

郑荣坤

rkzheng99@gmail.com | (+86) 13611357895 | <https://rkzheng99.github.io/>

教育背景

清华大学

电子工程系 学士

08/2018至06/2022

香港大学

计算与数据科学学院 博士生

赵恒爽助理教授组

09/2022 至今

实习经历

商汤科技

上海人工智能实验室

10/2021至09/2022

10/2022至今

相关研究经历

STAR-R1: Improving Video Perception via Spatio-Temporal Aggregated Reinforcement (NeurIPS 2025 在投)

Rongkun Zheng, Lu Qi, Xi Chen, Yi Wang, Kun Wang, Hengshuang Zhao

- 基于 RFT 的推理视频分割模型
- 针对视频理解/分割提出对应的优化奖励

ViLLa: Video Reasoning Segmentation with Large Language Model (ICCV 2025)

Rongkun Zheng, Lu Qi, Xi Chen, Yi Wang, Kun Wang, Yu Qiao, Hengshuang Zhao

- 基于大模型的推理视频分割模型
- 针对复杂的推理和长视频理解/分割提出优化方法
- 提出了包含更多视频推理分割场景的数据集

DisCo: Towards Distinct and Coherent Visual Encapsulation in Video MLLMs (ICCV 2025)

Jiahe Zhao, Rongkun Zheng, Yi Wang, Helin Wang, Hengshuang Zhao

- 基于视频多模态大模型中线性投影器导致的语义模糊与问题的改进方案
- 设计了视觉概念判别器和时间校准的模块，显著提升视频理解性能

SyncVIS: Synchronized Video Instance Segmentation (NeurIPS 2024)

Rongkun Zheng, Lu Qi, Xi Chen, Yi Wang, Kun Wang, Yu Qiao, Hengshuang Zhao

- 基于复杂视频场景的即插即用的视频分割模型，在多个 VIS 模型上取得显著性能提升
- 针对视频和单帧的物体建模提出更平衡和鲁棒的建模策略和优化方法

INTERVIDEO2: SCALING VIDEO FOUNDATION MODELS FOR MULTIMODAL VIDEO UNDERSTANDING (ECCV 2024)

Yi Wang, Kunchang Li, Xinhao Li, Jiashuo Yu, Yinan He, Chenting Wang, Guo Chen, Baoqi Pei, Ziang Yan, Rongkun Zheng, Jilan Xu, Zun Wang, Yansong Shi, Tianxiang Jiang, Songze Li, Hongjie Zhang, Yifei Huang, Yu Qiao, Yali Wang, Limin Wang

- 新的 SOTA 视频基础模型
- 在 VIS 任务上进行了测试和优化

TMT-VIS: Taxonomy-aware Multi-dataset Joint Training for Video Instance Segmentation (NeurIPS 2023)

Rongkun Zheng, Lu Qi, Xi Chen, Yi Wang, Kun Wang, Yu Qiao, Hengshuang Zhao

- 在 Mask2Former-VIS 基础上联合多数据集进行训练
- 提出了基于不同数据集类别的联合训练优化策略

Causal Prompts for Open-vocabulary Video Instance Segmentation (TPAMI 在投)

Rongkun Zheng, Lu Qi, Xi Chen, Yi Wang, Kun Wang, Hengshuang Zhao

- 开放世界视频实例分割
- 引入 prompting 策略对视频分割进行增强

TMT-VIS++: Improved Semi-supervised Taxonomy-aware Multi-dataset Joint Training for Video Instance Segmentation (TPAMI 在投)

Rongkun Zheng, Lu Qi, Xi Chen, Yi Wang, Kun Wang, Hengshuang Zhao

- 在 TMT-VIS 基础上结合 semi-supervised 策略在更大规模视频数据集上进行联合训练优化
- 提出了结合 bounding box 的半监督训练策略