

Nhận dạng hành vi và cảnh báo bạo lực trong đám đông

Nguyen-Anh-Cuong, Vu-Khanh-Hoan, Nong-Trung-Hieu

Khoa Công nghệ thông tin - Dainam University, Hanoi, Vietnam

Github: https://github.com/NguyenAnhCuong/Violence_Detection

Giới thiệu

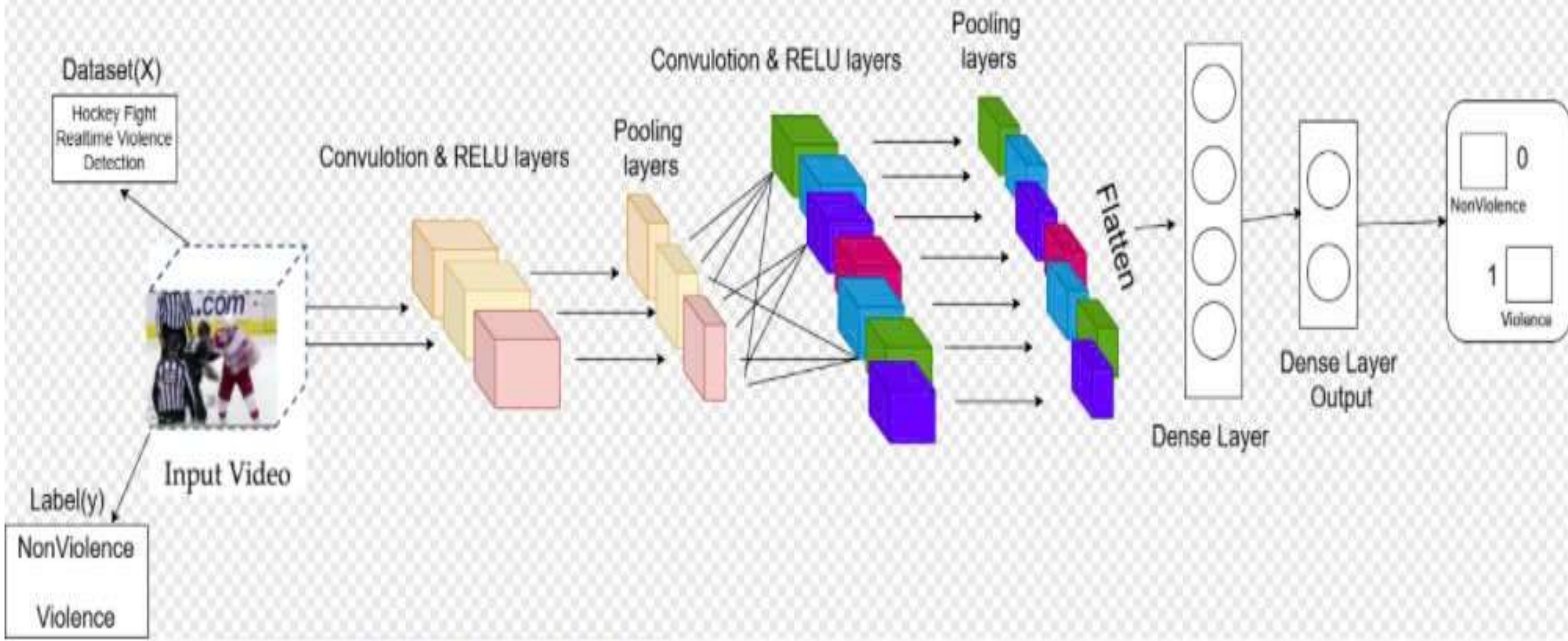
Nhận diện hành vi và cảnh báo bạo lực trong đám đông

- Nhằm phát hiện và phân loại các hành vi bạo lực trong đám đông từ dữ liệu video theo thời gian thực.
- Có tiềm năng ứng dụng lớn trong nhiều lĩnh vực như giám sát an ninh, kiểm soát đám đông và hỗ trợ cơ quan thực thi pháp luật.

Đóng góp của nghiên cứu

- Đề xuất ứng dụng **mạng nơ-ron tích chập 3D (3D CNN)** để trích xuất đặc trưng không gian-thời gian từ dữ liệu video, giúp nâng cao độ chính xác trong việc nhận diện hành vi bạo lực.
- Kết hợp mô hình **IoT với cảm biến thông minh** để thu thập và xử lý dữ liệu thời gian thực, từ đó cải thiện hiệu quả hệ thống cảnh báo bạo lực.

Sơ đồ 3D CNN

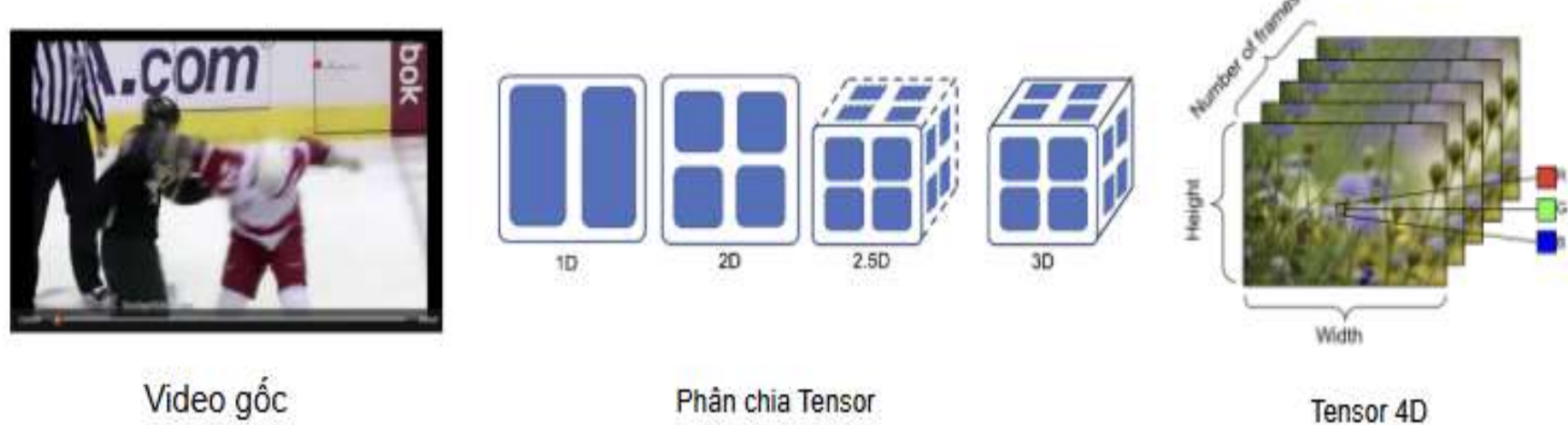


Phương pháp sử dụng

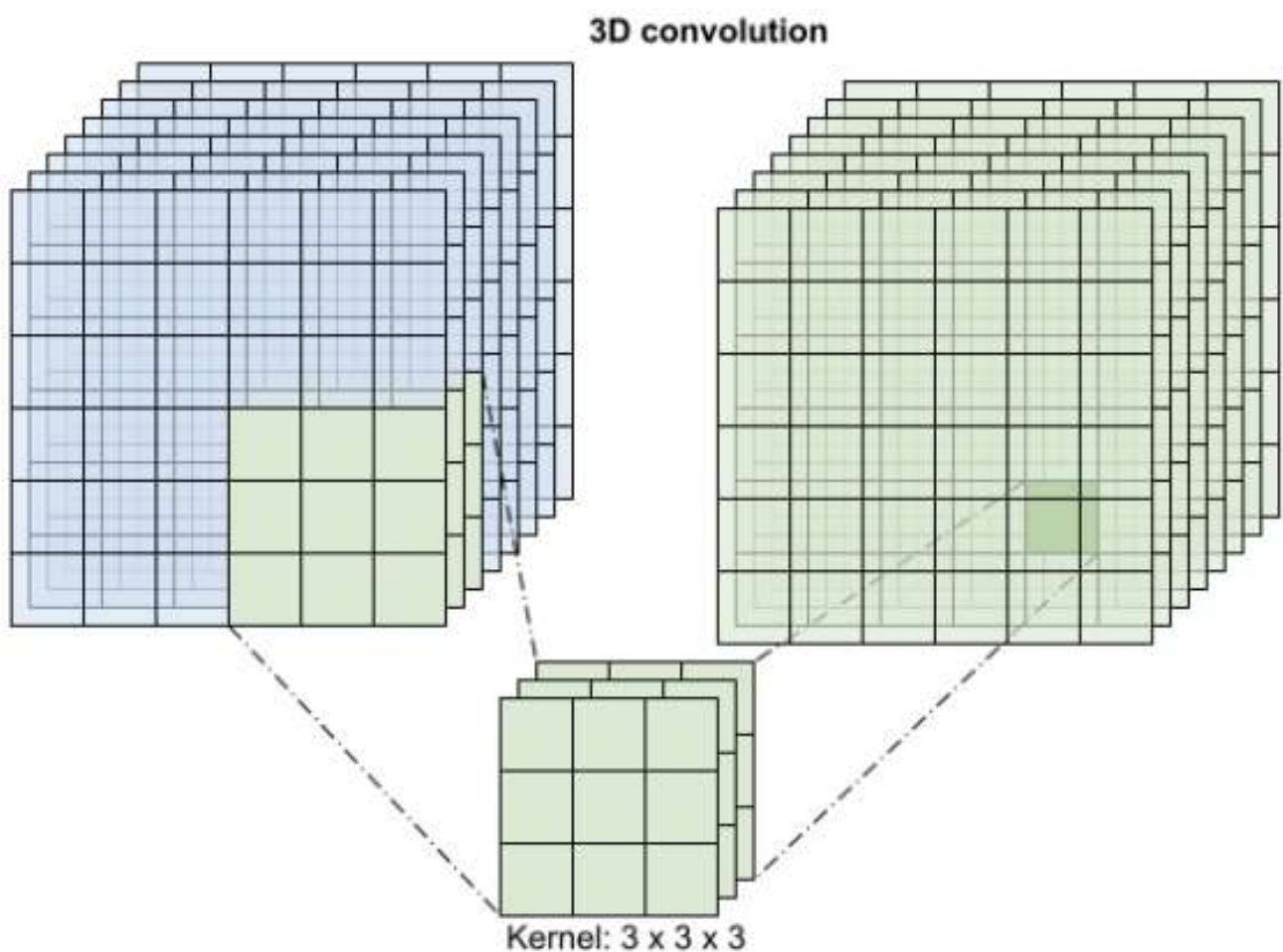
Hệ thống đề xuất bao gồm bốn thành phần chính:

- Tiền xử lý dữ liệu: Lọc nhiễu và chuẩn hóa dữ liệu video để cải thiện độ chính xác của mô hình.
- Trích xuất đặc trưng không gian-thời gian: Sử dụng 3D CNN để tạo biểu diễn đặc trưng từ các khung hình liên tiếp.

Chuyển đổi video thành tensor:



Video được chuyển đổi thành tập hợp các frame liên tiếp. Dữ liệu đầu vào được biểu diễn dưới dạng Tensor với cấu trúc: [batch_size, number_of_frames, height, width, channels]



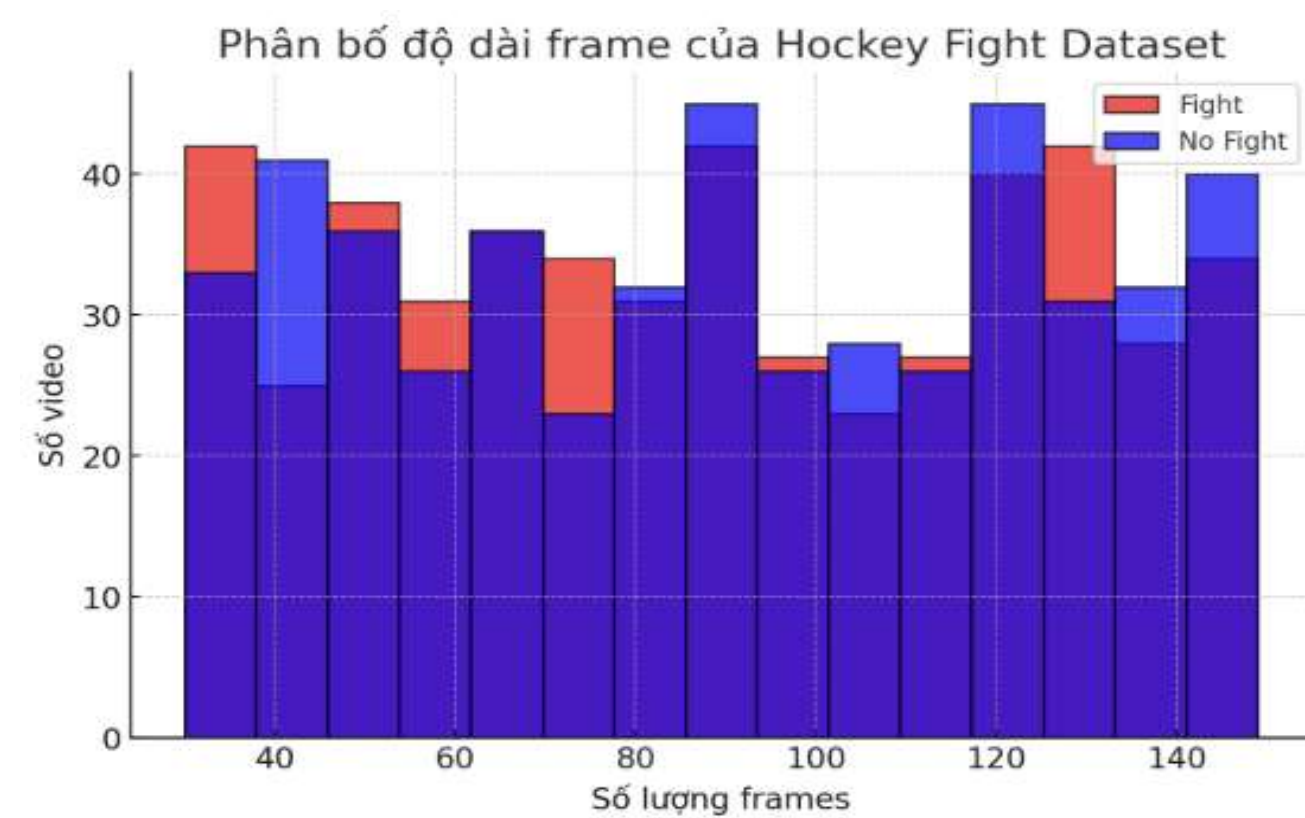
Áp dụng 3D-CNN sử dụng các kernel 3D để học đặc trưng không gian và thời gian từ video. Kernel có kích thước 3x3x3 giúp mô hình nhận diện hành vi dựa trên sự thay đổi của các frame theo thời gian.

Tập dữ liệu

Hockey Fight dataset:

- Trận đấu khúc côn cầu trên băng, nơi có các hành vi va chạm, đánh nhau giữa các cầu thủ.
- Phân loại video chứa hành vi bạo lực và không bạo lực.

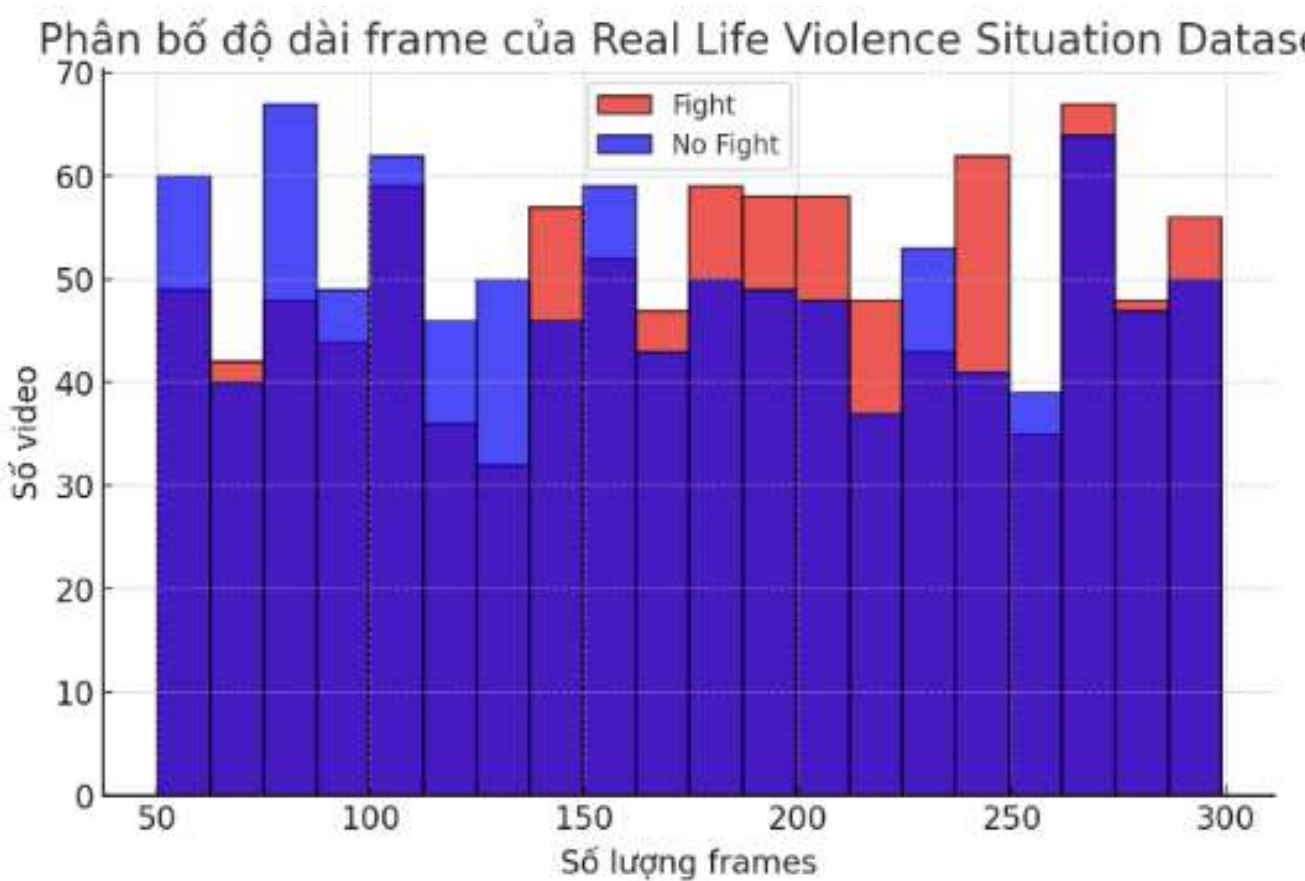
Attribute	Quantity
Classes	2
Frames	~30
Training set	800
Testing set	200



Real Life Violence Situation dataset:

- Gồm các video quay trong môi trường thực tế (đường phố, trung tâm thương mại, vv).
- Chứa cả cảnh bạo lực (đánh nhau, xô xát) và không bạo lực (đám đông bình thường, vv)

Attribute	Quantity
Classes	2
Frames	~50
Training set	1600
Testing set	400

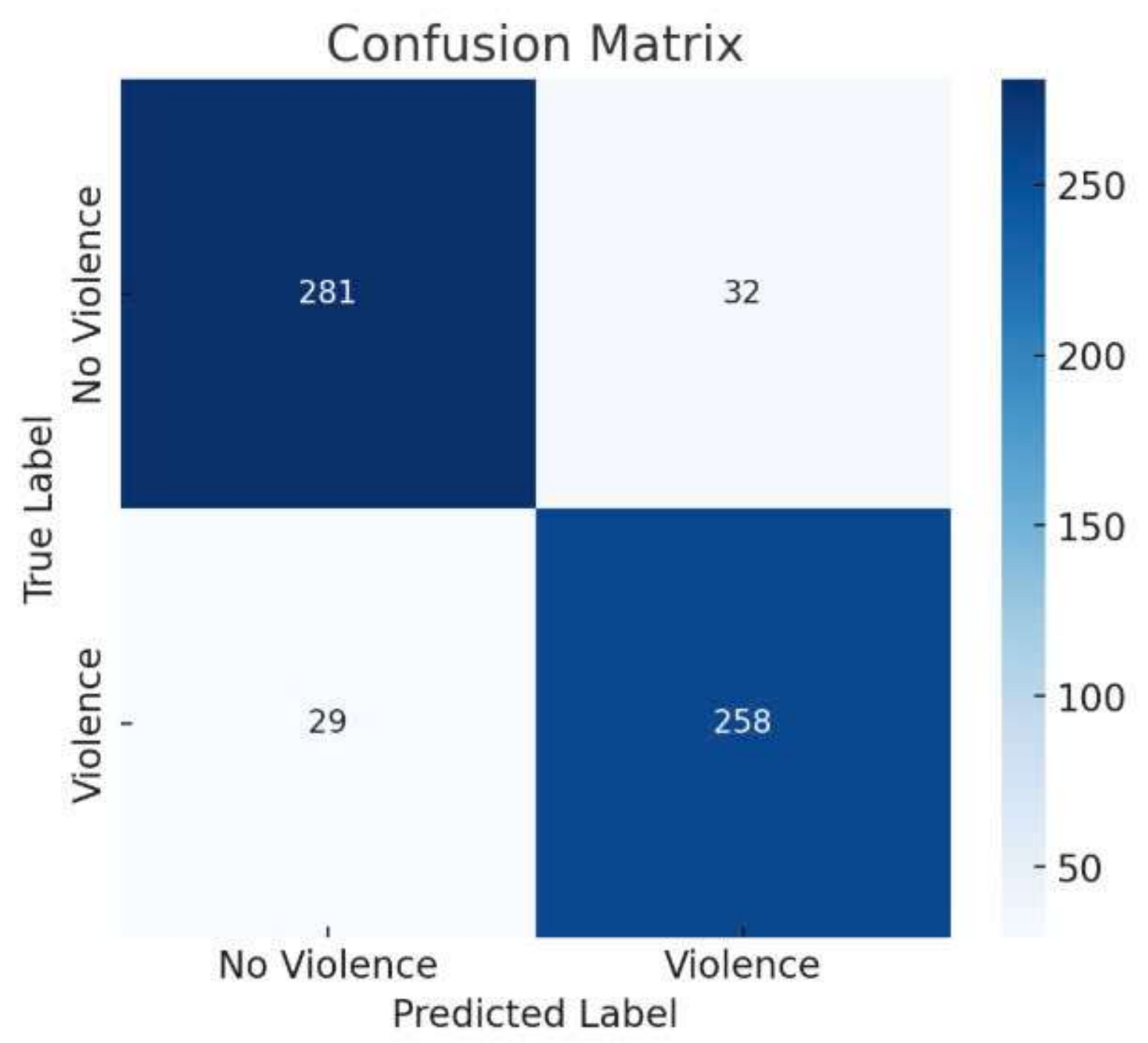


Kết quả

Đánh giá mô hình

Accuracy (%)	Precisson (%)	Recall (%)	F1-score (%)
90.00	89.00	89.60	89.00

Ma trận nhầm lẫn



Kết luận và các công việc trong tương lai

Kết luận:

- Kết quả cho thấy mô hình 3D CNN giúp cải thiện độ chính xác trong nhận diện hành vi, trong khi IoT và điện toán biên giúp hệ thống phản hồi nhanh hơn. Hệ thống có tiềm năng hỗ trợ giám sát an ninh hiệu quả và giảm thiểu rủi ro trong các tình huống bạo lực.

Trong tương lai:

- Tối ưu mô hình AI để tăng tốc độ xử lý và giảm tài nguyên tính toán. Ứng dụng vào thực tế trong các hệ thống giám sát giao thông, sự kiện đông người.