

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI



ĐỒ ÁN TỐT NGHIỆP

Phát hiện trẻ em sử dụng Generalized Attention

TRƯỜNG VĂN THỂ

The.tv219245@sis.hust.edu.vn

Ngành Công nghệ thông tin

Giảng viên hướng dẫn: PGS.TS. Phạm Văn Hải

Chữ ký của GVHD

Bộ môn:

Viện: Đào tạo liên tục

HÀ NỘI, 01/2024

1. Thông tin sinh viên

Họ và tên sinh viên: Trương Văn Thế

Điện thoại liên lạc: 0979717979 Email: The.TV219245@sis.hust.edu.vn

Lớp: VB2 CQ CNTT – K66

Hệ đào tạo: Đại học chính quy

Đề án tốt nghiệp được thực hiện tại: Trường Đại học Bách Khoa Hà Nội.

Thời gian làm đề án tốt nghiệp: Từ ngày 20/11/2023 đến 20/02/2024

2. Mục đích nội dung của đề án tốt nghiệp

Phát hiện trẻ em sử dụng Generalized Attention, trong đó có chức năng tìm trẻ em hoặc người lớn lạc trong đám đông.

3. Các nhiệm vụ cụ thể của đề án tốt nghiệp

- Thu thập dữ liệu ảnh trên mạng hoặc trên thực địa về trẻ em.
- Gán nhãn ảnh trẻ em bằng LabelImg hoặc LabelMe hoặc CVAC.
- Huấn luyện chuyển tiếp GeneralizedAttention trên COCO với tập dữ liệu gán nhãn sử dụng khung làm việc mmdetection
- Phân tích và thiết kế hệ thống.
- Cài đặt hoàn chỉnh hệ thống.

4. Lời cam đoan của sinh viên:

Tôi – Trương Văn Thế - cam kết đề án tốt nghiệp này là công trình nghiên cứu của bản thân tôi dưới sự hướng dẫn của PGS.TS Phạm Văn Hải.

Các kết quả được trình bày trong đề án tốt nghiệp là trung thực, không phải là sao chép toàn văn của bất kỳ công trình nào khác.

5. Xác nhận của giáo viên hướng dẫn về mức độ hoàn thành của đề án tốt nghiệp và cho phép bảo vệ:

.....
.....
.....
.....

Giáo viên hướng dẫn

PGS.TS Phạm Văn Hải

Lời cảm ơn

Trước hết, em xin cảm ơn chân thành và sâu sắc thầy giáo hướng dẫn PGS.TS Phạm Văn Hải, giảng viên viện Công Nghệ Thông Tin và Truyền Thông trường Đại Học Bách Khoa Hà Nội, người đã hướng dẫn em rất tận tình, chỉ cho em hướng tiếp cận phù hợp với bài toán, cung cấp cho em nhiều tài liệu tham khảo phù hợp trong suốt quá trình làm đồ án.

Đồng thời, em cũng xin được gửi lời cảm ơn đến các thầy cô trong trường Đại Học Bách Khoa Hà Nội, trong Viện Đào tạo Liên tục đã dạy bảo, truyền đạt cho em tất cả những kiến thức từ cơ bản đến nâng cao trong suốt 2 năm học, để em có đủ kiến thức để hoàn thiện được đồ án.

Và em cũng gửi lời cảm ơn sâu sắc đến gia đình và những người bạn đã giúp đỡ em rất nhiều trong học tập và trong cuộc sống.

Do thời gian cũng như kiến thức còn hạn chế, nên đồ án của em không thể tránh khỏi những thiếu sót, khiếm khuyết, rất mong thầy cô và các bạn đóng góp để em có thể hoàn thiện tốt hơn đồ án của mình.

Tóm tắt nội dung đồ án

Trẻ em đi lạc tại các nơi đông người có thể xảy ra. Việc tìm kiếm trẻ em trong các trường hợp này không hề dễ dàng, gây lo lắng không nhỏ cho các bậc phụ huynh. Việc phát hiện trẻ em trong các lễ hội đông người là cần thiết.

AI đang được ứng dụng khắp nơi, nhất là trên các camera quan sát để có thể thực hiện phát hiện trẻ em trên không gian rộng, thay cho con người. Hàng năm, có rất nhiều mô hình phát hiện vật thể ra đời vì đây là tác vụ rất quan trọng trong thị giác máy tính trong đó có mô hình GeneralizedAttention.

Hệ thống thông tin tìm kiếm trẻ em thất lạc có thể được thể hiện như trong hình sau: Các camera quan sát được đặt khắp nơi trong khu vực đông người, chẳng hạn siêu thị hay lễ hội và truyền thông tin thời gian thực về hệ thống trung tâm nơi chứa CSDL ảnh. Khách có nhu cầu tìm kiếm trẻ lạc thực hiện gửi một tấm ảnh gần nhất của bé lên ứng dụng để tìm kiếm vị trí hiện tại chính là tọa độ khu vực quan sát của camera chứa hình ảnh cập nhật của bé thông qua CSDL ảnh.

Trong Module nhận dạng mẫu thực hiện các thao tác xử lý ảnh bao gồm: trích chọn đặc trưng ảnh đầu vào, nhận dạng trẻ em trong các ảnh cung cấp bởi CSDL ảnh, đối sánh giữa đặc trưng ảnh đầu vào với ảnh của trẻ em trong CSDL ảnh.

Đồ án phát hiện trẻ em sử dụng Generalized Attention tập trung giải quyết vấn đề tìm trẻ em bị thất lạc tại các nơi tập trung đông người thông qua việc nhận dạng trẻ em bị lạc trong các ảnh cung cấp bởi CSDL ảnh, giúp phụ huynh các bé hoặc bảo vệ nhanh chóng xác định vị trí các bé để tiếp cận tìm kiếm đảm bảo an toàn cho các bé.

Sinh viên thực hiện

Trương Văn Thế

Mục lục

LIỆT KÊ HÌNH ẢNH	vii
DANH MỤC CÁC TỪ VIẾT TẮT VÀ GIẢI THÍCH CÁC THUẬT NGỮ	viii
CHƯƠNG 1 GIỚI THIỆU	1
1.1 Tổng quan.....	1
1.2 Bài toán	1
1.3 Bố cục đồ án.....	2
CHƯƠNG 2 TỔNG QUAN DỮ LIỆU	3
2.1 Tổng quan về gán dữ liệu.....	3
2.2 LabelImg là gì?	4
2.3 Chức năng chính	4
2.3.1 Giao diện thân thiện:.....	4
2.3.2 Hỗ trợ nhiều định dạng ảnh:	4
2.3.3 Tạo bounding box:	5
2.3.4 Gán nhãn:	5
2.3.5 Lưu trữ dữ liệu:	5
2.3.6 Tích hợp với các dự án máy học:.....	5
2.4 Cài đặt LabelImg :	5
2.5 Giao diện LabelImg:.....	6
2.6 Các bước gán nhãn dữ liệu.....	6
2.7 Ảnh đã thu thập	9
CHƯƠNG 3 LỰA CHỌN MÔ HÌNH.....	11
3.1 Mô hình nhận dạng vật thể Faster RCNN	11
3.1.1 Faster R-CNN	14
3.2 Mạng xương sống Resnet	14
3.3 Các khối trích chọn tập trung sử dụng Generalized Attention	16
CHƯƠNG 4 CÀI ĐẶT VÀ KẾT QUẢ	18
4.1 MMDetection	18
4.1.1 Tính năng của MMDetection.....	18
4.1.2 Cấu trúc.....	18
4.1.3 Cài đặt :	19

4.2 Cấu hình mô hình Generalized Attention trong mmdetecion	20
4.3 Huấn luyện mô hình Generalized Attention trong mmdetecion	30
4.4 Cấu hình thực nghiệm	32
4.5 Kết quả thực nghiệm	32
4.6 Đánh giá tổng quát	36
CHƯƠNG 5 KẾT LUẬN	37
5.1 Kết quả đạt được	37
5.2 Hướng phát triển	37
TÀI LIỆU THAM KHẢO	38

LIỆT KÊ HÌNH ẢNH

Hình 1 Công cụ gán nhãn labelImg	6
Hình 2 Thao tác trong công cụ LabelImg.....	6
Hình 3 Quá trình gán nhãn	7
Hình 4 Quá trình gán nhãn	8
Hình 5 Hoàn tất quá trình gán nhãn.....	8
Hình 6 Hình ảnh data đã thu thập	10
Hình 7 Mô hình Faster RCNN.....	11
Hình 8: Region proposal network(RPN)	12
Hình 9 Feed forward ảnh qua DNN để thu được convolutional features.....	13
Hình 10 : Region Proposal Network (RPN)	13
Hình 11 VGG-19, ResNet,	15
Hình 12 Kiến trúc mạng ResNet	16
Hình 13 Minh họa cấu hình mô-đun attention cho nghiên cứu thực nghiệm. Các mô-đun có màu xanh lam mới được thêm vào các khối hiện có.	17
Hình 14 Google colab.....	20
Hình 15 Cấu hình file requirements	21
Hình 16 Cấu hình file cocosplit.....	22
Hình 17 Quá trình huấn luyện	31
Hình 18 Kết quả sau khi huấn luyện	31
Hình 19 Quá trình chạy thực nghiệm	32
Hình 20 Các kết quả thu được	34
Hình 21 Thư mục lưu trữ kết quả	35

DANH MỤC CÁC TỪ VIẾT TẮT VÀ GIẢI THÍCH CÁC THUẬT NGỮ

DATN: Đồ án tốt nghiệp

AI: Trí tuệ nhân tạo

RPN :Region Proposal Network (một thành phần quan trọng trong hệ thống thị giác máy tính)

GA – RPN : Genetic Algorithm-based Region Proposal Network

IoU : Intersection over Union (là một phương pháp đo lường sự chồng lấn giữa hai khu vực hay đối tượng)

Fast R-CNN: Fast Region-based Convolutional Neural Network (là một phương pháp nhận diện đối tượng nhanh chóng trong lĩnh vực thị giác máy tính)

SSD -Single Shot MultiBox Detector

YOLO : You Only Look Once

CHƯƠNG 1 GIỚI THIỆU

1.1 Tổng quan

Tình trạng trẻ em bị lạc tại nơi công cộng đang ngày càng gia tăng trong những năm gần đây. Nguyên nhân chủ yếu là do sơ suất của cha mẹ trong việc quản lý con cái, đặc biệt là ở những nơi đông người. Bên cạnh đó, sự hiếu động, tò mò của trẻ em cũng khiến các bé dễ bị lạc khỏi tầm kiểm soát của người lớn.

Hậu quả của tình trạng này là gây ra lo lắng, hoảng loạn cho phụ huynh, đồng thời tiềm ẩn nhiều nguy hiểm cho trẻ như bị kẻ xấu bắt cóc, lạm dụng, ...

Để giải quyết vấn đề trên, công nghệ AI được kỳ vọng sẽ đóng vai trò quan trọng. Với khả năng xử lý dữ liệu lớn và phân tích hình ảnh, AI có thể được ứng dụng trên hệ thống camera giám sát để phát hiện và nhận dạng trẻ bị lạc, từ đó hỗ trợ tìm kiếm nhanh chóng.

Trong đề án này, mô hình Generalized Attention được đề xuất áp dụng. Đây là mô hình có khả năng "Tập trung" vào các vùng quan trọng của bức ảnh để nâng cao độ chính xác trong phát hiện và nhận dạng đối tượng. Nhờ đó, hy vọng mô hình sẽ giúp xác định chính xác vị trí của trẻ lạc nhanh chóng ngay cả trong môi trường đông người.

1.2 Bài toán

Bài toán cần giải quyết trong đề án này là ứng dụng mô hình AI có khả năng phát hiện và nhận dạng chính xác vị trí của trẻ em bị lạc trong không gian rộng với đám đông người.

Cụ thể, đầu vào của mô hình là các hình ảnh trực tiếp từ hệ thống camera giám sát được bố trí tại khu vực công cộng (ví dụ: khu vui chơi, ga tàu, sân bay...). Đầu ra mong muốn là xác định chính xác tọa độ vị trí hiện tại của đứa trẻ đang lạc trong không gian giám sát đó.

Yêu cầu của hệ thống là phải đảm bảo tốc độ xử lý nhanh chóng để hỗ trợ tìm kiếm kịp thời, đồng thời độ chính xác phát hiện và nhận dạng phải cao dưới nhiều điều kiện ánh sáng, góc nhìn khác nhau.

Để đáp ứng yêu cầu trên, mô hình phải có khả năng xử lý hiệu quả lượng dữ liệu hình ảnh lớn từ nhiều camera, trích xuất được các đặc trưng phân biệt của từng đối tượng, đặc biệt là khuôn mặt và các chi tiết nhận dạng khác. Bên cạnh đó, mô hình cũng cần có cơ chế so khớp chính xác với cơ sở dữ liệu ảnh để nhận diện đúng đứa trẻ đang được tìm kiếm.

Mục đích chính của đề án là:

- Ứng dụng AI vào thực tế, cụ thể là phát hiện trẻ em từ camera tìm được trẻ lạc cho cho các bậc phụ huynh (có thể dùng drone để quay).
- Cập nhật mô hình phát hiện đối tượng GeneralizedAttention

1.3 Bố cục đề án

Phần còn lại của báo cáo đề án tốt nghiệp này được tổ chức như sau

- Chương 2 - Thiết kế dữ liệu: Mô tả về quá trình thu thập và cách sử dụng công cụ để gán nhãn dữ liệu.
- Chương 3 - Mô hình đề xuất: Mô tả chi tiết về cách thức xây dựng mô hình Generalized Attention, cách thu thập và xử lý dữ liệu từ camera quan sát, và các phương pháp đánh giá hiệu suất của mô hình.
- Chương 4 - Cài đặt và kết quả: Trình bày cách cài đặt và kết quả thử nghiệm của mô hình, đánh giá hiệu quả và đề xuất hướng phát triển trong tương lai. Phần này cũng sẽ bao gồm các trường hợp thực tế và phản hồi từ người dùng.
- Chương 5 - Kết Luận: Tóm tắt những phát hiện chính và đề xuất các hướng can thiệp, cũng như hướng nghiên cứu tiếp theo. Phần này cũng sẽ đề cập đến tầm quan trọng của việc hợp tác giữa các bên liên quan để triển khai hệ thống một cách hiệu quả.

Đề án này sẽ tập trung vào việc áp dụng mô hình Generalized Attention để giải quyết bài toán nhận dạng trẻ em thất lạc trong các sự kiện đông người, qua đó góp phần đảm bảo an toàn cho trẻ em và hỗ trợ cộng đồng.