

王九牛



✉ wangjiuniu99@163.com

🌐 <https://wangjiuniu.github.io/> 📅 1993.08

教育背景

香港城市大学 (CityU) 计算机系, 博士在读 2019.09 - 2022.06

- 导师: Prof. Antoni B. Chan 研究方向: 多模态融合 (图像描述、零样本学习) 学业成绩: A+

中国科学院大学 (UCAS) 空天信息研究院, 直博在读 2016.09 - 2022.06

- 导师: 吴一戎院士 研究方向: 自然语言处理 (机器阅读理解) 学业成绩: 85.9/100

北京理工大学 (BIT) 信息与电子学院, 本科 2012.09 - 2016.06

- 学业成绩: 92.6/100 国家奖学金 2 次 北京市-三好学生; 电子赛二等奖; 数学、物理竞赛三等奖

研究概述

- 研究生期间从事计算机视觉和自然语言处理相关研究, 主要关注视觉和文本的多模态融合方向。
- 已发表论文 12 篇, 包括 ACM MM、ECCV、NeurIPS、CVPR、TPAMI、IJCV 等; 担任 NeurIPS、ICML、ICLR、AAAI 等顶会审稿人。

项目经历

有辨别性的图像描述 (2019.09 - 至今) 导师 Prof. Antoni B. Chan, 香港城市大学

- 项目简介: 图像描述 (Image Captioning) 是根据图像生成文本, 来描述图像内容。本项目构建相似图像集合并进行比较。生成的描述更具辨别性, 能体现出目标图像与其他相似图片的区别。
- 技术实现:
 - 1) 本项目采用 Encoder-Decoder 框架, 使用 Faster R-CNN 提取图像特征, 再训练 RNN 或 Transformer 解码器生成文本。提出的模块适用多种典型模型, 在正确率和辨别性方面产生累加效果。
 - 2) 针对 MSCOCO 数据集, 使用图文检索、文本相似度等方式衡量图像内容相似性, 构建相似图像组。在相似图像组内强调生成描述的辨别性。
 - 3) 提出 CIDErBtw、CIDErRank 等评价指标衡量图像描述的辨别性, 并根据辨别性调整句子、词语在训练时的权重, 使得模型生成的描述在正确率和辨别性方面均显著提升。成果发表于 ECCV 会议 (oral), 扩展为期刊论文 TPAMI 在投。
 - 4) 针对 Transformer 解码器提出相似性注意力模块, 为目标图像中具有可辨别性的局部特征增加权重, 并且利用辨别性词语引导模型训练。生成的描述中辨别性词语的覆盖率显著增加。成果发表于 ACM MM 会议 (oral)。

基于属性原型的可解释零样本图像分类 (2019.11 - 至今) 导师 Prof. Bernt Schiele, 德国马普所

- 项目简介: 零样本学习 (Zero-shot Learning) 是一种特殊的图像分类场景, 训练过程中部分类别无训练样本 (不可见类)。训练时通过可见类图像来学习语义属性和视觉特征的关联; 测试时判定不可见类图像的属性, 进而预测该图片的类别。本项目构建属性原型来优化图像特征提取, 同时具备属性定位能力。
- 技术实现:
 - 1) 基于 ResNet 构建零样本分类网络, 构建属性原型向量来学习图像属性的视觉空间特征。
 - 2) 比较图像特征和原型特征, 构建图像属性注意力图, 实现属性的弱监督定位。
 - 3) 提出正则化函数进行属性解耦, 减少耦合属性的消极影响, 提高图像分类准确率。
 - 4) 本模型在零样本图像分类数据集 CUB、AWA、SUN 上得到验证, 大幅提升分类准确率和弱监督属性定位准确率。成果发表于 NeurIPS 会议, 扩展为期刊论文 IJCV 在投。

对抗训练和多层注意力的机器阅读理解 (2017.09 - 2019.09) 导师 吴一戎院士, 中国科学院大学

- 项目简介: 机器阅读理解 (Machine Reading Comprehension) 实现对非结构化文本的自动问答。能够根据用户提问, 在对应篇章中抽取答案。本项目旨在提升机器阅读理解的稳定性和正确率。
- 技术实现:
 - 1) 采用多层注意力机制的神经网络对篇章和提问进行特征表示与融合。
 - 2) 在训练模型时采用对抗训练的方法增强模型稳定性。
 - 3) 在 WebQA 中文数据集中取得最佳结果, 成果发表于 NLPCCC 会议和 KBS 期刊。

竞赛获奖与实习

第五届百度 & 西安交大大数据竞赛暨 IKCEST 首届“一带一路”国际大数据竞赛 国际三等奖, 2019.05 - 2019.09

背景: 竞赛以“基于卫星遥感影像和用户行为的城市功能分类”为题, 预测城市区域的 9 种功能类别。

难点: 55 万的数据样本, 数据量较大; 图像样本像素较低只有 100×100 ; 类别的类间差距小, 类内差距大。

方案: 1) 对图像和数据进行预处理, 使用 SEResNeXt+DPN98 将图像和数据进行深度融合, 准确率 68%。

2) 针对数据从人数、假期、工作日等角度提取特征, 使用 Xgboost+LightGBM 等模型, 正确率为 73%。

改进: 1) 采用 5 折交叉、AMSGrad、Warm up、余弦学习率衰减、TTA 等, 方案 1 正确率可达 73%。

2) 从单个用户角度, 构建每个用户分别到访过各个区域的次数字典, 加入此特征后方案 2 正确率可达 81%。

成果: 两种方案融合, 最后以正确率 84.5% 获得竞赛第 9 名 (9 / 2312)。

2017 年华为网络技术大赛

全国三等奖, 2017.03 - 2017.06

背景: 大赛以“云聚天下, 联接未来”为主题, 研究和制定信息与通信技术 (ICT) 解决方案。

内容: 1) 由华为全球技术服务部主办, 考察信息通信技术的基础技能, 包括云计算、大数据、数据中心架构等。

2) 基于网络拓扑和协议构建基本网络。为客户设计私有云网络解决方案, 包括硬件采购及搭建。

广告监测系统回传数据质量监测

基础架构部研发实习生, AdMaster/明略科技, 2016.03 - 2016.08

背景: 统计广告监测系统中数千个广告位的日志数据的时间和数量分布, 每日数据增量为亿级。

内容: 1) 使用 Pig 等查询语言, 在 Hadoop 系统中监测日志的健康状况, 对数据缺失、激增等情况作出警报。

2) 记录广告位日志数据的统计量, 支撑客户服务部门提出的数据查询、统计等任务。

学术论文

一作论文 (已发表):

- Wang, J., Xu, W., Wang, Q., Chan, A. B. (2022). On Distinctive Image Captioning via Comparing and Reweighting. In IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI, IF: 16.4).
- Wang, J., Xu, W., Wang, Q., Chan, A. B. (2021). Group-based Distinctive Image Captioning with Memory Attention. In ACM International Conference on Multimedia (ACM MM, oral).
- Wang, J., Xu, W., Wang, Q., Chan, A. B. (2020). Compare and Reweight: Distinctive Image Captioning Using Similar Images Sets. In European Conference on Computer Vision (ECCV, oral).
- Wang, J., Xu, W., Fu, X., Xu, G., Wu, Y. (2020). ASTRAL: adversarial trained LSTM-CNN for named entity recognition. In Knowledge-Based Systems (KBS, IF: 5.9).
- Wang, J., Xu, W., Fu, X., ... Wu, Y. (2020). SRQA: Synthetic Reader for Factoid Question Answering. In Knowledge-Based Systems (KBS, IF: 5.9).
- Wang, J., Fu, X., Xu, G., Wu, Y., Chen, Z., Wei, Y., Jin, L. (2018). A3NET: Adversarial-and-attention network for machine reading comprehension. In CCF International Conference on Natural Language Processing and Chinese Computing (NLPPCC).

合作论文 (已发表):

- Xu, W., Xian, Y., Wang, J., Schiele, B., Akata, Z. (2022). VGSE: Visually-Grounded Semantic Embeddings for Zero-Shot Learning. In Conference on Computer Vision and Pattern Recognition (CVPR).
- Xu, W., Xian, Y., Wang, J., Schiele, B., Akata, Z. (2022). Attribute Prototype Network for Any-Shot Learning. In International Journal of Computer Vision (IJCV, IF: 7.4).
- Xu, W., Xian, Y., Wang, J., Schiele, B., Akata, Z. (2020). Attribute prototype network for zero-shot learning. In 34th Conference on Neural Information Processing Systems (NeurIPS).
- Wang, Q., Wang, J., Chan, A. B., Huang, S., Xiong, H., Li, X., Dou, D. (2020). Neighbours Matter: Image Captioning with Similar Images. In British Machine Vision Conference (BMVC).
- Xu, W., Wang, J., Wang, Y., ... Wu, Y. (2020). Where is the Model Looking At? - Concentrate and Explain the Network Attention. In IEEE Journal of Selected Topics in Signal Processing (JSTSP, IF: 4.5).