

# CCF

# 计算机视觉 专委会简报

## COMPUTER VISION NEWSLETTER

2019/04 期  
总第 20 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](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 上海科技大学

#### 专委动态

主 编 毋立芳 北京工业大学  
编 委 马占宇 北京邮电大学  
王瑞平 中国科学院计算技术研究所  
虞晶怡 上海科技大学

#### 科技前沿

主 编 王金甲 燕山大学  
编 委 崔海楠 中国科学院自动化研究所  
邓 成 西安电子科技大学  
任传贤 中山大学  
任桐炜 南京大学  
苏 航 清华大学  
杨巨峰 南开大学

#### 委员风采

主 编 余 焯 合肥工业大学  
编 委 黄 岩 中国科学院自动化研究所  
刘海波 哈尔滨工程大学  
张汗灵 湖南大学

#### 资源平台

主 编 李 策 兰州理工大学  
编 委 樊 鑫 大连理工大学  
贾 同 东北大学  
蹇木伟 山东财经大学  
金 鑫 北京电子科技学院  
刘 丽 国防科学技术大学  
沈沛意 西安电子科技大学

# COMPUTER VISION NEWSLETTER

## CONTENTS

### 目录

#### 专委动态

走进高校	走进高校系列活动	04
视界无限	视界无限系列研讨会	06
专委新闻	CCF-CV 秘书处工作会议	08
专委通知	2019 年度 CCF-CV 奖评选	09
	2019 年度学术新锐奖评选	10
会议推介	PRCV 2019 活动征集	11
	PRCV 2019 博士生论坛征集	13
	PRCV 2019 技术挑战赛	15

#### 科技前沿

专题综述	三维人脸表情识别研究进展	18
热点追击	面向智能无人机的计算机视觉	24
	多媒体检索技术中的主动学习方法	25

#### 委员风采

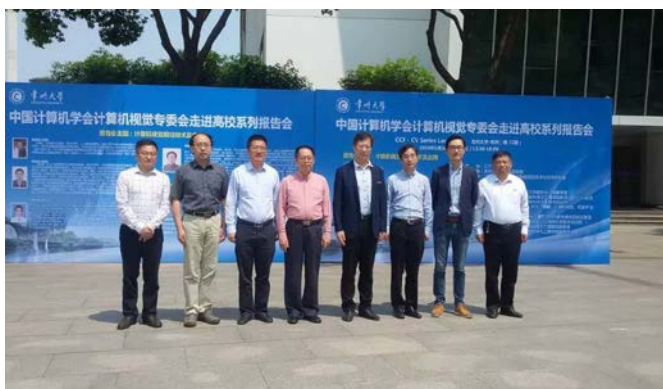
委员访谈	中国科学院大学赫然教授访谈	26
委员好消息		29

#### 资源平台

开源代码	图像描述	31
数据集	水下图像增强	35
招聘信息		38
征文通知		43

# CCF-CV 走进高校系列报告会

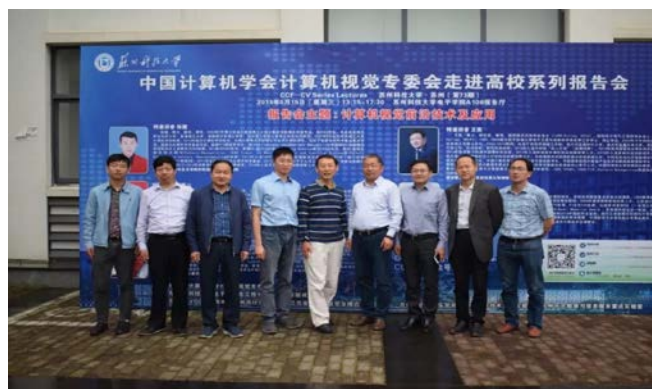
第 72 期 常州大学



2019年5月10日,在欢庆常州大学41周年校庆之际,由中国计算机学会计算机视觉专委会(CCF-CV)主办,常州大学信息科学与工程学院承办,江苏省人工智能学会智能感知技术与应用专委会、CCF常州、常州市计算机学会协办的第72期CCF-CV走进高校系列报告会活动——“计算机视觉前沿技术及应用”在常州大学报告厅举行。本期报告会邀请了南开大学方勇纯教授,中山大学赖剑煌教授,北京大学彭宇新教授,北京邮电大学马占宇副教授,东南大学耿新教授五位专家学者做特邀报告。

本次CCF-CV走进高校系列报告持续了四个多小时。五位专家详实地介绍了自己在计算机视觉领域地研究成果与心得,他们所展示地研究成果极大地激发了同学的学习热情。在报告会提问环节中,在座的校内外学生和老师们纷纷向五位专家提出自己科研方面的问题,专家们也对这些提问做出了详尽的回答,大家获益匪浅。最后,此次报告会在热烈的掌声中圆满结束。

第 73 期 苏州科技大学



2019年5月15日,由中国计算机视觉专委会主办、苏州科技大学电子与信息工程学院和苏州智慧城市研究院承办的第73期CCF-CV走进高校系列报告会活动——“计算机视觉前沿技术及应用”报告会在苏州科技大学石湖校区电子与信息工程学院报告厅A108成功举行。本期报告会邀请了南京理工大学杨健教授、中国科学院自动化所王亮研究员、南京信息工程大学刘青山教授、江南大学吴小俊教授、中山大学郑伟诗教授五位专家学者做特邀报告。苏州科技大学电子与信息工程学院副院长胡伏原教授、常熟理工学院计算机科学与工程学院院长龚声蓉教授以及苏州科技大学电子与信息工程学院夏振平博士为执行主席。

在报告会提问环节,在座的老师和同学们积极地提出自己在科学研究中所遇到的问题,专家们耐心并详细地作了回答,使大家受益匪浅。最后,此次报告会在久久不息的掌声中圆满结束。

## 第 74 期 河北大学



2019年6月10日,由中国计算机视觉专委会主办、河北大学电子与信息工程学院承办的第74期CCF-CV走进高校系列报告会活动——“计算机视觉前沿技术及应用”报告会在河北大学新图书馆多功能厅211报告厅成功举行。本期报告会邀请了西北工业大学王鹏博士,中科院自动化所李兵博士,上海交通大学马超博士,天津大学的任冬伟博士四位专家学者做特邀报告。河北大学电子与信息工程学院刘帅奇博士、河北大学电子信息工程学院院长赵杰博士为执行主席。来自河北大学、华北电力大学、河北农业大学、天津工业大学、河北软件职业技术学院等高校的100多名师生聆听了四位专家在该领域的前沿技术成果和最新动态介绍。

本次CCF-CV走进高校系列报告持续了四个多小时,报告内容丰富多彩,专家学者们的讲解精彩纷呈,提问环节互动热烈,所有参会人员享受了一场学术盛宴。四位专家精心准备的报告内容分别介绍了自己在计算机视觉领域的研究成果,同学们开阔了学术视野,拓展了学术思维。四位专家在提问环节对师生在科研中遇到的疑问做出了耐心详尽的解答,大家受益匪浅。最后,此次报告会在热烈的掌声中圆满结束。

## 第 75 期 天津工业大学



2019年6月28日,由中国计算机学会计算机视觉专委会主办、天津工业大学电气工程与自动化学院和人工智能学院共同承办的第75期CCF-CV走进高校系列报告会活动——“视觉信息智能识别前沿技术及应用”报告会在天津工业大学电气工程与自动化学院大会议室4d322成功举行。本期报告会邀请了南开大学方勇纯教授、中国科学院自动化研究所王亮研究员、武汉大学王密教授、南京理工大学左超教授四位专家学者做特邀报告。天津工业大学电气工程与自动化学院副院长宋丽梅教授、天津大学人工智能学院院长胡清华教授为执行主席。来自天津工业大学、天津大学、南开大学、天津科技大学、中国民航大学等高校的100多名师生聆听了四位专家在该领域的前沿技术成果和最新动态介绍。

此次CCF-CV走进高校系列报告将近四个小时,讲解深入浅出,内容精彩纷呈,让所有参会人员享受了一场学术盛宴。四位专家与与会的师生进行良好的互动,对师生在科学研究和日常学习中遇到的问题进行了耐心细致的回答,极大地激发了与会师生做好科学研究的热情。最后,此次报告会在久久不息的掌声中圆满结束。

(责任编辑:马占宇)

# CCF-CV 视界无限系列研讨会

第3期 中山大学



2019年7月6日,由中国计算机学会计算机视觉专委会主办、中山大学机器智能与先进计算教育部重点实验室承办的第3期 CCF-CV“视界无限”系列活动——“行人重识别(ReID)的前沿进展与未来趋势”研讨会在中山大学数据科学与计算机学院成功举行。本期研讨会由计算机视觉专委会委员、中山大学郑伟诗教授组织和主持。研讨会邀请了华为诺亚方舟实验室田奇教授、武汉大学胡瑞敏教授、清华大学鲁继文副教授、中国科学院大学马丙鹏副教授做主题报告,中科院自动化所杨阳助理研究员、商汤科技伊帅研究员、高新兴中央研究院毛亮副院长参加了深度研讨环节。来自全国高校和企业的100余名参会者汇聚羊城,就 ReID 的理论与应用进行了热烈的交流和深入细致的探讨。

会议议程依次包括引导发言、主题报告、海报/系统展示、深度研讨四部分。首先,计算机视觉专委会副主任、中山大学赖剑煌教授致辞。赖

老师对与会嘉宾和师生表示欢迎,然后分别介绍了专委会的组织结构、职能目标、举办的主要活动,以及中山大学承担的项目情况、科研平台、ReID 在中山大学的最新研究进展。最后,赖老师祝愿本次研讨会如雨后初晴的广州天气一样清新、热烈,希望通过深入研讨推动 ReID 同行之间的交流、促进合作、提升影响力,并祝愿研讨会圆满成功。



郑伟诗教授的引导发言题目是“弱监督行人重识别”。他指出,虽然行人重识别在过去几年有了跨越式发展,已经在许多标志性的数据库上取得非常高的识别结果,但目前这些算法大量依赖标注数据以训练复杂的学习模型。当前视频监控数据量巨大,且绝大部分缺乏标注;在很多新的场景下,非常缺乏新的标注数据。为了解决上述问题,郑老师介绍了近几年针对视频监控下的无标注、弱标注数据的代表性工作,包括基于非对称距离嵌入的深度聚类方法、深度多标签跨域匹配方法、基于弱视频标注的跨域适应方法、基于局部区域特征学习的无监督特征提取方法等。

田奇教授以“行人重识别：挑战与最新进展”为题，深入浅出地介绍了行人重识别的任务定义、主要难点和基准数据集。然后依次介绍了全监督、弱监督、无监督的行人重识别方法。最后，田老师从学术与工业双重视角出发，给出了行人重识别当前的开放问题以及下一阶段值得关注的重点。

在主题报告环节，胡瑞敏教授的报告题目是“面向社会安全的视频长程群体行人重识别与多元分析”。不同于主要面向个体查找的传统行人重识别应用，安全场景具有单次跨越多摄像机（长程）、多人合作作案（群体）特点，行人重识别技术面临新的挑战。胡老师团队基于街区视频和手机轨迹数据开展行人重识别多元分析技术研究。在人体 3D 精细模型表达与协同建模方法、个体/群体轨迹序列组合差异、简单社会活动条件下个体时间运动和粗粒度社会活动规律及身份辨识技术等方向开展研究。有效提升街区群体长程活动行人重识别的准确率，探索并发现了从行人、空间、时间、社会多元数据角度开展交叉研究的新途径。

马丙鹏副教授的报告题目是“视频行人重识别研究”。马老师指出，传统方法更多是针对基于单帧行人图像的行人重识别问题，但是通常单帧图像的信息是有限的。最近几年，随着行人数据的增加，基于视频片段的行人重识别也越来越受到重视。在此基础上，马老师介绍了基于行走周期的视频片段分割和行人特征提取，视频片段中神经网络模块的设计，视频片段中低质量行人图像的补全等最新技术。

鲁继文副教授的报告题目是“基于深度度量学习的行人重识别”，重点介绍了清华大学自动化系智能视觉实验室近年来提出的多个面向行人重识别的深度度量学习方法，包括深度迁移度

量学习、多视图深度度量学习、深度对抗度量学习等，以及它们在图像行人重识别、视频行人重识别、多模态行人重识别、跨场景行人重识别、摄像机网络行人重识别中的应用。

在海报/系统展示环节，共有来自企业和高校的 12 篇论文海报和 2 个系统进行了演示，为与会者提供了更多细致交流和深入探讨问题的机会。



在 Panel 环节，与会嘉宾与现场师生就“ReID 的理论价值与应用场景”、“ReID 的数据集现状与挑战”、“ReID 落地的难点”、“ReID 研究的东西方差异”等问题展开热烈讨论，参会者也就自己关心的问题与嘉宾交换了观点。



最后，第 3 期“视界无限”研讨会在热烈的掌声中圆满结束。

（责任编辑：王瑞平）

# CCF-CV 秘书处工作会议召开



2019年5月17日，中国计算机学会计算机视觉专委会（CCF-CV）秘书处2019年度工作会议于北京召开。专委秘书处全体成员参会，会议由秘书长王亮研究员主持。

本次会议主要讨论如何落实专委会常务委员会第四次会议的决议，并制定秘书处下一阶段的工作计划。过去的2018年，专委各项学术活动和组织建设稳步推进，走进高校、走进企业、视界无限、专委简报与网站的活动形式与内容不断推陈出新，继续展示CCF-CV专委的组织活力与特色引领性。新一年度，各项活动如何更有实效地开展，如何更加广泛地激发委员们参与专委建设的积极性，如何办好专委年度学术大会PRCV2019与学术论坛RACV2019，如何筹备年底专委换届工作，大家围绕这些议题展开了热烈讨论，并形成了具体可行的执行方案。

PRCV2019的会议筹备工作已经启动，秘书处将积极配合组委会需要，利用专委会网站、微信公众号、简报等不同方式跟进宣传；专委会品牌活动RACV2019论坛的举办目标与形式将进行改

革，定位为国内计算机视觉领域的小规模精品研讨会，拟定于暑期期间举行。聘请专业团队更新设计专委宣传彩页和视频宣传片，宣传片将增加配音解说与相关录像片段，继续在专委会今后举办的各类活动中采用，以扩大专委会的影响力。进一步充实和壮大专委会简报和网站的编辑队伍，积极调动新任委员的参与度，对有突出贡献的志愿服务人员进行表彰和奖励。走进高校活动将延续过去3年多70余期报告会的火爆，进一步优化活动流程，在活动申请、讲者征集、活动形式等方面将更为开放，为当地师生与讲者创造更多沟通机会，让计算机视觉前沿报告切实惠及更多地方院校。走进企业活动持续推进专委与业界龙头企业的产学研交流合作，在学术成果落地与企业技术需求之间搭建更为便捷的桥梁。2019年初启动的“视界无限”全新品牌活动已成功举办3期专题研讨会，后续将进一步围绕视觉研究前沿与热点问题提升活动质量与影响力，形成独具特色的品牌效应。



（责任编辑：王瑞平）

# 2019 年度 CCF-CV 奖评选

为发挥中国计算机学会计算机视觉专委会（简称 CCF-CV）的学术评价作用，激励、调动计算机视觉及相关领域专业工作者的积极性，促进计算机视觉及相关领域的创新和进步，自 2017 年起，CCF-CV 设立中国计算机学会计算机视觉专委会奖（简称 CCF-CV 奖），旨在奖励计算机视觉及其相关领域中，为领域发展做出突出贡献的老前辈、取得公认成就的科研人员、意义重大的科技成果、影响深远的学术成绩，以及为专委会工作提供卓越服务的优秀个人。

## 评选要求

### 一、候选人提名

CCF-CV 奖项实行推荐制，两名及以上 CCF-CV 专委会委员即可联合提名推荐。每个奖项的评选标准等事宜如下：

#### 1. CCF-CV 终身学术贡献奖

授予在计算机视觉及相关领域做出了突出贡献的中国计算机科技工作者（年龄在 70 岁以上）。每年评选一次，获奖人数不超过二名。获奖者可获得奖励证书、奖杯及奖金。获奖者到颁奖地领奖所需旅费由 CCF-CV 提供。

#### 2. CCF-CV 杰出成就奖

最高标准评选出计算机视觉领域取得了公认的杰出成就的计算机科技工作者，其获奖主要工作应在中国境内完成，国籍不限。每年评选一次，获奖人数不超过一名。获奖者可获得奖励证书、奖杯及奖金。获奖者到颁奖地领奖所需旅费由 CCF-CV 提供。

#### 3. CCF-CV 服务贡献奖

专委会活动、会议组织、秘书处工作、工委小组工作有重要贡献者。每年评选一次，获奖人数不超过三名。获奖者可获得奖牌、奖励证书和奖金。获奖者到颁奖地领奖所需旅费由 CCF-CV 提供。

## 推荐方法

由推荐人联合填写推荐表 1 份，将其电子版扫描件，于截止日期之前提交至电子邮箱 [ccfcvaward@sina.com](mailto:ccfcvaward@sina.com)，并抄送提名与奖励工作组组长查红彬 [zha@cis.pku.edu.cn](mailto:zha@cis.pku.edu.cn) 和秘书马占宇 [mazhanyu@bupt.edu.cn](mailto:mazhanyu@bupt.edu.cn)。

## 评奖流程

### 一、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，具体为：

工作组负责征集提名，提名材料在指定日期之前提交，由工作组对提名材料进行资格审查。工作组投票确定最终人选，名单提交常务委员会表决，过半数通过。

### 二、评奖时间

1. 报名时间：即日起-2019 年 8 月 31 日；
2. 初审阶段：2019 年 9 月 1 日-9 月 20 日；
3. 复审阶段：2019 年 10 月 31 日前。

三、其他未尽事宜详见《中国计算机学会计算机视觉专委会奖励条例》。

（责任编辑：马占宇）

# 2019 年度学术新锐奖评选

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。“学术新锐奖”于 2018 年首次开始评选，第一届获奖者为来自清华大学的董胤蓬、华中科技大学的廖明辉和中科院自动化所的张士峰。（2018 年评选具体情况请见下方链接）

## 冠名征集

为了提升“学术新锐奖”的影响力，推动企业与学术界和合作，吸引优秀青年人才参与活动，助力企业扩大在国内计算机视觉领域的影响力，现诚征“学术新锐奖”的冠单位名称。有意向者请与 CCF-CV 提名与奖励工作组秘书马占宇（邮箱：mazhanyu@bupt.edu.cn）联系。

## 参选要求

### 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）
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。

## 评奖流程

### 一、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。

## 二、评奖时间

1. 报名时间：2019 年 5 月 20 日-2019 年 8 月 31 日；
2. 初审阶段：2019 年 9 月 1 日-9 月 20 日；
3. 复审阶段：2019 年 10 月 31 日前。

三、其他未尽事宜详见中《CCF-CV 学术新锐奖评选条例》

(责任编辑：马占宇)

# PRCV2019 活动征集

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

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

## 诚邀申请讲习班

诚挚邀请广大模式识别与计算机视觉领域的从业者在 PRCV 2019 会议上组织与承办计算机视觉、模式识别与机器学习方面的各种讲习班活动。讲习班针对模式识别、计算机视觉与机器学习领域的某一特定方向 (举例：稀疏编码、行为识别、目标检测、图像细分类、跨媒体分析、生成对抗学习等)，由若干名讲者从不同维度、不同

视角或者不同侧面等对特定方向的理论、方法与应用进行详细讲解。

有意愿在 PRCV 2019 组织或承办讲习班活动者，请将讲习班申请书以电子邮件的形式提交给讲习班主席。申请结果将会通知组织者。所有被采纳的讲习班将会在 PRCV 2019 网站上公布。

讲习班申请书需包括以下信息：

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

重要日期：

- 申请截止：2019 年 8 月 31 日
- 申请结果通知：2019 年 9 月 15 日
- 完整材料提交：2019 年 10 月 15 日

讲习班主席：

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

## 诚邀申请专题论坛

面向广大科技工作者公开征集专题论坛的组织者和论坛题目,论坛范围包括(但不局限于):

1. 模式识别与计算机视觉前沿论坛
2. 顶会、顶刊交流论坛
3. 高水平学术论文写作交流论坛

论坛申请应包括:论坛题目,论坛简介、目的与意义,组织者个人简介,拟邀请的讲者题目与个人简介等。

重要日期:

- 申请截止:2019年8月31日
- 录用通知:2019年9月15日
- 完整信息提交:2019年10月15日

论坛单元主席:

- 刘青山,南京信息工程大学  
qslu@nuist.edu.cn
- 王蕴红,北京航空航天大学  
yhwang@buaa.edu.cn
- 卢湖川,大连理工大学  
lhchuan@dlut.edu.cn

## 诚邀申请成果展示

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

有意参与 PRCV2019 成果展示者需填写申请表(见会议主页 <http://www.prcv2019.com>)并发送邮件至 [xirunping@126.com](mailto:xirunping@126.com),提供展示内容名称、作者、单位、成果简介、图片示例等信息。申请展示的成果将由组委会组织专家进行评选,大会将为每个入选的展示成果提供展位一个,每个展位需要至少全注册两名参会人员(不得与参会论文作者注册共享)。

展示环境及要求:

- 展位(会务提供):标准展位(含一张展台、两把凳子和电源等)
- 展示设备(自备):笔记本、显示器及其它设备等
- 成果简介:制作展示期间用海报1页,推荐尺寸:宽80cm,高120cm。(参展单位也可以根据展位大小结合需求自备展架或自行布展)

重要日期:

- 申请截止:2019年9月15日
- 评选结果:2019年9月30日
- 会期展示:2019年11月8-11日

联系方式:

- 郗润平 西北工业大学  
[xirunping@126.com](mailto:xirunping@126.com)

会议邮箱:

- [prcv2019@nwpu.edu.cn](mailto:prcv2019@nwpu.edu.cn)

会议网址:

- <http://www.prcv2019.com>  
(责任编辑:毋立芳)

# PRCV2019 博士生论坛征集通知

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

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

## 论坛概况

本论坛为最近获得或即将获得博士学位的学生提供与计算机视觉和模式识别领域内经验丰富的研究者沟通交流的机会。每位学生都会被分配一位领域内的资深成员作为导师, 具体分配方案会基于学生的个人偏好或两者研究兴趣的相似程度而制定。投稿 PRCV2019 且论文被录用的学生可以优先选择导师, 获得跟导师面对面一对一交流的机会。导师团队包括但不限于如下著名专家学者:

谭铁牛 中科院自动化所研究员  
王 亮 中科院自动化所研究员  
陈熙霖 中科院计算所研究员  
赖剑煌 中山大学教授  
王 涛 爱奇艺公司首席科学家  
查红彬 北京大学教授

白 翔 华中科技大学教授  
纪荣嵘 厦门大学教授  
贾云得 北京理工大学教授  
林宙辰 北京大学教授  
刘青山 南京信息工程大学教授  
卢湖川 大连理工大学教授  
鲁继文 清华大学副教授  
苗启广 西安电子科技大学教授  
山世光 中科院计算所研究员  
王蕴红 北京航空航天大学教授  
虞晶怡 上海科技大学教授  
张艳宁 西北工业大学教授  
高新波 西安电子科技大学教授  
黄庆明 中国科学院大学教授  
赵 耀 北京交通大学教授  
薛建儒 西安交通大学教授  
杨小康 上海交通大学教授  
刘成林 中科院自动化所研究员  
王井东 微软亚洲研究院研究员  
姜育刚 复旦大学教授  
程明明 南开大学教授  
高盛华 上海科技大学副教授、研究员  
程 健 中科院自动化所研究员  
彭宇新 北京大学教授  
孟德宇 西安交通大学教授  
左旺孟 哈尔滨工业大学教授  
欧阳万里 悉尼大学副教授  
董伟生 西安电子科技大学教授  
方玉明 江西财经大学教授  
丛 杨 中科院沈阳自动化所研究员  
毋立芳 北京工业大学教授

韩军伟 西北工业大学教授

耿 新 东南大学教授

刘 偲 北京航空航天大学副教授

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

### 参会资格

参会学生投稿时须为在读博士研究生，且研究方向为计算机视觉或模式识别领域。

### 申请步骤

满足参会资格的学生需把申请材料发送至论坛主席邮箱，申请材料须为单个 pdf 文件或单个 zip/rar 压缩文件（内含多个 pdf 文件），材料中须包括内容如下：

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

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

### 审核流程

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

### 海报格式

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

### 重要时间节点

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

### 论坛主席

有任何问题，请联系以下论坛主席：

韩军伟, junweihan2010@gmail.com, 西北工业大学

耿 新, xgeng@seu.edu.cn, 东南大学

刘 偲, liusi@buaa.edu.cn, 北京航空航天大学

（责任编辑：毋立芳）

# PRCV2019 技术挑战赛

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

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

本次会议共设置 7 项竞赛简介如下 (总奖金 70 万元)。详细内容请看链接: <http://www.prcv2019.com/竞赛通知.html>。

## 1. 车纹识别技术挑战赛

车辆作为交通场景中的主要行为个体, 对其进行全方位感知和智能化分析是当前智慧交通发展战略的重要环节之一。从外观上来说, 车辆具有车型车款、车身颜色、车牌类型、车牌号码、天窗、年检标、挂件、摆件、喷涂以及驾乘人员等属性特征, 这些所有的属性信息组合起来形成了车辆独一无二的“车纹”, 可以作为车辆身份的唯一性判定依据, 比单纯的车牌号码更加可靠。车纹识别即从车辆图片外观中提取车辆的唯一性判定特征, 并从图片库中找到与查询车辆身份一致的车辆图片的过程。该技术在公安刑侦、违章监测、收费稽查等方面均有着重要的研究价值

和广泛的应用前景。该竞赛旨在提高车纹识别技术的研究水平, 推动计算机视觉学科与智慧交通产业发展, 促进人工智能技术在交通领域的应用。

### 奖项设置

- 一等奖 1 名 (奖金 10 万元人民币)
- 二等奖 1 名 (奖金 5 万元人民币)
- 三等奖 3 名 (奖金各 1 万元人民币)

报名截止日期: 2019 年 8 月 1 日

## 2. 大田作物病害图像挑战赛

为提高作物病害图像智能识别技术的研究水平, 推动计算机视觉学科与农业植保产业发展, 促进智能图像识别技术在农业病虫害防治领域的应用, 特举办此次挑战赛。比赛数据集全部数据来源于中国科学院合肥智能机械研究所智能认知研究组构建的农业病虫害研究图库 IDADP, 包含了水稻、小麦两种作物各 3 类、总共 6 类病害图像。

### 奖项设置

- 一等奖 1 名 (奖金 1.5 万元人民币)
- 二等奖 2 名 (奖金各 1 万元人民币)
- 三等奖 3 名 (奖金各 0.5 万元人民币)

报名截止日期: 2019 年 7 月 15 日

## 3. 短视频内容制作技术挑战赛

短视频内容智能制作技术挑战赛鼓励多媒体内容智能生产技术的研究, 推动计算机视觉学科与视频社交产业发展, 促进智能内容生成技术在短视频、直播领域的应用。引导视频内容智能生成领域的算法朝多元化、轻量化、模块化、可共享、可解释方向发展。

### 奖项设置

- 一等奖 1 名(奖金 5 万元人民币)
- 二等奖 1 名(奖金 2 万元人民币)
- 三等奖 3 名(奖金各 1 万元人民币)

报名截止日期: 2019 年 7 月 10 日

#### 4. 高速低功耗视觉理解挑战赛

当前, 计算机视觉技术已经取得了长足发展, 各类任务的性能指标不断被刷新, 其成果也被广泛应用到人类生活的各个领域。然而, 在实际部署时, 系统和产品对算法速度和功耗都有诸多要求, 随着人工智能应用规模的扩大, 这类问题日益突显。当前的研究主要关注视觉算法的性能提升, 对速度、功耗的考虑依然欠缺, 为解决实际问题, 算法功耗、速度、性能都是需要关注的重要因素。因此, 为了综合评估当前视觉领域的研究成果, 构建一个在功耗、速度、性能三个方面评测的通用基准, 我们发起了本次高速低功耗视觉理解挑战赛。本次竞赛旨在发掘高速低功耗的视觉理解技术, 基于 CPU、GPU 和 AI 芯片三个平台, 综合提高视觉理解技术的研究水平, 促进智能视觉理解技术的实际应用。

奖项设置

- 一等奖 1 名(奖金 4 万元人民币)
- 二等奖 1 名(奖金 2 万元人民币)
- 三等奖 1 名(奖金 1 万元人民币)

报名截止日期: 2019 年 7 月 31 日

#### 5. 面向自动阅卷的 OCR 技术挑战赛

近年来, 随着人工智能技术的飞速发展, 计算机视觉领域中目标检测、图像分类、物体识别、对象分割等诸多任务取得突破性进展, 并逐渐获得实际应用。其中, OCR, 即光学字符识别, 在理想条件下的文档与书籍电子化、车牌识别等方面已经非常成熟, 然而, 在自动阅卷这种相对开放的场景中, 虽然也已经有越来越多的研究, 但仍

然难以解决试卷场景中面临的版面复杂、中英文、数字、符号混杂、图像畸变等一系列挑战性难题。

针对 OCR 中当前面临的这些挑战, 好未来教育集团与中国科学院计算技术研究所拟在 PRCV2019 组织“面向自动阅卷场景的 OCR 技术挑战赛”(简称挑战赛), 涵盖试卷版式分析与文本行检测、文本识别两个代表性任务。通过这一挑战赛, 期望能促进业界对面向自动阅卷场景的 OCR 研究中相关重要问题和关键技术的深入理解以及相关方法与技术的发展与突破, 从而使得 OCR 技术能向自动阅卷等典型场景落地更进一步。

奖项设置

竞赛任务 1: 版面分析及文本行检测

- 一等奖 1 名(奖金 3 万元人民币)
- 二等奖 1 名(奖金 1 万元人民币)

竞赛任务 2: 文字识别

- 一等奖 1 名(奖金 3 万元人民币)
- 二等奖 1 名(奖金 1 万元人民币)

报名截止日期: 2019 年 7 月 15 日

#### 6. 同方威视箱包再识别技术挑战赛

在各类安检场景中, 经常看到旅客携带箱包排队等候对全部箱包进行检查, 严重影响出行体验。如果能够在某个位置通过各种技术手段事先锁定有嫌疑的箱包, 当旅客携带箱包通过另一位置时, 利用箱包外观进行识别后仅检查这些被锁定的箱包, 将能够在保证安全的前提下为旅客带来更美好的出行体验。我们把这个愿景中的研究课题称作箱包的再识别。

目标再识别(Re-Identification, ReID)技术是现在计算机视觉领域的热门研究方向, 主要应用于监控场景下行人的识别与检索问题, 若将该技术应用于安检领域实现箱包的再识别, 可有效提升安检效率, 降低漏检率, 进一步推动安检

产业发展。然而，相比于行人再识别，存在很多外观相似只有细节不同的箱包，在实际应用场景中箱包图像还易受到诸如遮挡、姿态、光照、背景、运动模糊等不同因素的影响。为了应对箱包再识别任务的上述挑战，本竞赛首次公开发布一个大规模的多视角箱包再识别数据集（Multi View Baggage, MVB），旨在探索目标再识别技术在安检领域的应用。参赛者可利用 MVB 数据集，研究开发高效的计算机视觉算法，实现对箱包的再识别。

#### 奖项设置

- 一等奖 1 名(奖金 5 万元人民币)
- 二等奖 2 名(奖金各 3 万元人民币)
- 三等奖 3 名(奖金各 1 万元人民币)

报名截止日期：2019 年 7 月 31 日

### 7. 鞋印图像数据年龄挖掘挑战赛

足迹是人们在行走、跑动、攀登、跳跃时在踩踏物体表面留下的痕迹，有裸足、穿袜、穿鞋之分。我国在足迹生物特征的分析 and 认定方面的

研究早在 19 世纪就已经开始，并形成了理论体系，一直处于世界领先地位。但足迹不同于人脸图像，人容易分辨，存在明显的人可认知的特征。基于解剖学的多个足部结构特征进行人身分析和认定，准确性不高，主观差异性大，且需要专家级的鉴定水平。

在此背景下，我们希望通过本次竞赛可以凝聚各方力量和资源，通过机器学习、深度学习等相关算法，对使用恒锐足迹样本采集仪获取的鞋印图像（即穿鞋足迹）数据进行挖掘分析，以获取足迹对应的样本年龄信息，实现鞋印图像数据挖掘年龄信息的目的。

#### 奖项设置

- 一等奖 1 名(奖金 4 万元人民币)
- 二等奖 1 名(奖金 2 万元人民币)
- 三等奖 2 名(奖金各 1 万元人民币)

报名截止日期：2019 年 8 月 12 日

（责任编辑：毋立芳）

# 三维人脸表情识别研究进展

北京航空航天大学 黄迪

## 一、引言

人脸表情是人类交流情感的主要方式，自动人脸表情识别一直以来都是模式识别和计算机视觉领域的重要研究问题。在过去十几年中，随着便携式智能设备（例如智能手机和智能手表）的普及，人脸表情识别相关的应用层出不穷，并在人机交互、医疗健康、游戏娱乐等方面体现出巨大潜力，也因此受到了工业界的广泛关注。

对人脸表情识别的研究可以追溯到二十世纪七十年代。Ekman[1, 2]等人定义了六种普适的基本表情（即愤怒、厌恶、恐惧、高兴、悲伤与惊讶）和数十种面部动作单元，并提出将表情表示为若干面部动作单元的组合进行分析。自此之后，表情识别的研究得到大幅发展。早期的人脸表情识别方法主要基于二维图像或视频，然而受限于固有成像的特点，这些方法的准确率和鲁棒性并不理想。进入本世纪后，三维数据采集设备不断革新，三维人脸表情识别研究开始受到重视。同时，BU-3DFE、Bosphorus、BU-4DFE 等数据集的发布实质性地推动了相关研究的开展。

表情本质上是由人类面部肌肉运动引发的形状变化。三维数据能够直接记录这些形变。相比于二维数据，三维数据对于环境光照和头部姿态变化更具鲁棒性。近年来，动态三维（四维）人脸表情识别开始进入视野，序列信息的利用使三维人脸表情识别的准确率得到了进一步提升。

本文主要介绍三维人脸表情（六种基本表情）识别的核心问题及一般流程，重点阐述基于静态和动态三维数据的人脸表情识别的最新进展，并对主要方法进行比较分析和定性评价，最后指出三维人脸表情识别未来的一些可能的研究方向。

## 二、三维人脸表情识别背景知识

### 1. 处理流程

与其他三维人脸分析任务类似，传统三维人

脸表情识别方法的一般处理流程包含四个关键技术环节（如图 1）：即数据获取（Data Acquisition）、数据预处理（Data Pre-processing）、形状表示（Shape Representation）和特征分类（Feature Classification）。

数据获取是通过特定的三维采集设备记录人脸的原始形状信息。数据预处理是对原始人脸模型进行毛刺移除和孔洞填充等操作，尽可能减少噪声的影响；此外还会根据需要进行标志点定位和人脸配准等步骤。形状表示是从人脸曲面上提取几何特征，其特征的判别能力直接影响识别精度。特征分类是基于曲面表示结合特定的分类器输出表情标签。

深度学习相关理论和技术最近被成功应用于三维人脸表情识别。其流程基本沿袭了传统方法的处理框架，将曲面表示和表情建模进行整合，以数据驱动的方式面向三维人脸表情识别任务建立更为有效的几何特征。

### 2. 核心问题

形状表示（几何特征提取）和表情建模是三维人脸表情识别的两个核心问题。形状表示目的是获得与表情相关的几何特征。这些特征不仅要不同表情有高区分性，还要对身份等其他属性变化有高鲁棒性。区分性和鲁棒性相互制约，使这一问题极具挑战。此外，三维人脸数据一般保存为点云或网格形式，与二维图像和视频不同，呈现分布无规则的特点，这为几何特征提取增加了额外难度。表情建模旨在建立表情类别预测模型。这是典型的机器学习问题，一般被建模为分类任务求解。而根据数据类型不同，该问题又可以进一步细分为面向静态特征的建模与面向动态特征的建模。前者偏重分析不同表情的全局或局部形状差异，而后者则偏重分析不同表情面部曲面形变的时空特性。

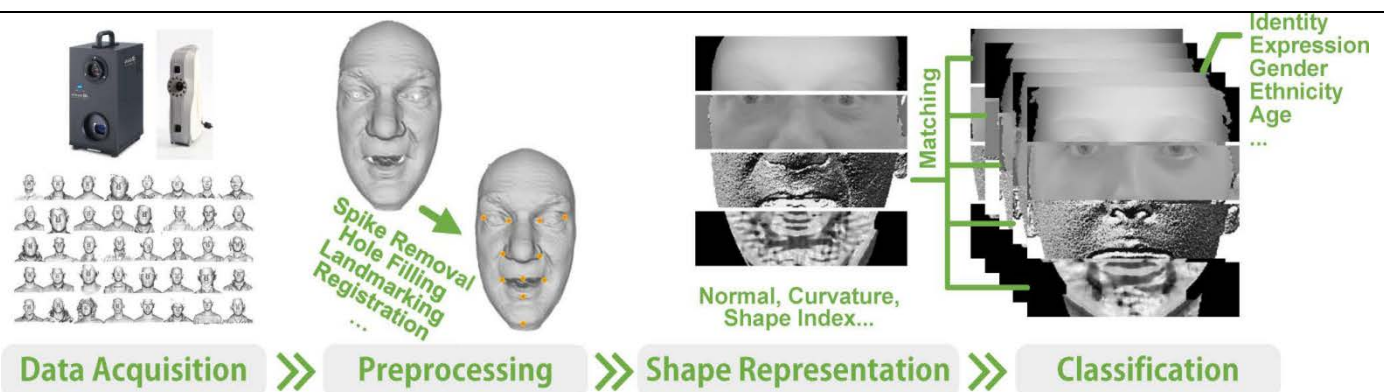


图 1 三维人脸表情识别的一般处理框架

基于深度学习的方法尝试突破传统手工设计特征的束缚，采用端到端的结构，利用随机梯度下降等优化技术，使模型同时改善特征提取和分类过程。初步的研究结果表明了这些方法的有效性，其中核心问题在于如何设计网络结构和损失函数，使模型具有更强的任务适应性和泛化性能。

### 三、静态三维人脸表情识别进展

在近十几年的研究中，静态三维人脸表情识别方法逐渐形成了两大类：基于模型的方法和基于特征的方法。

#### 1. 基于模型的方法

基于模型的方法建立通用人脸形状表示模型，并拟合给定三维人脸样本将得到的系数直接作为特征或进一步构建特征用于类别表情预测。

Mpiperis 等人[3]提出了一种基于双线性模型的表情识别方法。该方法首先基于标志点将三维人脸曲面重新网格化，使所有样本拓扑相同建立点对应关系，并进一步采用主成分分析（PCA）方法将身份和表情进行分解得到双线性模型。这样每个三维人脸模型都可以计算身份系数和表情系数，其中表情系数用来进行表情识别。

Gong 等人[4]提出了一种基于共性形状表示的表情识别方法。该方法建立统一的基础面部形状部件模型，将三维人脸曲面表示成基础面部形状部件和表情形状部件，并基于原始人脸曲面和基础面部形状部件，在眼睛和嘴巴等区域提取几何特征，之后通过计算这两部分特征的差值得到表情形状部件特征，并将其用于表情分类。

Zhao 等人[5]提出了一种基于贝叶斯置信网

络的表情识别方法。该方法基于吉布斯-玻耳兹曼分布在选定的十几个人脸标志点的一系列纹理和几何特征空间建立统计学表征模型。该模型可以在拟合过程中自动定位关键点并提取相应特征进行表情识别。

Zhen 等人[6]提出了一种基于肌肉运动模型的表情识别方法。该方法结合人脸面部肌肉的组成和运动特点，采用迭代最近法相量点方法对人脸曲面进行分割。在每一个肌肉区域提取多阶几何特征，并通过遗传算法选择最优的子集作为特征表示用于表情标签预测。

基于模型的方法理论体系完备，更易于人类理解，但它们对于表情相关的形状属性（尤其是对局部细节特点）刻画能力有限，因此并未体现出明显的优势。

#### 2. 基于特征的方法

基于特征的方法直接从人脸曲面（或相应的人脸深度图像）上提取点、线、面特征用于表情分类。

基于点特征的方法在以特定的标志点为中心的区块上计算局部形状特征，并将所有区块的特征融合形成最终的曲面表示作为表情识别的全局描述。Maalej 等人[9]在关键点周围通过同心测地圆环来构造区块，并将每个区块映射为高维流形上的样本点。他们通过计算样本点间最短的测地距离来衡量相应区块的差异，而后将人脸对应点附近的区块差异进行融合描述不同人脸形状的整体差异并用于分类。Derkach 和 Sukno[10]同样以测地同心圆环的方式提取人脸

关键点特征，由于每个区块具有完全相同的网格拓扑结构，该方法可方便地采用图拉普拉斯算子计算特征基向量，并将每个区块的三维坐标在特征基向量上进行分解，得到在该基向量下的谱域特征，最终用费舍尔线性判别（LDA）和支持向量机（SVM）进行分类。

基于线特征的方法计算人脸曲面上表情相关的距离，并通过特定点间距离的变化或特定线段夹角的变化等信息判别表情。Soyel 和 Demirel[7]手动在人脸上标注了十余个关键点，并根据这些点计算眼睛开合程度、眼眉高度、嘴巴左右开合程度、嘴巴上下开合程度、唇形拉伸程度及人脸左右边缘最外侧点间跨度等六种距离特征，并将这些距离输入人工神经网络进行表情分类。Tang 等人[8]采用类似方法，但大幅增加了特征点个数（68 点），并计算了多达 24 个距离特征和 24 个角度特征用于表情预测。

基于面特征的方法通常并非只关注人脸上某些局部区块，而是将三维人脸完整划分为若干子区域，并综合考虑每个子区域的特性。Wang 等人[11]在人脸的每个点计算主曲率方向及最大最小主曲率值，进而根据这些值逐点进行曲率类型划分。他们随后将人脸根据关键点划分成七个区域，并在每个区域计算曲率方向直方图及曲率类型直方图作为形状特征。最后采用线性判别分析、二次判别分类器、朴素贝叶斯分类器及径向基核支持向量机等四种方法进行了标签预测。Sha 等人[12]利用人脸上的几何线段将曲面分成不同区域，并计算各区域的曲率信息。他们提取曲率强度较为突出的点，并将其组成一个图，通过基于图的归一化裁剪过滤器对特征点进行选择，最后采用选择的特征进行判别。Lemaire 等人[13]采用变化邻域估计平均曲率，并将其映射在二维平面生成微分平均曲率图，而后提取方向梯度直方图（HOG）特征用于表情识别。Yang 等人[14]利用散射算子在人脸深度图、一阶微分法向量图及二阶微分形状索引（Shape Index）图上计算散射特征提取更为细节的几何变化，最后基于支持向量机进行分类。

基于特征的方法处理流程直观，其精度取决于人脸曲面几何表征的鉴别性，尤其是基于深度学习的方法近期体现出了巨大潜力。例如，Li 等人[15]通过预训练卷积神经网络（VGG 模型）在[14]中的多种几何图上提取深度特征，并采用支持向量机进行分类。随后，他们设计了深度融合卷积神经网络，实现了端到端训练，进一步提升了表情识别精度。

#### 四、 动态三维人脸表情识别进展

动态三维人脸表情识别除静态身份无关特征提取外，更强调视频序列中时序特性的建模。

在首个动态三维人脸数据集 BU-4DFE 发布时，Sun 和 Yin[17]提出了一个尝试性的动态三维人脸表情识别方法。该方法根据三维人脸曲率类别，将人脸曲面划分为六个区域，并采用线性判别分析对区域内特征进行降维。之后，该方法基于二维隐马尔科夫模型（HMM）在特征空间对时序变化进行建模，将空间区域作为第一维状态，将时序信息作为第二维状态。该方法针对每一种表情训练一个模型，并采用似然概率下的贝叶斯决策规则进行判别。

Le 等人[18]将视频各三维人脸曲面帧以鼻尖点为中心划分为一系列同心圆环，并将与参考人脸对应圆环间的倒角距离（Chamfer Distance）作为特征，经主成分分析和线性判别分析降维后训练隐马尔科夫模型建立表情识别分类器。

Sandbach 等人[19]将三维人脸帧序列在六个时空方向（ $x-y$ 、 $x-z$ 、 $y-z$ 、 $x-t$ 、 $y-t$  及  $z-t$ ）上进行投影并计算其二叉树分解模式作为表情特征，之后采用隐马尔科夫模型进行识别。

Fang 等人[20]使用基于网格数据的方向梯度直方图（MeshHOG）方法进行曲面配准，并采用三个正交平面上的局部二值模式（LBP-TOP）特征对三维人脸序列的形状变化进行表示，最后采用支持向量机进行判别。

Hayat 等人[21]将不同位置可变长度的三维人脸序列区块映射为格拉斯曼（Grassmannian）流形上的点，并利用局部二值模式（LBP）等特征空间的谱聚类对不同表情的流形点分别计算中

心。通过匹配测试三维人脸序列的流形点与上述聚类中心得到表情标签。

Danelakis 等人[22]基于关键点构造了三种特征,包括热核特征、法向量及角度、面积和距离等几何特征,并采用动态时间对齐算法度量两个时间序列的相似性,将不同特征的加权结果生成评分,最后通过K近邻(KNN)方法进行预测。

Drira 等人[23]和 Ben Amor 等人[24]将三维人脸曲面沿以鼻尖点为中心的放射曲线重新网格化,计算帧间稠密标量场(Dense Scalar Field)作为特征,经线性判别分析降维后采用随机森林和隐马尔可夫模型识别表情。Zhen 等人[25]采用同样的特征表示和相似的分类器,提出了空间域高斯金字塔分解方法,并在时域上对形变细微的表情进行增强。此外,他们还提出了基于聚类的表情周期检测算法,增强了方法适用性。

Li 等人[26]率先提出了一种基于深度学习技术的动态三维人脸表情识别方法,采用动态几何图网络捕获三维人脸曲面序列的时序形变信息和由粗到精的二阶段划窗采样技术进行数据扩增。通过将多几何特征图的预测结果进行加权融合,提升了识别精度。

## 五、未来可能的研究方向

基于静态数据和动态数据的三维人脸表情识别在过去的十几年间都取得了长足的进步,但算法的准确性和实用性仍需进一步提升。

在**准确率**方面,关键挑战在于部分表情难以区分。具体原因一方面体现在一些表情(如愤怒和悲伤)驱动面部肌肉的模式有相似性,导致几何特征易混淆;另一方面体现在一些表情程度较弱,面部形变轻微,导致几何特征不明显。

在**实用性**方面,关键挑战在于应对姿态和遮挡等真实环境的变化。其中,如何从连续三维人脸视频流中检测到表情信号并正确识别成为必

须。此外,如何提升对于头部姿态变化和外部遮挡变化的鲁棒性愈发重要。

针对这些挑战,可以考虑的方案如下:

1. **通过深度学习技术提升人脸曲面(序列)的表征能力**。目前一些工作[14-16]尝试利用深度学习技术提升三维人脸表情识别的精度,并初步展示出了不错的前景。但是这些方法都是将三维人脸的深度图作为输入,本质上是复制二维数据的处理方法,因此(部分)丢掉了三维的优势。

目前,已经有一些研究者尝试提出面向三维点云和网格数据的深度学习方法[27-29],也取得了一些进展。在三维人脸表情识别问题上,Chen 等人[30]新近提出了基于快速轻量化流形卷积神经网络的表情识别方法,不但取得了非常有竞争力的结果,还体现出了对于姿态变化的良好稳定性。然而由于三维数据点分布的不规则性,对三维人脸曲面的深度表征还存在很多困难,是需要重点探索的方向。

2. **通过融合多模态数据引入辅助信息提升识别准确率**。最直接和常用的多模态融合方式就是整合二维与三维人脸数据,这也是三维人脸表情识别的一个分支。Oyedotun 等人[16]最新提出的方法针对形状信息和纹理信息,分别设计了不同的深度神经网络进行学习,之后采用全连接的方式将两种模态的深度特征进行融合,其表情识别精度明显优于只基于单一模态的方法。

另外,表情的产生常会伴随着语音、手势等信息,还会引起人体一些生理指标(如血压、心跳)的变化。一些研究表明,通过采集更多模态的数据、提取更为丰富的信息可以进一步提升表情识别的准确率。所以,如何将不同模态数据进行融合也是一个很有必要的研究方向。

(责任编辑:任传贤)

## 参考文献

- [1] Ekman P and Oster H. Facial Expressions of Emotion. Annual Review of Psychology, 1979(30):527-554
- [2] Ekman P. Strong Evidence for Universals in Facial Expressions: A Reply to Russell's Mistaken Critique. Psychological Bulletin,

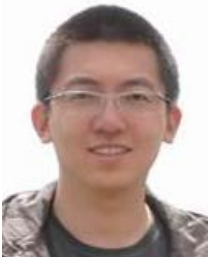
1994, 115(2):268–287

- [3] Mpiperis I, Malassiotis S, Srinivasan M G. Bilinear Models for 3-D Face and Facial Expression Recognition. *IEEE Transactions on Information Forensics and Security*, 2008, 3(3): 498-511
- [4] Gong B, Wang Y, Liu J, et al. Automatic Facial Expression Recognition on a Single 3D Face by Exploring Shape Deformation. *ACM International Conference on Multimedia*. ACM, 2009: 569-572
- [5] Zhao X, Huang D, Dellandréa E, et al. Automatic 3D Facial Expression Recognition based on a Bayesian Belief Net and a Statistical Facial Feature Model. *IEEE International Conference on Pattern Recognition*. IEEE, 2010: 3724-3727
- [6] Zhen Q, Huang D, Wang Y, et al. Muscular Movement Model-based Automatic 3D/4D Facial Expression Recognition. *IEEE Transactions on Multimedia*, 2016, 18(7): 1438-1450
- [7] Soyel H, Demirel H. Facial Expression Recognition using 3D Facial Feature Distances. *International Conference Image Analysis and Recognition*. Springer, Berlin, Heidelberg, 2007: 831-838
- [8] Tang H, Huang T S. 3D Facial Expression Recognition based on Automatically Selected Features. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. IEEE, 2008: 1-8
- [9] Maalej A, Amor B B, Daoudi M, et al. Shape Analysis of Local Facial Patches for 3D Facial Expression Recognition. *Pattern Recognition*, 2011, 44(8): 1581-1589
- [10] Derkach D, Sukno F M. Local Shape Spectrum Analysis for 3D Facial Expression Recognition. *IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, 2017: 41-47
- [11] Wang J, Yin L, Wei X, et al. 3D Facial Expression Recognition based on Primitive Surface Feature Distribution. *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2006, 2: 1399-1406
- [12] Sha T, Song M, Bu J, et al. Feature Level Analysis for 3D Facial Expression Recognition. *Neurocomputing*, 2011, 74(12-13): 2135-2141
- [13] Lemaire P, Ardabilian M, Chen L, et al. Fully Automatic 3D Facial Expression Recognition using Differential Mean Curvature Maps and Histograms of Oriented Gradients. *IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, 2013: 1-7
- [14] Yang X, Huang D, Wang Y, et al. Automatic 3D Facial Expression Recognition using Geometric Scattering Representation. *IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, 2015, 1: 1-6
- [15] Li H, Sun J, Xu Z, et al. Multimodal 2d+ 3d Facial Expression Recognition with Deep Fusion Convolutional Neural Network. *IEEE Transactions on Multimedia*, 2017, 19(12): 2816-2831
- [16] Oyedotun O K, Demisse G G, Shabayek A E R, et al. Facial Expression Recognition via Joint Deep Learning of RGB-Depth Map Latent Representations. *IEEE International Conference on Computer Vision Workshops*. IEEE, 2017: 3161-3168
- [17] Sun Y, Yin L. Facial Expression Recognition based on 3D Dynamic Range Model Sequences. *European Conference on Computer Vision*. Springer, Berlin, Heidelberg, 2008: 58-71
- [18] Le V, Tang H, Huang T S. Expression Recognition from 3D Dynamic Faces using Robust Spatio-Temporal Shape Features. *IEEE International Conference on Automatic Face and Gesture Recognition*. IEEE, 2011: 414-421
- [19] Sandbach G, Zafeiriou S, Pantic M, et al. Recognition of 3D Facial Expression Dynamics. *Image and Vision Computing*, 2012, 30(10): 762-773
- [20] Fang T, Zhao X, Shah S K, et al. 4D Facial Expression Recognition. *IEEE International Conference on Computer Vision Workshops*. IEEE, 2011: 1594-1601
- [21] Hayat M, Bennamoun M, El-Sallam A A. Clustering of Video-Patches on Grassmannian Manifold for Facial Expression Recognition from 3D Videos. *IEEE Workshop on Applications of Computer Vision*. IEEE, 2013: 83-88
- [22] Danelakis A, Theoharis T, Pratikakis I, et al. An Effective Methodology for Dynamic 3D Facial Expression Retrieval. *Pattern Recognition*, 2016, 52: 174-185
- [23] Drira H, Amor B B, Daoudi M, et al. 3D Dynamic Expression Recognition based on a Novel Deformation Vector Field and Random Forest. *IEEE International Conference on Pattern Recognition*. IEEE, 2012: 1104-1107
- [24] Amor B B, Drira H, Berretti S, et al. 4D Facial Expression Recognition by Learning Geometric Deformations. *IEEE Transaction*

---

on Cybernetics, 2014, 44(12): 2443-2457

- [25] Zhen Q, Huang D, Dira H, et al. Magnifying Subtle Facial Motions for Effective 4D Expression Recognition. IEEE Transactions on Affective Computing, 2017
  - [26] Li W, Huang D, Li H, et al. Automatic 4D Facial Expression Recognition using Dynamic Geometrical Image Network. IEEE International Conference on Automatic Face and Gesture Recognition. IEEE, 2018: 24-30
  - [27] Riegler G, Ulusoy A O, Geiger A. Octnet: Learning Deep 3D Representations at High Resolutions. IEEE Conference on Computer Vision and Pattern Recognition. 2017, 3
  - [28] Qi C R, Yi L, Su H, et al. PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. Conference on Neural Information Processing Systems. 2017: 5099-5108
  - [29] Monti F, Boscaini D, Masci J, et al. Geometric Deep Learning on Graphs and Manifolds using Mixture Model CNNs. IEEE Conference on Computer Vision and Pattern Recognition. 2017, 1(2): 3
  - [30] Chen Z, Huang D, Wang Y, et al. Fast and Light Manifold CNN based 3D Facial Expression Recognition across Pose Variations. ACM International Conference on Multimedia. ACM, 2018: 229-238
- 



黄迪

博士、北京航空航天大学计算机学院院长聘副教授、博士生导师。主要研究方向涉及生物特征识别，智能图像/视频处理及情感计算等。Email: dhuang@buaa.edu.cn

# 面向智能无人机的计算机视觉

天津大学 朱鹏飞 胡清华 京东数科 文珑银 纽约州立大学奥尔巴尼分校 独大为

无人机云台由于搭载了不同类型的传感器，可实现对周围环境的大视角多维度感知。天津大学机器学习与数据挖掘团队依托天津市机器学习重点实验室建立了 VisDrone 无人机数据平台，覆盖了 120 米以下的无人机目标检测、目标跟踪、目标计数等任务，可见光和热红外等模态，单机和多机等配置。目前目标检测和跟踪数据采集遍布中国 14 个城市，共包含 189473 张视频帧/图像，约 250 万个标注框。图 1 为 VisDrone 检测和跟踪数据集的样例。研究团队在 ECCV2018 举办了第一届 VisDrone 无人机视觉数据竞赛，来自全球百余家大学、科研结构和企业参加了此次竞赛，目前正在 ICCV2019 举办第二届比赛。VisDrone 数据平台未来将不断拓展，服务于计算机视觉等研究领域。

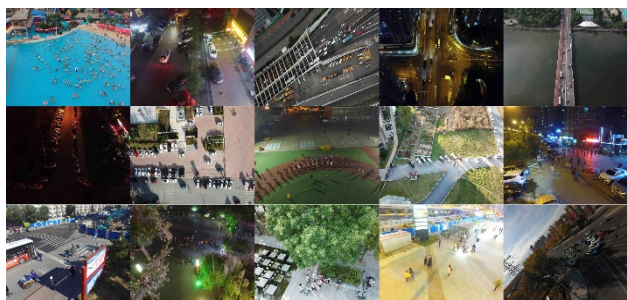


图 1 VisDrone 数据集样例

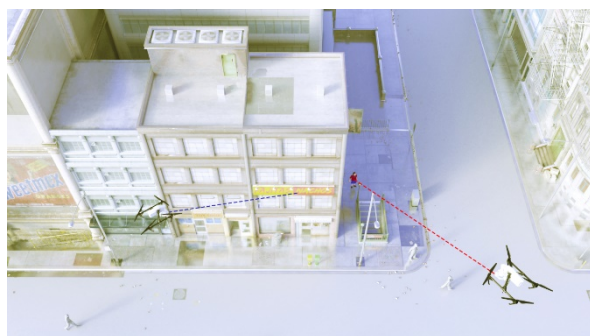


图 2 多无人机协同感知

未来的智能无人机面对复杂环境的挑战，同时环境中同时存在多源异构的多类型智能体，如何实现无人机的自主感知、学习和决策，并联合其他智能体协同学习，是值得期待的重要研究方向。图 2 展示了多无人机协同跟踪的示例。同时 5G 时代的到来使得无人机在通信和数据传输方面形成新的突破，如何开展 5G 时代无人机端的边缘计算以及云端边协同等研究也是无人机视觉研究可以关注的方向。此外无人机视觉的研究可考虑控制学、多智能体学习等结合，在交叉学科领域寻求新的突破。

(责任编辑：杨巨峰)



**朱鹏飞**

天津大学副教授，主要研究方向是无人机视觉、机器学习。

Email: zhupengfei@tju.edu.cn



**胡清华**

天津大学教授，主要研究方向是机器学习。

Email: huqinghua@tju.edu.cn



**文珑银**

京东数科美国研究院高级研究员，主要研究方向是计算机视觉和机器学习。

Email: longyin.wen@jd.com



**独大为**

纽约州立大学奥尔巴尼分校研究助理，主要研究方向是计算机视觉。

Email: ddu@albany.edu

# 多媒体检索技术中的主动学习方法

新加坡资讯通信研究院 张立宁

近年来，随着 Weibo、Facebook 等社交媒体的流行，图像、视频等异构数据的数量每天都在以惊人的速度增长。当人们面对一个超大型多媒体信息库时，单单凭借关键字很难对多媒体数据库进行有效的描述和检索。如何有效地帮助人们快速而准确地找到所需信息，成为多媒体信息领域需要解决的核心问题。多媒体信息检索可以分析图像内容，提取图像颜色、形状、纹理以及图像空间关系等信息，并对图像数据库建立索引。

目前各大主流网站有一些针对具体场景的应用，但依然存在算法处理速度慢、漏检率高、检索效果差等问题。多媒体检索技术的核心是系统既可以从图像自身提取图像特征，又可以通过与用户交互获得更多的信息，并将其用于图像间相似性的计算。系统和用户之间的关系是双向的：用户可以向系统提出查询要求，系统针对用户的查询要求返回相应的查询结果；系统可以通过用户查询结果的相关反馈来改进系统检索的性能。

我们提出了一种新颖的主动学习方法用于多媒体检索系统中选择最具信息量的样本数据由用户标记反馈，称为流形正则实验设计方法。

传统的主动学习方法，目前最大的挑战是如何有效地度量无标签样本的信息量。支撑向量机主动学习借助自身最优超平面来选择最不确定的样本由用户去标记。而要获得支撑向量机的最优超平面，系统往往需要用户大量的反馈信息，

通常用户会厌倦和疲劳，因此系统的检索准确度很难提高。

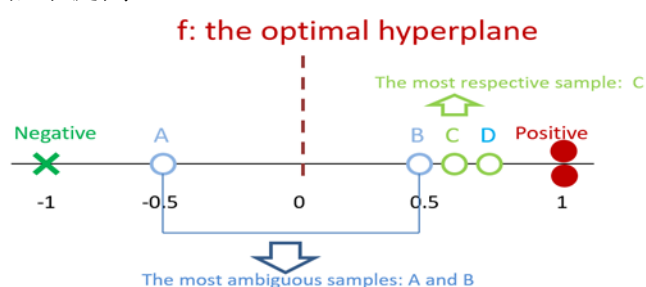
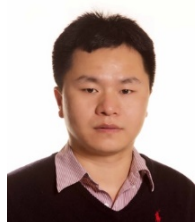


图 1 多长度哈希联合学习框架图

如图 1，新方法可以使检索系统选择多媒体数据库中最具代表性的样本集 B 和 C，而不是选择最不确定的样本集 A 或 B。我们的方法通过极小化目标函数在测试样本上的期望平均预测方差，可以使得被选择的样本具有清楚的几何空间解释，即可以选择数据库中最具代表性的样本，同时还可以全局最优地迭代选择最具信息量的样本进行给用户标记。因此可以极大地减少用户的标记工作，同时还可以保持系统更高的检索精确度。提出方法并不依赖于任何预选取的样本标签信息，可以有效地避免传统支撑向量机主动学习方法中由不充分、不准确样本标签数据引起的潜在问题。在大量真实数据库上的实验结果验证了流形正则实验设计方法在多媒体检索应用中的有效性。以上工作发表于 TIP 2017。

(责任编辑：邓成)



张立宁

新加坡资讯通信研究院研究员。主要研究方向为机器学习、计算机视觉、多媒体计算、新材料器件设计，电路设计等人工智能应用领域。

Email: zhangl1@i2r.a-star.edu.sg

# 中国科学院大学赫然教授访谈

2019年6月15日，专委秘书处采访了中国科学院大学国家优秀青年科学基金获得者赫然教授。下面是采访实录。

赫教授，您在图像模式识别理论研究方面取得了一些成果，获得了国家优秀青年科学基金、北京市首届杰出青年科学基金等多项人才项目资助。特别在损失函数（或最优值函数）等价优化基础上的稀疏学习、信息理论学习和生成对抗学习三个方面做出很多优秀工作。您能否向我们介绍一下您如何选择现在的研究方向？以及怎么沿着当前的研究方向慢慢做深的？

模式识别是一个智能任务，是人工智能的重要组成部分。正是在人脸识别和物体分类等经典模式识别问题的深度突破和广泛应用推动了人工智能的快速发展。但另一方面，模式识别基础理论的研究依旧薄弱。初次接触科研，我的导师推荐给我“先见森林，后见树木”的思想。在接触信息理论学习和稀疏学习的初期，我花了半年左右时间仔细阅读了信息理论的经典教材，并且旁听了中国科学院数学所开设的数学课程。这些学习使我对相关研究领域有了较全面的了解，并打下一定的数学基础。因此，在博士期间以及职业生涯刚刚开始时期，我主要研究了一簇常用的最优值函数。从共轭函数角度分析了他们之间的区别和联系，提出了它们的半二次展开方式和优化方法，进而建立了信息理论学习和稀疏学习的统一半二次框架，为一些图像模式识别问题的研究提供了不同的解决思路，并取得了一些成果。

随着存储能力和计算能力不断提升，如何使现有理论和深度学习技术相结合，以更低的成本创造出更高的价值，是模式识别领域亟待解决的挑战之一。2014年 Goodfellow 等提出了生成对抗网络，它可以通过对抗博弈的策略从低维白噪

声生成高维图像。这促进了生成式机器学习模型的飞速发展，为图像模式分析的研究带来了新的思路。图像生成，一方面可以给人类视觉带来更美好、随心的体验，另一方面可以推动模式分类模型在小样本学习上的发展。人脸图像生成技术是图像生成的一个重要分支，相比于自然场景图像，人脸具有更稳定的拓扑结构，更便于在深度学习技术上验证；恰巧我们团队在人脸方向已经有了一些探索，生成模型又与信息理论学习密不可分，之前研究的最优值函数也可以广泛应用于深度学习。综合以上两方面原因，我们选择了人脸图像生成（合成）作为课题组的一个重要研究方向，并在国内较早地实现百万像素（1024分辨率）的人脸图像无条件生成和人脸图像属性变换。由于大规模人脸图像生成的理论和方法研究涉及到很多研究领域的多方面内容，并且人脸图像生成又有着众多具体应用场景，因此我们本着“目标有限、重点突出”的原则，结合华为、百度、地平线等公司的应用需求，从“生成对抗网络理论”以及“高维图像生成方法”两个方面深入研究人脸图像生成。

您在国际顶级期刊和会议上发表了学术论文140篇，其中第一作者 IEEE TPAMI 论文4篇、IEEE TIP 论文3篇、IEEE TNNLS 论文2篇、IEEE CVPR 论文3篇并被多次引用，在您取得的这些成果中，有哪些工作让您印象特别深刻？

做研究一定要有严谨的科研态度与习惯，它不是与生俱来的，或者一朝一夕可以养成的，需要经历长期培养和自我规范。对于学术论文的写作，要不断从失败中汲取教训，从而实现锻炼和提升。例如，初次投稿计算机视觉及模式识别领域顶尖期刊 IEEE TPAMI 时，我经历了四次论文小修（minor），历时1年半才最终被接收。这真

的是一次十分磨人耐性的修改经历，但同时更加提醒了我“做科研要非常严谨，语言表达要准确”。自那以后，这句话也常被挂在嘴边，用来督促自己和学生。

科研项目的申请对于广大科研工作者，尤其是刚刚步入科研行业的青年科技工作者，其重要性不言而喻。您获得了很多国家级和省部级项目，科研项目的选题和撰写应该注意哪些问题？能否分享一下您申请科研项目的成功经验？

科研项目的选题首先要结合自身的研究方向和研究兴趣，切不可做“墙头草”人云亦云；其次，要淡泊名利，以解决国家和社会面临的重要问题为目标，敢于进行学术界的前沿性探索。此外，项目申请书的结构与写作也是十分重要的，尽量做到结构清晰、图文并茂、高度凝练，并听取不同领域专家的意见，反复修改校正。每一次写完项目申请书，我们会请学生仔细校对，一方面检查申请书是否还存在文字错误，另一方面也让学生了解未来的工作重点，共同探讨更好的解决方案和技术路线。

您近年多次被受邀到一些国际学术机构进行学术访问和交流，并作相关主题报告，如美国马里兰大学（计算机视觉权威学者、IEEE Fellow、多届 ICCV 和 CVPR 大会主席 Larry Davis 教授），美国普林斯顿大学（信息理论权威学者、IEEE TIT 前主编、美国三院院士 Vincent Poor 教授）等。请您分享一下如何做好学术报告，特别在这个方面是如何训练您的学生的？

我们主要以谭铁牛院士的 PE3C 模型（Passionate, Enthusiastic, Confident, Considerate, Clear）为目标来提高自己的学术报告水平。为了提升学生的演讲能力，我们主要通过以下几个方面进行培养：鼓励学生在日常组会报告中使用英文演讲，并对其演讲效果从内容、结构、表达、时长、声音、热情和互动等方面进行打分；鼓励学生参加国际会议，并和国际知名学者面对面交

流；鼓励学生参加国内的 PRCV 会议和 VALSE 论坛并了解模式识别领域的前沿动态；鼓励学生积极参与 CCF-CV 专委举办的 ICCV、CVPR、ECCV 中国预会议并介绍学生自己的研究工作，并根据专家的反馈意见进一步提高报告水平。另外，在谭老师的倡导下，我们研究中心每半年举行一次内部研讨会，为每位同学提供总结、展示、交流的舞台。

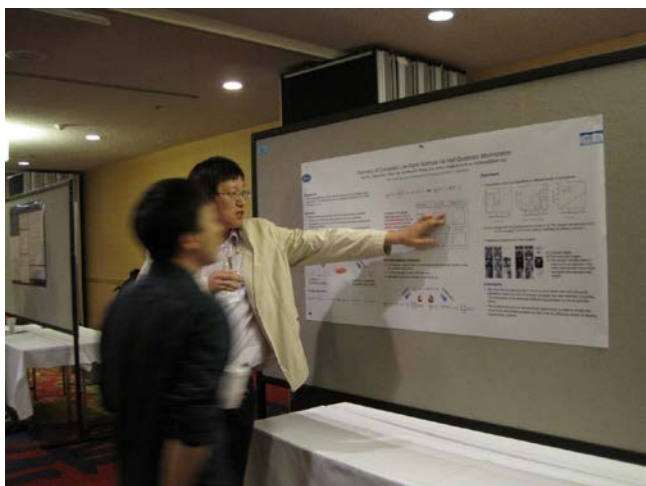


科学研究需要团队的合作，请问您的团队是如何协作的？分享一下您在团队建设方面的成功经验？

我所在的团队是谭铁牛院士带领的智能感知与计算研究中心团队，我的课题组是团队下的研究方向之一。我们把“立德树人”作为课题组建设的宗旨，在科研工作的同时努力实现全程育人和全方位育人。我们主要根据每位成员自身的能力和兴趣分配不同的任务，使每位成员能充分发挥其特长，也能在其他成员的帮助下弥补不足。团队建设要格外注重及时有效地沟通，每一次的周组会都是我们发现问题、讨论问题和解决问题的时候，大家坐在一块面对面沟通效率往往是最高的。一个默契的团队不仅需要科研上的相互配合，也需要情感上的相互关心。我们中心也会不定期为大家组织丰富的娱乐活动，例如春游、秋游、新年茶话会等，在科研之余让大家放松身心，享受生活。

请问您是如何选拔博士生和硕士生？能否和大家分享一下您在研究生管理和培养方面的成功经验？

在学生的选择和培养方面，我们以“德才兼备、以德为先”为首要标准，重点考察学生的思想水平和道德品质，然后考虑学生是否具有“动脑、用心、尽责”三个特质。在选拔研究生的时候，我们会让高年级博士生参与进来，与候选人讨论具体的科研问题，考察候选人的科研兴趣、数学基础和协作能力。学术研究有个入门的过程，在培养新生时，我们采用“传、帮、带”方式，指定一到两位高年级学生来帮助新生，尽可能缩短新生跟上学术前沿的时间。另外，我们会根据每位同学自身的能力和兴趣因材施教，为其提供发展建议，鼓励学生互相合作，一起探索前沿挑战科学问题。除了科研交流之外，我们还定期邀请其他领域的知名学者进行讲座，拓展学生视野，丰富学生生活。最近，我们邀请了中医领域的博士来讲解现代养生观念，希望以此来培养学生合理安排作息，提高科研效率。



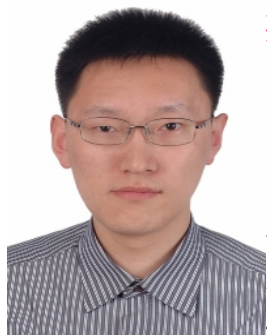
现在有很多大学或研究机构开始鼓励本科生开展学术研究，您对此有什么看法，您有什么样的建议提供给他们吗？

为贯彻落实国家关于深化本科教改的精神，

着力培养学生创新精神和科技创新能力，中国科学院自 2015 年开始实施了“中国科学院大学生创新实践训练计划”，旨在提高在学大学生科技创新能力，帮助大学生了解中国科学院和我国科技的最新进展。多名本科生依托该计划参与到我们的研究工作中，并在模式识别领域的权威国际会议 ICCV、AAAI、SIGIR 等上发表了论文，其中三名本科生（中国地质大学 黄睿 2017；北京科技大学 陈禹嘉 2018；北京工业大学 宋荣 2019）被美国卡耐基梅隆大学（Carnegie Mellon University, CMU）计算机系录取深造。黄睿同学在 ICCV 2017 以第一作者身份发表的论文到目前为止已经被引用 153 次（Google Scholars）。大学生参与中国科学院的创新训练计划有助于提高大学生的科技创新能力，但是凡事都要有度。从事过多的科研项目或辗转过多的实习单位会过犹不及，甚至适得其反。大学生最重要的还是要夯实基础，学好数学和编程，有课余时间的同学最好选择一个研究团队对一个前沿课题进行长期研究。

您最想和众多博士生、学术研究路上的青年学者们分享的是什么？

如今人工智能研究蓬勃发展，每年都有数以万计的相关学术论文发表。但在这种光鲜的表面之下，我们要明白“发表一篇文章，不一定能够解决一个实际问题；解决一个实际问题，一定能够发表一篇好文章”的道理。做科研的最终目标并不应该是发表文章，而应该是解决实际问题。如果能从解决国家和社会面临的关键问题出发，踏踏实实做科研，努力做出能够写入教科书的工作，有影响力的科研成果自然会水到渠成。另外，做科研切忌朝三暮四，盲目跟风。“生也有涯，而知也无涯。以有涯随无涯，殆已”；每个人的时间和精力都是有限的，哪个都想做只会哪个都做不好。反之，科研中要盯住重要问题，扎实积累，追求卓越，有时甚至需要有甘做无名英雄的勇气和信念。



## 赫然

中国科学院自动化研究所模式识别国家重点实验室研究员/博士生导师，中国科学院脑科学与智能技术卓越创新中心年轻骨干，中国科学院大学人工智能学院岗位教授。2001、2004 年在大连理工大学获计算机科学学士和硕士；2009 年于中国科学院自动化研究所，获模式识别与智能系统工学博士。2010 年至今，在模式识别国家重点实验室任助理研究员、副研究员、项目研究员、研

究员。从事模式识别、信息理论学习和稀疏学习基础理论研究，并应用到生物特征识别和计算机视觉。近期主要聚焦在生成式概率深度学习以及图像生成中遇到的瓶颈问题，展开图像模式分析基础理论研究。出版信息理论学习专著 1 部；在 IJCV、IEEE TPAMI、TIP、TIFS、TNNLS、TKDE、TBD、TSMCS 等权威国际期刊以及 NIPS、ICCV、CVPR、IJCAI、AAAI、SIGIR、ACM MM 等权威国际会议发表论文 130 篇，研究工作获得国家自然科学基金优秀青年科学基金和北京自然科学基金杰出青年科学基金资助。

（责任编辑：张汗灵 余焯 黄岩）

## 委员好消息

✪ 2019 年 5 月 23 日，广东省科学技术协会公布了第十五届广东省丁颖科技奖获得者名单，CCF-CV 专委会副主任、中山大学赖剑煌教授喜获殊荣，全省共 20 名科技工作者荣获第十五届广东省丁颖科技奖。广东省丁颖科技奖是 1989 年经广东省政府批准并以著名科学家丁颖院士名义设立的科技奖项，表彰奖励为广东经济建设、社会发展和科技进步做出突出贡献的中青年人才。

✪ 2019 年 6 月 6 日，从希腊召开的 2019 国际生物特征识别大会上获悉，CCF-CV 专委会委员、中科院自动化所模式识别国家重点实验室雷震研究员，因其在人脸生物特征识别领域对人脸识别方法、系统和数据等方面的突出贡献，获得国际模式识别学会(IAPR)颁发的"国际生物特征识别青年学者奖"。该奖项由 IAPR 在 2012 年设立，用于表彰和鼓励国际上在生物特征识别界做出突出贡献的中青年科学家，每两年评审一次，每次从全球范围内评选出 40 岁以下的中青年学者一名。

✪ 2019 年 6 月 16 日，NTIRE 2019 (New Trends in Image Restoration and Enhancement workshop

and Challenges on image and video restoration and enhancement in conjunction with CVPR 2019) 公布了图像和视频恢复与增强挑战赛结果，CCF-CV 专委会委员、厦门大学曲延云教授课题组的团队在“图像去雾”竞赛和 UG2+挑战赛 Track 2 的“雾霾场景下的（半）监督目标检测”竞赛中夺得两项 Runner-Up 奖。

✪ 2019 年 6 月 16 日，CVPR2019 举办的 Workshop on Bridging the Gap between Computational Photography and Visual Recognition 宣布了本次 UG<sup>2</sup>+计算机视觉算法竞赛（低光照环境下的人脸检测）的比赛结果。由 CCF-CV 专委会委员、中科院自动化研究所模式识别国家重点实验室生物识别与安全技术研究中心(CBSR)雷震研究员带领的团队以测试集上 62.25% 的 AP 成绩取得了该项赛事的冠军。

✪ 2019 年 6 月 17 日，CVPR 2019 公布了今年的监测事件检测竞赛（Activity in Extended Video Prize Challenge）最终结果，由 CCF-CV 专委会委员、北京邮电大学人工智能研究院（北邮多媒体通信与模式识别实验室）苏菲等老师带领的团

---

队获得本次监控事件检测竞赛冠军。这是该团队在连续获得 2015 年美国联邦标准技术局的 TRECVID 监控事件检测事件冠军、CVPR 的 2016 多媒体事件检测、监控事件检测两项世界冠军、2017 年监控事件检测世界冠军、2018 年监控事件检测世界亚军后，再次夺得的全球第一名。

✪ 2019 年 6 月 17 日，CVPR 2019 举办的 Workshop on Long-Term Visual Localization under

Changing Conditions 公布了本年度 Long-Term Visual Localization Challenge（大时间跨度视觉定位竞赛）的比赛结果，由 CCF-CV 专委会委员、中科院自动化所模式识别国家重点实验室机器人视觉课题组申抒含副研究员所带领的团队提交的定位结果在 5 个评测数据集上获得 4 个第一，1 个第二，总分排名第一，夺得冠军。

（责任编辑：刘海波）

# 图像描述开源代码

北京电子科技学院 李熹桥 金鑫

图像描述 (Image Captioning) 是从图像中自动生成一段描述性文字。可以认为是一种动态的目标检测, 由全局信息生成图像摘要。对于人来说, 图像描述是一件简单的事, 但对于机器来说, 这项任务却充满了挑战性。原因在于机器不仅要能检测出图像中的物体, 而且要理解物体之间的相互关系, 最后还要用贴近人类的语言表达出来。

早期的研究例如《Baby talk》, 《Every picture tell a story》等都是利用图像处理的一些算子提取出图像的特征, 经过 SVM 分类等等得到图像中可能存在的目标。根据提取出的目标以及它们的属性利用 CRF 或者是一些人为规则来生成图像的描述。这种做法非常依赖于图像特征的提取和生成句子时所需要的规则。

随着深度学习技术的发展和 COCO 等图像标注数据集的出现, 图像描述相关的技术得到了快速的发展。目前图像描述的方法主要是在编码器-解码器框架上进行改进与不断尝试。



图 1 图像描述示意图

本文着重介绍几个基于深度学习的图像描述技术的开源代码, 包括训练深度学习模型的常

用数据集和几个经典的图像描述模型。

## 1. Flickr 数据集

**介绍:** Flickr 数据集, 是雅虎发布的一个大型数据集, 这个数据集由 1 亿张图像和 70 万个视频的 URL 以及与之相关的元数据(标题、描述、标签)所组成的。Flickr8k、Flickr30K 分别包含 8000 和 31000 张图像, 每一张图像都包含人工标注的 5 句话。

**主页:** <http://shannon.cs.illinois.edu/DenotationGraph>

## 2. SBU captioned photo 数据集

**介绍:** SBU captioned photo 数据集是 Flickr 数据集的一个子集, 是从 Flickr 数据集种经过筛选出 100 万张图像以及相关的描述。

**主页:** <http://www.cs.virginia.edu/~vicente/sbucaptions>

**论文:** [http://tamaraberg.com/papers/-generation\\_nips2011.pdf](http://tamaraberg.com/papers/-generation_nips2011.pdf)

## 3. MS COCO Caption 数据集

**介绍:** MS COCO 数据集是微软构建的一个数据集, 图像主要从复杂的日常场景中截取, 数据集中一共包括 91 类别, 328,000 张图像和 2,500,000 个标签, 以及通过精确的语义分割进行标定的位置。

**主页:** <http://cocodataset.org>

## 4. nocaps 数据集

**介绍:** nocaps 包含了来自 Open Images 平台, 一共 15100 张图像, 以及相应的 166100 个人工描述。数据集包括 MS COCO 缺失的 400 多种类别。

**主页:** <http://places2.csail.mit.edu>

**论文:** <https://nocaps.org/paper.pdf>

## 5. AI CHALLENGER 数据集

**介绍:** AI CHALLENGER 图像中文描述是场景最丰富、语言描述最丰富、规模最大的图像中文描述图像数据集。每一张图像都有对应的五句话的中文描述。描述句子符合自然语言习惯，点明了图像中的重要信息，涵盖主要人物、场景、动作等内容。数据集包含 30 万张图像，150 万句中文描述。

**主页:** <https://challenger.ai/dataset/caption>

**论文:** <https://arxiv.org/abs/1711.06475>

### 6. VISUAL GENOME 数据集

**介绍:** VISUAL GENOME 数据集是一个致力于将语言描述与结构化图像联系在一起的数据集。图像包括三种描述属性：地区、属性和关系。数据集一共包括 108,077 张图像，540 万地域性描述，3.8 百万个对象实例，2.8 百万个属性，230 万关系以及 170 万视觉问答对。

**主页:** <http://visualgenome.org>

**论文:** [http://visualgenome.org/static/paper/-Visual\\_Genome.pdf](http://visualgenome.org/static/paper/-Visual_Genome.pdf)

### 7. IAPR TC-12 数据集

**介绍:** IAPR TC-12 数据集包含 20000 张真实图像，这些图像包括不同运动和动作的照片、人物、动物、城市、风景和当代生活的许多其他方面的照片。每个图像都有最多三种不同语言（英语、德语和西班牙语）的描述。

**主页:** <https://www.imageclef.org/photodata>

**论文:** <http://nla.gov.au/anbd.bib-an43036734>

### 8. Adaptive attention

**论文:** Gao L L, Li X P, Song J K, *et al.* Hierarchical LSTMs with adaptive attention for visual captioning[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015: 14(8): 1-19.

**工作:** 目前大多数基于注意力机制的图像描述采用的都是编码器-解码器框架，然而解码器应该对不同的词有不同的注意策略。一些英文代词连词等不需要视觉信息的词汇生成更多依赖语义信息而不是视觉信息。这种非视觉词的梯度会影响视觉信息的有效性。本文提出了一种带有视

觉标记的自适应注意力模型，模型可以决定在一个时间点依赖于视觉信息还是语义结构。具体架构如图 2 所示：

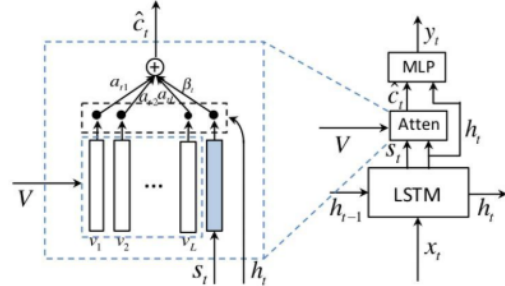


图 2 自适应注意力模型

**代码:** <https://github.com/jiasenlu/AdaptiveAttention>

### 9. Neural Baby Talk

**论文:** Lu J, Yang J, Batra D, *et al.* Neural baby talk[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7219-7228.

**工作:** 该论文认为当前典型的图像描述模型过分依赖于语言模型，导致生成的图像描述与图像内容关联不够。论文的基本思想是采用物体检测器检测图像中的物体，然后借鉴了自适应注意力模型的做法，学习了一个哨兵参数来决定在生成每个词的时候是选取数据集中的词汇还是检测到的词汇。图 3 为具体架构：

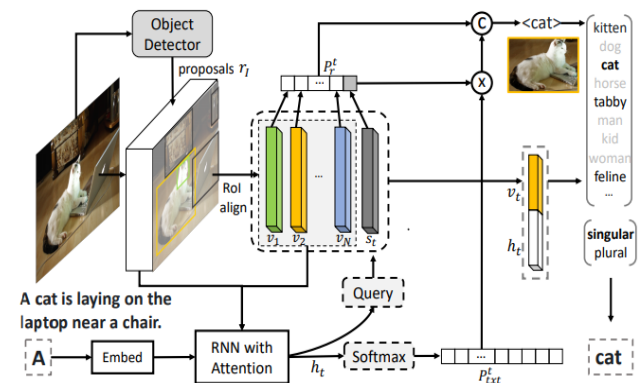


图 3 Neural Baby Talk

**代码:** <https://github.com/jiasenlu/NeuralBabyTalk>

### 10. Self-critical Sequence Training for Image Captioning

论文: Rennie S J, Marcheret E, Mroueh Y, *et al.* Self-critical sequence training for image captioning[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7008-7024.

工作: 该论文认为当前图像描述主要存在的问题是模型训练时, 输入 RNN 的上一时刻的单词是来自训练集的单词。而在测试的时候依赖的是自己生成的单词, 一旦生成得不好就会导致误差的积累, 导致后面的单词也生成得不好。模型训练的时候用的是交叉熵损失函数, 而评估性能的时候却用的是 BLEU、ROUGE、METEOR、CIDEr 等衡量指标, 存在不对应的问题。由于生成单词的运算是不可微的, 本文使用了强化学习来优化误差的思路。

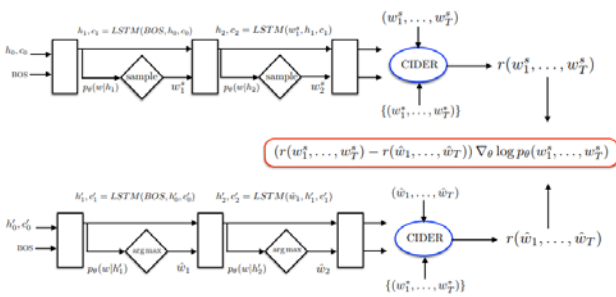


图 4 强化学习网络结构图

代码: <https://github.com/ruotianluo/self-critical.pytorch>  
 11. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention

论文: Xu K, Ba J, Kiros R, *et al.* Show, attend and tell: neural image caption generation with visual attention[C]. International Conference on Machine Learning. 2015: 2048-2057.

工作: 该方法提出了将注意力机制与传统的 encoder-decoder 框架结合的方法。在编码器端, 模型使用一些预训练的模型来提取特征, 本方法提取的特征向量来自于低层, 这样解码器可以通过选择特征向量的子集来聚焦图像的某些部分。在解码器端采用 RNN 来解码并生成词序列, 与传统方法不同的地方在于本文采用两种注意力机制的结合并增加一个跟时间有关的维度因子, 从而让模型的注意力更加集中在目标图像区域, 进而解码得到更加合理的词序列。

两种注意力机制分别是 Stochastic “Hard”

Attention 和 Deterministic “Soft” Attention。“硬”注意力机制是计算将不同图像区域输入解码器的信息的概率, “软”注意力机制是计算将图像区域输入解码器所占的比例。图 5 为该方法的网络结构图:

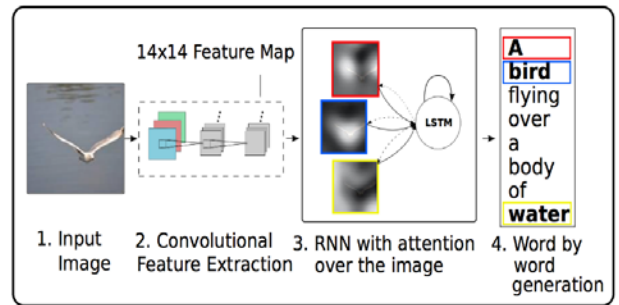


图 5 Visual Attention 网络结构图

代码: <https://github.com/sgrvinod/a-PyTorch-Tutorial-to-Image-Captioning>

## 12. DenseCaps: Fully Convolutional Localization Networks for Dense Captioning

论文: Johnson J, Karpathy A, Fei-Fei L. Densecap: fully convolutional localization networks for dense captioning[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4565-4574.

工作: 这篇文章的主要贡献在于提出了新的密集定位层, 它完全可微并且可以将其单独插入其他任何图像处理的网络中来实现区域定位的训练和预测。该层预测了图像中的一系列感兴趣的区域 (ROI) 并使用了双线性插值来较为平滑得提取区域中的激活值。其网络模型有三部分组成: 卷积网络、密集定位层和 RNN 语言模型。其模型结构图 6 所示:

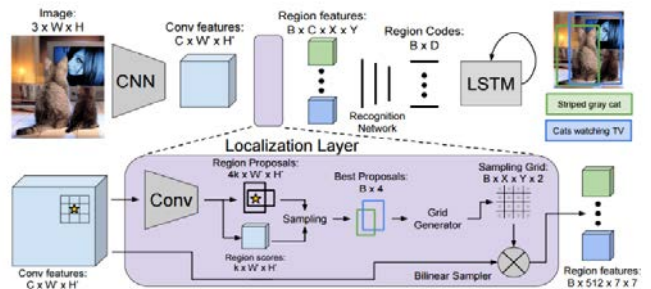


图 6 DenseCaps 模型结构图

代码: [https://github.com/deepsemantic/image\\_captioning](https://github.com/deepsemantic/image_captioning)

### 13. Towards Diverse and Natural Image Descriptions via a Conditional GAN

论文: Dai B, Fidler S, Urtasun R, *et al.* Towards diverse and natural image descriptions via a conditional GAN[C]. IEEE International Conference on Computer Vision. 2017: 2970-2979.

**工作:** 当前的 Image Captioning 工作由于大多使用极大似然估计来训练模型。这种类似的方法会带来模型产生的图像描述与 ground truth 高度一致, 缺失了多样性, 显得生硬。该论文提出利用 Conditional GAN 让模型产生的图像描述更加贴合人类的表达, 改善句子的多样性。

Generator 方面, 以 CNN 提取的图像特征及噪声作为输入, 用 LSTM 生成句子, 随后通过蒙特卡洛方法从 Evaluator 得到反馈, 并通过策略梯度算法来更新参数。

Evaluator 方面, 使用 LSTM 对句子编码, 然后与图像特征向量合并再处理, 得到一个概率值。Evaluator 既要判别句子是否像是人类生成的, 同时要判别句子和图像是否相关。在训练时, 作者把对句子是否自然和对与图像内容是否相关的区分开。

该模型结构如图 7 所示:

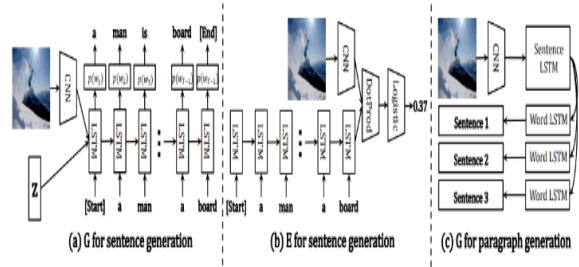


图 7 Conditional GAN 示意图

代码: <https://github.com/msracver/Deep-Exemplar-based-Colorization>

### 14. Contrastive Learning for Image Captioning

论文: Dai B, Lin D. Contrastive learning for image captioning[C]. Advances in Neural Information Processing Systems, 2017: 898-907.

**工作:** 本文是为了解决一些图像描述算法的问题, 由于采用极大似然估计来训练模型, 使得生成的描述不具备独特性的问题。针对此问题, 本文提出了通过构造损失函数利用负样本参与训练, 来提高模型的独特性。

最终的损失函数:

$$J(\theta) = \frac{1}{k} \frac{1}{T_m} \sum_{k=1}^k L(\theta; X, Y_k, \phi)$$

代码: [https://github.com/doubledaibo/clcaption\\_nips2017](https://github.com/doubledaibo/clcaption_nips2017)



李熹桥

北京电子科技学院在读研究生, 研究方向为计算机视觉。

Email: azxcvgy@outlook.com



金鑫

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

个人主页: jinxin.me

# 水下图像增强数据集

大连理工大学 樊鑫 侯岷君

海洋蕴藏着巨大的资源，被人类视为可以利用的“第六大洲”，开发和保护海洋资源是一项影响深远、面向二十一世纪的战略选择。海洋信息的获取、传输和处理等理论技术对合理开发和利用海洋资源至关重要。水下图像是海洋信息的重要载体，然而，由于光在水中特殊的传输特性，直接的水下成像质量往往很差。导致水下图像退化的主要因素有两个：其一，水下场景的反射光在到达相机之前被悬浮在介质中的粒子吸收和散射，造成低对比度低和水下雾效应。其而，由于在水下波长较长的红光比绿色和蓝色的光被吸收更多，水下图像大部分呈现蓝色或绿色。这些降至特征限制了水下图像和视频在海洋生物学和考古学、水下机器人领域中的实际应用，因此，提升水下图像的质量有着重要的实际意义，水下图像增强算法的发展也受到越来越多的关注。

而水下图像增强算法的综合研究和深刻分析与真实水下图像数据集密不可分。近年来，由于硬件设备的进步和算法研究的需要，水下图像数据库也得到了更多的完善。本文着重介绍几个水下图像增强相关的数据库，包括 RUIE、UIEBD、MHL、Jamaica。

## 1. RUIE 数据集

Real-world Underwater Image Enhancement (RUIE) 数据集是一个专用于评价水下图像增强算法的大型数据库，由大连理工大学的研究者提出，其包含超过 4000 张水下图像。数据库分为 UIQS、UCCS 和 UHTS 三个子数据集。其中，UIQS 为衡量水下图像质量而设定，共 3630 张图像。综合参考无参考水下图像质量评价 underwater image quality measure (UIQM) 和 underwater color image quality evaluation (UCIQE) 的评分，UIQS 被平均划分为 5 个子集，用于评估不

同增强算法在不同质量图像上的性能。UCCS 为衡量水下色偏而设定，包含 300 张图像。其根据色偏程度分为“蓝”、“蓝绿”、“绿”三个子集，用于衡量算法的色偏校正能力。UHTS 包含 300 张图像，用于评价增强后结果用于目标检测的性能。其依照图像质量被划分为 5 个子集，含海参、海胆、扇贝三种检测物。

该数据集采集于大连獐子岛，拍摄环境为一个带有 22 个防水摄像机的多视图水下图像捕获系统。摄像机沿 10 米×10 米的方形框架安装，位于靠近黄海章子岛的海床上方 0.5 米处，其地理坐标位于 (N39: 186; E44: 625)。相机视角可调节，最大场景深度可以在 0.5 到 8 米之间，拍摄时段为上午 8 点至 11 点和下午 1 点至下午 4 点。水深由于周期性潮汐而在 5 至 9 米之间变化。不断变化的照明和水深会产生不同的色调。拍摄的视频超过 250 小时，涵盖了各种照明，模糊度和色偏多种水下图像降质特征。此外，该地区保持着包括鱼类、海胆、海参、扇贝等丰富的海洋生物在内的自然海洋生态系统。这种生态系统使我们能够为水下目标探测任务的标注提供了便利。

成功的水下图像增强算法必须解决水下成像中的一个或所有以下问题，包括可见性降低，色偏和针对高级检测任务的准确度。这些多个目标需要在水下图像增强的基准测试中提供多种测试示例组合。而用于评估可见度改进能力的图像通常需要更大的场景深度，以便使得由水散射引起的退化效果明显。另一方面，包含多种色调的数据集可以评估算法校正色偏的性能。此外，检测/分类准确度的计算要求对象/目标标签作为基准中的基础。然而，大多数现有的水下图像数据集通常旨在评估水下图像增强算法的三个目标中的一个或两个。因此一个大规模、多样化、任务特定的数据库对于水下图像增强算法的公

平和全面的比较具有重要意义。

该数据集包含丰富的水下场景及由低到高的多种降质程度，适合评估水下增强算法的性能，尤其适合评估基于物理成像模型的算法的水下除雾能力。此外，由于水下图像增强算法常用于水下生物检测、抓取等

实际应用中，应用于高级视觉任务的性能对于算法有效性的评估非常重要。该数据的 UHTS 按照不同降质程度划分，有利于衡量增强算法在不同降质程度下的针对高级视觉任务的性能。

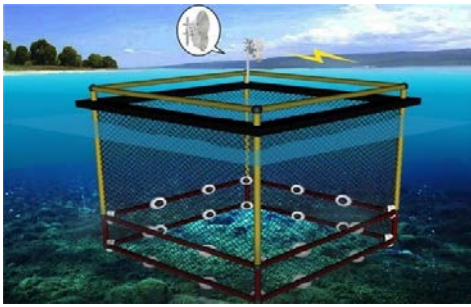
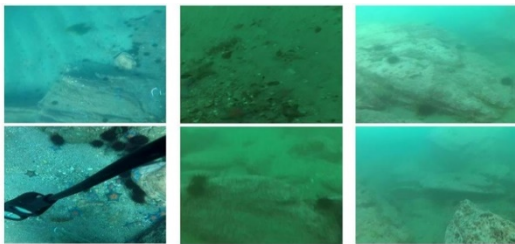


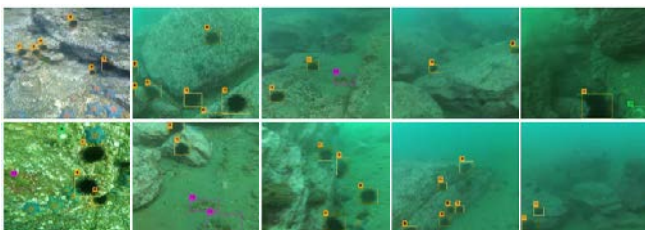
图 1 RUIE 数据集图像采集系统



(1) UIQS 子集，从左至右图像质量依次降低



(2) UCCS 子集，依次为蓝、绿、蓝绿色偏



(3) UHTS 子集，从左至右图像质量依次降低

图 2 RUIE 数据集，从上到下依次为 UIQS、UCCS、UIQS 子集

数据库网址：<https://github.com/dlut-dimt/Realworld-Underwater-Image-Enhancement-RUIE-Benchmark>

## 2. UIEBD 数据集

Underwater Image Enhancement Benchmark Dataset (UIEBD) 数据库由香港城市大学的研究者提出，该数据集包含 950 幅分辨率为  $183 \times 275$ — $1350 \times 1800$  的真实水下图像，其中 890 幅具有相应的高质量参照图像（即清晰水下图像），其余 60 幅为具有挑战性的数据。该数据库涵盖了多种水下场景、不同的质量退化特征、广泛的图像内容。水下图像退化程度较大，并提供相应的高质量参考图像，使成对的图像能够进行公平的图像质量评价和端到端学习。这也是首个含有原始水下图像和相应的高质量参考图像的大型水下图像数据库。UIEBD 提供了评估不同水下图像增强的性能的平台。

数据库图像获得方法如下：研究者从网络收集大量的水下图像，经过数据细化后，大部分采集到的图像被剔除，剩下 950 幅候选图像。为获得高质量参考图像，研究者使用 12 种水下图像增强算法对降质图像进行处理后，邀请 50 名志愿者（25 名有图像处理经验的志愿者和 25 名无相关经验的志愿者），在同一台显示器下对每幅原始水下图像的 12 个增强结果进行两两比较。增强的图像对是从所有比较方法中随机抽取的，较好的结果将在下一轮再次进行比较，直到选出最佳的一个。志愿者不受时间限制，可进行放大操作。对于每一对增强结果，志愿者首先需要以原始水下图像为参考，独立判断哪一个效果更好。因此，对于每一位志愿者，经过 11 对对比，选出一幅原始水下图像的最佳结果。此外，志愿者需要再次检查最佳结果，然后将最佳结果标记为满意或不满意。最后，对原始水下图像进行配对比较，以多数票的方式选出参考图像，如果选中的参考图像中标注不满意的投票数超过一半，则将其对应的原始水下图像视为具有挑战性的图像，并将该参考图像丢弃，由此获得了 890 张可用的

高质量参考图像和 60 张无参考的挑战性图像。



图 3 UIEBD 数据库

数据库网址: [https://li-chongyi.github.io/proj\\_benchmark.html](https://li-chongyi.github.io/proj_benchmark.html)

### 3. MHL 数据库

MHL 数据库拍摄于密歇根大学海洋流体力学实验室 (MHL) 的一个试验池中的人造岩石平台。岩石平台尺寸为 4 英尺 x 7 英尺, 岩石离平台约 3 英尺。平台底座覆盖着纹理网格。本数据集收集超过 7000 幅水下影像, 其中大部分图像具有低光照、绿色色偏、晕影、较低雾效应等特点, 适于评估不基于物理成像模型的增强算法或水下低光图像增强算法。



图 4 MHL 数据库

数据库地址: <http://www.umich.edu/~dropopen/MHL.tar.gz>

### 4. Jamaica 数据集

该数据集由 6500 张水下图像组成, 采集于在牙买加皇家港 (Port Royal) 的一个水下城市遗址, 该城市位于地表以下 3-4 米, 包含自然和人工建筑。这些图像是用安装有高分辨率立体摄像机的手持潜水器在当地不同潜水点采集的。与海底的最大距离约为 1.5m。该数据集中图像具有低光照、蓝色色偏、纹理丰富、较低雾效应等特点, 适于评估不基于物理成像模型的增强算法或水下低光图像增强算法。

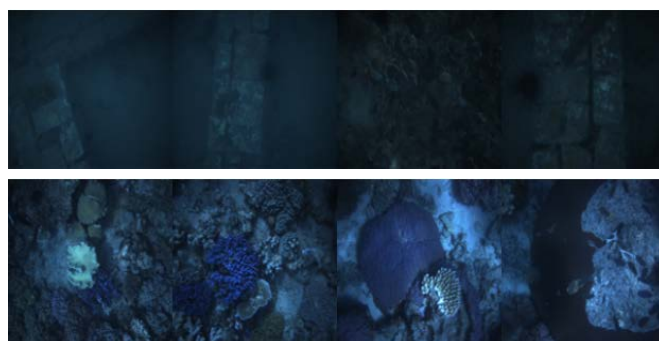


图 5 Jamaica 数据库

数据库地址: <http://www.umich.edu/~dropda/Jamaica.tar.gz>

(责任编辑: 沈沛意, 金鑫)



#### 樊鑫

大连理工大学教授、博士生导师, 现任大工-立命馆国际信息与软件学院院长。研究方向为计算几何和机器学习, 以及它们在低级图像处理和 DTI-MR 图像分析中的应用。

个人主页: <http://dutmedia.org>



#### 侯岷君

大连理工大学在读研究生, 研究方向为低级图像处理。

Email: [houminjun1995@gmail.com](mailto:houminjun1995@gmail.com)

# 招聘信息

## 一、 字节跳动实习生

### 1、 HRBP 实习生-商业化-北京:

#### 1) 职位描述:

- 协助招聘, 电话预约、协调候选人、面试官时间, 并安排面试;
- 基础岗位招聘, 网络筛选简历、并进行电话面试;
- 协助进行员工入职、离职、绩效的管理工作;
- 协助进行招聘渠道维护的相关工作;
- 协助进行人力资源相关的数据统计工作。

#### 2) 职位要求:

- 在读 19 届准备读研/出国的同学, 20 届及以后毕业的本科或研究生。(19 届毕业准备就业的同学请勿投递, 请直接参加校招);
- 熟练运用 excel、word、ppt 等办公软件;
- 为人正直、亲和力好、耐心、细致、认真、负责, 抗压;
- 善与人沟通、逻辑思维能力及应变能力较好, 富于团队协作精神;
- 能够连续实习 3 个月及以上, 每周出勤 4 天及以上者优先。

#### 3) 简历投递:

投递邮箱: [guomingxuan@bytedance.com](mailto:guomingxuan@bytedance.com)

简历及邮件命名格式: 姓名+学校+电话+来自互联镖局。

### 2、 HRBP 实习生-武汉:

#### 1) 工作内容

- 协助招聘, 电话预约、协调候选人、面试官时间, 并进行面试安排;
- 负责后端岗位招聘, 根据岗位 JD 筛选简历、并进行电话面试;
- 负责每日面试情况的数据记录和相关反馈;
- 协助 HRBP 进行招聘项目相关数据的统计分析工作, 及时发现招聘过程的问题, 提出优化建议;

- 完成 HRBP 上级安排的其它支持性工作。

#### 2) 岗位要求:

- 大三、研一、研二、人力资源、工商管理等相关专业优先, 可实习三个月以上;
- 具有一定的数据分析能力和数据敏感度, 熟练掌握 Office 软件操作(Excel、Word 等);
- 热爱互联网行业和人力资源工作, 有相关实习工作经验优先;
- 学习能力强, 能够迅速适应头条工作节奏, 抗压力强, 接受加班;
- 沟通能力强, 善于团队协作、思路清晰、逻辑思维强、有判断力。

#### 3) 地址: 武汉东湖高新技术开发区凌家山南路 1 号武汉光谷企业天地。

#### 4) 投递邮箱: [aceoffer2018@vip.163.com](mailto:aceoffer2018@vip.163.com)

邮件&简历命名:

【今日头条】【HRBP】【武汉】【姓名】【年级】【学校】【电话】【一周到岗几天】【实习时长】【最早到岗时间】【爱思益推荐】(注: 邮件&简历命名有误不进行推优)

### 3、 抖音运营实习生:

#### 1) 工作内容:

- 协助培训师进行培训课件资料(管理训、在职训)的准备;
- 协助培训师开展新员工培训, 后期可独立进行授课;
- 协助培训师进行新员工学员的管理;
- 完成上级安排的其他工作。

#### 2) 职位要求:

- 大学本科及以上学历(2020 届毕业生优先);
- 具备良好的沟通及抗压能力, 工作认真、细致、负责;
- 具有丰富的演讲、表达能力以及文案写作能力, 对互联网/新媒体行业有浓厚兴趣;
- 熟练掌握办公软件, 如 Word、Excel、PowerPoint 等;

- 具有良好的服务意识与团队精神；
- 一周可工作天数至少为 4 个工作日，并可长期实习（3 个月及以上）。

3) 投递邮箱：[aceoffer2018@vip.163.com](mailto:aceoffer2018@vip.163.com)

邮件&简历命名：

【今日头条】培训部实习生【武汉】【姓名】  
【年级】【学校】【电话】【一周到岗几天】【实习时长】  
【最早到岗时间】【爱思益推荐】（注：邮件&简历命名有误不进行推优）

#### 4、前端实习生（北京）：

##### 1) 职位描述：

负责社区产品各垂直领域的内容增长及运营，提高兴趣垂类 UGC 的质量和活跃。

##### 2) 职位要求：

- 有较强的信息获取及筛选整合能力，好奇心强、乐于探索；
- 对网络热门话题敏感，话题活动组织能力强；
- 有特定擅长的兴趣领域，包括但不限于饭圈、二次元、游戏、体育等话题；
- 有互联网社区或自媒体运营经验者优先（贴吧吧主、豆瓣组长、即刻主理人）；
- 可尽快到岗，实习时间至少 3 个月以上，每周不少于 4 个工作日。

##### 3) 职位亮点：

大厂经历，包三餐下午茶，租房补贴，有竞争力薪酬。

##### 4) 投递邮箱：

[zhangqianying@bytedance.com](mailto:zhangqianying@bytedance.com)

（岗位+姓名+实习时长+每周实习天数+来自互联镖局） 欢迎投递。

## 二、商汤科技 UI 视觉设计实习生（深圳）

### 1、工作职责：

- 负责参与公司产品的视觉设计；
- 与团队成员合作共同完成设计任务；
- 配合主设计师完成各项设计工作任务；
- 跟进项目开发的实现，保证开发对设计的高还原度；
- 协助完成大中型设计项目。

### 2、任职资格：

- 2020 年及以后艺术专业本科或以上学历毕业，艺术相关专业（美院及 985 院校优先）；
- 具备优秀的审美能力；
- 熟练的工具使用，如 Ps/Ai/C4D 等设计工具；
- 具有优秀的设计理论与娴熟的设计技巧；
- 具备良好的合作态度及团队精神，富有工作激情；
- 性格开朗，做事有责任心；
- 有一定互联网产品项目设计经验者优先。

### 3、工作地址：

深圳南山区深圳湾南海大道 1052 号海翔广场。

### 4、福利待遇：

300-400/天+免费零食、健身房人工智能平台公司，超 nice leader~。

### 5、联系方式

简历附上本人近期设计作品集备注“姓名+设计实习生”发送至 [dingliyong@sensetime.com](mailto:dingliyong@sensetime.com)（姓名+电话+邮箱+来自互联镖局）。

## 三、虎牙直播 AI 专项算法工程师

虎牙（HUYA）作为“中国游戏直播第一股”，坚持科技驱动创新，是首家上线 HTML5、规模化使用 H265、实现 5G+4K 直播、采用 AI 技术理解视频内容的直播平台。

2019 年初，虎牙成立 AI 基础中心，以研发行业领先的核心 AI 技术为使命，专注计算机视觉、语音识别、自然语言处理、机器学习等人工智能领域的研究和应用。全球化的战略布局、多样化的业务形态、上亿量级的产品数据、年轻创新的工作氛围，AI 算法中台的业务视角，让你快速成为团队核心骨干。加入虎牙，你将参与推动直播行业的发展与变革，创建直播+AI 的未来。

### 1、AIOPS 算法工程师：

#### 1) 职位描述：

- 基于虎牙直播的海量业务&技术数据，在时序数据异常检测、告警根源分析、运维决策等场景负责算法实现和智能运维产品落地；
- 负责 AIOPS 系统的设计与研发；
- 参与算法和模型的实现，丰富算法库。

## 2) 任职要求:

- 2020 届硕士学历，人工智能、深度学习等相关专业；
- 熟练掌握深度学习的基本方法（如 LSTM, RNN 等）和框架（TensorFlow, PyTorch 等其中一种或多种）；
- 熟悉机器学习的常见算法，包括但不限于回归，分类，聚类，关联规则等算法；
- 有较强的算法实现能力，熟练掌握 Python/C++ 编程；
- 能主动寻找运维效率问题，愿意用算法解决实际问题，愿意迭代算法和产品，提高运维效率；
- 获得过重要技术比赛、或参与过相关 aiops 产品落地的同学优先；
- 此次招聘优先发放暑期实习 offer，针对特别优秀的同学直接发放秋招提前批录用意向。

## 2、机器学习算法工程师

### 1) 岗位职责:

- 负责机器学习基本算法的开发与性能提升，涉及的问题包括但不限于：特征工程，大数据分析，分类聚类，数据挖掘，搜索排序，个性化推荐等；
- 推动机器学习在众多实际应用场景的落地。

### 2) 任职要求:

- 2020 届硕士或博士学历（优秀本科生亦可），人工智能、机器学习等专业，有顶会论文或实际业务落地经验者优先；
- 熟练掌握机器学习基本方法（如 SVM, LR, 分类，回归，聚类，决策树，特征工程等）和应用（如数据挖掘，搜索排序，个性化推荐等）；
- 有较强的算法实现能力，熟练掌握 Python/C/C++ 编程，熟悉 Linux 系统，Shell 编程；
- 优秀的分析问题和解决问题的能力，对解决具有挑战性的问题充满激情；
- 有较强的学术比赛或工程项目经验，获得过重要代码竞赛（如 ACM-ICPC, NOI/IOI,

- Google Code Jam, 百度之星）的荣誉，在重要数据集的 Leaderboard 上排名靠前，或开源代码在 github 上产生较大影响者优先；
- 此次招聘优先发放暑期实习 offer，针对特别优秀的同学直接发放秋招提前批录用意向。
- \*本岗位同时面向优秀的 2021 及 2022 届研究生、博士生同学开放日常实习通道。

## 3、NLP 算法工程师

### 1) 职位描述:

- 负责 NLP 和深度学习基本算法的开发与性能提升，涉及的问题包括但不限于：文本、音频、视频等；
- 推动 NLP 和深度学习在众多实际应用场景的落地。

### 2) 任职要求:

- 2020 届硕士或博士学历，人工智能、自然语言、深度学习等专业，有顶会论文或实际业务落地经验者优先；
- 熟练掌握机器学习算法（如 GBDT, Random forest, CNN, RNN 等）和框架（TensorFlow, PyTorch, Caffe 等其中一种或多种）；
- 有较强的算法实现能力，熟练掌握 Python/C/C++/java 编程，熟悉 Linux 系统，Shell 编程；
- 学习能力强，优秀的分析问题和解决问题的能力，对解决具有挑战性的问题充满激情；
- 有较强的学术比赛或工程项目经验，获得过重要代码竞赛（如 ACM-ICPC, NOI/IOI, Google Code Jam, 百度之星）的荣誉，在重要数据集（如 ImageNet, Coco, KITTI, Kaggle）的 Leaderboard 上排名靠前优先；
- 此次招聘优先发放暑期实习 offer，针对特别优秀的同学直接发放秋招提前批录用意向。

## 4、计算机视觉算法工程师

### 1) 岗位职责:

- 负责计算机视觉和深度学习基本算法的开发与性能提升，涉及的问题包括但不限于：人脸人体关键点检测/分割/跟踪/重建，图像视频理解/编辑，3D 视觉，AR 等；

- 推动计算机视觉算法和深度学习在众多实际应用场景的落地。
- 2) 任职要求:
- 2020 届硕士或博士学历 (优秀本科生亦可), 人工智能、计算机视觉、深度学习等专业, 有顶会论文或实际业务落地经验者优先;
  - 熟练掌握计算机视觉和深度学习的基本方法 (如 CNN, RNN 等) 和框架 (TensorFlow, PyTorch, Caffe 等其中一种或多种);
  - 有较强的算法实现能力, 熟练掌握 Python/C/C++ 编程, 熟悉 Linux 系统, Shell 编程;
  - 优秀的分析问题和解决问题的能力, 对解决具有挑战性的问题充满激情;
  - 有较强的学术比赛或工程项目经验, 获得过重要代码竞赛 (如 ACM-ICPC, NOI/IOI, Google Code Jam, 百度之星) 的荣誉, 在重要数据集 (如 ImageNet, Coco, KITTI, Kaggle) 的 Leaderboard 上排名靠前, 或开源代码在 github 上产生较大影响者优先;
  - 此次招聘优先发放暑期实习 offer, 针对特别优秀的同学直接发放秋招提前批录用意向。
  - \*本岗位同时面向优秀的 2021 及 2022 届研究生、博士生同学开放日常实习通道。
- ### 5、语音识别/合成算法工程师
- 1) 岗位职责:
- 负责语音分析/识别/合成等基本算法的开发与精度性能提升, 涉及的问题包括但不限于: 语音分析处理, 语音识别, 说话人风格转换, 语音合成 (TTS), 自动对话等;
  - 推动语音分析/识别/合成在众多实际应用场景的落地。
- 2) 任职要求:
- 2020 届硕士或博士学历 (优秀本科生亦可), 人工智能、语音识别合成、深度学习等专业, 有顶会论文或实际业务落地经验者优先;
  - 熟练掌握语音分析识别和深度学习的基本方法 (如 RNN, CNN 等) 和框架 (TensorFlow, PyTorch, Caffe 等其中一种或多种);
  - 有较强的算法实现能力, 熟练掌握 Python/C/C++ 编程, 熟悉 Linux 系统, Shell 编程;
  - 优秀的分析问题和解决问题的能力, 对解决具有挑战性的问题充满激情;
  - 有较强的学术比赛或工程项目经验, 获得过重要代码竞赛 (如 ACM-ICPC, NOI/IOI, Google Code Jam, 百度之星) 的荣誉, 在重要数据集的 Leaderboard 上排名靠前, 或开源代码在 github 上产生较大影响者优先;
  - 此次招聘优先发放暑期实习 offer, 针对特别优秀的同学直接发放秋招提前批录用意向。
  - \*本岗位同时面向优秀的 2021 及 2022 届研
- ### 6、深度学习算法工程师
- 1) 岗位职责:
- 负责深度学习算法的开发和模型加速压缩, 目标是在移动平台的实时模型;
  - 推动深度学习在众多实际应用场景的落地。
- 2) 任职要求:
- 2020 届硕士或博士学历 (优秀本科生亦可), 人工智能、深度学习等专业, 有顶会论文或实际业务落地经验者优先;
  - 熟练掌握计算机视觉和深度学习的基本方法 (如 CNN, RNN 等) 和框架 (TensorFlow, PyTorch, Caffe 等其中一种或多种);
  - 有较强的算法实现能力, 熟练掌握 C/C++ 编程, 熟悉 Linux 系统, Shell 编程;
  - 优秀的分析问题和解决问题的能力, 对解决具有挑战性的问题充满激情;
  - 有较强的学术比赛或工程项目经验, 获得过重要代码竞赛 (如 ACM-ICPC, NOI/IOI, Google Code Jam, 百度之星) 的荣誉, 在重要数据集的 Leaderboard 上排名靠前, 或开源代码在 github 上产生较大影响者优先;
  - 此次招聘优先发放暑期实习 offer, 针对特别优秀的同学直接发放秋招提前批录用意向。
  - \*本岗位同时面向优秀的 2021 及 2022 届研

究生、博士生同学开放日常实习通道。

## 7、编解码 AI 算法工程师

职位描述：

- 负责直播场景下视频质量评测、场景聚类、目标检测、图像增强、超分辨率以及编解码优化等开发；
- 负责流媒体实时处理、音画质优化与 AI 技术结合预研。

2) 任职要求：

- 2020 届硕士及以上学历，计算机相关专业；
- 熟练掌握一种或者多种编程语言：C/C++/Python/Java；
- 熟悉 h.264/h.265 编解码原理以及常用编解码器的使用；
- 了解常见的图像处理算法、了解机器学习/深度学习相关框架；
- 实习表现优秀者，有机会提前获得 2020 届校园招聘 offer。

## 8、简历投递

[http://hr.huya.com/campus\\_apply/huya/4112#/jobs?zhineng=18100&k=pqig4j](http://hr.huya.com/campus_apply/huya/4112#/jobs?zhineng=18100&k=pqig4j)

## 四、网易前端开发实习生（杭州）

### 1、岗位职责

- 负责阴阳师双微的日常原创内容撰写与运营；
- 研究游戏产品与网络新动态，结合产品需求，对微博、微信公众号内容进行把控调整；
- 参与线上粉丝活动、相关营销活动方案策划、创意提出、执行；

- 组织平台的推广。

### 2、岗位要求

- 本科或以上学历，熟悉《阴阳师》IP 相关内容，《阴阳师》资深玩家优先；
- 熟知双微运营机制，对网络热点具备高度敏感度，能够结合游戏内容产出相应内容；
- 掌握 PS、PR 等软件的基础操作，能够完成配图的制作，掌握手绘者加分；
- 对二次元文化有广泛了解，有日系同人内容产出经验者优先；
- 思维活跃，耐心负责，具备优秀的团队协作能力及良好的抗压能力。

### 3、岗位福利

- ¥100 元/天，猪厂三餐全包，学习平台大！！
- 活泼逗趣上进的团队氛围，还有大神开车带飞；
- 不定期下午茶福利，各种内部福利，免费健身房；
- 表现优异可提供转正机会。

### 4、简历投递

- [wangyiyouxizp@126.com](mailto:wangyiyouxizp@126.com)
- 简历及邮件命名格式：姓名+学校+电话+来自互联镖局
- PS：投递请附上个人游戏经历。

责任编辑（樊鑫，刘丽）

# 征文通知

## 1 会议征文

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

## 2 期刊征文

计算机视觉领域近期相关期刊专刊的征文通知如表 2 所示,包括 Artificial Intelligence, IEEE Trans. On PAMI 特刊等。

## 3 会议简介 (WACV)

WACV (Winter Conference on Applications of Computer Vision) 是计算机视觉领域最权威的学术会议之一,为全球模式识别

领域的研究人员及参与者提供一个交流平台,其内容涉及近期在计算机视觉、机器学习、模式识别方面的最新进展,会议论文集代表了上述领域最新的发展方向和研究水平。

WACV 2020 将于 2020 年 3 月 2 日至 3 月 5 日在美国科罗拉多州斯诺马斯村举办,论文截止时间为 2019 年 7 月 26 日。征稿范围涉及机器学习在图像压缩、机器视觉、人机交互、行人识别、3-D 建模等方向的研究进展。

责任编辑 (贾同)

**表 1 计算机视觉领域相关国内外会议**

会议名称	会议时间	会议地点	截稿日期	会议网站
WACV 2020	2020.03.02-05	美国 Colorado	2019.07.26	<a href="http://wacv20.wacv.net">http://wacv20.wacv.net</a>
EWRE 2019	2019.10.18-20	中国 Shanghai	2019.07.20	<a href="http://www.ewre2019.net">http://www.ewre2019.net</a>
GCA 2019	2019.11.26-29	日本 Nagasaki	2019.08.10	<a href="https://is-candar.org/GCA19">https://is-candar.org/GCA19</a>
ACPR 2019	2019.11.26-29	新西兰 Auckland	2019.08.01	<a href="https://www.acpr2019.org">https://www.acpr2019.org</a>
VBS 2020	2020.01.05-06	韩国 Daejon	2019.09.06	<a href="http://www.videobrowsershowdown.org">http://www.videobrowsershowdown.org</a>
ISBDAI 2020	2020.04.28-30	中国 Hong Kong	2020.03.30	<a href="http://www.isbdai.org">http://www.isbdai.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
Artificial Intelligence	Special Issue on Explainable Artificial Intelligence	2020.03.01
IEEE Trans. On PAMI	Vision Computing in Environmental Perception of Intelligent Vehicles	2020.01.01