

个人简历

PERSONAL RESUM

基本信息

姓名：沈飞鸿

手机：18851653218

年龄：25

邮箱：feihongshen@seu.edu.cn

学历：硕士在读

个人主页：<https://fhshen2022.github.io>



教育经历

软件工程

东南大学（硕士）

2022.09-至今

➤ 研究方向：计算机视觉，AIGC，多模态

计算机科学与技术（人工智能）

东南大学（本科）

2018.09-2022.07

➤ GPA: 3.8/4.0 (专业前 30%)

实习经历

阿里集团通义实验室

计算机视觉算法工程师-实习

2024.07-至今

➤ 研究内容：1. 基于扩散模型 SVD，利用 motion lora 实现图生视频的运镜控制。平面运镜控制成功率均达到 0.95 以上，旋转成功率达 0.49，超越 MotionCtrl 平均 0.71 和 0.37 的控制率。2. 利用单个 lora 结合 motion bucket 实现不同运动幅度的控制。3. ID 一致性约束提高主体的 ID 保持。

项目论文

基于 stable diffusion 的图片重定向

一作

2023.11-2024.05

图片重定向旨在保留关键物体的情况下调整图像的纵横比，使其适应不同的显示设备。针对传统重定向方法缺失语义导致的关键区域形变残缺等问题，提出带语义先验与结构先验的接缝雕刻算法以防止显著区域丢失；针对重定向后像素移位导致的裁剪痕迹，自适应重绘模块找出像素不连续区域，并在极端宽高比下扩图使得关键信息在任意宽高比下保留完整；最后采用 global+local 的图像引导重绘的方式保证图中物体的连续性。

论文：Prune and Repaint: Content-Aware Image Retargeting for any Ratio, **Advances in Neural Information Processing Systems**, 2024. (JCR Q1, CCF A)

基于注意力机制的 RGB-D 显著物体检测

学生一作

2022.03-2024.01

针对现有 RGB-D 显著物体检测任务中模态特征融合时模态差异与空间差异同时存在的问题，提出了跨模态自注意力与局部互注意力，并将两模态特征解耦为高相似度与高差异度特征，显式地挖掘一致性与互补性。在 8 个数据集上达到 SOTA，并将融合算法扩展至 RGB-D 语义分割任务。

论文：Disentangled Cross-modal Transformer for RGB-D Salient Object Detection and Beyond, **IEEE Transactions on Image Processing**, 2024. (JCR Q1, CCF A)

平衡模态的 RGB-D 对比学习框架

学生一作

2022.06-2023.06

提出了一种针对 RGB 图像与深度图像的自监督对比学习框架，通过分阶段预训练，强模态掩码和梯度调整三个策略，解决模态学习不平衡问题。预训练后的编码器为下游 RGB-D 显著物体检测带来更强的 RGB 与深度模态的特征提取与融合能力，并将框架扩展至 RGB-T 数据。

论文：Modality-Balanced Contrastive Learning for RGB-D Salient Object Detection.

荣誉证书

- 思特奇未来之星教育基金一等奖（年级前 10%；2024）
- 东南大学二等学业奖学金（2022, 2023）
- “中国光谷·华为杯”第十九届中国研究生数学建模竞赛二等奖（2022）
- CET 4 (567); CET 6 (545)