

NGHIÊN CỨU VÀ XÂY DỰNG HỆ THỐNG ẨN DANH HÓA ĐỐI TƯỢNG CÓ CHỌN LỌC TRONG VIDEO THỜI GIAN THỰC SỬ DỤNG MẶT NẠ KỸ THUẬT SỐ

Nguyễn Thiên Nhân¹

Bùi Ngọc Thiên Thanh¹

¹ Trường Đại Công nghệ Thông tin, Đại học Quốc gia TP. Hồ Chí Minh

Vấn đề nghiên cứu ?

- Nghiên cứu đề xuất giải pháp **ẩn danh hóa có chọn lọc** cho video phát trực tiếp.
- Tập trung vào việc **che mặt người lạ** xuất hiện vô tình trong khung hình mà **không làm ảnh hưởng đến chủ thể chính**.
- Thay thế các phương pháp làm mờ truyền thống bằng **mặt nạ kỹ thuật số** để tăng tính thẩm mỹ.
- Xây dựng hệ thống hoạt động **theo thời gian thực**.

Tại sao cần nghiên cứu ?

- Bối cảnh:** Sự bùng nổ của livestream ngoại cảnh đe dọa quyền riêng tư của người xung quanh.
- Hạn chế hiện tại:**
 - Công cụ an ninh: Bảo mật tốt nhưng gây mất thẩm mỹ, giảm trải nghiệm người xem.
 - Filter giải trí: Chỉ làm đẹp chủ thể, bỏ qua việc xử lý/bảo vệ người ở hậu cảnh.
- Mục tiêu:** Cân bằng giữa sự tự do sáng tạo nội dung và bảo vệ quyền riêng tư bằng giải pháp thẩm mỹ hơn.

Tổng quan



Mô tả chi tiết

1. Cơ sở lý thuyết & Mô hình

- Công nghệ cốt lõi:** Sử dụng Deep Learning để giải quyết bài toán phát hiện đối tượng thời gian thực.
- Mô hình đề xuất:**
 - Tập trung vào các giải pháp hạng nhẹ để tối ưu hiệu năng, đảm bảo chạy mượt mà trên thiết bị cá nhân.
 - Áp dụng các phương pháp SOTA về Face Detection và Trích xuất đặc trưng.
- Theo dõi đa đối tượng:** Giúp hệ thống ghi nhớ và theo dõi vị trí của từng người trong khung hình liên tục.



Hình 1. Phát hiện khuôn mặt.

2. Giải pháp ẩn danh & xử lý

- Phân loại tự động:** Hệ thống tự động định danh để phân biệt đâu là chủ thể (không che) và đâu là người lạ (cần che).
- Kỹ thuật mặt nạ số:**
 - Xây dựng module thay thế khuôn mặt người lạ bằng các hình ảnh đồ họa.
 - Sử dụng thuật toán biến đổi hình học để xác định vị trí, kích thước và góc nghiêng khuôn mặt.
 - Ứng dụng kỹ thuật Image Blending để ghép mặt nạ tự nhiên.



Hình 2. Phương pháp ẩn danh truyền thống.

3. Thực nghiệm & kết quả mong đợi

- Tập dữ liệu:** Tự thu thập dữ liệu video thực tế tại môi trường công cộng (quán cafe, công viên) với điều kiện ánh sáng và mật độ người khác nhau.
- Tiêu chí đánh giá:**
 - Định lượng: Tốc độ khung hình, độ trễ, độ chính xác phân loại.
 - Định tính: Tính thẩm mỹ và trải nghiệm người xem so với phương pháp làm mờ truyền thống.
- Kết quả hướng tới:** Hệ thống vận hành ổn định, mặt nạ bám sát chuyển động khuôn mặt, tạo cảm giác thoải mái cho người xem.



Hình 3. Kết quả dự kiến.