

Kiến trúc phát hiện phương tiện hiệu quả trong hệ thống giám sát giao thông

Đặng Đình Quốc Trung, Chế Vũ Gia Hy, Đinh Bá Tiến

Khoa Công nghệ Thông tin,

Trường Đại học Khoa học Tự Nhiên, ĐHQG-HCM

ddqtrung@apcs.vn, cvghy@apcs.vn, dbtien@fit.hcmus.edu.vn

Tóm tắt

Hệ thống giám sát giao thông có thể giải quyết nhiều vấn đề ở Việt Nam. Bên cạnh đó, việc phát triển xe tự lái cũng cần một phương pháp phát hiện xe cộ hiệu quả. Các kiến trúc hiện đại có độ chính xác cao, nhưng chúng thường rất chậm nên ta khó áp dụng cho những ứng dụng thời gian thực hoặc nhúng. Trong báo cáo này, chúng tôi đề xuất một kiến trúc phát hiện phương tiện vô cùng nhanh và tiết kiệm bộ nhớ so với những thiết kế phức tạp, nhưng đem lại kết quả tốt hơn các phương pháp thường dùng trong hệ thống nhò. Kiến trúc này dựa trên mạng MobileNetV2 sử dụng cho bài toán phân loại đối tượng. Bộ dữ liệu của nghiên cứu được xây dựng từ video giao thông ở thành phố Hồ Chí Minh và Đà Nẵng, cùng với tập dữ liệu DETRAC. Kết quả thí nghiệm cho thấy giải pháp của chúng tôi có độ chính xác tương đương kiến trúc phức tạp (75%) dù sử dụng ít tham số (3.4 triệu) và mất 100ms để xử lý một ảnh.

Từ khóa: phát hiện vật thể, theo dõi vật thể, giám sát giao thông, ...

Efficient vehicle detector in traffic surveillance systems

Dang Dinh Quoc Trung, Che Vu Gia Hy, Dinh Ba Tien

Faculty of Information Technology, University of Science, VNU-HCM

ddqtrung@apcs.vn, cvghy@apcs.vn, dbtien@fit.hcmus.edu.vn

Abstract

Traffic surveillance is the key to solve various traffic problems in Vietnam. In addition, self-driving cars are being developed in some countries, so the needs for a fast and reliable vehicle detector are rising. Although recent deep neural architectures have shown a dramatic boost in performance, they are often too slow and burdensome for embedded and real-time applications. In this paper, we describe a new detection architecture that is more efficient than state-of-the-art detectors while improving the accuracy of small mobile models. Moreover, our main setting is traffic in Vietnam, which has distinctive characteristics compared to that of big cities. In more detail, our architecture is based on MobileNetV2, whose building blocks consist of depthwise convolution layers. We train this network with smooth localization loss and weighted softmax loss in tandem with hard negative mining. Both training and test sets are built with recorded videos of Ho Chi Minh and Da Nang traffic or selected from DETRAC dataset. The experimental results show that our proposed solution can still achieve a solid performance (75%) comparable to state-of-the-art architectures on the test set while using significantly fewer parameters (3.4 million) and running faster (100ms) on a cheap machine.

Key words: object detection, object tracking, traffic surveillance, ...