

郭志成

男 24岁 西安 硕士 | 学生
18483676630 zchguo@stu.xidian.edu.cn



教育经历

2019.09 - 2022.06 硕士 西安电子科技大学 计算机科学与技术 (保研) 专业排名1%~5%
2015.09 - 2019.06 本科 成都理工大学 电子信息科学与技术 专业排名1%~5%

求职意向

北京,上海,深圳,杭州,成都,西安 视觉-文本任务; 问答系统; 视频理解; 多模态学习

项目经历

2019.10 - 2021.04 人工智能实验简明教程 成员

参与写书,我负责图像超分辨率这一章节的内容。先介绍了图像超分辨率这个任务的背景和发展历史,然后具体描述了SRCNN这个网络,最后阐述了这个模型的代码实现过程以及实验结果。这本书由清华大学出版社正式出版,现在京东可购买。

2019.07 - 2019.10 眼球定位系统 成员

眼球定位,即找到人眼看屏幕的位置。我们在电脑端实现了眼球定位功能,应用于我们的产品。通过摄像头捕捉到人脸信息,根据眼睛瞳孔位置和脸部信息训练模型,就能得到人眼看屏幕的位置。

校内经历

2020.02 - 2021.05 百度校园菁英俱乐部 成员

作为俱乐部成员负责程序设计竞赛的出题。

2017.06 - 2019.06 ACM-ICPC 竞赛队 队长

本科期间,组建了学校的ACM竞赛队,作为队长指导并带领同学参加各个程序设计竞赛,并搭建了本校的OJ。

学术经历

西电人工智能学院研二在读,导师是院长焦李成教授 (IEEE Fellow, 欧洲科学院外籍院士, 俄罗斯自然科学院外籍院士, 中国人工智能学会副理事长)。在导师的指导下,研究方向包括视觉问答, 视频问答, 视频理解, 多模态学习, 超图网络等。曾面试通过阿里达摩院的实习, 方向是多模态理解。

论文发表

1. 一篇 ACL 2021 (CCF A) 录用, 一作。

论文标题 Multi-Scale Progressive Attention Network for Video Question Answering, 针对视频的时序信息提出了多尺度渐进注意力, 模型在视频问答基准数据集上取得了SOTA精度, 代码即将开源。

2. 一篇 IEEE Transactions on Multimedia 即将录用, 一作。

论文标题 A Universal Quaternion Hypergraph Network for Multi-Modal Video Question Answering, 利用四元数和超图分别表示视频的多模态和结构信息, 理论推导的四元数超图卷积算子极大的提高了视频多模态特征之间关系推理的能力, 通用的推理结构拓展到了其它video-text任务。

3. 一篇 NeurIPS 2021 (CCF A) 在投, 一作。

arxiv链接: <https://arxiv.org/abs/2106.07501>

论文标题 Balanced Coarsening for Multilevel Hypergraph Partitioning via Wasserstein Discrepancy, 提出了一种多级超图划分的平衡粗化方案。此外, 为了提高k-way超图划分的质量, 设计了一种初始划分算法。工程C++代码和docker即将开源, 致力于超图划分算法的实际应用。

4. 一篇 AAAI 2022 (CCF A) 在投, 二作。

论文标题 Statistical and Structural Domain Adaptation in Semantic Segmentation

5. 一篇 ICCV 2021 Workshop 录用。

论文标题 The Multi-Modal Video Reasoning and Analyzing Competition

6.一篇 CVPR 2021 (CCF A) 被拒，一作。

感谢审稿人的宝贵意见，这次被拒极大的提高了我的论文写作和rebuttal水平，论文已改投。

■ 获奖情况

2020.11 研究生学业奖学金（一等），优秀研究生荣誉称号
2020.10 IEEEExtreme 14.0 全球极限编程挑战赛 全球第153名（153/4132），国内第39名
2019.06 2019 四川省大学生程序设计竞赛——铜牌
2019.05 2019年ACM-ICPC全国邀请赛——铜牌
2019.04 团体程序设计天梯赛——全国团队三等奖
2018.06 2018 ACM-ICPC 中国大学生程序设计竞赛——铜牌
2018.05 2018年中国大学生程序设计竞赛全国邀请赛——铜牌
2019.03 牛客网一战到底编程挑战赛——排行榜第5名，获牛客战神大礼包

■ 算法竞赛

ICCV 2021：Multi-Modal Video Reasoning and Analyzing Competition 第一名 一作
CVPR 2021：Chart Question Answering Challenge 第二名 一作
CVPR 2021：FloodNet Challenge @ EARTHVISION 2021-VQA 第一名和第二名 指导师兄
CVPR 2020：VizWiz 2020 Grand Challenge 第二名 一作
比赛方向为视觉问答（VQA）。根据一张图片的视觉信息和一个问题的文本信息，回答这个问题。不同于VQA模型的基准数据集COCO 2014，这个比赛的数据集图片质量差且问答关系复杂。我们选择MCAN模型作为Baseline，对图片提取多种类型的视觉特征，问题嵌入使用的是Glove和BERT，逐步调整分类器，最后一共集成了14个模型。
ECCV 2020：DramaQA Challenge 2020 第二名 一作
比赛方向为视频问答（VideoQA）。根据一部电影的视觉信息和文本字幕，回答多个问题。我们提出了由电影人物引导的多模态融合网络，解决了传统电影视频问答模型中信息缺失和信息不匹配的问题。

■ 个人技能

- 1.熟练掌握C++和Python，以及深度学习框架PyTorch，能够快速实现论文算法和C++工程项目。
- 2.会一点CUDA编程知识，能够理解PyTorch底层代码。
- 3.工作环境为Ubuntu，并熟练使用Docker，在Docker Hub上有开源项目。
- 4.通过人工智能相关算法竞赛，学会了許多调参技能以及代码规范，可以训练出高性能和高效率的模型。
- 5.Github链接：<https://github.com/gzcsudo>

■ 自我评价

踏实能干，不怕困难。写代码能力强，思维敏捷，很快想到论文创新点。