

CCF

计算机视觉 专委会简报

COMPUTER VISION NEWSLETTER

2018/03 期
总第 13 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	06
专委新闻	CCF-CV 常委会年度工作会议	07
专委通知	CCF-CV 2018 年委员增选申请	08

科技前沿

专题综述	图像检索年度进展概述	09
热点追击	非接触式掌纹识别：系统、数据、方法	15
	可解释机器学习：从知识引导到特征可视	16

委员风采

委员访谈	北京航空航天大学陆峰教授访谈	17
委员好消息		20

资源平台

开源代码	图像美学质量评价技术	21
数据集	计算机视觉中的医学图像	24
招聘信息		26
征文通知		28

CCF-CV 走进高校系列报告会

第 48 期 重庆大学

时间：2018 年 3 月 28 日

第 48 期走进高校系列活动“人工智能：跨媒体视觉大数据分析与应用”在重庆大学成功举办。本期报告会邀请了澳门大学 C.L.PhilipChen 教授、中科院自动化所王亮研究员、电子科技大学朱策教授、厦门大学纪荣嵘教授以及哈尔滨工业大学左旺孟教授五位专家做特邀报告。报告会由重庆大学通信工程学院承办，通信工程学院张磊研究员为执行主席。



报告会由重庆大学通信工程院副院长谭晓衡教授主持，副校长王时龙教授致欢迎辞。中国计算机学会计算机视觉专委会秘书长王亮研究员向在座的全体师生，对中国计算机学会计算机视觉专委会走进高校系列报告背景和现状进行了介绍。来自重庆大学、西南大学、重庆邮电大学等高校约 300 名师生聆听了计算机视觉领域的专家学者在该领域的最新前沿技术及应用报告。

第 49 期 华东理工大学

时间：2018 年 3 月 30 日

第 49 期走进高校系列活动“计算机视觉与深度学习前沿技术及应用”报告会于华东理工大学成功举办。本次报告会由华东理工大学信息科学与工程学院电子与通信工程系、计算机系联合承办，共邀请南京信息工程大学刘青山教授、上海交通大学熊红凯教授、浙江大学李玺教授及浙江大学章国锋教授四位学者做特邀报告。活动由电子与通信工程系朱煜教授及计算机系陈志华教授共同主持。来自华东理工大学、华东师范大学、上海大学、上海理工大学等多所高校的近 160 名师生及相关企业同行聆听了四位专家的精彩报告。



报告会现场，学生、企业嘉宾都珍惜宝贵机会与四位教授进行了充分的交流。在近四个小时的报告会中，师生们始终保持极佳的精神状态，会议现场讨论学习气氛十分浓厚。

第 50 期 华北理工大学

时间：2018 年 4 月 24 日

由中国计算机学会计算机视觉专委会主办，华北理工大学信息工程学院承办的第 50 期 CCF-CV 走进高校系列报告会“智能计算技术及应用”在华北理工大学 A3 阶 1 成功举行。本期报告邀请了天津大学胡清华教授、中国科学院沈阳自动化研究所丛杨研究员、北京交通大学桑基韬教授出席活动并作精彩报告。报告会由执行主席华北理工大学信息工程学院李志刚教授主持。共有 200 多位师生参加此次会议，聆听了计算机视觉领域的前沿技术成果和最新动态介绍。



此次 CCF-CV 走进高校系列报告会，为现场听众奉献了一场学术盛宴，会场互动热烈，精彩纷呈。会后，与会人员对三位专家和举办方示以由衷的感谢，报告会在热烈的掌声中圆满结束。

第 51 期 江西财经大学

时间：2018 年 4 月 28 日

2018 年 4 月 28 日上午，由中国计算机学会计算机视觉专委会主办，江西财经大学信息管理学院承办的第五十一期 CCF-CV 走进高校系列报告会“计算机前沿技术及应用”在江西财经大学荟庐四楼报告厅成功举行。

本期报告邀请了电子科技大学李宏亮教授、武汉大学陈震中教授、上海交通大学林巍峭教授以及中山大学特聘研究员谢晓华出席活动并作精彩报告。报告会由江西财经大学信息管理学院方玉明教授、杨勇教授和袁非牛教授担任执行主席。



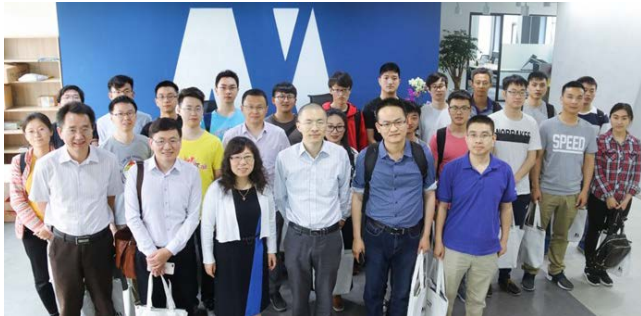
来自南昌大学、江西师范大学、南昌航空大学、东华理工大学、江西农业大学、江西科技师范大学、江西财经大学等单位共 200 多位师生参加本次活动，师生们济济一堂，满怀期待前来聆听计算机视觉领域的专家学者就有关该领域最新前沿技术及应用做报告。

此次 CCF-CV 走进高校系列报告会，内容精彩，互动热烈，高潮迭起，为现场听众奉献了一次学术盛会，一场知识盛宴，会场互动热烈，精彩纷呈，报告会在热烈的掌声中圆满结束。

(责任编辑：马占宇)

CCF-CV 走进企业系列交流会

2018年5月6日，中国计算机学会计算机视觉专委会走进企业系列交流会第十四期活动——走进 Momenta，在北京市海淀区中关村东路东升大厦 C 座 Momenta 会议室成功举办。



活动由 CCF-CV 秘书处召集，Momenta 校园关系负责人侯苏苏和黄娜负责现场组织协调。活动通知发出后，参会申请邮件纷至沓来，大家对这家自动驾驶明星企业与其年轻的创业团队表现出了浓厚的兴趣。由于主办方场地席位所限，最终选择了来自企业、高校、科研院所等三十余名代表参加了此次活动。

在活动开始前，CCF-CV 专委秘书长王亮研究员表达了对主办方的诚挚感谢，向参会人员简要介绍了专委会概况和走进企业活动的宗旨意义。Momenta 公司 CEO 曹旭东领衔，合伙人兼研发总监任少卿博士、合伙人兼研发总监夏炎博士也参加了本次活动，并分别就发展历程与研发概况、自动驾驶的技术及解决难点、深度学习在自动驾驶中的应用与挑战等作了专题报告与 Panel 讨论。

首先，公司 CEO 曹旭东以“可量产的自动驾驶”为题分享了公司的成长历程和发展航路，全面介绍了 Momenta 围绕自动驾驶行业的技术、平台、数据、应用场景所进行

的方向布局和技术研发。之后，公司合伙人兼研发总监任少卿博士带来了无人驾驶技术及解决难点的主题报告。任博士的报告干货满满、技术内涵丰富。听众真切感受到了公司在无人驾驶领域的卓越研发实力和丰硕技术成果。接下来，公司合伙人兼研发总监夏炎博士作为主持人，特邀北京邮电大学邓伟洪教授、清华大学马惠敏副教授、北京理工大学沈建冰教授、中科院自动化所王亮研究员四位嘉宾，围绕深度学习在自动驾驶中的应用与挑战这一主题，在现场展开了 panel 讨论。嘉宾分别介绍了自己的研究领域，并且结合自己的研究课题分享了对于在深度学习、自动驾驶等领域的进展与难点的独到观点。参会的老师同学们纷纷与嘉宾展开了热烈地互动，问答环节异彩纷呈。



活动间歇，Momenta 工作人员为参会人员精心准备了茶点，大家边讨论交流边观摩公司相关技术演示，充分体验了自动驾驶领域的快速进展与技术魅力！活动结束后，公司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 常委会年度工作会议

2018年5月12日,中国计算机学会计算机视觉专委会(CCF-CV)常务委员会2018年度工作会议于扬州会议中心圆满召开。除3名委员因事请假之外,专委会现任常务委员全部出席了本次会议,专委会秘书处成员列席会议。



本次常委会工作会议系统地总结了一年以来专委会各项工作的进展以及存在的不足,并针对专委会今后的发展方向,提出了多项富有建设性的建议。



作为 PRCV2018 大会程序委员会主席、专委会副主任、中山大学的赖剑煌教授代表大会组委会汇报了“2018 第一届中国模式识别与计算机

视觉大会(PRCV2018)”的各项筹备情况。PRCV2018 是本年度专委会最重要的学术活动之一,将于11月份在广州举行。届时大会将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)、中国图象图形学学会(CSIG)联合主办,由中山大学承办。会议目前处于论文征稿的环节,国内相关领域的师生们投稿踊跃。本届 PRCV 大会的论文集将由 Springer 出版社出版,并被 EI 和 ISTP 检索;优秀论文经扩展后将被推荐至国内外高质量期刊的特刊发表。会议的其它各项组织工作目前也正在有条不紊地推进中。



西安紧随广州之后,成为 PRCV2019 的大会举办地,届时将由西北工业大学承办。本着提前筹备、早做规划的精神,PRCV2019 的组织筹办工作也提上了日程,大会承办主席、专委会常委、西北工业大学张艳宁教授向常委会通报了会议场地规划、会议日期、论文征稿计划等相关工作的进展情况。

随后,常委会针对秘书处所整理的专委会发展相关议题展开讨论,专委会委员们积极建言献策,

逐一形成了切实可行的指导性建议，为专委下一步发展明确了方向。秘书处对相关议题进行了分类汇总，涉及专委日常工作、系列活动、组织建设、CV 领域发展 4 个方面共 26 项议题。其中，值得广大委员们期待的专委发展新举措包括：举办区域性专题方向 workshop 以促进学术深度交流、举办青年学者寒、暑期学校以推广领域前沿知识并惠及广大青年学者、后续筹划 CV 领域中期刊建设等等。



会议最后由谭主任作了总结发言。会议结束后，全体与会委员合影留念。

（责任编辑：王瑞平）

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）

图像检索年度进展概述

新加坡科技设计大学 郑良

一、引言

近十年来，随着互联网和计算机技术的迅猛发展，人们进入了大数据的时代。其中，呈几何级数增长的图像数据是大数据时代重要的组成部分。合理利用图像检索技术，对海量图像数据进行有效的管理，是一项非常重要并且具有挑战性的课题。在图像检索中，如何对图像进行鉴别

型在图像检索中的工作流程如图 1 所示。局部特征，例如 SIFT，是一个高维向量，为了将图像中的众多 SIFT 特征向量化，一般会在离线阶段训练一个视觉词典，并将所有局部特征量化成词典中的一个或多个视觉单词。由于词典的规模是有限的，当局部特征被转化为单词后，便可以对一幅图像生成一个词典长度的直方图。在离线过程

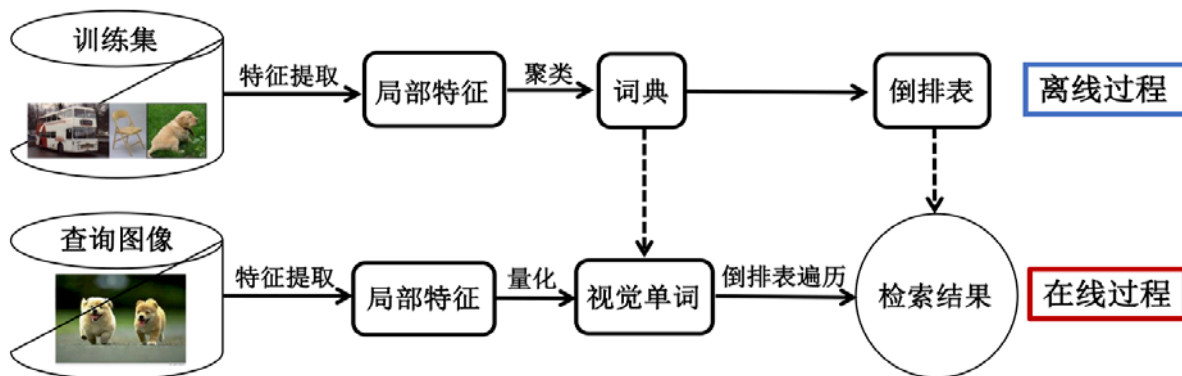


图 1 图像检索框架图

性视觉表达，以及如何建立高效的索引结构，从而实现对图像数据库快速而准确的检索，是当前计算机视觉领域一项重要的前沿研究课题。

鉴于全局特征对图像描述能力的欠缺，当今的图像检索研究主要集中在基于局部特征的视觉模型上。目前，广泛应用的局部特征是 SIFT[1]，与全局特征不同之处在于，它具有较优异的对选择、尺度、亮度等方面的不变性，因此可以减轻图像的变化对检索性能的负面影响。因此，在计算机视觉和多媒体应用等领域，基于局部特征的视觉表达被普遍使用。

局部特征为图像检索研究打下了坚实的基础，使得图像检索可以充分借鉴文本检索的框架。在文本检索中，由于单词的存在，使得一个文本可以方便地被向量化，表达成单词的直方图。沿着这个思路，为了对图像也进行有效而紧凑的表达，研究人员提出了词袋模型[2]，该模

中，首先对训练图像进行特征提取，并聚类生成词典，然后将测试数据库图像利用词典生成倒排表。在在线过程中，给定查询图像，在特征提取之后，利用词典进行量化，然后进行倒排表遍历，得到图像排序结果。至此，图像检索便从形式上与文本检索实现了统一。

局部特征是词袋模型的基础，尽管它对图像变化具有良好的鲁棒性，比如实现了尺度、光照、旋转不变，但局部特征对图像的描述能力依然有限，导致对图像中很多特性的描述不够。同时，在量化过程中，局部特征的大量信息发生丢失，视觉单词只继承了原特征很小一部分的鉴别信息，因此，词袋模型的鉴别能力进一步下降。实际上，基础的词袋模型的鉴别力是比较弱的，需要加入更多鉴别性信息。这种信息可以从局部特征本身挖掘，通过恢复原始特征的一些鉴别信息来实现；另一方面，也可以从其他互补特征上挖

掘，通过引入与局部特征互补的描述子，从侧面增强词袋模型的鉴别力。在这个问题上，存在几种潜在的研究方向，例如，如何对视觉单词间的几何关系进行建模？如何将局部的互补鉴别信息融合到原模型中？如何将更大尺度上的鉴别信息与局部鉴别信息有效地融合到一起？这些都是非常具有挑战性的问题。

此外，图像检索的时间和内存消耗与检索准确度往往有着直接互相制约的关系。在词袋模型中，当视觉词典较大时[3]，由于图像中的单词数有限，图像的单词直方图往往是非常稀疏的。此时，模型一般采用倒排表这种索引策略对图像进行索引（如图 1 所示），从而避免了直接存储稀疏直方图带来的时间和内存消耗。然而，当视觉表达发生改变（对原始特征上下文关系进行建模、引入新特征等）时，如何对倒排表索引进行相应的调整，以降低内存和时间开销是一个重要的问题。同时，一幅图像可能由成百上千个视觉单词组成，而文本检索中，查询文本往往只有几个单词。因此，较大的查询特征数很可能是制约图像检索速度的瓶颈。那么，如何在保证检索精度的前提下，有效地降低查询图像中的特征数，将是提高检索效率的一个有效措施。

二、图像检索研究现状

1. 深度嵌入特征融合

导致 SIFT 视觉单词鉴别力低有两个主要因素，一是量化带来的误差，二是 SIFT 只能描述局部梯度特征。第一个因素是词袋模型固有的；而第二个因素则是 SIFT 特征固有的。为提高视觉单词鉴别力，Zheng 等人[2]在特征融合上做工作，同时考虑局部、区域和全局特征的融合，这样可以在更大的程度上提高 SIFT 视觉单词的鉴别力。基于这个想法，他们提出一个全新的匹配策略，即两个关键点是匹配的，当且仅当它们在局部、区域和全局三个层次上都匹配，如图 2 所示；而另一方面，如果在三个层次上，有一个层次不匹配，则认为两个关键点是不匹配的。

特别地，为了描述区域和全局特征，作者采用卷积神经网络（Convolutional Neural Network, CNN）[3]来提取特征。该网络结构已被应用在多种计算机视觉问题中，如图像分类[4]和目标检测[5]，并取得了优异的性能。可是在图像检索领域，CNN 特征还使用的比较少，尤其是如何将 CNN 融入词袋模型的方法。为此，首先将 CNN 特征进行归一化，再将其送入概率融合模型。实验证明，CNN 特征可以为 SIFT 提供互补性信息，而且相比于颜色直方图特征，对图像变化具有更好的鲁棒性。



图 2 正确匹配的关键点

由于 CNN 特征是 4096 维的，当把它存储在倒排表中，会导致时间和内存开销大大增加。比如，存储 100 万张图像会消耗 1 200 G 内存，而且对于这种长度的浮点数向量，计算相似度也是非常耗时的。考虑到时间和内存的高效性，作者将浮点数的高维 CNN 向量转化成低维的二进制向量。在这一步，使用了局部敏感哈希（Locality Sensitive Hashing, LSH）算法。在 LSH 中，每个哈希码是将特征向量与一个随机平面点积之后的四舍五入结果。其每一维是从一个 0 均值高斯 $N(0, I)$ 中随机选取的。对每个 CNN 向量 x ，通过重复投影 b 次，可以用 b 个哈希函数产生 b 个哈希码。在实验中， b 设为 128，因此将每个 4096 维的 CNN 特征转化为一个 128 维的二进制特征。在匹配两个关键点时，由于每个关键点耦合了两个区域特征，因此计算四组区域特征间的汉明距离，并将最小距离作为衡量标准，这样便可以实现一定程度上的尺度不变性。

2. 多维索引

为了增强视觉单词的鉴别力，对每个关键点，Zheng 等人提取两个特征[6]，即 SIFT 和颜色特征。在多种特征存在的情况下，建立多维索引表是重要的技术。Babenko 等[7]提出将 SIFT 特征切割为 n 段，每段训练一个词典，这样 n 个词典就对应了一个 n 维倒排表。但是，这个多维倒排表是用于近似最近邻搜索（Approximate Nearest Neighbor, ANN）任务的，与图像检索有本质不同。这是因为，最近邻搜索只需要找到查询特征的最近邻，这个最近邻只有一个；而图像检索则需要找到所有查询特征的所有近邻，这些近邻可能会有很多。因此，在图像检索中，不仅要保证足够的近邻查询精度，也要保证一定的召回率。因此，文献[8]中提出的多维倒排表，在图像检索中实际上是不适用的。

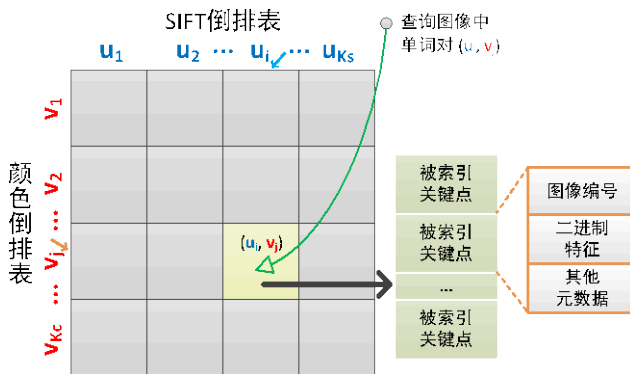


图3 耦合二维倒排表结构

为了解决这个问题，Zheng 等人[6]提出了耦合式多维索引表，如图3所示。该索引表的每一维表示一种特征。令 $x=[x_s, x_c] \in \mathbb{R}^{\hat{D}(s+c)}$ 是共同描述关键点 p 的一对描述子，其中 $x_s \in \mathbb{R}^{\hat{D}_s}$ 、 $x_c \in \mathbb{R}^{\hat{D}_c}$ 分别是维度为 D_s 和 D_c 的 SIFT 和颜色特征。对 SIFT 和颜色特征，作者训练两个词典 $U=\{u_1, \dots, u_{(K_s)}\}$ 和 $V=\{v_1, \dots, v_{(K_c)}\}$ ，词典大小分别为 K_s 和 K_c 。因此，耦合二维倒排表包含 $K_s \times K_c$ 个词条，

这些词条表示为 $W=\{W_{11}, W_{12}, \dots, W_{(K_s K_c)}\}$ ， $i=1, 2, \dots, K_s$ ， $j=1, 2, \dots, K_c$ 。

在构建耦合二维倒排表时，作者使用词典 U 和 V ，将所有特征对 $x=[x_s, x_c]$ 都量化为视觉单词对 (u_i, v_j) ， $i=1, 2, \dots, K_s$ ， $j=1, 2, \dots, K_c$ ，其中 u_i 和 v_j 分别是词典 U 和 V 中与特征 x_s 、 x_c 距离最近的单词。然后，就可以将关键点的元数据（图像编号，颜色二进制特征等）存在词条 W_{ij} 中。在查询时，给定一个特征对 $x=[x_s, x_c]$ ，首先将它量化为单词对 (u_i, v_j) ，然后，在耦合二维倒排表中找到对应的条目 W_{ij} ，并找到候选图像列表进行加分。实际上，对于两个特征对 $x=[x_s, x_c]$ 和 $y=[y_s, y_c]$ ，它们的相似性度量耦合二维倒排表结构。SIFT 和 CN 特征的词典大小分别为 K_s 和 K_c 。

因此，一对正确匹配的局部关键点，必须满足在 SIFT 和颜色特征空间中都相似。在二维的情况下，即表示两种特征在倒排表层次的融合。实验中，耦合二维倒排表的检索准确性达到了国际先进水平，同时其检索时间也大大降低，这是因为，二维倒排表缩短了每个倒排列表的长度，从而减少了倒排表遍历的时间损耗。在后续工作中，作者又进一步将耦合二维倒排表扩展到多维，实现了多词典、多倒排表的联合检索，从而大大增强了检索系统的准确性。

3. 深度学习端到端训练图像检索

由于 CNN 是在分类任务的数据集 ImageNet[8]上训练的，故其特征虽能区分出不同类别的目标或场景，但有可能存在难以描述同一类别的不同实例的情况。例如巴黎圣母院和埃菲尔铁塔具备相同建筑结构，但预训练的 CNN 特征难以将两者区别开。因此针对具体任务去 fine-tune 预训练的 CNN 以得到更好的特征是很重要的。

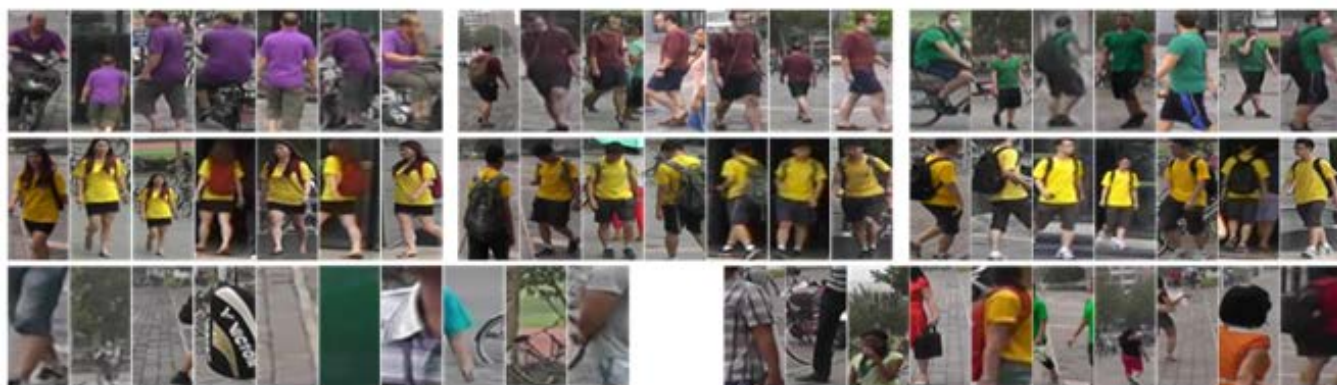


图 4 Market-1501 数据集中的图像样例

通过 fine-tune 预训练的 CNN 能够端到端地得到图像的特征表达，也就是说 CNN 经过 fine-tune 之后不需要其他编码或池化处理。用于 fine-tune 的 CNN 网络结构可以归纳为两种：分类结构和验证结构。分类结构是指针对当前任务的类别去训练一个分类网络，并将训练好的分类网络的某一层特征用于图像检索。文献[9]就采用这种分类结构来获取图像的特征。验证结构是指采取 Siamese 网络并通过 pairwise loss 或 triplet loss 来训练网络。在文献[10]中，Radenović等引入 3D architecture model 为网络选取样对，并将网络最后一层修改为 MAC[11]，最后采用 contrastive loss 来 fine-tune 网络。在同时期的文献[12]中，Gordo 等采取 triplet loss 来 fine-tune 网络并引入 region proposal network 来选取图像的区域。文献[10]和文献[12]通过端到端 fine-tune 以获取图像特征的方式在地标检索任务上取得了比较好的性能。

4. 行人重识别——图像检索的新应用

行人重识别是图像检索的一个子任务，具有重要的学术和应用价值。Zheng 等人[13]提出一个新的行人再识别数据集，即 Market-1501 数据集。在数据采集阶段，将六个摄像机摆放在清华大学超市附近，其中包括 5 个 1280×1080 高清相机和一个 720×576 标清相机。这些相机之间存在一定的场景重叠。共标定了 1501 个不同行人，共 32643 个检测框。由于数据采集环境是开

放式的，每个人最多被 6 个相机拍摄到，而且在标定过程中，保证每个被标定的行人至少在两个相机中出现过，这样就可以进行跨摄像头的识别任务。在每个相机下，同一个人呈现各种不同的外表和姿态，如图 4 所示。总体而言，Market-1501 数据集具有如下几个特点。

首先，针对大多数现有数据集使用手标的边界框的问题，Market-1501 数据集采用了国际先进的检测器，即可变形部件模型 (Deformable Part Model, DPM) [8] 来检测行人。手标边界框的对齐效果是比较好的，但使用检测器后，行人对齐往往存在一定问题，而这正是实际中会遇到的。在检测到的行人图像中，对齐出错和身体部件缺失是非常常见的。

第二，除正确检测的边界框之外，作者还提供了虚警检测结果。作者注意到，在 CUHK03 数据集[9]也使用了 DPM 检测器，但该数据集提供的边界框大多数都是较好的检测结果(对齐较好)。但实际中，检测到的边界框可能很差；也就是说，边界框无法对行人真值提供有效的描述。考虑到这个情况，对每个需要标定的边界框，作者首先手标一个边界框，然后，计算手标边界框和 DPM 检测边界框之间的覆盖程度，即两个边界框交集与并集的面积比。在 Market-1501 数据集中，如果该面积比大于 50%，当前 DPM 边界框被标为“好”；如果该面积比小于 20%，将该边界框标为“干扰”图像；如果面积比大于 20%而小于 50%，该边界框被标为“中性”，即这幅图像对

行人再识别性能不产生影响。同时，一些明显背景上的边界框也被标为干扰图像。在图 4 中，前两行为“好”的图像，第三行左边和右边分别是干扰图像和中性图像。可以看到，在该数据集中，行人的姿态、分辨率等都有较大程度的类内差异，这使得 Market-1501 数据集非常具有挑战性。

第三，对于每个被标定的行人，他在每个相机下可能会有多幅图像。因此，在跨相机查询时，每个行人可能会有多幅查询图像，或者多幅相关真值图像。一方面，这与实际应用是相符的，特别是可以从多个查询图像中得到更多的关于该行人的鉴别信息；另一方面，对于一个性能优良的行人再识别系统来说，对于一个查询行人，一个理想的识别结果是将该个体所有相关图像都找出。从这个意义上，该数据集为开放环境下的行人再识别研究提供了数据保障。

为搭建图像检索与行人再识别间的桥梁，[13]提出了基于词袋模型的行人描述子，算法如下。第一，特征提取。使用颜色名称(CN)描述子。给定一幅归一化到 128×64 的行人图像，对其稠密采样，每个图像块的大小为 4×4 ，采样步长为 4，所以图像块之间是没有重叠的。对每个图像块，计算其中所有像素的 CN 特征向量，每个向量被一范数归一化并开根号。计算图像块中的 CN 特征的平均值，作为该图像块的特征。第二，训练词典。对于 Market-1501 数据集，在其训练集上训练词典；而对于其他数据集，则在独立的 TUD-Brussels 数据集上训练。使用经典的 k 均值

聚类算法，其中词典大小为 k。第三，特征量化。给定一个局部特征，使用多赋值法(MA)，在词典中找到它的近邻单词。MA 设为 10，即一个特征由 10 个视觉单词表示。

三、未来发展方向

图像检索是大数据时代对海量数据进行分析的关键工具，具有重要的研究意义和应用价值。随着局部不变性特征的提出与发展，基于词袋模型的图像检索成为计算机视觉领域最为活跃的研究方向之一。本文主要关注基于词袋模型的图像检索及其相关应用，并分别从鉴别性的视觉表达、高效性的索引结构，以及图像检索在行人再识别中的应用三个方面进行了讨论。

图像检索是一项非常具有挑战性的研究课题。在计算机视觉和多媒体领域，随着图像数量的指数级增长，实时准确的图像检索技术及其相关应用，已受到越来越多研究人员的关注。目前，虽然基于词袋模型的图像检索技术已经取得了一定的进展，但是，由于图像的多样性、局部特征在描述能力上的局限性，以及词袋模型在特征鉴别力上的损失，图像检索依然面临着很多亟待解决的问题。因此，我们认为未来发展方向包括倒排表压缩与安全中止技术、基于深度学习的图像检索，以及小物体检索和语义层图像检索等。

(责任编辑：杨巨峰)

参考文献

- [1] Lowe D G. "Distinctive image features from scale-invariant keypoints." *International Journal of Computer Vision*, 2004, 60(2):91-110.
- [2] Zheng L et al. "Accurate image search with multi-scale contextual evidences." *International Journal of Computer Vision*, 120(1), 1-13, 2016.
- [3] LeCun Yet al. "Gradient-based learning applied to document recognition." *Proceedings of the IEEE*, 1998, 86(11):2278-2324.
- [4] Krizhevsky A et al. "Imagenet classification with deep convolutional neural networks." *Proceedings of Advances in neural information processing systems*, 2012: 1097-1105.
- [5] Girshick R et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*.2014:580-587.

-
- [6] Zheng L et al. "Packing and padding: Coupled multi-index for accurate image retrieval." IEEE Conference on computer vision and pattern recognition. 2014.
 - [7] Babenko, Artem, and Victor Lempitsky. "The inverted multi-index." Computer Vision and Pattern Recognition, 2012.
 - [8] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in CVPR, 2009.
 - [9] A. Babenko, A. Slesarev, A. Chigorin, and V. Lempitsky, "Neural codes for image retrieval," in ECCV, 2014.
 - [10] F. Radenović, G. Tolias, and O. Chum, "Cnn image retrieval learns from bow: Unsupervised fine-tuning with hard examples," in ECCV, 2016
 - [11] G. Tolias, R. Sivic, and H. Jégou, "Particular object retrieval with integral max-pooling of cnn activations," in ICLR, 2016.
 - [12] A. Gordo, J. Almazán, J. Revaud, and D. Larlus, "Deep image retrieval: Learning global representations for image search," in ECCV, 2016
 - [13] Zheng L et al. "Scalable person re-identification: A benchmark." In Proceedings of the IEEE International Conference on Computer Vision. 2015: 1116-1124.
-



郑良

新加坡科技设计大学助理教授。于 2010 年和 2015 年从清华大学获本科和博士学位，之后在美国德州大学和澳大利亚悉尼科技大学从事研究工作，并于 2018 年加入新加坡科技设计大学。主要研究方向包括：计算机视觉、机器学习、行人重识别等。已在 IEEE PAMI、IJCV 等国际期刊和国际会议发表论文 20 余篇。Google Scholar 引用 1800 余次，最高单篇他引 350 余次。

Email: zheng_liang@sutd.edu.sg

非接触式掌纹识别：系统、数据、方法

同济大学 张林

掌纹识别是生物特征识别技术的一种，通过在可见光下采集到的手掌内侧固有的皮肤纹理信息来鉴别人的身份，在很多领域有着广泛的应用前景。

根据掌纹图像采集设备的类型，掌纹识别系统可以分为接触式掌纹识别系统和非接触式掌纹识别系统。接触式采集系统会设计某些机械支撑部件来固定手的位置、限制手掌的自由度。这种设备存在固有的缺陷，比如使用者会担心有疾病的传播，同时，用户体验也不是很好。

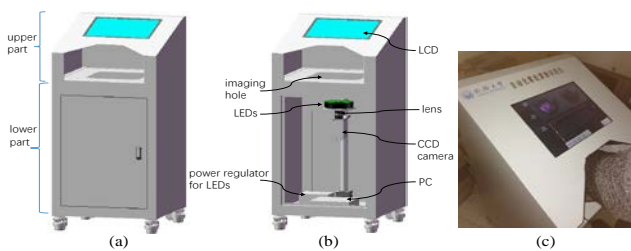


图 1 (a) (b) 是掌纹采集 CAD 模型；(c) 是实际系统外观

近年来，很多学者也在探索非接触式掌纹识别技术，但在这个领域依然存在很多没有解决的问题：1) 现有的非接触掌纹采集设备都设计的比较简陋，处于实验室原型状态，很难投入实际使用，设计上也很少考虑用户接受度问题；2) 采集的图像质量普遍不高，不能回答“非接触掌纹到底能达到多高识别率”这个问题；3) 公开的非接触掌纹图像数据库规模普遍很小，难以对识别算法进行有效的验证；4) 没有有效地针对非接触掌

纹进行分类的方法。

针对上述问题，我们开展了如下几方面工作：

1) 设计了一款非接触式掌纹采集设备，该设备易于使用、用户接受程度较高，符合人体工学特点。图 1 展示了此设备的三维 CAD 模型以及实际制作出来的外观效果。

2) 采集了一个较大规模的非接触掌纹数据集。包含了来自 300 个人、600 个手掌、2 次采集、每次每只手 10 个样本。共计包含 12000 个图像样本。

3) 首次对掌纹图像的质量做了客观化分析和度量。

4) 针对非接触式掌纹的特点，提出了基于协同表示 (Collaborative Representation, CR) 的快速识别方法 CR_CompCode。该方法提取掌纹 ROI 部分的 CompCode (竞争编码) 来作为特征向量；然后，用 CR 作为分类框架。在 Matlab 实现下，该方法可以在 13ms 左右的时间内完成一次 identification 操作 (注册数据集包含 600 个类 6000 个样本)。

该研究的相关成果发表在 IEEE Trans. PAMI、Pattern Recognition 等国际期刊。有关的数据和代码已经公开发布在 github 网页 <https://cslinzhang.github.io/ContactlessPalm/>。

(责任编辑：任传贤)



张林

CCF 计算机视觉专委会委员，同济大学副教授，主要研究方向是生物特征识别、多

媒体质量评价、智能驾驶中的视觉感知等。

Email: cslinzhang@tongji.edu.cn

可解释机器学习：从知识引导到特征可视

清华大学 苏航 董胤蓬 王宇龙

目前以深度学习为代表的机器学习算法在很多任务上的性能都获得了突破性进展，但这些模型通常因为很难被理解而被当做“黑盒”使用，从而给机器学习模型的应用带来了巨大的障碍。在诸多场景中，不但要求系统具有良好的性能，同时要保证其行为是可解释、可预测、符合测试的基本假设和原理的，这是机器学习系统大规模部署的一个重要前提。这就迫切需要研究可解释机器学习模型，将“黑箱”模型转化为“白箱”，使得机器学习模型在给出预测结果的同时，可以提供其决策的依据，为更加有效的人机沟通和协作奠定基础。

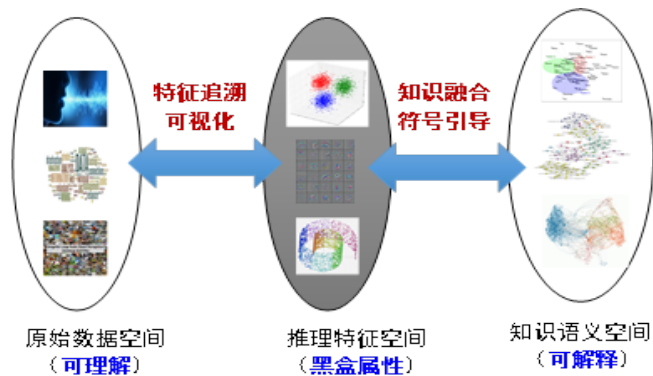


图1 可解释机器学习

从本质上来说，可解释机器学习的核心难点是跨越数据特征空间和人类语义空间之间的鸿沟。当代机器学习算法的核心思想是将观测样本映射到特征空间当中，进而在特征空间进行分析，而特征空间通常是人类所无法解析的。人类可以

理解的是原始数据空间和知识语义空间，因此机器学习模型实现可解释的核心是把特征空间和人类可理解的另外两个空间进行关联。

对于大量感知类的任务来说(例如图像分类、语音识别等)，由于通常无法对其结果给出明确的原因解释，我们可以通过将决策的结果追溯到原始数据空间，利用数据进行解析。针对这一问题，我们利用可视化的方法，将深度学习模型的隐层特征和数据样本进行关联，从而帮助专家更好地理解与诊断深度学习模型的训练和决策过程。

而对于人类更加高层的推理、决策等认知层面的任务来说，我们则是通过将人类的背景知识引入到深度学习模型的决策过程中，利用知识嵌入、知识后验等方法，将符号化的人类知识引入以深度学习为代表的驱动系统，研究知识在统计模型中的应用机制，使得知识的利用与数据优化的目标能够有效的结合在一起，建立知识语义空间和特征空间之间的联系，进而提升机器学习模型的可解释性。

可解释机器学习模型可以使用户更加正确地使用机器学习模型，逐步改进和完善机器学习模型，尤其是在一些高风险的任务中(如工业、金融、医疗、军事等)。近两年来，课题组在机器学习的可解释性等方面发表包括 CVPR、IJCAI、TVCG 等多篇论文，相关工作得到了包括 Google、CMU 等国内外学者的广泛关注。



苏航

清华大学计算机系助理研究员，CCF 计算机视觉专委会委员，CSIG 视觉大数据专委会

委员。在 CVPR、ECCV 等国际会议和期刊发表论文 50 余篇，曾获得 MICCAI “青年学者”、AVSS “最佳论文” 等奖项。

Email: suhangss@mail.tsinghua.edu.cn

北京航空航天大学陆峰教授访谈

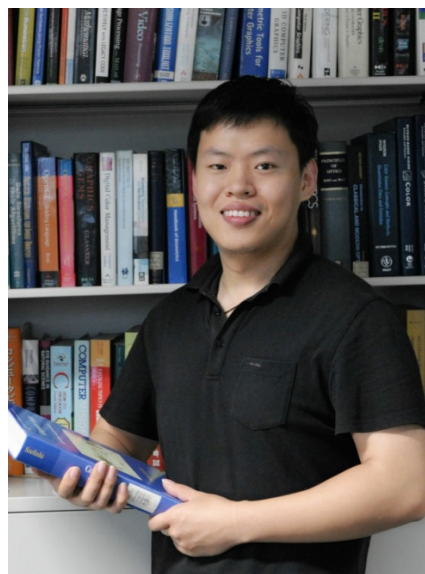
2018年5月3日,专委秘书处采访了北京航空航天大学计算机学院青年千人计划陆峰教授。下面是采访实录。

陆老师,您在虚拟现实、增强现实和计算机视觉领域都取得了突出的成就,入选“千人计划”青年项目,入围国家第十二批海外高层次人才引进名单,能否具体介绍一下您现在主要的研究聚焦点,以及您在这个聚焦点上的研究历程呢?

我们团队的主要研究方向包括虚拟现实、增强现实、计算机视觉、人机交互、人工智能等。目前的重点研究思路是,发展智能视觉感知与多通道人机交互两条主线技术,并通过二者的结合,探索人机协同的增强智能的实现,并应用于航空航天、医疗卫生、公共安全等重要领域中。截止目前,在物体三维重建、材质属性解析、视觉场景理解、人眼眼动计算、手势体感捕捉、语音信息利用等方面,我们均获得了一定的研究成果或正在积极探索。同时,我们也在将人机协同的增强智能这一概念及其相关的核心技术推广应用于民用航空、智能医疗等具体场景中。

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?

在清华大学自动化系 BBNC 实验室,我的主要研究方向是视觉多媒体信息处理。因此,在出国读博时,选择了比较相关的计算机视觉方向,并进一步涉足人机交互等方向。在回国后,随着人工智能的飞速发展,我意识到计算机视觉与人机交互技术的融合将更好地推动人工智能、虚拟现实等领域的发展,于是确定了新的研究思路。我认为在科研中,明确自己的研究方向很重要。只要大方向把握住,具体的困难都很容易克服。



您在计算机视觉等领域顶级期刊 TPAMI、TIP, 顶级会议 CVPR、ICCV 上均发表了系列文章,请问您对发表高水平论文有什么建议呢?这些文章里面的工作,您印象最深的是什么呢?

我认为能够发表高水平论文,代表个人的研究工作得到了学术界的公开认可,还是较为不易的,因此需要各方面的努力做到极致,包括研究问题的意义、提出方法的创新性、实验求证的严谨性、以及写作表达的质量等。我感觉,作为研究者,如果能够具有一种“完美主义”的自我要求,会很有帮助。我个人印象最深的是 2011 年第一次在 ICCV 上发表论文,从毫无经验到最终成功,导师给了我巨大帮助,这一过程让我对如何做高质量研究有了直观认识,算是一种跃迁。

您对以后的研究有什么规划呢?是否会坚持走目前的聚焦点,还是会往其他的方向拓展呢?

在通过感知与交互技术实现增强智能这一总体研究路线上,我会继续坚持下去,因为我对

这方面比较有兴趣，而且它也足以涵盖计算机视觉、人机交互、人工智能等领域的众多具体方向。而在具体的研究内容上，会不断拓展思路，因为目前计算机视觉及其相关领域发展速度都很快，很多新技术新思想都可以被快速吸纳和融合。

您是在日本获得博士学位的，能否和众多研究生们分享一下您读博士期间的奋斗历史？在日本读博士和在中国读博士的区别有哪些呢？各自有什么优缺点？

我在国外读博的经历还是比较顺利的。其实在国外最大的困难往往还是在生活上，需要具备一定的自主生活能力，能够处理生活中的各种不便和意外。比如我在国外时，碰到的最大问题经常是突然生病和找房搬家之类。在研究方面，对于博士生来说，在国内或国外读博的差别主要是一些外部因素，比如师生合作模式、环境氛围等。虽然对于读博的过程有不小的影响，但其实最终一个博士需要培养具备的能力和品质都大同小异，因此近年来国内博士的水平同样非常高。

在日本读完博士后，您为何选择了回国工作？又是什么样的机遇让您留在了北京航空航天大学呢？

对我本人来说，出国学习从一开始就是一个过渡状态，因此在学成回国这一点上是自然而然的，可能主要还是考虑家庭因素吧。而加入北航是有一定偶然性的，因为在计算机视觉方面，国内很多高校和课题组实力都很强。最终综合考虑城市、学科、平台等因素，加上北航对于引进人才的重视，最终选择加入北航。

您现在已经是博士生导师了，请问您是如何选拔博士生，以及培养博士生的？能否和大家分享一下您的经验？

在选拔博士研究生时，我的经验是重点考察学生是否具有成为合格博士的基本素质，例如反应敏捷、敢于尝试、不怕出错、心理承受力强等。

这些素质是攻读博士学位，挑战未知过程中不可缺少的条件。事实上，我在去年招收的直博生，虽然本科成绩并非最好，但是基本具备上述特点，因此科研工作开展很顺利，在遇到困难和失败时也能够积极寻找解决方法。我认为这种积极性和不畏困难的心态是博士培养中的关键。

作为 80 后的新秀，您是如何平衡家庭和事业的？您又是如何安排业余生活，如何在繁忙的工作之余给自己充电的呢？

作为 80 后，在人生的当前阶段，不可避免要承担来自社会、个人和家庭各方面的责任和压力。坦白说，近些年来，我的主要精力都放在了工作上，业余生活相对单调，对家庭的照顾也非常欠缺。因为我家住在顺义，距离学校接近 30 公里，每天回家后，往往已经过了午夜，家人孩子也都已经休息，感觉很有亏欠，也很感谢家人的支持。这可能也是和我同龄的很多人面临的共同难题，需要做出更好的平衡。

您最想和众多博士生、学术研究路上的菜鸟们分享的是什么呢？

说一些我近年来自身的感悟吧。首先，在人的的一生中，能够投身科研，成为一个合格的博士，是一笔巨大的财富。这种收获不仅是学术上、事业上，更体现心态上、意志上。其次，现在从事计算机视觉领域的研究，恰逢其时，赶上了很好的历史机遇。相比从前任何时候，都有可能获得更多的高质量研究成果，并且形成真正落地的应用。最后，也是最重要的，这个时代切忌浮躁，大量研究工具的出现使做出成果变得看似容易，也导致越来越多的人忽略了对理论基础的真正掌握。作为博士生，我建议不要过于急功近利，多和真理较劲、多和自己较劲，把基础打牢，仍然是成为一个合格博士的必由之路。



陆峰

北京航空航天大学计算机学院青年千人计划教授、虚拟现实技术与系统国家重点实验室成员。2016年通过中组部“千人计划”青年项目评选，入围国家第十二批海外高层次人才引进名单。于清华大学

自动化系获学士、硕士学位，于日本东京大学信息学院获博士学位。2013年起任东京大学特任助理教授。曾在松下新加坡实验室、微软亚洲研究院等科研院所从事科研工作。主要研究方向：虚拟现实、计算机视觉、人工智能、人机交互等。具体研究内容：基于视觉的三维模型重建、基于视觉的物体属性理解、人眼视线跟踪交互、图像视频处理和编解码、以及上述技术在虚拟现实与增强现实中的实际应用。长期研究目标：让计算机“精准快速”理解视觉场景，并与人类“自然高效”进行交互，在此基础上达成真正沉浸式、可交互、高智能的虚拟现实和增强现实。理论

研究与具体实践并重。近期在计算机视觉等领域顶级期刊 IEEE Trans. Pattern Analysis and Machine Intelligence (TPAMI)、IEEE Trans. Image Processing (TIP)、顶级会议 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)、IEEE International Conference on Computer Vision (ICCV) 等国际知名期刊杂志上发表系列文章。参与多项国际合作科研项目，深入研究人类交互行为、材质感知等科学问题；主持包括 2 项国家自然科学基金项目在内的多项科研项目；担任 IEEE TPAMI、IEEE TIP、IJCV、IEEE TIP 等二十余种顶级和重要国际期刊长期审稿人，及 IEEE CVPR、ICCV、ECCV、ICCP 等十余种顶级和重要国际会议的审稿人；担任多个国际期刊的编委和客座编辑；担任多个领域重要国际会议的程序委员会委员；担任中国计算机学会 (CCF) 虚拟视觉专业委员会委员、计算机视觉专业委员会委员。

(责任编辑：余焯、黄岩、张汗灵)

委员好消息

☛ 2018年3月9日获悉，中共中央组织部办公厅下发《关于印发第三批国家“万人计划”入选人员名单的通知》。CCF-CV专委会有8位委员入选科技创新领军人才：北京交通大学**赵耀**教授、北京理工大学**黄华**教授、东南大学**孙长银**教授、南京理工大学**唐金辉**教授、上海交通大学**熊红凯**教授、西安交通大学**薛建儒**教授、中国科学技术大学**张勇东**教授、中国科学院计算技术研究所**山世光**研究员；7位委员入选青年拔尖人才：华中科技大学**白翔**教授、南开大学**程明明**教授、西安电子科技大学**董伟生**教授、厦门大学**纪荣嵘**教授、复旦大学**姜育刚**教授、中山大学**林惊**教授和西安交通大学**孙剑**教授。第三批国家“万人计划”入选人员名单包括两个层次的五类人才，共计1635人，其中科技创新领军人才720人，科技创业领军人才367人，哲学社会科学领军人才215人，教学名师195人；青年拔尖人才138人。

☛ 2018年3月11日，由CCF-CV专委会主任、中国科学院院士、中科院自动化所**谭铁牛**研究员，CCF-CV专委会秘书长、中科院自动化所**王亮**研究员与中科院自动化所吴书副研究员合著的《Context-aware Collaborative Prediction》一书由Springer出版。该书介绍了三种最新的情境感知协同预测方法（情境表达、分层表达和情境感知递归神经网络）及其相关应用。

☛ 2018年3月14日，CCF-CV专委会主任、中国科学院院士、中科院自研所**谭铁牛**研究员当

选中国人民政治协商会议第十三届全国委员会常务委员。

☛ 2018年3月23日，科技部正式发布《2017年中国独角兽企业榜单及趋势研究报告》，共有164家企业上榜，由CCF-CV专委会委员**孙剑**出任首席科学家的旷视科技位列其中。该报告是中国最权威、含金量最高的独角兽认证。上榜独角兽企业必须是在中国境内注册的、具有法人资格、成立时间不超过10年、获得过私募投资且尚未上市、企业估值超过10亿美元的企业。

☛ 2018年4月28日，中国计算机学会公布了“YOCSEF二十周年最具影响力的论坛及主持人奖”获奖名单，该奖项颁给10个论坛及18位专业人士。CCF-CV专委会副主任、爱奇艺公司首席科学家**王涛**等主持的“科研人员为什么成了贼？”、CCF-CV专委会常务委员、中科院计算所**山世光**研究员等主持的“从屠呦呦获诺奖看中国科研评价体系”获奖。

☛ 2018年5月10日，中国计算机学会公布了“YOCSEF二十周年纪念”贡献最大荣誉委员奖，共有8人获奖。CCF-CV专委会委员、北京理工大学**黄华**教授因长期关心和支持YOCSEF的发展、在对YOCSEF文化和制度的传承方面做出了非凡的贡献而获此殊荣。

（责任编辑：刘海波）

图像美学质量评价技术开源代码

北京电子科技学院 金鑫

图像美学质量评价是使得计算机能够从美学的角度给出图像的质量评价，模拟人类的审美能力，其主要任务包括：(1) 判断一幅图像的美学质量是“高”还是“低”（美感分类），(2) 给出一幅图像的美学质量数值评价（美感评分），(3) 给出一幅图像的美学质量评价分数分布（美感分布预测），(4) 给出一幅图像的美学质量语言评论（美学描述）等。

图像美学质量评价是计算美学研究中的热点问题，从 2004 年开始，先后经历了内容无关人工设计美学特征的图像美感分类与评分、内容敏感人工设计美学特征的图像美感分类与评分、基于深度学习的图像美感分类与评分、图像美学质量多样化评价几个发展阶段。

本文着重介绍几个基于深度学习的图像美学质量评价技术开源代码，包括：图像美感分类、图像美感评分、图像美学质量评价分数分布预测、图像美感语言评论四个方面。

1. 图像美感分类

1.1. RAPID

论文： Lu X, Lin Z, Jin H, et al. RAPID: Rating pictorial aesthetics using deep learning [C]// Proceedings of the ACM International Conference on Multimedia. New York: ACM, 2014: 457-466.

工作： 美国宾夕法尼亚州立大学的这项工作首次将深度学习技术应用于图像美学质量评价，提出了一种双通道深度神经网络将局部和全局的图像块分开处理，如图 1 所示。其在 AVA 数据集上能够达到 74.46% 的美感分类准确率。

主页：

<http://infolab.stanford.edu/~wangz/project/imsearch/Aesthetics/ACMMM2014/>

代码：

<https://github.com/indivisibleatom/aestheticsPerception>

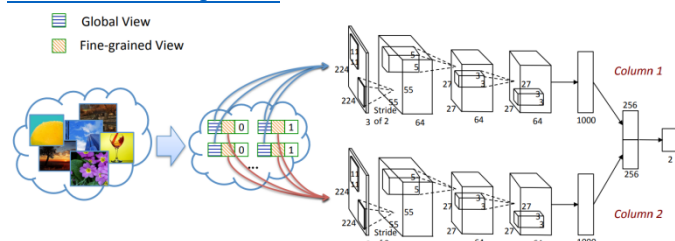


图 1 RAPID 网络结构图

1.2. ILGNet

论文： Jin X, Wu L, Li X D, et al. ILGNet: Inception Modules with Connected Local and Global Features for Efficient Image Aesthetic Quality Classification using Domain Adaptation [J]. arXiv.org, 2018, arXiv:1610.02256v3.

工作： 北京电子科技学院的这项工作提出了一种基于深度卷积神经网络的图像美感分类方法，能够给出图像美学质量的分类结果：“高”与“低”或者“好”与“不好”，如图 2 所示。该方法结合局部与全局特征，设计了一种新的深度卷积神经网络：ILGNet，能够自动区分出美学质量高与美学质量低的两类图像，其在公开的 AVA 图像数据集上取得了当前最高的美感分类准确率 (82.66%)。

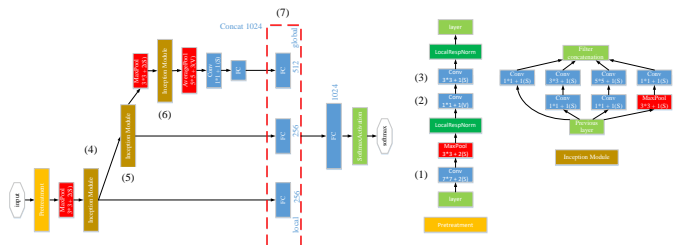


图 2 ILGNet 网络结构图

主页： <http://kislab.besti.edu.cn/victory/?p=242>

代码： <https://github.com/BestiVictory/ILGnet>

1.3. Triplet-CNN

论文： Schwarz K, Wieschollek P, Lensch H P

A. Will people like your image? [C]//Proceedings of IEEE Winter Conference on Applications of Computer Vision (WACV), 2018.

工作：德国蒂宾根大学的这项工作首先利用 Flickr 的数据构建了一个新的美学质量评价数据集：AROD (multi-user agreements and assemble a large dataset)，包含了 38 万幅图像，利用 Flickr 上收藏数与查看数的比值作为每一幅图像的美学评分，提出了一个基于三元损失函数的神经网络 Triplet-CNN，如图 3 所示，能够输出图像的美感分类。

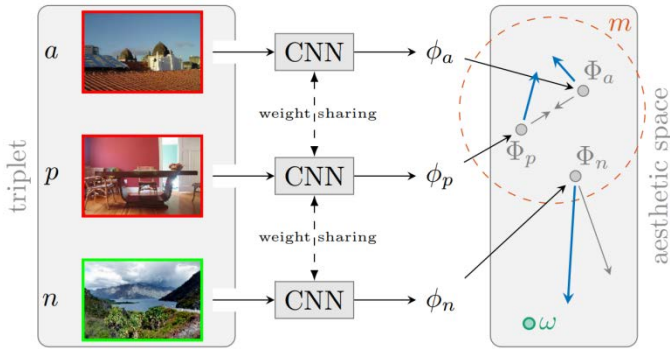


图 3 Triplet-CNN 网络结构图

主页：

http://patwie.com/pdf/schwarz_wacv2018.pdf

代码：

<https://github.com/cgtuebingen/will-people-like-your-image>

2. 图像美感评分

论文：Kong S, Shen X, Lin Z, et al. Photo aesthetics ranking network with attributes and content adaptation [C]//Proceedings of 14th European Conference on Computer Vision. Heidelberg: Springer-Verlag, 2016: 662-679

工作：美国加州大学欧文分校的这项工作首先构建了一个新的美学数据集：AADB (aesthetics and attributes database) 包含了图像的美学总体评分与多美学因素/属性的评价。提出了基于回归损失、排序损失、属性损失的深度卷积神经网络，如图 4 所示，能够给出图像的美感数值评价，即，

美感评分。

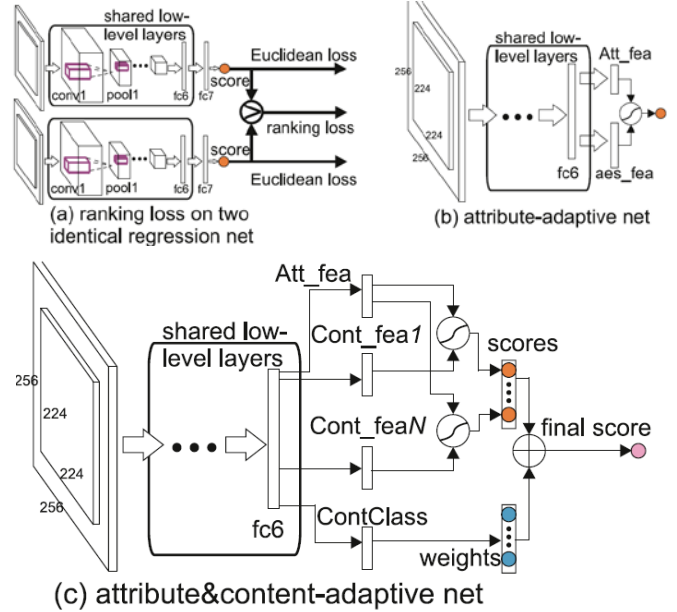


图 4 CNN 的回归损失、排序损失、属性损失

主页：

<http://www.ics.uci.edu/~skong2/aesthetics.html>

代码：

<https://github.com/aimerykong/deepImageAestheticsAnalysis>

3. 图像美感分布预测

论文：Jin X, Wu L, Li X, et al. Predicting aesthetic score distribution through cumulative Jensen-Shannon divergence [C]//Proceedings of AAAI Conference on Artificial Intelligence (AAAI), New Orleans, Louisiana, USA. 2-7 Feb. 2018.

工作：北京电子科技学院的这项工作提出了基于深度学习的图像美学质量评价分数分布（向量）预测方法，不同于以往的单标量评价，该方法能够给出一幅图像的美学质量评价分数分布，能够更好的描述人类美感的多样性。传统的深度学习最小化预测标量或一般向量与人类标记的差异，无法直接应用于美学质量评价分数分布这种有序分布向量的预测中。提出了一种基于累积 JS 散度的卷积神经网络 (CNN based on the Cumulative distribution with Jensen-Shannon divergence: CJS-CNN) 用于预测图像美学质量评

价分数分布。并且，还提出了一种基于峰度统计量的可靠性敏感的学习算法，该算法只需要归一化的评价分数分布数据就可以进行可靠性敏感的学习，而之前基于评价人数的可靠性敏感学习算法需要归一化之前完整的评价分数分布数据。

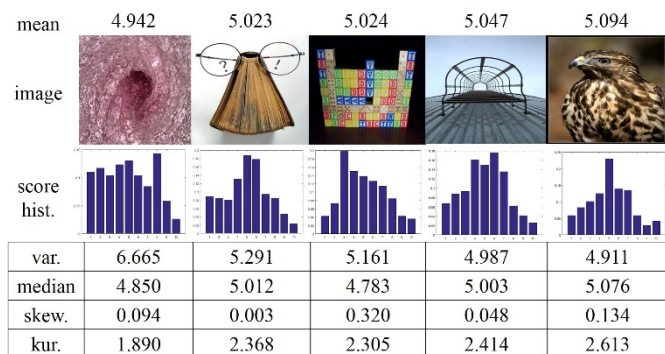


图 5 图像美感分布

主页:

<http://kislab.besti.edu.cn/victory/?p=825>

代码:

<https://github.com/BestiVictory/CJS-CNN>

4. 图像美学语言评论

论文: Chang K Y, Lu K H, Chen C S. Aesthetic critiques generation for photos [C]//Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV). Piscataway NJ: IEEE, 2017: 3534-3543.

工作: 台湾信息科学研究所的这项工作首先构建了一个新的美学图像数据集: PCCD (photo critique captioning dataset), 包含了图像的美学语

言评论, 提出了一个基于 CNN 和 LSTM 的神经网络, 如图 6 所示, 能够输出图像的美学语言评论,

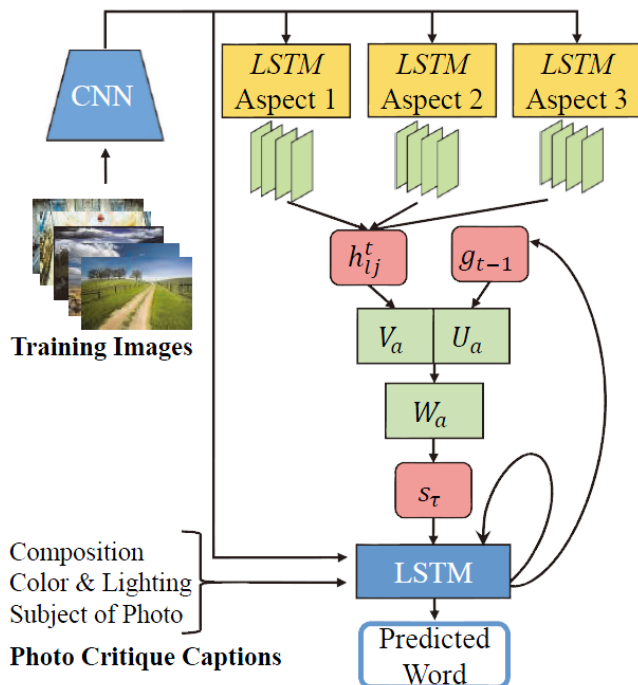


图 6 CNN-LSTM 网络结构图

主页:

http://www.iis.sinica.edu.tw/~kuangyu/icc_v17_aesthetic_critiques.pdf

代码:

<https://github.com/kunghunglu/DeepPhotoCritic-ICCV17>

(责任编辑: 贾同)



金鑫:

现于北京电子科技学院计算机科学与技术系、中共中央办公厅信息安全重点实验室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。

主要研究方向包括: 计算美学、计算机视觉、虚拟现实、可视媒体安全、人工智能安全等。兼任 CCF-CV 专委会

委员、CAA-PRMI 专委会委员、中国电子学会虚拟/增强现实技术与产业分会副秘书长、国际人工智能与机器人协会理事。

在 TIP、CVPR、AAAI、ECCV、CGF、ICME、MTAP 等相关领域重要国际期刊与会议上发表学术论文十余篇。个人主页: www.jinxin.me

计算机视觉中的医学图像数据集

兰州理工大学 李策 张栋

计算机辅助诊断(CAD)主要是指基于医学影像学的计算机辅助技术。CAD 被视为医生的“第三只眼”，其广泛应用有助于提高诊断的敏感性和特异性。近年来，基于深度学习技术基础并利用海量医学影像数据的完备性优势，使得临床诊断精度得到很大的提高。本文针对肺结节数据集、乳腺 MG 数据集等进行介绍，文中列出的一些常用的不同类型医学影像数据库地址链接包括：

- ①<http://www.cancerimagingarchive.net>
- ②<https://grand-challenge.org>
- ③<https://github.com/beamandrew/medical-datasets>

1. 肺结节检测数据集

1.1 天池医疗 AI 大赛数据集

数据集地址：<https://tianchi.aliyun.com/competition/introduction.htm?raceId=231601>

数据由天池医疗 AI 大赛与合作机构 LinkDoc 授权提供。全部数据集有 3000 例病人，数据集样本示例如图 1 所示。数据集中所有 CT 影像的层厚小于 2mm，全部都有肺结节，其结节大小大致分布如表 1 所示。

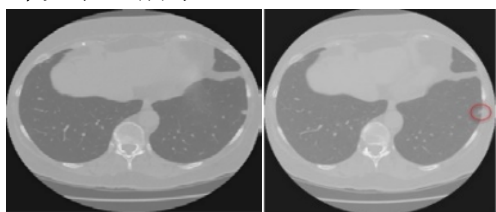


图 1. 天池医疗 AI 大赛数据集样本示例

表 1. 肺结节大小分布

结节大小	5-10mm	10-30mm
1000 例	50%	50%

2000 例	70%	30%
--------	-----	-----

全部数据集的数据格式有 CT 影像 mhd 格式（mhd 文件具有包含关于患者 ID 的必要信息，以及诸如切片厚度的扫描参数等）和 CSV 结节标注文件，标注文件信息如表 2 所示。每个图像中的结节都给出了中心坐标和半径。难点之一在于结节的判断容易和其他血管的影像混淆。

表 2. CSV 文件中的结节标注信息示例

seriesuid	LKDS_00001
coordX	-100.56mm
coordY	67.26mm
coordZ	-231.81mm
diameter_mm	6.44mm

1.2 LIDC-IDRI 数据集

数据集地址：<https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>

该数据集是由美国国家癌症研究所(National Cancer Institute)发起，由七个学术中心和八家医学影像公司合作创建。全部数据集由临床胸部 CT 等扫描的图像和相关的 XML 标注文件组成。该数据集中，共收录了 1018 个研究实例，数据集统计说明如表 3 所示。

表 3. 数据集统计说明

Images (DICOM)	124GB
Modalities	CT DX CR
Number of Images	244527
Number of Patients	1010
Number of Series	1018CT 290CR/DX
Number of Studies	1308

对于每个实例中的图像，都由 4 位专业的放射科医师进行两阶段的诊断标注。在第一阶段，每位医师分别独立诊断并标注病患位置，其中会标注三中类别：①结节 $\geq 3\text{mm}$ ；②结节 $< 3\text{mm}$ ；③非

结节 $\geq 3\text{mm}$ 。在第二阶段，医师都独立的复审其他三位医师的标注，并给出自己最终的诊断结果。

2. 乳腺 MG 数据集

2.1 INbreast 数据集

数据集地址 http://medicalresearch.inescporto.pt/breastre-search/index.php/Get_INbreast_Database

INbreast 数据集是一个乳房 X 射线造影图像数据集。数据集来自葡萄牙波尔图乳腺中心医院。INbreast 数据集中包含 115 病例（共 410 张图像），其中 90 病例是双乳房数据（每例 4 张图像），25 病例是乳房切除患者的数据（每例 2 张图像）。INbreast 数据集包括 4 种类型的病变（肿块，钙化，不对称和扭曲），数据集样本示例如图 2 所示。专家标记的这些数据样本，其精确病症轮廓信息都记录在 XML 格式文件中。

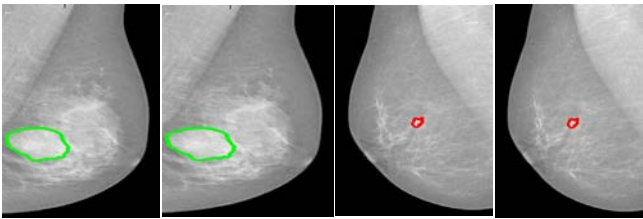


图 2. INbreast 数据集样本示例。左：绿色是良性的肿块，右：红色是恶性的肿块

2.2 DDSM 数据集

数据集地址：<http://marathon.csee.usf.edu/Mammography/Database.html>

该数据集包含大约 2500 个研究病例，每个病例包括每个乳房的两幅图像，以及一些相关的患者信息（年龄，ACR 乳房密度等级，异常的精确评级和异常的 ACR 关键字描述）以及图像信息（扫

描仪，空间分辨率等）。包含可疑区域的图像已经关联了包含可疑区域位置和类型的像素级“Ground Truth”信息。用于筛查乳房 X 射线造影的数字数据库被编组为“病例”和“卷”。“病例”是与一名患者的一次乳房 X 射线检查相对应的图像和信息的集合。

每个病例由 6 到 10 个文件组成，分为四类：

- ①“ics”文件：包含图像相关信息，如患者年龄，图像的大小，是否存在用于异常轮廓叠加的文件等；
- ②“16-bit PGM”文件：真实乳房 X 射线造影图像文件；
- ③“1jpeg”文件：包含四个使用无损 JPEG 编码进行压缩的图像文件；
- ④“overlay”文件：给出每个视图中给定异常的关键字描述，而正常情况下不会有任何叠加的文件。

2.3 MIAS 数据集

数据集地址：<http://peipa.essex.ac.uk/pix/mias/all-mias.tar.gz>

该数据集中的乳房 X 射线造影图像是由英国国家乳腺筛查计划中心精细挑选，并进行数字化（分辨率为 $50\mu\text{m}$ ，这是一种线性光学密度描述）。该数据集包含 161 名患者的左右乳房图像，由 322 个图像组成，其包含三种类型：正常 208 个，良性 63 个和恶性（异常）51 个。该数据集包括放射科专业医师对任何可能存在异常的位置标记“Ground Truth”。对于每个病例文件给出它们的类型，位置，大小等其它有用信息。根据这些专家标注的描述，该数据库包括四种异常（结构畸变，星状病变，腺瘤肿块和钙化病变）。该数据集标注文件为 info.txt，其标注信息包含以下信息：

- ①类型：四种异常；
- ②分类：异常是恶性还是良性；
- ③位置和大小：异常的原始坐标和直径。

（责任编辑：金鑫）



李策

教授，博士生导师，兰州理工大学电气工程与信息工程学院，主要研究方向为计算机视觉、医学图像处理、机器学习。



张栋

硕士研究生，兰州理工大学电气工程与信息工程学院，主要研究方向为计算机视觉、目标检测、医学图像处理。

招聘信息

一、上海交通大学人工智能研究院全球诚聘(新)

上海交通大学人工智能研究院于2018年1月18日成立。由中国科学院院士毛军发教授担任院长，杨小康教授担任常务副院长，已获批人工智能教育部重点实验室。

研究院现诚邀海内外人工智能领域的英才加盟，研究领域包括（但不限于）：机器学习、计算机视觉、自然语言处理、多媒体智能、人工智能芯片设计、智能无人系统、智能网联汽车、智能医疗等，各个岗位人数不限：

1、岗位名称：

— “长聘正教授”

(1) 海外世界一流大学、研究机构任副教授或相当职务，或在国内一流高校、科研院所担任教授；

(2) 把握人工智能学科的发展方向，能够引领学科前沿，掌握关键技术，具有凝聚多学科协同攻关，赶超或保持国际领先水平的能力，在本学科领域取得国内外同行公认的突出成就；

(3) 学风优良、师德高尚，并具有创新、探索精神和团结协作。

2、岗位名称：

— “长聘教轨副教授”

(1) 具有世界一流大学博士学位并有连续3年以上（含）海外科研工作经历，年龄一般不超过40周岁；

(2) 在学术前沿领域开展创新性研究工作，取得重要的研究成果，表现出很强的学术潜力，达到世界一流大学助理教授水平；

(3) 对特别优秀海外博士，可突破任职年限限制。

3、岗位名称：

— “长聘教轨助理教授”

(1) 国内外知名高校及研究机构取得相关专业博士学位，并有连续3年以上海外科研工作经历，年龄一般不超过35周岁；

(2) 在学术前沿领域开展创新性研究工作，取得

重要的研究成果，表现出很强的学术潜力；

(3) 对特别优秀海外博士，可突破任职年限限制。

4、岗位名称：

— “专职科研”

(1) 国内外知名高校及研究机构取得相关专业博士学位，年龄不超过35周岁；

(2) 在相关领域期刊和会议上已发表高水平研究论文或有较强的工程开发能力，对本领域基础知识和前沿方向有良好把握；

(3) 具备良好的道德素养和团队合作精神。

5、岗位名称：

— “博士后”

(1) 国内外知名高校及研究机构取得博士学位，相关专业背景，年龄一般不超过35周岁；

(2) 深入了解人工智能相关领域最新发展动态有较好的研究工作经验；

(3) 较强的项目研发和组织能力，富有创新意识，具有独立的科研能力和良好的团队合作精神。

上述所有海外工作经历包括博士后经历。

薪资福利：

(1) 基本工资：长聘正教授薪资面议，长聘教轨副教授年薪40万起，助理教授30万起，专职科研20万起，博士后18万起；

(2) 科研配套：博士生或硕士生招生资格；根据学科方向和工作需要，提供科研启动经费，长聘教轨副教授60万，助理教授15万；

(3) 安家补贴：提供安家补贴及过渡性的教师公寓，长聘教轨副教授安家费与住房津贴50万，助理教授30万；博士后可入住博士后公寓或享受租房补贴；

(4) 其他支持：按照国家和上海市相关政策提供各类保险和住房公积金；依托强大的附属医院资源，提供良好的医疗服务；协助落实配偶工作；为子女入学提供优惠政策和条件；

(5) 研究院视情况提供额外的激励措施和必要的工作开展条件；

(6) 若成功入选青年千人计划, 年薪 45 万起, 安家补贴 150 万 (含国家和上海市补贴), 与国家按照 1:1 配套科研启动经费。

联系方式:

有意者请发送电子版申请材料至联系人邮箱, 请在邮件主题上注明: 姓名-学校-学历-应聘岗位名称。

所需提供的材料:

(1) 个人简历一份 (包含姓名、出生年月、联系方式、主要研究和教学领域、论文列表等相关资料);

(2) 发表论文清单及代表作 5 篇全文 (其中申请博士后者需代表作 3 篇及博士论文全文);

(3) 研究和教学概述;

(4) 推荐信 3-5 封 (博士后申请者推荐信其中一封需来自读博期间的导师)。

联系人: 都晓慧

联系电话: +86-21-34206041

联系邮箱: duxiaohui2010@sjtu.edu.cn

二、山东省人工智能研究院海外人才招聘信息

1、面向对象:

— “海内外优秀人才”

需求人数: 10 人

岗位要求:

(1) 热爱中华人民共和国, 学风严谨, 为人正派, 年龄不超过 40 岁的青年学者;

(2) 海外博士, 需毕业于 QS 世界大学排名前 200 名;

(3) 国内博士的应聘人员, 要求毕业于双一流的高校;

(4) 发表不少于两篇 CCF 推荐 A 类会议或期刊论文 (一作长文);

(5) 研究方向为人工智能的子方向, 其中下列方向优先: 机器学习、多媒体计算、计算机视觉、信息检索、数据挖掘、自然语言处理以及模式识别。

2、享受待遇:

(1) 同时满足上述条件的应聘人员, 原则上一人一议;

(2) 应聘人员的综合个人收入, 高于 BAT 等一线公司;

(3) 科研启动经费 200-1000 万;

(4) 有编制;

(5) 解决子女入学 (山东省最好的教育资源), 解决配偶工作;

(6) 办公地点在国家 4A 级风景名胜区济南市千佛山脚下。

3、招聘日程:

(1) 新加坡 5 月 26 日-29 日

地点: NUS, NTU, SMU, SUTD

(2) 日本横滨 6 月 11 日-14 日

地点: Yokohama Media and Communication Center (ACM ICMR 2018 会场)

(3) 美国密西根州安娜堡 7 月 8 日-11 日

地点: Michigan League in University of Michigan (ACM SIGIR 2018 会场)

(4) 韩国首尔 10 月 22 日-26 日

地点: Lotte Hotel, Seoul (ACM MM 2018 会场)

4、联系方式:

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

简历投递截止时间: 招满即止。

联系人: 李老师

简历接收邮箱: shandongai@sdas.org

(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 ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)、IEEE Access 特刊 Advanced Optical Imaging for Extreme Environments 等。

3 会议简介

第 24 届国际模式识别大会 ICPR (24th International Conference on Pattern Recognition) 是模式识别领域的权威会议，今年将首次在中国举办。ICPR 是一个国际性的学术会议，将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，其内容涉及近期在模式识别、机器学习、计算机视觉方面的进展，以及这些技术在多个领域的应用。

本届 ICPR 会议由国际模式识别协会 (IAPR: International Association for Pattern Recognition) 主办，由中科院自动化研究所，模式识别国家重点实验室承办，将于 2018 年 8 月 20-24 日在北京市中国国家会议中心举行 (全文截止时间为 2018 年 5 月 28 日)。涉及模式识别与机器学习、计算机视觉、语音、图像、视频、多媒体、生物信息、人机交互、文本分析与识别、生物成像等方向的研究进展。同时将举办相关的研讨会 (Workshop)、教程 (Tutorials)、竞赛 (Contests)。

(责任编辑: 贾同、沈沛意)

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
SSIST 2018	2018.07.02-04	中国 上海	2018.05.31	http://ssist.shanghaitech.edu.cn
PRCV 2018	2018.11.22-25	中国 广州	2018.05.28	http://prcv-conf.org/2018/
PRICAI 2018	2018.08.27-31	中国 南京	2018.05.31	http://cse.seu.edu.cn/pricai18/
ACML 2018	2018.11.14-16	中国 北京	2018.08.01	http://www.acml-conf.org/2018
FFER 2018	2018.8.20	中国 北京	2018.05.21	https://ffer.aau.dk
CMVIT 2019	2019.02.22-24	中国 广州	2018.06.10	http://www.cmvit.org/
IIP 2018	2018.10.19-22	中国 南京	2018.05.31	http://www.intsci.ac.cn/iip2018/

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
TOMM	Cross-Media Analysis for Visual Question Answering	2018.06.30

Momenta

打造自动驾驶大脑

Momenta 成立于 2016 年，是世界顶尖的自动驾驶公司。其核心技术是基于深度学习的环境感知、高精度地图、驾驶决策算法。产品包括不同级别的自动驾驶方案，以及衍生出的大数据服务。

Momenta 致力于“打造自动驾驶大脑”，拥有世界顶尖的深度学习专家，如图像识别领域最先进的框架 Faster R-CNN 和 ResNet 的作者，ImageNet 2015、ImageNet 2017、MS COCO Challenge 2015 等多项比赛冠军。团队成员主要来源于清华大学、麻省理工学院、微软亚洲研究院等高校及研究机构，以及百度、阿里、腾讯、华为、商汤等知名高科技公司，拥有深厚的技术积累、极强的技术原创力和丰富的行业经验。

截至目前，Momenta 已获得亿级美元融资，投资方包括蔚来资本、戴姆勒集团（梅赛德斯-奔驰母公司）、创新工场、顺为资本、GGV 纪源资本、凯辉基金、蓝湖资本、真格基金和九合创投等。

Momenta 的技术领先性及产品落地能力除了为资本市场广泛认可外，也深受行业合作方的认可。

核心技术：基于深度学习的环境感知、高精度地图和驾驶决策算法。

环境感知包括道路识别、行人感知、车辆感知等核心技术。**道路识别：**不论是正常情况，还是在黑暗、逆光、恶劣天气和缺乏清晰车道线的情况下，都可以做到高性能地识别多个车道、交通标志和信号以及可行驶区域。**行人感知：**通过检测行人和识别人体特征点，可以做到理解行人

姿势和行为的意图，同时也可以准确估计行人与汽车的距离。**车辆感知：**路面上的每一辆车都可以还原其 3D 边界框，鲁棒地检测车的方向，精确地估算距离，并在高精度地图上进行实时定位。

通过提取从多辆汽车拍摄的 2D 图像语义点来重建道路、交通标志、信号及周围环境的 3D 位置，融合来自 GPS 和 IMU 的数据，创建更高精度的地图。Momenta 的高精度地图方案更具扩展性和商业化落地可能性，其成本仅为 LIDAR 数据收集方案的 1/10 到 1/100。

驾驶决策由数据推动，类似于建立一个拥有 1000 亿公里驾驶经验的智能司机。通过众包路测，Momenta 获得了高精度语义地图中海量的驾驶轨迹。通过对海量驾驶轨迹的学习，Momenta 的算法可以根据当前环境感知和高精度地图信息，做出驾驶决策规划。



全球顶级人才持续招募中！登陆 www.momenta.ai 了解更多精彩！[简历投递至 talentoverflow@momenta.ai](mailto:talentoverflow@momenta.ai)。

（责任编辑：王瑞平）