

ỨNG DỤNG XỬ LÝ ẢNH SỐ VÀ VIDEO SỐ SEMINAR

Lưu Nam Đạt
22127062
ln-dat22@clc.fitus.edu.vn

Nguyễn Bá Công
22127046
nbcong22@clc.fitus.edu.vn

Nguyễn Huỳnh Hải Đăng
22127052
nhhdang22@clc.fitus.edu.vn

Đặng Trần Anh Khoa
22127024
dtakhoa22@clc.fitus.edu.vn

LỜI GIỚI THIỆU

None

MỤC LỤC

| | |
|--|----------|
| 1 NHẬN DIỆN HÀNG HOÁ BÁN LẺ | 1 |
| 1.1.1 phương pháp dựa theo đặc trưng | 1 |
| 1.1.2 Deep learning | 1 |
| 2 HỆ THỐNG TRUY VẾT ĐỐI TƯỢNG DỰA VÀO CÂU MÔ TẢ | 2 |
| 3 Phát hiện bất thường trong giao thông | 2 |
| Tham khảo | 2 |

1 NHẬN DIỆN HÀNG HOÁ BÁN LẺ

AMAZON GO

(cần xếp thành category, cần có những tác vụ gì)

1.1.1 PHƯƠNG PHÁP DỰA THEO ĐẶC TRƯNG

SIFT - scale invariant feature extraction

1.1.2 DEEP LEARNING

A deep learning pipeline for product recognition on store shelves

Detection

2 HỆ THỐNG TRUY VẾT ĐỐI TƯỢNG DỰA VÀO CÂU MÔ TẢ

1. GIỚI THIỆU

1.1. BỐI CẢNH CHUNG

- xe tự hành, giao thông, an ninh

thách thức: hạn chế ngôn ngữ phân biệt đối tượng mục tiêu theo vết trong điều kiện phức tạp

2. PHÁT BIỂU BÀI TOÁN

3. CÁC CÔNG TRÌNH LIÊN QUAN

phải nói rõ về cách thức theo vết đối tượng

- TP-GMOT: Tracking Generic Multiple Object by Textual Prompt with Motion Appearance Cost SORT
- DTLLM-VLT:

tại 1 frame bất kỳ, có 2 trường hợp: 1 là đối tượng đang theo vết bị biến mất, 2 là đối tượng xuất hiện; khi đó câu mô tả phát huy như thế nào?

3 PHÁT HIỆN BẤT THƯỜNG TRONG GIAO THÔNG

Input: một đoạn video từ camera hành trình / camera an ninh

Output: Xác suất xảy ra tai nạn trong frame đang xét

Threshold: Một ngưỡng cảnh báo mức độ nguy hiểm

- MEDAVET: Traffic Vehicle Anomaly Detection Mechanism based on

spatial and temporal structures in vehicle traffic

- chưa giải thích được cơ chế tìm chiều di chuyển và vận tốc của phương tiện
- cần nói rõ ý chung trước khi đi sâu vào những biểu đồ và thuật toán, tuy có rất nhiều những neural network nhưng việc giải thích chưa đáng kể
- trong khung cảnh mà camera bắt được,
- dữ liệu không gian - thời gian (spatial - temporal)

cần hiểu “thế nào là tai nạn”

- đối với mỗi frame, cần quan tâm đến object nào để tính toán ra xác suất?

→ liệt kê 11 vật thể nó quan tâm:

- từ hình ảnh, rút ra đối tượng ra sao, từ đối tượng rút ra xác suất thế nào

Dùng YOLOv7 để phát hiện

-

THAM KHẢO