

PHÁT HIỆN BẤT THƯỜNG TRONG CÁC VIDEO THU LẠI ĐƯỢC TỪ CAMERA GIÁM SÁT

Phạm Thanh Bình - 230201003

Trường Đh Công nghệ Thông tin – Đh Quốc gia TP HCM

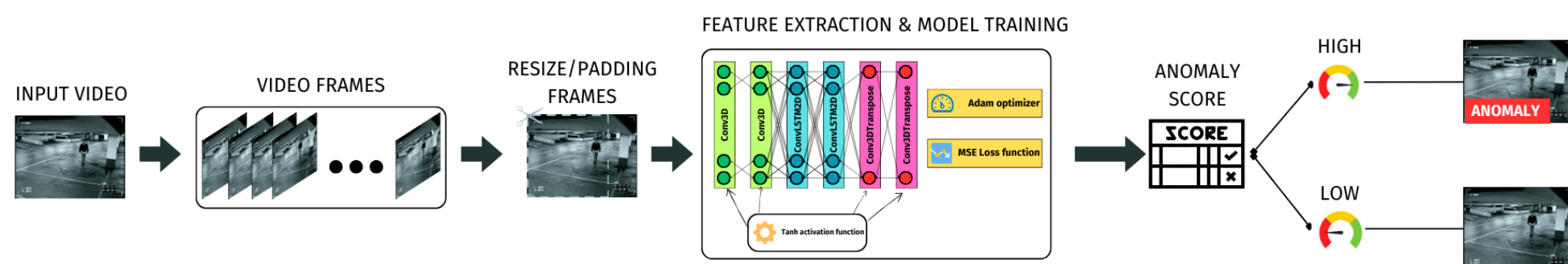
Giới Thiệu

- Phát hiện bất thường trong video là một chủ đề thiết thực trong cuộc sống của chúng ta. Tuy nhiên, các mô hình phổ biến tới nay đã trở nên lỗi thời và không còn hiệu quả.
- Do vậy nên chúng tôi đã nghiên cứu mô hình mạng **Convolutional LSTM Autoencoder**.
- Mô hình kết hợp khả năng trích xuất đặc trưng không gian của mạng **CNN** với khả năng trích xuất đặc trưng thời gian của mạng **LSTM** trong kiến trúc của mạng **Autoencoder**.

Mục tiêu

- Phát triển thành công mô hình phát hiện bất thường trong video với độ chính xác từ 80% trở lên.
- Cải thiện quá trình training của mô hình sao cho không bị tốn quá nhiều tài nguyên và thời gian.
- Nếu hoàn thành 2 mục tiêu trên, chúng tôi sẽ nghiên cứu mô hình có thể phát hiện bất thường với input là video trong thời gian thực thay vì video có sẵn.

Sơ đồ mô hình



Chi tiết

1. Dataset

Chúng tôi sử dụng 2 bộ dataset sau:

Bộ dataset	Số lượng video	Nội dung
Avenue Dataset	16 video huấn luyện, 21 video thử nghiệm	Các video giám sát được chọn lọc kỹ càng
UCF-Crime	1900 video	13 loại bất thường khác nhau, chủ yếu là hoạt động phi pháp và tai nạn

2. Tổng quan mô hình

- Chúng tôi chọn lọc các video đầu vào để đầu ra của mô hình đạt hiệu quả tốt nhất có thể.
- Các video đầu vào sẽ được resize hoặc padding sao cho phù hợp với chuẩn của mô hình.
- Mô hình của chúng tôi sẽ có cấu trúc chính như sau:

1. **Conv3D**: là lớp đầu tiên của mô hình, các lớp 3 chiều này sẽ học đặc trưng từ từng frame của video.

2. **ConvLSTM2D**: các lớp này là một lớp kiểu của Mạng nơ ron tái phát triển (RNN), chúng sẽ học các dependencies tạm thời trong dữ liệu.

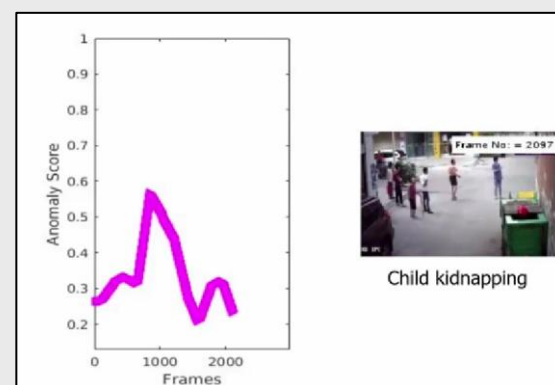
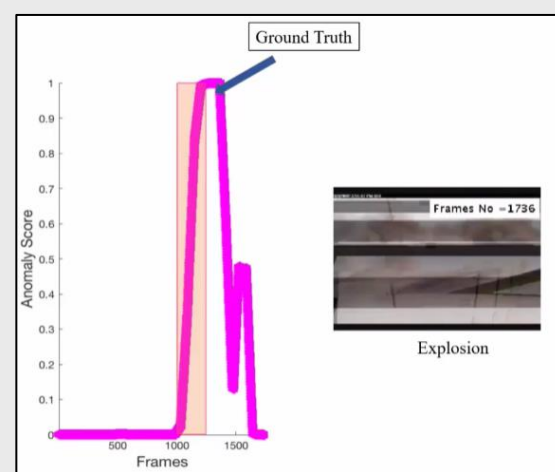
3. **Conv3DTranspose**: các lớp này sẽ thực hiện nghịch đảo của thao tác tích chập từ lớp 3D trước đó. Chúng còn được sử dụng để lấy mẫu lại từ bản đồ đặc trưng.

4. **Hàm kích hoạt**: mô hình sẽ sử dụng hàm "tanh", đảm bảo rằng các giá trị đầu ra trong phạm vi mà ta mong muốn, đây là một điều cần thiết đối với các tác vụ như tạo ảnh hoặc khử nhiễu ảnh.

5. **Hàm mất mát và thuật toán tối ưu**: để tối ưu đầu ra của mô hình chúng tôi sử dụng hàm mất mát MSE và thuật toán tối ưu Adam để đầu ra của mô hình có thể chính xác hơn.

6. Mô hình duyệt qua từng frame của video và tính điểm bất thường dựa vào hàm mất mát.

3. Kết quả dự kiến



Khi phát video tới frame có điểm bất thường cao, mô hình sẽ hiển thị cảnh báo.