

AENHANCING AMODAL COMPLETION FOR VIDEO: PROGRESSIVE MIXED CONTEXT DIFFUSION APPROACH

Dương Phạm Hoàng Anh - 22520042

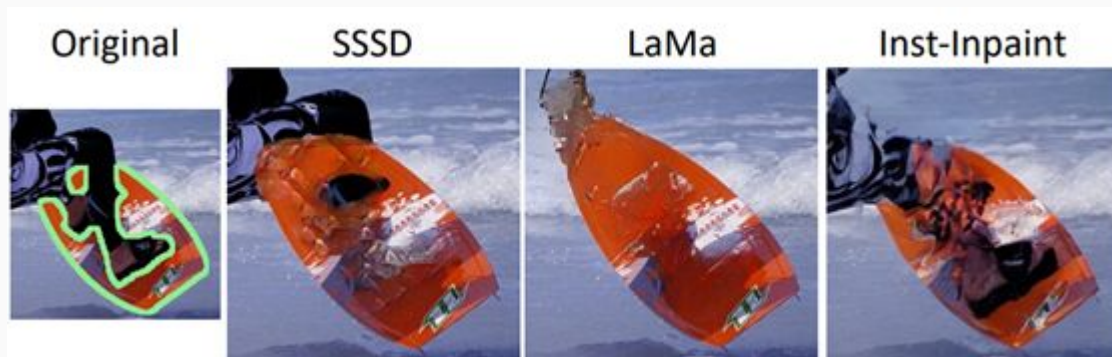
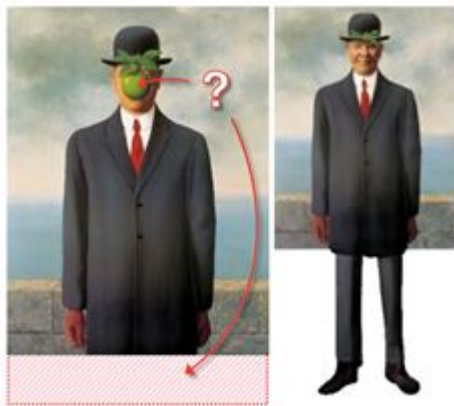
Tóm tắt

- Lớp: CS519.021.KHTN
- Link Github: [CS519.021.KHTN/AMODAL COMPLETION CHO VIDEO: ÁP DỤNG PROGRESSIVE MIXED CONTEXT DIFFUSION at main · HoangAnh109/CS519.021.KHTN \(github.com\)](https://github.com/CS519.021.KHTN/AMODAL_COMPLETION_CHO_VIDEO:_ÁP_DUNG_PROGRESSIVE_MIXED_CONTEXT_DIFFUSION)
- Link YouTube video: https://youtu.be/I9_dJlo5k6U
- Ảnh + Họ và Tên: Dương Phạm Hoàng Anh
- Tổng số slides không vượt quá 10



Giới thiệu

- Amodal Completion: Nhận diện và hoàn thiện đối tượng bị che khuất
- Progressive Mixed Context Diffusion
- Mở rộng từ hình ảnh tĩnh sang chuỗi video



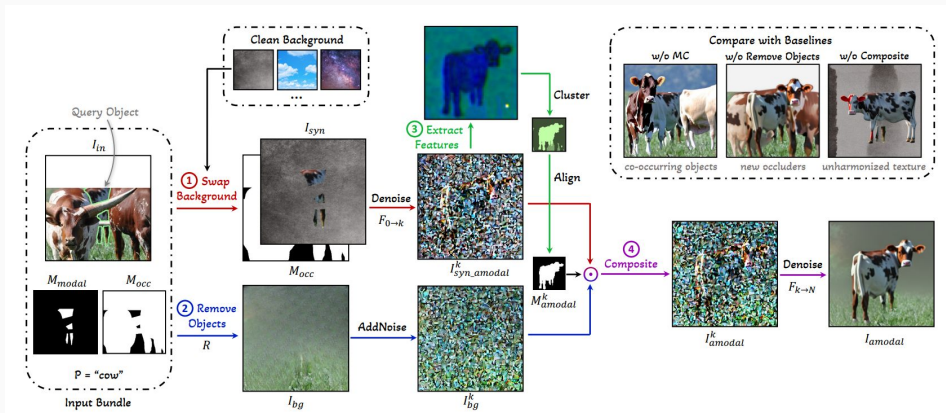
Mục tiêu

- **Hoàn thiện amodal:** Áp dụng mô hình khuếch tán để hoàn thiện amodal cho từng khung hình.
- **Duy trì liên tục:** Đảm bảo tính nhất quán và liên tục của đối tượng qua chuỗi khung hình.
- **Giảm co-occurrence bias:** Sử dụng kỹ thuật Mixed Context Diffusion Sampling để giảm thiểu đối tượng không mong muốn.

Nội dung và Phương pháp

Khuếch Tán Ngữ Cảnh Hỗn Hợp (Mixed Context Diffusion):

- Swap Background:** Thay nền gốc bằng nền sạch tạm thời, tạo hình ảnh tổng hợp (I_{syn}) và áp dụng khuếch tán đến bước k ($I_{k,syn}^{amodal}$).
- Create Object-removed Background Image:** Loại bỏ đối tượng chính và occluders từ khung hình gốc, thêm noise để đạt bước khuếch tán k ($I_{k,bg}$).
- Segment Query Object in Noisy Image:** Tạo cluster từ đặc trưng khuếch tán và chọn mask amodal ở bước k ($M_{k,amodal}$).
- Composite:** Sử dụng $M_{k,amodal}$ để gắn đối tượng vào nền sạch, tạo hình ảnh tổng hợp $I_{k,amodal}$, hoàn thiện hình ảnh cuối cùng (I_{amodal}).



Nội dung và Phương pháp

1. **Phân Tích và Phân Đoạn Video:**
2. **Khuếch Tán Ngữ Cảnh Hỗn Hợp (Mixed Context Diffusion):**
3. **Duy Trì Tính Liên Tục và Nhất Quán:**

Thực hiện hoàn thiện đối tượng từng bước, sử dụng đầu ra của bước trước làm đầu vào.

Kiểm tra và tiếp tục khuếch tán nếu đối tượng còn bị che khuất.

Đảm bảo tính nhất quán và liên tục của đối tượng qua các khung hình.

4. **Progressive Context Control:**

Giảm thiểu thiên vị ngữ cảnh bằng cách thay nền sạch.

Xác định và loại bỏ occluders mới để đảm bảo đối tượng không bị méo mó.

5. **Kiểm Tra và Đánh Giá:**

Sử dụng PSNR, SSIM, và LPIPS để đo lường chất lượng video hoàn thiện.

Kết quả dự kiến

Hệ thống hoàn thiện amodal cho video:

Xây dựng một mô hình có khả năng phân tích và xử lý các video đầu vào, hoàn thiện các đối tượng bị che khuất trong từng khung hình.

Khả năng duy trì tính liên tục và nhất quán:

- **Đối tượng liên tục:** Hệ thống đảm bảo các đối tượng trong video được hoàn thiện một cách liên tục và không bị đứt đoạn qua các khung hình.
- **Chất lượng cao:** Sử dụng các kỹ thuật tiên tiến như Mixed Context Diffusion để đảm bảo tính chân thực và giảm thiểu sự thiên vị không mong muốn.

Tài liệu tham khảo

- Katherine Xu, Lingzhi Zhang, Jianbo Shi: Amodal Completion via Progressive Mixed Context Diffusion. CoRR abs/2312.15540 (2023)