



## 师资队伍

Teaching staff

## 学术科研

您所在的位置: 教师科研 > 师资队伍

## 详细资料

Details



## 李锡荣

个人主页: <http://lixirong.net/>

电子邮箱: 名下划线姓@126.com

**教育经历** 2007.08 - 2012.03, 博士, 荷兰阿姆斯特丹大学 Intelligent Systems Lab Amsterdam

2005.09 - 2007.06, 硕士, 清华大学 计算机系

2001.09 - 2005.07, 本科, 清华大学 计算机系

## 工作经历

2016.08 - 至今, 副教授, 中国人民大学

2012.05 - 2016.08, 讲师, 中国人民大学

## 研究方向

图像/视频搜索

图像自然语言理解

跨媒体表示学习

基于深度学习的医学影像分析

学术报告 / 讲座:

- + 2020.08 面向眼底病识别的多模态深度学习, 北京医师协会眼科专科医师分会AI分委会高峰论坛, 北京
- + 2019.11 Learn to Represent Queries and Videos for Ad-hoc Video Search, TRECVID 2019 workshop, Gaithersburg
- + 2019.10 Deep Learning for Video Retrieval by Natural Language, Keynote talk at the 1st International Workshop on Fairness, Accountability, and Transparency in MultiMedia
- + 2019.09 眼科影像理解的人工智能方法, 浙江工商大学, 杭州
- + 2019.08 面向眼底影像的深度学习技术, 2019 北京眼科大会 (协和眼科论坛), 北京
- + 2019.04 Dual Encoding for Zero-Example Video Retrieval, 第十四届图像图形技术与应用学术会议, 北京
- + 2019.01 Deep Models for Zero-Example Video Retrieval, ITI-CERTH, Greece
- + 2018.11 Recent Advances in Zero-Example Video Retrieval, YOCSEF成都 “智能视觉前沿发展技术报告会”, 成都
- + 2018.11 Word2VisualVec++ for Ad-hoc Video Search, TRECVID 2018 workshop, Gaithersburg
- + 2018.06 人工智能与影像识别, 中央财经大学文化与传媒学院, 北京

- + 2018.05 人工智能在眼科的应用, 眼科E20新设备新技术高峰论坛, 杭州
- + 2018.04 基于深度学习的图像内容识别, CFDA医疗器械技术审评中心, 北京
- + 2017.11 Multi-Scale Word2VisualVec for Video Caption Retrieval, TRECVID 2017 workshop, Gaithersburg
- + 2017.10 人工智能机遇与挑战: 以影像内容理解为例, 2017眼科影像与信息高峰论坛, 上海
- + 2016.11 Word2VisualVec for Video-To-Text Matching and Ranking, TRECVID 2016 workshop, Gaithersburg
- + 2016.10 Tag Embeddings for Multimedia Retrieval and Description, SIGMM Raising Stars Symposium 2016, Amsterdam
- + 2016.04 图片句子生成的新进展, 北京大学语言、逻辑、认知及计算论坛 (LLCC), 北京

## 讲授课程

1. 模式识别
2. 图像内容分析 (数字图像处理)
3. 实用Python编程

## 科研项目

- [8] 北京市自然科学基金 (面上项目): 面向常见眼底病识别的多模态可解释深度学习研究, 2020.01-2022.12 (No. 4202033)
- [7] 中国人民大学决策咨询及预研委托项目: 多媒体内容的中文语言自动描述, 2018.01-2020.12 (No. 18XNLG19)
- [6] 国家自然科学基金 (面上项目): 面向中文的看图造句若干关键问题研究, 2017.01-2020.12 (No. 61672523)
- [5] 上海市智能信息处理重点实验室开放基金: 基于相关样本的图像标签相关性计算研究, 2014.01-2015.12 (No. IIPL-2013-002)
- [4] 国家自然科学基金 (青年基金项目): 基于网上弱标注数据的个性化图像标注研究, 2014.01-2016.12 (No. 61303184)
- [3] 教育部高等学校博士点专项科研基金(新教师类): 基于分类的社会化标签与图像相关度估计方法研究, 2014.01-2016.12 (No. 20130004120006)
- [2] 教育部留学回国人员科研启动基金项目: 社会网上图像检索若干关键问题研究, 2014.01-2015.12
- [1] 中国人民大学新教师启动金项目: 基于社会化媒体的图像检索新方法研究, 2013.01-2015.12 (No. 13XNLF05)

## 科研成果

\*\*\* 论文 \*\*\*

- [46] Jianfeng Dong, Xirong Li, Chaoxi Xu, Xun Yang, Gang Yang, Xun Wang, Meng Wang (2021): Dual Encoding for Video Retrieval by Text. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2021
- [45] Jie Wang, Kaibin Tian, Dayong Ding, Gang Yang, Xirong Li (2021): Unsupervised Domain Expansion for Visual Categorization. *ACM Transactions on Multimedia Computing Communications and Applications (TOMM)*, 2021
- [44] Xueyao Zhang, Juan Cao, Xirong Li, Qiang Sheng, Lei Zhong, Kai Shu (2021): Mining Dual Emotion for Fake News Detection. In: *The Web Conference 2021 (WWW)*, 2021
- [43] Xirong Li, Fangming Zhou, Chaoxi Xu, Jiaqi Ji, Gang Yang (2021): SEA: Sentence Encoder Assembly for Video Retrieval by Textual Queries. *IEEE Transactions on Multimedia (TMM)*, 2021
- [42] Aozhu Chen, Xinyi Huang, Hailan Lin, Xirong Li (2020): Towards Annotation-Free Evaluation of Cross-Lingual Image Captioning. In: *ACM Multimedia Asia (MMAsia)*, 2020.
- [41] Xirong Li, Wencui Wan, Yang Zhou, Jianchun Zhao, Qijie Wei, Junbo Rong, Pengyi Zhou, Limin Xu, Lijuan Lang, Yuying Liu, Chengzhi Niu, Dayong Ding, Xuemin Jin (2020): Deep Multiple Instance Learning with Spatial Attention for ROP Case Classification, Instance Selection and Abnormality Localization. In: *25th International Conference on Pattern Recognition (ICPR)*, 2020
- [40] Qijie Wei, Xirong Li, Weihong Yu, Xiao Zhang, Yongpeng Zhang, Bojie Hu, Bin Mo, Di Gong, Ning Chen, Dayong Ding, Youxin Chen (2020): Learn to Segment Retinal Lesions and Beyond. In: *25th International Conference on Pattern Recognition (ICPR)*, 2020
- [39] Jakub Lokoč, Tomáš Souček, Patrik Veselý, František Mejzlík, Jiaqi Ji, Chaoxi Xu, Xirong Li (2020): A W2VV++ Case Study with Automated and Interactive Text-to-Video Retrieval. In: *ACM Multimedia*, 2020
- [38] Zhengxiong Jia, Xirong Li (2020): iCap: Interactive Image Captioning with Predictive Text. In: *ACM International Conference on Multimedia Retrieval (ICMR)*, 2020
- [37] Yutong Liu, Jingyuan Yang, Yang Zhou, Weisen Wang, Jianchun Zhao, Weihong Yu, Dingding Zhang, Dayong Ding, Xirong Li, Youxin Chen (2020): Prediction of OCT Images of Short-term Response to Anti-VEGF Treatment for Neovascular Age-related Macular Degeneration using Generative Adversarial Network. In: *British Journal of Ophthalmology*, 2020
- [36] Jianfeng Dong, Xun Wang, Leimin Zhang, Chaoxi Xu, Gang Yang, Xirong Li (2019): Feature Re-Learning with Data Augmentation for Video Relevance Prediction. In: *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 2019
- [35] Xirong Li, Chaoxi Xu, Gang Yang, Zhineng Chen, Jianfeng Dong (2019): W2VV++: Fully Deep Learning for Ad-hoc Video Search. In: *ACM Multimedia*, 2019

- [34] Zhuoya Yang, Xirong Li, Xixi He, Dayong Ding, Yanting Wang, Fangfang Dai, Xuemin Jin (2019): Joint Localization of Optic Disc and Fovea in Ultra-Widefield Fundus Images. The 10th International Workshop on Machine Learning in Medical Imaging (MLMI), 2019
- [33] Chaoxi Xu, Xiangjia Zhu, Wenwen He, Yi Lu, Xixi He, Zongjiang Shang, Jun Wu, Keke Zhang, Yinglei Zhang, Xianfang Rong, Zhennan Zhao, Lei Cai, Dayong Ding, Xirong Li (2019): Fully Deep Learning for Slit-lamp Photo based Nuclear Cataract Grading. International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2019 (early accept)
- [32] Weisen Wang, Zhiyan Xu, Weihong Yu, Jianchun Zhao, Jingyuan Yang, Feng He, Zhikun Yang, Di Chen, Dayong Ding, Youxin Chen, Xirong Li (2019): Two-Stream CNN with Loose Pair Training for Multi-modal AMD Categorization. International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI), 2019 (early accept)
- [31] Jianfeng Dong, Xirong Li, Chaoxi Xu, Shouling Ji, Yuan He, Gang Yang, Xun Wang (2019): Dual Encoding for Zero-Example Video Retrieval. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019
- [30] Xirong Li, Chaoxi Xu, Xiaoxu Wang, Weiyu Lan, Zhengxiong Jia, Gang Yang, Jieping Xu (2019): COCO-CN for Cross-Lingual Image Tagging, Captioning and Retrieval. IEEE Transactions on Multimedia (TMM), 2019
- [29] 蓝玮毓, 王晓旭, 杨刚, 李锡荣 (2019): 标签增强的中文看图造句, 计算机学报, 2019
- [28] Xin Lai, Xirong Li, Rui Qian, Dayong Ding, Jun Wu, Jieping Xu (2019): Four Models for Automatic Recognition of Left and Right Eye in Fundus Images. the 25th International Conference on MultiMedia Modeling (MMM), 2019
- [27] Qijie Wei, Xirong Li, Hao Wang, Dayong Ding, Weihong Yu, Youxin Chen (2018): Laser Scar Detection in Fundus Images using Convolutional Neural Networks. Asian Conference on Computer Vision (ACCV), 2018
- [26] Xirong Li, Jianfeng Dong, Chaoxi Xu, Jing Cao, Xun Wang, Gang Yang (2018): Renmin University of China and Zhejiang Gongshang University at TRECVID 2018: Deep Cross-Modal Embeddings for Video-Text Retrieval. TRECVID 2018 Workshop, 2018
- [25] Jianfeng Dong, Xirong Li, Chaoxi Xu, Gang Yang, Xun Wang, Feature Re-Learning with Data Augmentation for Content-based Video Recommendation, ACM Multimedia, 2018 (Grand challenge paper)
- [24] Gang Yang, Jinlu Liu, Jieping Xu, Xirong Li, Dissimilarity Representation Learning for Generalized Zero-Shot Recognition, ACM Multimedia, 2018
- [23] Bin Liang, Hongcheng Li, Miaoqiang Su, Pan Bian, Xirong Li, Wenchang Shi (2018): Deep Text Classification Can be Fooled. IJCAI, 2018
- [22] Gang Yang, Jinlu Liu, Xirong Li (2018): Imagination Based Sample Construction for Zero-Shot Learning. SIGIR, 2018
- [21] Jianfeng Dong, Xirong Li, Cees G. M. Snoek (2018): Predicting Visual Features from Text for Image and Video Caption Retrieval. IEEE Transactions on Multimedia (TMM), 2018

- [20] Jianfeng Dong, Xirong Li, Duanqing Xu (2018): Cross-Media Similarity Evaluation for Web Image Retrieval in the Wild. IEEE Transactions on Multimedia (TMM), 2018
- [19] Cees G. M. Snoek, Xirong Li, Chaoxi Xu, Dennis C. Koelma (2017): University of Amsterdam and Renmin University at TRECVID 2017: Searching Video, Detecting Events and Describing Video. TRECVID Workshop, 2017
- [18] Weiyu Lan, Xirong Li, Jianfeng Dong (2017): Fluency-Guided Cross-Lingual Image Captioning. ACM Multimedia, 2017
- [17] Qijie Wei, Xiaoxu Wang, Xirong Li (2017): Harvesting Deep Models for Cross-Lingual Image Annotation. CBMI, 2017
- [16] Xirong Li (2017): Tag Relevance Fusion for Social Image Retrieval. In: Multimedia Systems, 23 (1), pp. 29–40, 2017
- [15] Cees G. M. Snoek, Jianfeng Dong, Xirong Li, Xiaoxu Wang, Qijie Wei, Weiyu Lan, Efstratios Gavves, Noureldien Hussein, Dennis C. Koelma, Arnold W. M. Smeulders (2016): University of Amsterdam and Renmin University at TRECVID 2016: Searching Video, Detecting Events and Describing Video. TRECVID Workshop, 2016
- [14] Jianfeng Dong, Xirong Li, Weiyu Lan, Yujia Huo, Cees G. M. Snoek (2016): Early Embedding and Late Reranking for Video Captioning. ACM Multimedia, 2016
- [13] Xirong Li, Yujia Huo, Qin Jin, Jieping Xu (2016): Detecting Violence in Video using Subclasses. ACM Multimedia, 2016
- [12] Xirong Li, Qin Jin (2016): Improving Image Captioning by Concept-based Sentence Reranking. PCM, 2016
- [11] Masoud Mazloom, Xirong Li, Cees G. M. Snoek (2016): TagBook: A Semantic Video Representation Without Supervision for Event Detection. In: IEEE Transactions on Multimedia (TMM), 18 (7), pp. 1378-1388, 2016
- [10] Xirong Li, Weiyu Lan, Jianfeng Dong, Hailong Liu (2016): Adding Chinese Captions to Images. ICMR, 2016
- [9] Xirong Li, Tiberio Uricchio, Lamberto Ballan, Marco Bertini, Cees G. M. Snoek, Alberto Del Bimbo (2016): Socializing the Semantic Gap: A Comparative Survey on Image Tag Assignment, Refinement, and Retrieval. ACM Computing Surveys (CSUR), 49 (1), pp. 14:1-14:39, 2016
- [8] Xirong Li, Tiberio Uricchio, Lamberto Ballan, Marco Bertini, Cees G. M. Snoek, Alberto Del Bimbo (2015): Image Tag Assignment, Refinement and Retrieval. ACM Multimedia, 2015
- [7] Jianfeng Dong, Xirong Li, Shuai Liao, Jieping Xu, Duanqing Xu, Xiaoyong Du (2015): Image Retrieval by Cross-Media Relevance Fusion. ACM Multimedia, 2015
- [6] Qin Jin, Xirong Li, Haibing Cao, Yujia Huo, Shuai Liao, Gang Yang, Jieping Xu (2015): RUCMM at MediaEval 2015 Affective Impact of Movies Task: Fusion of Audio and Visual Cues. In: Working Notes Proceedings of the MediaEval 2015 Workshop, 2015
- [5] Xirong Li, Qin Jin, Shuai Liao, Junwei Liang, Xixi He, Yujia Huo, Weiyu Lan, Bin Xiao, Yanxiong Lu, Jieping Xu (2015): RUC-Tencent at ImageCLEF 2015: Concept Detection, Localization and Sentence Generation. CLEF (Working Notes), 2015

- [4] Xirong Li, Shuai Liao, Weiyu Lan, Xiaoyong Du, Gang Yang (2015): Zero-shot Image Tagging by Hierarchical Semantic Embedding. SIGIR, 2015
- [3] Shuai Liao, Xirong Li, Heng Tao Shen, Yang Yang, Xiaoyong Du (2015): Tag Features for Geo-Aware Image Classification. In: IEEE Transactions on Multimedia (TMM), 17 (7), pp. 1058-1067, 2015
- [2] Junwei Liang, Qin Jin, Xixi He, Gang Yang, Jieping Xu, Xirong Li (2015): Detecting semantic concepts in consumer videos using audio. In: International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 2279–2283, 2015
- [1] Svetlana Kordumova, Xirong Li, Cees G.M. Snoek (2015): Best Practices for Learning Video Concept Detectors from Social Media Examples. In: Multimedia Tools and Applications (MTAP), 74 (4), pp. 1291–1315, 2015

\*\*\* 国际评测 \*\*\*

- [9] Runner-up of the TRECVID 2020 Ad-hoc Video Search (AVS) task
- [8] Runner-up of the TRECVID 2019 Ad-hoc Video Search (AVS) task
- [7] Winner of the ACM Multimedia 2018 Hulu Content-based Video Relevance Prediction Challenge
- [6] Top performer of the TRECVID 2018 Ad-hoc Video Search (AVS) task
- [5] Top performer of the TRECVID 2018 Video-to-Text (VTT) Matching and Ranking task
- [4] Top performer of the TRECVID 2016 Video-to-Text (VTT) task
- [3] Top performer of the ImageCLEF 2015 Image Sentence Generation task
- [2] Top performer of the MSR Bing Image Retrieval Challenge at ACM Multimedia 2015
- [1] Top performer of the TRECVID 2013 Video Semantic Indexing with No Annotation task

\*\*\* 专利 \*\*\*

- [1] 发明专利，具备跨语言学习能力的图像自然语言描述生成方法和装置，专利号：ZL 201710657104.3

## 社会兼职

<http://lixirong.net/prof>

- + Associate Editor (July 2020 - June 2022), ACM TOMM
- + Associate Editor (Feb 2020 - ), Multimedia Systems
- + PC co-chair, MMM 2021 (<https://mmm2021.cz/>)
- + Area chair, ICPR 2020

- + Workshop co-chair, ACM Multimedia Asia 2019
- + Area chair, ACM Multimedia 2019
- + Senior PC, ACM ICMR 2019
- + Area chair, ACM Multimedia 2018
- + Area chair, ICPR 2016
- + Demo / short paper co-chair, PCM 2015
- + Publication co-chair, ICMR 2015
- + Publicity co-chair, ICMR 2013

### 荣誉获奖

- [9] Winner of The Content-Based Video Relevance Prediction (CBVRP) Challenge, ACM Multimedia 2018 (Feature Re-Learning with Data Augmentation for Content-based Video Recommendation)
- [8] 2017 中国多媒体大会优秀论文奖 (标签增强的中文看图造句)
- [7] ACM Multimedia 2016 Grand Challenge Award (Early Embedding and Late Reranking for Video Captioning)
- [6] PCM 2016 Best Paper Runner-Up (Improving Image Captioning by Concept-based Sentence Reranking)
- [5] PCM 2014 Outstanding Reviewer Award
- [4] SIGMM 2013 Best PhD Thesis Award (Content-based Visual Search Learned from Social Media)
- [3] IEEE Transactions on Multimedia 2012 Prize Paper Award (Learning Social Tag Relevance by Neighbor Voting)
- [2] 2011 国家优秀自费留学生奖学金
- [1] ACM International Conference on Image and Video Retrieval 2010 Best Paper Award (Unsupervised multi-feature tag relevance learning for social image retrieval)



Copyright © 2015 School of Information, Renmin University of China. All rights reserved.  
版权所有中国人民大学信息学院

地址：北京市海淀区中关村大街59号  
邮编：100872