问题都只有一个单词的答案, 且仅有 435 个的独立的答案。

主页: http://www.cs.toronto.edu/~mren/imag_eqa/data/cocoqa/

COCO-QA: What does an intersection show on one side and two double-decker buses and a third vehicle?

Ground Truth: Building



图 2. COCO-QA 数据集

1.3 VQA 数据集系列

介绍: VQA 数据集包含来自 COCO 的图像。关于这个数据集的大部分工作都集中在包含来自 COCO 的现实图像部分, 一般称其为 COCO-VQA。而数据集的合成部分称为 SYNTH-VQA。COCO-VQA 每个图片包含三个问题, 每个问题有十个答案。COCO-VQA 包含的问题相对较多 (共 614, 163 个, 有 248, 349 个用于训练集, 121, 512 个用于验证集, 244, 302 个用于测试集)。

主页: <http://visualqa.org/>

2. 视觉问答技术开源算法

视觉问答技术的基本思路是：使用 CNN 从图像中提取图像特征，用 RNN 从文字问题中提取文本特征，之后设法融合视觉和文本特征，最后通过全连接层进行分类。图像和问题特征是分类系统的输入，每个唯一答案被视为不同的类别。该任务的关键是如何融合这两个模型网络的特征。最为简单直接的融合方法是将视觉和文本特征在某个维度上拼成一个向量，或者让视觉和文本特征向量逐元素相加或相乘。

2.1 基线模型(Baseline Models)

论文： H. Noh, P. H. Seo, and B. Han, “Image question answering using convolutional neural network with dynamic parameter prediction,” in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

工作： 通过学习一个动态参数层的卷积神经网络 (CNN)，进行网络权值的自适应预测。在自适应参数预测中，我们采用了一个单独的参数预测网络，该网络由以问题为输入的门控递归单元和生成一组候选权重作为输出的全连通层组成。通过使用哈希技术来减少这个问题的复杂性，使用预定义的哈希函数选择参数预测网络给出的候选权重，以确定动态参数层中的各个权重，进而提高基线模型的准确率。

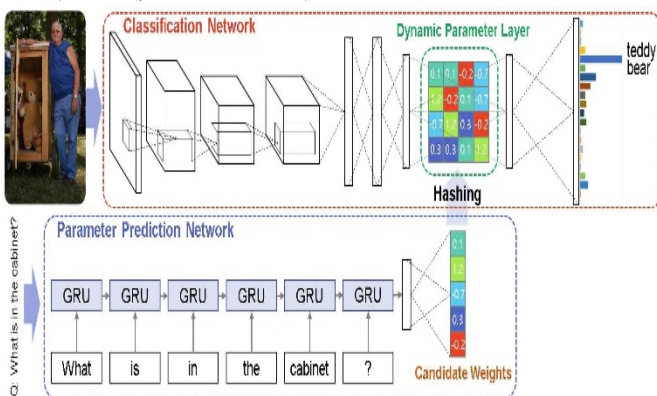


图 3. FCN 网络结构

代码： <https://github.com/MarvinTeichmann/te nsorflow-fcn>

2.2 注意力模型(Attention Based Models)

经典注意力模型：

论文： K. J. Shih, S. Singh, and D. Hoiem, “Where to look: Focus regions for visual question answering,” in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

工作： 美国伊利诺伊大学的这项工作使用 CNN 来从这些盒子中提取特征。VQA 系统的输入包括这些 CNN 特征，问题特征和一个选择性的答案。他们的系统被训练为每一个选择题的答案都能得到一个分数，并选出得分最高的答案。通过传递 CNN 区域特征点积以及问题，最后合并到一个全连接层中，可以简单地计算出权重的每个区域的加权平均得分。

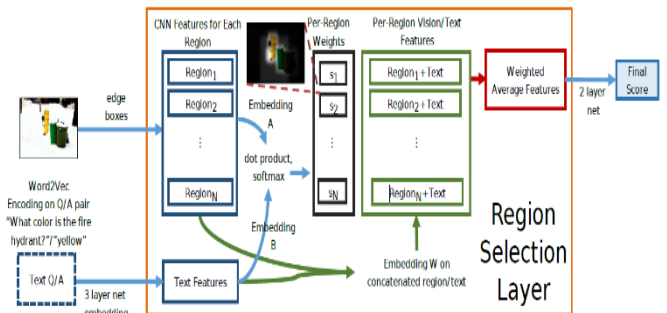


图 4. 经典注意力模型

代码： https://github.com/kevjjshih/wtl_vqa

堆积注意力网络 Stacked Attention Network

论文： Z. Yang, X. He, J. Gao, L. Deng, and A. J. Smola, “Stacked attention networks for image question answering,” in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

工作： 美国卡耐基梅隆大学和微软提出的动态注意力网络。其注意层由一个单独的权重层确定，该权重层用问题和带激活函数的 CNN 的特征图计算图像位置上的注意力分布。然后将该分布应用到 CNN 的特征层中，使用加权和在空间特征位置之间进行聚合，从而生成一个完整的图像，它比其他区域更强调某些空间的区域。将这个特征向量与问题特征向量结合，可得到预测答案。

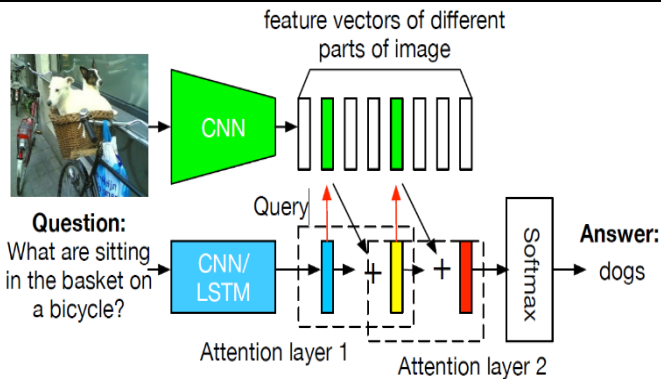


图 5. Stacked Attention Network

代码: <https://github.com/zcyang/imageqa-san>

空间记忆网络 Spatial Memory Network

论文: H. Xu and K. Saenko, “Ask, attend and answer: Exploring question-guided spatial attention for visual question answering”, in European Conference on Computer Vision (ECCV), 2016.

工作: 美国麻省理工学院提出的空间记忆网络, 该模型通过估计图像斑块与问题中单个单词的相关性来产生空间关联性。这种文字引导的注意力被用来预测注意力分布, 然后用来计算图像视觉特征的加权和。然后探索了两种不同的模型。在单跳模型中, 将编码整个问题的特征与加权视觉特征相结合并预测答案。在两跳模型中, 将视觉特征和问题特征的组合循环回到关注机制中, 用于细化注意力分配。

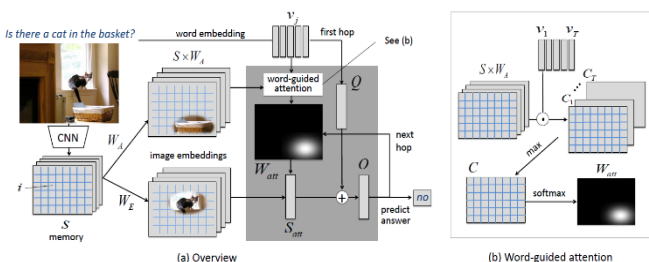


图 6. 空间记忆网络 Spatial Memory Network

代码: <https://github.com/VisionLearningGroup/Ask Attend and Answer>

DAN

论文: H. Nam, J. Ha, and J. Kim. Dual attention networks for multimodal reasoning and matching. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

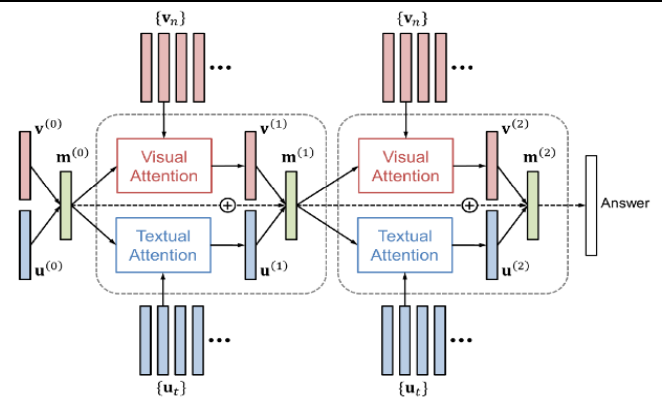


图 7. DAN 结构

代码: <https://github.com/iamrhelo/pytorch-vqa-dan>

2.3 双线性融合 (Bilinear Pooling Methods)

MLB

论文: J.-H. Kim, K.-W. On, J. Kim, J.-W. Ha, and B.-T. Zhang, “Hadamard product for low-rank bilinear pooling”, arXiv preprint arXiv:1610.04325, 2016.

工作: 首尔国立大学的工作者们使用多模态低秩双线性池 (MLB) 方案, 使用 Hadamard 乘积和线性映射来实现近似双线性池。当与空间视觉注意机制一起使用时, MLB 可以与 VQA 中的 MCB 相媲美, 但计算复杂度较低, 并且使用的神经网络参数较少。

代码: <https://github.com/jnhwkim/MuLowBiVQA>

MCB

论文: A. Fukui, D. H. Park, D. Yang, A. Rohrbach, T. Darrell, and M. Rohrbach, “Multimodal compact bilinear pooling for visual question answering and visual grounding,” in Conference on Empirical Methods on Natural Language Processing (EMNLP), 2016.

工作: 美国加州大学和日本索尼公司在众多的基线模型中发现, 由于在双线性融合的过程中, 计算高维双线性外积通常是不可行的, 而使用多模态双线性池 (MCB) 来高效地表达多模态特征。在视觉问答方面对 MCB 进行了范围较广的评价, 取得了目前所有基线模型中效果最好的结果。

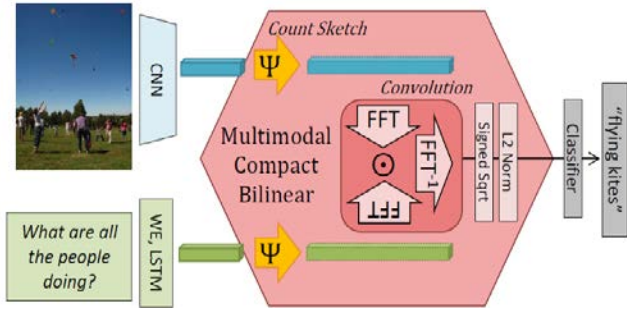


图 8. MCB-ensemble 的模型

代码: <https://github.com/akirafukui/vqa-mcb>

a-mcb

MFB

论文: Z. Yu, J. Yu, J. Fan, and D. Tao. Multi-modal Factorized Bilinear Pooling with Co-Attention Learning for Visual Question Answering. ArXiv e-prints, Aug. 2017.

工作: 本文提出了一种多模态分解双线性池化方法, 有效地结合了多模态特征, 使 VQA 性能优于其他双线性融合方法。对于细粒度的图像和问题表示, 我们开发了一种“共享注意力”机制, 使用端到端深度网络架构来共同学习图像和问题注意力。

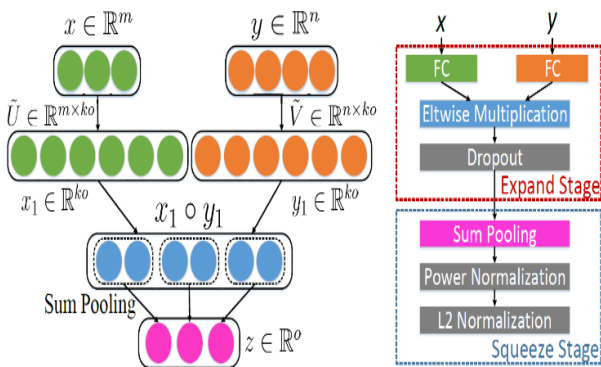


图 9. MFB 模型

代码: <https://github.com/yuzcccc/vqa-mfb>

MFH

论文: Z. Yu, J. Yu, C. Xiang, J. Fan, and D. Tao. Beyond bilinear: Generalized multimodal factorized high-order pooling for visual question answering. TNNLS, 2018.

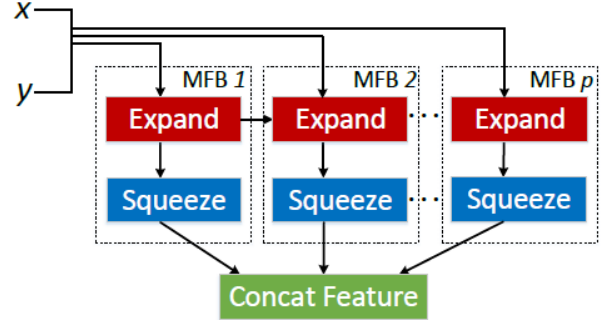


图 10. MFH 模型

工作: 针对多模态特征融合问题, 作者提出了一种广义多模态因子分解的高阶池化方法 (MFH), 通过充分利用多模态特征之间的相关性, 实现多模态特征的更有效融合, 进一步提高 VQA 性能。在答案预测中, 利用 KL (Kullback-Leibler) 散度作为损失函数, 可以更准确地表征具有相同或相似含义的多个不同答案之间的复杂关联, 从而使我们能够获得更快的收敛速度, 并在答案预测中获得略好的准确性。

代码: <https://github.com/yuzcccc/vqa-mfb>

(责任编辑: 沈沛意)



金鑫:

现于北京电子科技学院计算机科学与技术系从事教学与科研工作, 担任北京电子科技学院可视计算与安全实验室负责人。研究方向为计算美学、图像理解等。
个人主页: www.jinxin.me



周兴晖:

北京电子科技学院硕士生, 研究方向为计算美学。
Email: graydoo@163.com

异常行为检测数据集

北京航空航天大学 王田 乔美娜

随着人类社会的发展，城市化进程的不断推进，在广场、购物中心、车站等场合中人们的活动集中，使得公众场景的安全问题面临的挑战逐步加剧。同时，近几十年来电子技术飞速发展，摄像机、网络通信设备的硬件成本逐步降低，监控摄像机被广泛的应用于社会生活中的各个领域。针对公众领域的监控需求，需要检测异常行为，辅助人们对异常态势及时防控。

传统的安防方式主要包括安防人员在电视墙前做现场监控和在特殊事件发生后调阅视频数据对事件反演推算。但是这两种方式都需要人类对监控视频无间断、高强度体力支出，容易造成人的视觉疲劳、心理懈怠，并且都不能快速有效的对突发异常事件做出回应。异常行为检测可采用计算机视觉与模式识别的关键技术，分析理解监控视频，对公众场景的异常行为进行快速检测与识别，并为针对异常行为的智能优化的决策提供技术支持。

本文着重介绍几个异常行为检测的常用数据库，包括：局部异常行为检测数据库 UCSD、Avenue；全局异常行为检测数据库 UMN、PETS。

1. 局部异常行为检测数据库

1.1. UCSD 数据库

UCSD 数据库是 University of California San Diego 的研究者提出，用于局部异常事件 (Local Abnormal Event, LAE) 的检测，是以人行道为场景的数据库，由相机以俯拍的角度获得，场景中行人的密集程度由稀疏到拥挤。真值包括帧级别和像素级别，帧级别为二进制值，表示当前帧是否为异常；像素级别为二进制掩码，表示当前区域是否存在异常。数据库分为两个子集：Ped1，图片的分辨率为 238×158 ，34 个训练视

频样本、36 个测试样本，其中 10 个片段有帧级别的真值，人的运动方向与相机拍摄视角平行；Ped2，图片的分辨率为 360×240 ，16 个训练视频样本、12 个测试样本，其中 12 个片段有帧级别的真值，人的运动方向垂直于相机拍摄视角。在正常场景中，视频中只包括行走的人；在异常场景中，异常行为由人类的异常行为和异常目标闯入造成。常见的异常包括使用自行车、滑板、坐轮椅的人以及小汽车等。

数据库的网址为：<http://www.svcl.ucsd.edu/projects/anomaly/dataset.htm>



(a) 正常场景

(b) 异常场景

图 1. UCSD Ped1 数据库

1.2. Avenue 数据库

Avenue 数据库是香港中文大学的研究者提出，是以校园为场景的数据库。包括 16 个训练视频样本、21 个测试样本，图片分辨率为 360×640 ，与 UCSD 类似，正常场景下为行人在场景中徘徊，异常的行为包括奔跑、抛物等。

数据库的网址为：<http://www.cse.cuhk.edu.hk/leojia/projects/detectabnormal/dataset.html>



图 2. Avenue 数据库

2. 全局异常行为检测数据库

2.1. UMN 数据库

UMN 数据库为 University of Minnesota 大学的研究者提出。分为三个场景：草坪、室内和广场。图片分辨率为 320×240 。本数据库用于全局异常事件(Global Abnormal Event, GAE)的检测。全局异常行为检测只有帧级别的真值，没有像素级别的真值。此数据库中正常的行为是场景中的行人徘徊，异常行为是场景中的人突然奔跑，用于模拟恐慌场景下的全体异常事件。其中，在室内场景中，由于入口处的门会经常被行人打开或关闭，导致室内光线变化，这给异常行为的检测带来了挑战。

数据库的网址为：http://mha.cs.umn.edu/projects/abnormal_events.shtml



图 3. UMN 数据库

2.2 PETS 数据库

PETS 数据库为 Reading University 的研究者提出。是多相机获得的群体活动数据库，图片分辨率为 768×576 。此数据库被应用于多个方面，包括人群的计数和密度估计、单个(多个)个体的跟踪、群体动作识别等。针对不同的任务，数据库给出了对应的训练、测试视频片段。对于异常行为检测问题，根据需求可定义场景为：正常行为是人们沿不同方向行走，异常行为是大家沿同一个方向前进，用于模拟有组织的人群活动的异常行为；或正常行为是人们排成一个队列前进，异常行为是部分人离开队列，沿不同方向前进，用于模拟群体行为中行进队列发生变化的异常行为。

数据库的网址为：<http://www.cvg.reading.ac.uk/PETS2009/a.html>



图 4. PETS 数据库

(责任编辑：李策)



王田：

讲师，北京航空航天大学，自动化科学与电气工程学院，研究方向为计算机视觉、模式识别、机器学习等。



乔美娜：

硕士研究生，北京航空航天大学，自动化科学与电气工程学院，主要研究方向为计算机视觉、异常行为检测和识别。

招聘信息

一、DeepMotion 招聘深度学习研发和嵌入式开发工程师

DeepMotion 深动科技提供以高精地图为核心的感知、定位、决策自动驾驶解决方案。我们将最前沿的人工智能和三维重建技术融合到产品设计中，使消费级的传感器达到专业级设备的性能和精度，实现可量产的自动驾驶解决方案。我们相信，创新的技术可以使驾驶更加安全、可靠、智能，给人们的生活带来更好的出行体验，为社会创造更大价值。

深动科技是一支快速成长的年轻团队。公司主要创始人员来自微软研究院，有深厚的机器视觉、三维重建和深度学习的产品技术开发经验，曾在国际顶级会议、期刊上发表过百余篇论文。团队成员大部分来自清华、北大、中科院、微软、UC Berkeley、CMU 等顶尖的科研院所。我们欢迎不同行业有才华的您加入 DeepMotion 大家庭。

1、岗位名称：深度学习研发工程师

岗位职责：

- 研发无人驾驶中的视觉理解算法；
- 负责模型设计、训练、调优；
- 负责代码实现，文档/专利/论文撰写。

任职要求：

- 熟悉计算机视觉常用图像理解方法(如图像分类、分割，物体检测、跟踪等)；
- 熟悉机器学习、深度学习原理和常用模型；
- 有 CVPR / ICCV / ECCV 等顶会文章或竞赛者优先；
- 熟练使用 Python / C++ / Git 。

2、岗位名称：深度学习嵌入式开发工程师”

岗位职责：

- 负责深度学习模型的加速、量化、压缩、修剪；
- 负责深度学习模型在 ARM / FPGA / TX2 等嵌入式平台的算法优化和部署。

任职要求：

- 熟悉嵌入式系统编程及优化，及常见深度学习

模型及框架；

- 熟悉矩阵运算的加速；
- 精通 C/C++，熟练使用 Git；
- 熟悉 Caffe / TensorFlow / mxnet。

3、薪资福利：

DeepMotion 薪酬：全职薪资 15k - 40k/月；

实习薪资：平均 200-500 元 /日，实习 3 个月以上时间，优秀者公司将予以录用或免于试用期；

My Health：补充医疗保险计划；

My Life：弹性工作制，舒适、人性化的办公环境，带薪年假；

My Perks：零食无限量供应，咖啡、健康饮品无限畅饮。员工生日会、入职周年庆祝。年度 Outing；

My Money：国家社会保险及住房公积金足额缴纳。

联系方式：

投递邮箱：career@deepmotion.ai

公司地址：北京市海淀区中关村 SOHO

邮件主题：应聘岗位+姓名+来源

例如：应聘全栈开发工程师+姓名+来自微信公号

二、中国科学院深圳先进技术研究院智能仿生中心博士后招聘启事

中国科学院深圳先进技术研究院(以下简称先进院)由中国科学院、深圳市政府和香港中文大学三方 2006 年在深圳共同组建。现已形成多学科交叉、集成创新的特色与优势，成为海外高层次人才回国工作的理想发展平台。

智能仿生中心隶属于先进院先进集成技术研究所，由国际机器人著名专家徐扬生院士于 2006 年创建。中心致力于服务机器人的研究和开发，以服务机器人三要素感知、认知和动作中的前沿科学问题和核心技术为主要研究内容，以机械、信息、仿生学、医学等多学科交叉集成为特色，致力于家庭服务机器人、医疗和康复机器人、智能系统等方面的前沿研究和产业化。中心网页：

<http://www.siat.cas.cn/jgsz/kyxt/jcs/yjdy>

[/znfs 2/](#)

合作导师吴新宇，博士生导师，研究员，现为智能仿生中心执行主任。国家“万人计划”科技创新领军人才，十三五国家重点研发计划“智能机器人”专项专家组成员，十三五国家重点研发计划“公共安全风险防控与应急技术装备”重点专项指南撰写专家组成员。于中国科学技术大学获工学学士和硕士学位，博士毕业于香港中文大学。近5年承担国家自然科学基金，国家863项目，中科院知识创新工程重要方向项目等30余项。在机器人国际期刊和国际会议上发表论文120余篇，出版英文专著2本，申请国家专利和国际专利50余项。他是中国仪器仪表学会理事，广东省智能机器人重大专项专家组副组长，中国自动化学会机器人专业委员会委员，中国青年科技工作者协会委员。他是相关国际会议重要组织者，如ICIST2014组织委员会主席，ICIA2017程序委员会主席，ICARCV2018程序委员会主席等。

1、岗位要求：

- (1) 获工科院校博士学位，计算机、控制工程、模式识别与智能系统等相关专业毕业，成绩优秀；
- (2) 品行端正，学术作风良好；
- (3) 有机器人项目经验者优先；
- (4) 有在相关国际期刊上发表SCI检索论文者优先；
- (5) 年龄35周岁以下。

2、享受待遇：

- (1) 提供在深圳地区及全国范围内有竞争力的薪酬，博士后年薪30万以上(含深圳市每年补助12万)；
- (2) 出站在深工作满5年，或者在站期间申请到中国博士后基金或国自然等项目的，可以认定为深圳市后备人才，享有购房补贴160万及相关待遇；
- (3) 博士后出站留深工作，可享受深圳市有关规定的待遇，其中包括30万元科研启动经费(等同于生活补助)；
- (4) 在站期间可申请深圳市住房补贴3万元；
- (5) 进站报到后可落户深圳，其配偶及其未成年

子女可办理户口随迁；

- (6) 享有年度考核奖金、年终奖金、横向课题奖励、专利申请奖励、五险一金及伙食补贴；
- (7) 博士后在站工作期间计入我院工龄，可参加院职称评定，享受员工同等的其他福利待遇，出站优先留院工作；
- (8) 深圳市提供给每名博士后每月免费使用新能源小汽车90小时。

3、应聘程序：

- (1) 请有意向者通过电子邮件发送个人简历以及代表性学术成果，包括但不限于论文、专利、参与项目情况等资料；
- (2) 对于有进一步意向的应聘者，面试内容和程序另行通知。

联系方式：

简历接收邮箱：nan.lil@siat.ac.cn

本岗位常年招聘，欢迎有意向者投递简历。

三、腾讯AI Lab计算机视觉招聘实习生

腾讯AI Lab计算机视觉中心招聘实习生，参与3D计算机视觉项目。腾讯AILab网址：<http://ai.tencent.com>

1、岗位要求：

- (1) 熟悉计算机视觉和深度学习的常见算法和工具(Pytorch/Tensorflow/Caffe)，具有较强的编程能力，能承担项目子模块的开发工作；
- (2) 有Structure from Motion, SLAM, Non-rigid Mesh/Point Registration, Geometric CNN等方面研究/项目经验；
- (3) 能够自主阅读理解主流视觉和图形学论文(CVPR, ICCV, ECCV, SIGGRAPH, SIGGRAPH ASIA)；
- (4) 工作地点为深圳，能够实习6个月或以上者优先。学历硕博皆可，根据个人背景确定有市场竞争力薪酬。表现优秀者，可获推荐正式员工职位。

联系方式

简历接收邮箱：ethanygao@tencent.com

四、世界 500 强企业——碧桂园集团招募优秀博士人才

公司拥有完善的培训机制，为博士制定完善的职业规划，预计将博士 1-3 年内培养为上市企业的优秀领导。详情查阅：<http://suo.im/2MT3kf>

1、岗位分类：

— “地产类”

院校专业：土木、建筑、结构、工程管理、岩土工程、城市规划、水电暖通、经济、土地资源管理、工商管理专业的博士生。

海外背景：一年及以上海外/港澳台的工作与学习经历(若国内博士，需兼具政府和房企背景)优先考虑有房地产工作经验者优先。

— “人工智能产业”

院校专业：包括但不限于以下理工科专业：数学、机械工程、自动化、信息管理与信息技术、信号与信息处理、自动化与模式识别、电子电气工程、控制科学与工程、机械工程、自动控制、光电子工程、机器人、材料科学、计算机科学与技术、仪器设计与技术、应用物理、通信工程等。

工作经历：不限。

— “其他”

财务、新业务、SSGF、产城、设计、营销、人力、教育等方向的博士生。

联系方式：

详情咨询：陈小姐：18520376205(微信同号)

简历投递：chenruilin@bgy.com.cn

五、华为中央媒体技术院招聘

华为中央媒体技术院是华为公司媒体领域核心算法技术能力中心，已构建遍布全球的研发分部，与全球顶级研究机构建立广泛联系。中央媒体技术院以世界领先的研究课题为牵引，以充足的研究资源为支撑，为优秀人才快速成长提供了最坚实的平台。

1、岗位名称：“计算机视觉/深度学习/机器学习算法工程师(社招或特博)”

(1) 计算机视觉前沿技术研发，承担人脸识别、图像识别、目标检测、图像分割、自然场景文字识别等算法研发，落地华为旗舰机，达到业界领先水平。

(2) 与全球 TOP 高校、研究机构建立合作关系，开展联合研究。

2、岗位要求：

(1) 社招：本科以上学历。

(2) 特博：2018-2019 届毕业博士。

(3) 计算机视觉、深度学习、机器学习、计算数学、应用数学、最优化方法等相关技术方向；熟悉主流机器学习、深度学习框架，如 Tensorflow、caffe、Pytorch 等；熟悉 C/C++/Python 编程。

3、工作地点：北京，深圳。

联系方式：

咨询和简历请联系：yunsn.li@huawei.com

六、美丽联合集团招聘

美丽联合集团是女性时尚媒体和时尚消费平台，旗下包括蘑菇街、美丽说等产品，拥有超 2 亿注册用户。通过整合现在已有的资源，包括电商、社区、红人、内容等，来服务于不同的用户，打造时尚目的地。

1、图像算法团队的工作

采用业务驱动模式开发，为公司业务提供图像技术的支持，包括几个方面的业务技术研发。

(1) 前台的用户体验。利用图像搜索技术，打造“以图搜图”的手机实拍购物体验，帮助用户发现和浏览更多的内容和商品。在搜索场景中，针对 Query 提供符合用户需求的召回商品；比如：用户搜索“红色的格子连衣裙”，对商品按照颜色、纹理、类目进行分析，将符合 query 意图的商品，排列出来给用户；

(2) 后台的图像管理。利用人脸识别、OCR 技术，分析站内商品，将违规的图片进行过滤，保证公司图片的内容安全。更进一步，利用图像识别和分类的技术，为电商服务提供支持；

(3) 深度学习的研究和优化，是底层一直在做的工作，目标是让模型更快、更小、更容易部署；

(4) 计算机图形学研究：进行图像合成的研究，希望做出更好的营销创意图(比如：频道 banner 图)；研究自动的字体生成技术。总体上通过这些方式，将图像技术运用到业务中，改善用户体验，提高商家能力，促进运营效率。

2、岗位名称

— “资深图像算法工程师：计算机视觉和机器学习”

职位描述：

负责图像处理、模式识别的算法核心技术研发工作。参与互联网应用的图像技术的产品设计和开发工作。

职位要求：

(1) 计算机视觉、模式识别等相关专业硕士及以上学历。具备良好的沟通表达能力和团队合作意识；

(2) 熟悉计算机视觉和图像处理基本算法(目标检测、图像分割、图像识别、图像分类等)，有相关项目经验或者较好的研究成果。掌握机器学习领域的基本算法，在深度学习领域有实际项目经验；

(3) 具备良好的编程能力，熟练掌握 C/C++、python 等编程语言，有 Linux 开发经验为佳。

工作地点：杭州。

— “资深图像算法工程师：计算机图形学”

职位描述：

(1) 负责图像处理、计算机图形学的算法核心技术研发工作；

(2) 参与互联网应用的图像/图形技术的产品设计和开发工作。

职位要求：

(1) 图像处理、计算机图形学等相关专业硕士及以上学历。具备良好的沟通表达能力和团队合作意识；

(2) 熟悉计算机图形学、图像处理(图像增强、图像融合、图像抠图等)、计算摄影学的基础算法。

有相关项目经验或者较好的研究成果；

(3) 加分项：(a) 有摄像和摄影的爱好或经验；(b) 绿幕合成技术研发经验；(c) 或者有深度学习的项目经验。至少有其中一项，都是加分项；

(4) 具备良好的编程能力，熟练掌握 C/C++、python 等编程语言，有 Linux 开发经验为佳。

工作地点：杭州。

联系方式：

简历接收邮箱：minda@meili-inc.com

七、京东集团 AI 平台与研究部招聘信息

京东集团坚持以技术驱动公司成长，致力于将人工智能技术与商业场景应用相结合，以不断实现业务升级和创新。京东集团围绕其在零售电商、金融、物流等领域的核心业务，基于京东海量精准丰富的大数据基础和非常明确的应用场景，研发自有技术平台，以云计算能力为基础，致力于机器学习、自然语言处理、虚拟现实、计算机视觉和语音识别等人工智能技术方面的研究，在包括智能消费、智能供应、智能物流、金融科技、实体零售科技在内的多元领域持续投入，运用技术驱动第四次零售革命，进一步推动电商和互联网行业的发展，服务消费者，赋能合作伙伴，为制造业、零售业、服务业全面提升效率，创造社会价值。未来十二年，京东集团将快速完成智能商业体的战略转型，成为全球领先的零售基础设施服务商。

AI 平台与研究部以创建京东 AI 平台为核心，未来将成为京东生态系统大中台里重要的组成部分。垂直方向上，平台将人工智能打造的各种应用场景对内外赋能，横向上，将引导京东 AI 在技术和应用场景中持续创新。

1、Java 开发工程师

岗位职责：

(1) 负责京东人工智能相关服务系统的开发；
(2) 独立进行需求分析、技术架构设计、独立完成编码及单元测试；

(3) 负责相关设计文档编写；

(4) 对产品技术架构提出优化方案并推动实施；

(5)解决系统中的关键问题和技术难题。

任职资格:

- (1)踏实诚恳、责任心强,能接受较大的工作强度,具备良好的沟通能力和学习能力;
- (2)本科及以上学历,计算机相关专业;
- (3)精通 Java 并有较全的知识面,熟悉 Spring 等主流框架,具有相关开发经验。熟悉 Linux 平台下常用命令操作、环境部署。熟悉数据库技术(MySQL),数据库优化及 SQL 优化;
- (4)思维灵活,热爱技术,对服务端编程开发有丰富的经验。

2、计算机视觉算法工程师

岗位描述:

- (1)负责人脸检测/识别,物体检测/识别等相关计算机视觉算法在创新应用和产品落地的技术需求;参与项目和产品诉求的相关算法分析、集成和优化;
- (2)评估和测试算法可行性、及监控算法效果,根据用户需求进行分析改进;
- (3)负责计算机视觉算法前沿技术跟创新使用场景相结合的研发和工程实现;
- (4)负责算法模块在产品化集成中的整体性能优化。

任职资格:

- (1)计算机科学、应用数学、统计学、通信与信息等相关专业硕士以上学历,有两年以上相关模型训练经验;
- (2)精通常用机器视觉处理相关算法,精通 OpenCV 图像处理;熟练掌握 CNN, RCNN, ResNet 等深度学习相关算法和模型,熟悉深度学习计算框架包括 Tensorflow, MXNET, Caffe 在内的其中一种;精通 python/C++其中一种编程语言,有 CUDA 编程经验者优先;
- (3)有团队精神,良好的表达能力和沟通能力。

3、语音信号处理工程师

岗位职责:

- (1)负责麦克风阵列前端信号处理算法的设计和实现;

(2)跟踪国内外麦克风阵列最新的信号处理相关算法。配合相关同事优化系统性能。

任职资格:

- (1)硕士及以上学历,声学或信号处理相关专业毕业,或从事过声学,信号处理等与阵列相关工作;
- (2)熟练掌握麦克风阵列信号处理算法,包括 AEC、DOA、AGC、去混响、多麦克风噪声抑制、BeamForming 等;
- (3)有嵌入式编程或者语音识别项目经验优先。

4、C++软件工程师

岗位职责:

- (1)参与 NeuHub 开放 AI 平台的全栈研发,旨在提升 AI 智能化算法开发迭代效率;集成并优化目前主流机器学习模型训练及预测框架;
- (2)根据业务需求对相关工具,算法进行稳定性和性能优化。

一任职资格

- (1)计算机科学、电子信息、自动化、通信工程等信息技术相关专业硕士以上学历;
- (2)熟悉 Linux 开发环境,有云服务化软件基础架构开发及服务器后端系统开发经验;
- (3)熟悉 TensorFlow、PyTorch 等机器学习框架,或 Kubernetes、Docker、Yarn 等微服务分布式系统;具有扎实的编程基础,熟练掌握 C++、Python 等一种以上编程语言;
- (5)良好的团队合作精神和沟通能力。
具有以下条件者优先
- (6)熟悉常见机器学习算法,有分布式机器学习平台项目经验;参与过大型服务化软件项目开发;各类程序设计竞赛优胜者。

联系方式:

简历接收邮箱: airecruiting@jd.com, 邮件标题: 职位+姓名

(责任编辑:樊鑫 蹇木伟)

征文通知

1 会议征文

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

2 期刊征文

计算机视觉领域近期相关期刊专刊的征文通知如表 2 所示，包括 International Journal of Computer Vision (IJCV)、Signal Image and Video Processing (SIVP) 和 Computers & Electrical Engineering (CEE)。

3 会议推荐

第 28 届国际人工智能联合会议 (IJCAI-19) 将于 2019 年 8 月 10 日至 16 日在中国澳门举行。会议由加拿大 IJCAI 委员会主办。IJCAI 会议是享誉世界的人工智能顶级会议，旨在介绍人工智能的尖端科学成果，并通过会议录的形式传播人工智能科学的最新信息。

第 20 届 IEEE 多媒体国际会议暨展览会 (ICME 2019) 将于 2019 年 8 月 10 日至 16 日在中国上海举行。会议旨在从电路系统、通信、计算机、信号处理等多个角度交流多媒体技术研究的最新进展。2019 年，多媒体产品展览会将与会议

同时举行。ICME 2019 旨在提供高质量的学术论文，会议论文集将被 IEEE Xplore Library 收录。

第 10 届国际图形与图像处理国际会议 (ICGIP 2018) 将于 2018 年 12 月在成都举行，由电子科技大学通信与信息工程学院与四川大学计算机学院联合主办，中国海洋大学、英国朴茨茅斯大学、中国仿真学会数字娱乐与仿真专委会为协办单位。大会将面向基础与前沿、学科与产业的紧迫问题，探究学术界和产业界面临的机遇与挑战。在建立高端学术交流的平台的基础上，共享研究成果，推动相关研究与应用的发展与进步，推进学科的发展和促进人才培养。

第 11 届数字图像处理国际会议 (ICDIP 2019) 将于 2019 年 5 月 10 日到 13 日在广州中山大学举行。本次大会由中山大学，华东师范大学和新加坡计算机科学与信息技术协会联合主办。大会旨在为数字图像处理等相关领域的全球学术界研究人员提供一个交流平台。本次会议论文将出版为会议论文集，并提交 EI 核心，Scopus 及 Web of Science (CPCI) 检索。

(责任编辑：贾同、金鑫)

表 1 计算机视觉领域相关国内外会议

会议名称	会议时间	会议地点	截稿日期	会议网站
IJCAI-19	2019.08.10-16	中国 澳门	2019.02.25	http://www.ijcail9.org
ICME 2019	2019.07.08-12	中国 上海	2018.12.03	http://www.icme2019.org/
ICGIP 2018	2018.12.12-14	中国 成都	2018.09.25	http://www.icgip.org/
ICDIP 2019	2019.05.10-13	中国 广州	2018.12.15	http://www.icdip.org/

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
IJCV	Generating Realistic Visual Data of Human Behavior	2018.09.30
SIVP	Advances in Multimodal/Multivariate Image Processing and Analysis	2018.10.28
CEE	Computational Methods in Neuroscience - Challenges and Future Trends	2018.11.25

图森未来

领跑自动驾驶未来

图森未来 (TuSimple) 是一家专注于自动驾驶货运卡车技术研发与应用的人工智能公司，成立于 2015 年 9 月，总部位于中国北京和美国圣迭戈，并在中美多地设有研发中心，目前员工总数约 400 人。图森未来现已完成 C 轮融资，估值 4.5 亿美元（逾 30 亿元人民币）。投资方包括新浪、英伟达、治平资本以及复合资本。



自成立以来，图森未来一直致力于打造针对高速场景和限定场景的 L4 级别 (SAE) 自动驾驶卡车。目前已基于“仓到仓”货运和港内集装箱码头转运两大场景，形成以摄像头为主要传感器，并融合激光雷达和毫米波雷达的 L4 级自动驾驶技术解决方案，实现了卡车在两大场景内的安全、高效行驶。

具体来说，图森未来解决了以下三大技术难题：(1) 远距离感知：重卡要实现平稳、安全的高速行驶，至少需要 300 米-500 米的感知距离。目前，图森未来已实现最远可达 1000 米的物体检测、分类、测距、行为预测，保证对前方 500

米范围内的全部路况进行完整、精确的分析，能为卡车做应急处理留出充足的时间，让货运更安全。

(2) 高精度定位：以港口码头这样的低速场景为例，重卡需要精准定位，定点停车，以便同港区的岸桥、轮胎吊高效配合。图森未来目前主要通过视觉高精度定位、多传感器融合技术和本地计算，实现了厘米级定位。

(3) 精准车身控制：图森未来对卡车底层的控制精度，超过了人类司机。在 100 公里/时的高速行进中，图森未来自动驾驶卡车偏离道路中线的平均横向误差是 3cm，而有几十年经验的人类卡车司机，平均误差在 10cm 左右。

目前，图森未来 L4 级自动驾驶卡车正在中美两地持续进行路测，已能够全天候无人接管安全运行，且已累计完成超过 13000 小时的真实环境路测。无人内集卡车队在中国国内港区完成了超 150 天的安全运营，无人为介入自动运载超过 10000 标准集装箱，具备了在白天、夜晚以及雨天等特殊工况下的商用能力。

图森未来研发团队总人数超过 320 人，汇集来自加州理工学院、香港科技大学、卡内基梅隆大学等世界顶尖名校人才以及来自丰田、通用等公司的产业界精英，覆盖了自动驾驶全研发链条。目前，图森未来正在中美两地招募大量人才。欢迎登陆 www.tusimple.com 了解图森未来，通过官网招聘页面或 campus.tusimple.com 投递简历。

(责任编辑：王瑞平)