

《跨模态内容精准生成及高效检索理论与方法》项目信息

一、项目名称：跨模态内容精准生成及高效检索理论与方法

二、提名者及提名等级：

提名者：昆明理工大学

提名等级：云南省自然科学奖一等奖

三、项目简介：

面向图像、视频、文本等不同类型数据的跨模态内容精准生成与高效检索技术具有广泛的应用场景。然而在真实场景中不同模态数据质量低、模态差异大，跨模态内容理解还面临跨模态语义统一表征、跨模态内容精准生成以及跨模态精准高效检索等难点问题。依托国家自然科学基金及云南省科技重大专项等国家及省部级项目，通过长期持续研究，在跨模态表征学习、生成及检索三个方面取得了突破，主要科学发现如下：

(1) 提出了低质量、模态缺失及跨模态对齐的特征学习的理论与方法。揭示了不同模态数据的特征差异及关联对跨模态表征的作用机制，提出了几何结构一致性特征修复补偿关联驱动的低质量图像协同表征范式；构建了拓扑流形模态关联及想象推理的不完整多模态图像的全模态映射表征；提出了多粒度语义关联约束和协同演化的图文跨模态表征方法，实现了基于深度 Transformer 的跨模态对齐学习框架，解决了低质量图像、不完整模态及图文跨模态表征的难题。

(2) 构建了视觉-语义关系驱动的图像视频-文本高精度生成框架。揭示多层次语义关系建模与特征交互对于跨模态内容生成的核心机理，提出了基于语义差异特征与模态融合的图像视频-文本精准生成方法，攻克复杂场景下图像视频-文本内容理解、关系建模及语义对齐的关键难题。

(3) 建立了模态语义一致性的跨模态精准高效检索理论和技术框架。揭示了多模态数据的模态内和模态间差异对跨模态检索的影响机理，提出了模态语义一致性约束的精准高效检索技术与方法，攻克了模态差异大条件下跨模态一致性特征学习的技术瓶颈，解决了大规模跨域图像、跨模态图像及跨模态图文的高精度、高效率的检索难题。

四、代表性论文（专著）目录：

序号	论文名称	作者	来源出版物	年卷页码	出版年
1	Deep Multi-View Feature Learning for Person Re-Identification	Dapeng Tao*, Yanan Guo, Baosheng Yu, Jianxin Pang, Zhengtao Yu	IEEE Transactions on Circuits and Systems for Video Technology	2018年 28(10)卷 2657-2666页	2018年 10月 01日
2	Learning Deep Transformer Models for Machine Translation	Qiang Wang, Bei Li, Tong Xiao*, Jingbo Zhu, Changliang Li, Derek F. Wong, Lidia S. Chao	Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics	2019年 1810-1822页	2019年 07月 28日
3	Embedded Block Residual Network: A Recursive Restoration Model for Single-Image Super-Resolution	Yajun Qiu, Ruxin Wang, Dapeng Tao*, Jun Cheng	Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision	2019年 4180-4189页	2019年 10月 27日
4	Semantic Relation-aware Difference Representation Learning for Change Captioning	Yunbin Tu, Tingting Yao, Liang Li, Jiedong Lou, Shengxiang Gao*, Zhengtao Yu, Chenggang Yan	Proceedings of the 59th Annual Meeting of Association for Computational Linguistics	2021年 63-73页	2021年 08月 01日
5	Enhancing the Alignment Between Target Words	Yunbin Tu, Chang Zhou, Junjun	Pattern Recognition	2021年 111卷	2021年 03月

	and Corresponding Frames for Video Captioning	Guo, Shengxiang Gao, Zhengtao Yu*		1-11 页	01 日
6	I ² Transformer: Intra- and Inter-Relation Embedding Transformer for TV Show Captioning	Yunbin Tu, Liang Li*, Li Su, Shengxiang Gao*, Chenggang Yan, Zhengjun Zha, Zhengtao Yu, Qingming Huang	IEEE Transactions on Image Processing	2022 年 31 卷 3565-3577 页	2022 年 03 月 21 日
7	Triple Adversarial Learning and Multi-View Imaginative Reasoning for Unsupervised Domain Adaptation Person Re-Identification	Huafeng Li, Neng Dong, Zhengtao Yu*, Dapeng Tao*, Guanqiu Qi	IEEE Transactions on Circuits and Systems for Video Technology	2021 年 32(5)卷 2814-2830 页	2021 年 05 月 01 日
8	Intermediary-Guided Bidirectional Spatial-Temporal Aggregation Network for Video-Based Visible-Infrared Person Re-Identification	Huafeng Li, Minghui Liu, Zhanxuan Hu*, Feiping Nie, Zhengtao Yu	IEEE Transactions on Circuits and Systems for Video Technology	2023 年 39(9)卷 4962-4972 页	2023 年 09 月 01 日

五、主要完成人基本情况

序号	姓名	完成单位	工作单位	职称	职务
1	余正涛	昆明理工大学	云南省大数据有限公司	教授	实验室主任
2	肖桐	东北大学	东北大学	教授	人工智能系系主任
3	高盛祥	昆明理工大学	昆明理工大学	教授	实验室常务副主任
4	陶大鹏	云南大学	云南大学	教授	无
5	李华锋	昆明理工大学	昆明理工大学	教授	无
6	舒振球	昆明理工大学	昆明理工大学	教授	无
7	郭军军	昆明理工大学	昆明理工大学	副教授	无