About Me

- **谢伟迪**，上海交通大学人工智能学院长聘轨副教授，牛津大学视觉几何组访问研究员 (Visiting Researcher at Oxford VGG)，教育部 U40，国家优青（海外），科技部科技创新 2030 — “新一代人工智能”重大项目青年项目负责人，上海市（海外）高层次人才计划，上海市启明星计划，基金委面上项目，阿里巴巴创新研究计划（Alibaba Innovative Research, AIR），华为 EX 基金主持人。
- 博士毕业于牛津大学视觉几何组（Oxford VGG），首批 Google-DeepMind 全额奖学金获得者，China Oxford Scholarship Fund (Magdalen Award) 奖学金获得者，牛津大学工程系杰出奖 (Oxford Excellence Award) 获得者。
- 主要研究**计算机视觉，多模态自监督学习，AI4Science**。发表论文超 100 篇，Google Scholar 引用 $>$ 17000 次。开源多个领域标准数据集合，包括 VGGFace2, Voxceleb, VGGSound 等；获得多个国际顶级会议研讨会的最佳论文奖和最佳海报奖，最佳期刊论文奖；担任计算机视觉和人工智能领域的旗舰会议 CVPR，ECCV，NeurIPS Area Chair。
- 更多细节，请移步个人主页: `https://weidixie.github.io`
- **招本科实习同学，硕士，博士**

Weidi Xie

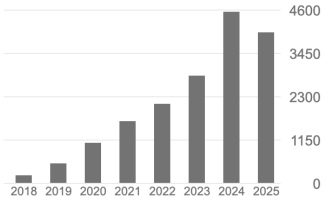
已关注

Shanghai Jiao Tong University | VGG, University of Oxford
在 robots.ox.ac.uk 的电子邮件经过验证 - 首页
Computer Vision AI for Healthcare AI for Science

标题	引用次数	年份
VGGFace2: A Dataset for Recognising Faces Across Pose and Age Q Cao, L Shen, W Xie, OM Parkhi, A Zisserman IEEE International Conference on Automatic Face & Gesture Recognition (FG …	3783	2018
Voxceleb: Large-scale Speaker Verification in The Wild A Nagrani*, JS Chung*, W Xie*, A Zisserman Computer Speech & Language 60, 101027	912	2020
VGGSound: A Large-Scale Audio-Visual Dataset H Chen, W Xie, A Vedaldi, A Zisserman International Conference on Acoustics, Speech, and Signal Processing (ICASSP …	802	2020
NeRF--: Neural Radiance Fields Without Known Camera Parameters Z Wang, S Wu, W Xie, M Chen, VA Prisacariu arXiv preprint arXiv:2102.07064	747	2021
Microscopy Cell Counting with Fully Convolutional Regression Networks W Xie, JA Noble, A Zisserman MICCAI Workshop	680 *	2017
PMC-LLaMA: Towards Building Open-source Language Models for Medicine C Wu, X Zhang, Y Zhang, Y Wang, W Xie Journal of the American Medical Informatics Association, 2024	610 *	2023
Prompting Visual-language Models for Efficient Video Understanding C Ju, T Han, K Zheng, Y Zhang, W Xie European Conference on Computer Vision (ECCV 2022)	542	2022

引用次数

	总计	2020 年至今
引用	17217	16367
h 指数	54	54
i10 指数	115	115



开放获取的出版物数量 查看全部

1 篇文章 65 篇文章

无法查看的文章 可查看的文章

根据资助方的强制性开放获取政策

合著作者 修改

Andrew Zisserman
University of Oxford

On-going Research Topics

Traditionally, computer vision research has mainly focused on solving individual task with supervised learning, for example, classifying images, detecting and tracking objects, recognizing human actions, *etc.* However, real-world problems are often complex, open-ended, infinitely fine-grained, the requirement for human annotations quickly becomes unsustainable and infeasible. **As a computer scientist, my long-term ambition is to develop intelligent agents (machines) that can perceive the world at the same level as humans do.**

To be specific, consider a question on the Harry Potter movie, "what does Harry trick Lucius into doing ?" To answer such question, an intelligent agent should be able to extract information from various sources, for instance, images, languages, and audios, to understand when, where, and what actions are being done by whom, to maintain long-term memory (a two-hour movie can have more than 180k frames), to infer relationships between characters and objects, and eventually to reason about the events. My research thus focuses on the following topics:

(I) Multi-modal Self-supervised Representation Learning refers to a new paradigm for acquiring effective visual representation from multimodal signals, for example, videos. There is almost an infinite supply available in videos (from Youtube etc.), image level proxy tasks can be used at the frame level; and, there are plenty of additional proxy losses that can be employed from the temporal information. In this area, we are one of the pioneers, and have proposed a number of influential works that are widely used as baselines for various tasks.

(II) AI4Heathcare. For a human physician, he/she is expected to see a limited number of patients in the lifetime, each of them with a unique body mass, blood pressure, family history, and so on —a huge variety of features I track in my mental model. Each human has countless variables relevant to their health, but as a human doctor working with a limited session window, he/she will only be able to focus on the several factors that tend to be the most important historically. In contrast, for AIs, they can tirelessly process countless features of every patient, give deep, vast insights, as an example, ChatGPT, Med-PaLM2 have passed the U.S. Medical Licensing Exam. I'm keen to contribute part of my research in revolutionising the medical community !

To Students

- 你需要对计算机视觉，自监督学习，AI4Science 有兴趣，热爱探索未知
- 你需要有极强的自我驱动力，能够应对时常出现的压力和竞争，追求极致的完美
- 你需要能够突破已有学术研究格局，拒绝低质量 paper 发表，宁缺毋滥
- 我们有友好开放的实验室环境，活泼开朗的学长学姐，导师细致耐心的科研指导
- 我们提供与国际顶级研究机构，实验室合作机会，甚至访学机会
- 感兴趣同学，请将简历，成绩单发邮件: weidi@sjtu.edu.cn
- 如果实验室对你提交的内容感兴趣，会尽快联系您，并组织面试