

Tìm kiếm có chọn lọc để phát hiện đối tượng | R-CNN

Cập nhật lần cuối: ngày 22 tháng 7 năm 2021

Vấn đề định vị đối tượng là phần khó khăn nhất trong việc phát hiện đối tượng. Một cách tiếp cận là chúng ta sử dụng cửa sổ trượt có kích thước khác nhau để định vị các đối tượng trong ảnh. Cách tiếp cận này được gọi là tìm kiếm toàn diện. Cách tiếp cận này rất tốn kém về mặt tính toán vì chúng ta cần tìm kiếm đối tượng trong hàng nghìn cửa sổ ngay cả đối với kích thước hình ảnh nhỏ. Một số tối ưu hóa đã được thực hiện như lấy kích thước cửa sổ theo các tỷ lệ khác nhau (thay vì tăng thêm một số pixel). Nhưng ngay cả sau này do số lượng cửa sổ nên nó không hiệu quả lắm. Bài viết này xem xét thuật toán tìm kiếm có chọn lọc sử dụng cả tìm kiếm toàn diện và phân đoạn (một phương pháp để tách các đối tượng có hình dạng khác nhau trong ảnh bằng cách gán cho chúng các màu khác nhau).

Thuật toán tìm kiếm có chọn lọc:

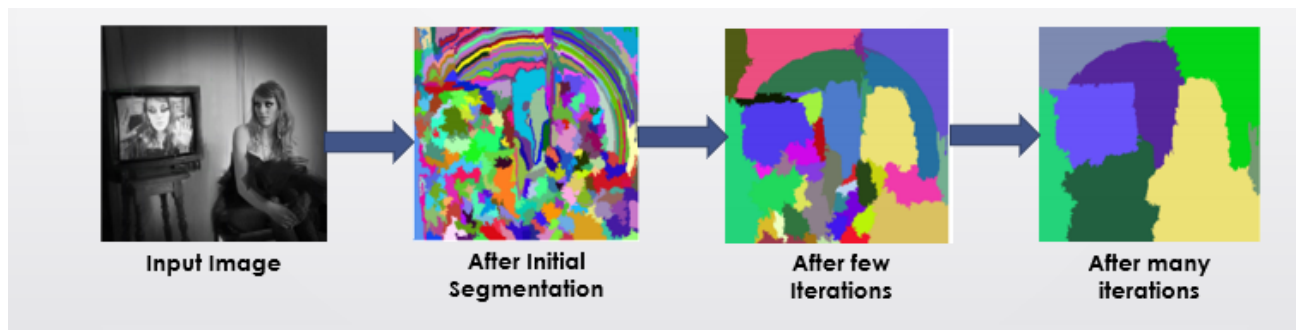
1. Tạo phân đoạn phụ ban đầu của hình ảnh đầu vào bằng phương pháp được mô tả bởi *Felzenszwalb* và cộng sự trong bài báo “Phân đoạn hình ảnh dựa trên biểu đồ hiệu quả”.



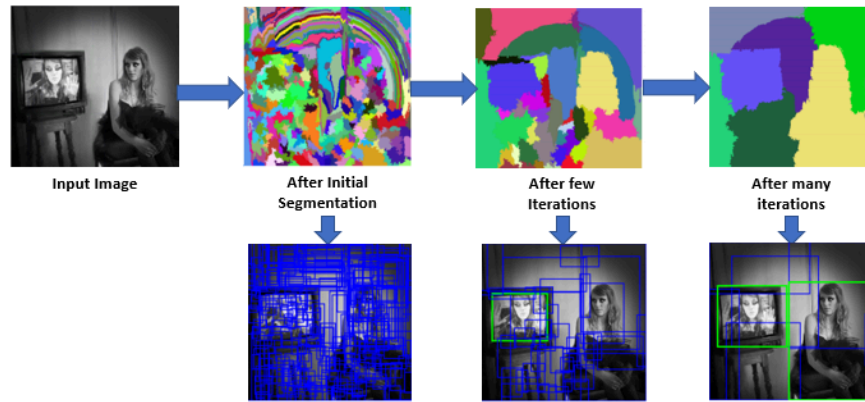
2. Kết hợp đệ quy các vùng tương tự nhỏ hơn thành các vùng lớn hơn. Chúng tôi sử dụng thuật toán Greedy để kết hợp các vùng tương tự nhau để tạo thành các vùng lớn hơn. Thuật toán được viết dưới đây.

Thuật toán tham lam:

1. Từ tập hợp các vùng, chọn hai vùng giống nhau nhất.
2. Kết hợp chúng thành một vùng lớn hơn.
3. Lặp lại các bước trên cho nhiều lần lặp.



3. Sử dụng các đề xuất vùng được phân đoạn để tạo ra các vị trí đối tượng đề xuất.



Sự tương đồng trong phân khúc:

Bài viết tìm kiếm có chọn lọc xem xét bốn loại điểm tương đồng khi kết hợp phân đoạn nhỏ ban đầu thành phân đoạn lớn hơn. Những điểm tương đồng này là:

- **Độ tương tự màu sắc:** Cụ thể đối với từng vùng, chúng tôi tạo ra biểu đồ của từng kênh màu có trong hình ảnh. Trong bài báo này, 25 thùng được lấy trong biểu đồ của mỗi kênh màu. Điều này mang lại cho chúng tôi 75 thùng (25 thùng cho mỗi R, G và B) và tất cả các kênh được kết hợp thành một vectơ ($n = 75$) cho mỗi vùng. Sau đó, chúng tôi tìm thấy sự tương đồng bằng cách sử dụng phương trình dưới đây:

$$S_{\text{color}}(r_i, r_j) = \sum_{k=1}^n \min(c_i^k, c_j^k)$$

$C_i^k, C_j^k = k^{\text{th}}$ value of histogram bin of region r_i and r_j respectively

- **Độ tương tự về kết cấu:** Độ tương tự về kết cấu được tính toán bằng cách sử dụng 8 đạo hàm Gaussian được tạo ra của hình ảnh và trích xuất biểu đồ với 10 thùng cho mỗi kênh màu. Điều này mang lại cho chúng ta vectơ $10 \times 8 \times 3 = 240$ chiều cho mỗi vùng. Chúng tôi rút ra sự tương tự bằng cách sử dụng phương trình này.

$$S_{\text{texture}}(r_i, r_j) = \sum_{k=1}^n \min(t_i^k, t_j^k)$$

$t_i^k, t_j^k = k^{\text{th}}$ value of texture histogram bin of region r_i and r_j respectively

- **Tương tự về kích thước:** Ý tưởng cơ bản của sự tương tự về kích thước là làm cho các vùng nhỏ hơn dễ dàng hợp nhất. Nếu sự tương đồng này không được xem xét thì khu vực lớn hơn sẽ tiếp tục hợp nhất với khu vực lớn hơn và các đề xuất khu vực ở nhiều quy mô sẽ chỉ được tạo tại vị trí này.

$$S_{\text{size}}(r_i, r_j) = 1 - (\text{size}(r_i) + \text{size}(r_j)) \div \text{size}(\text{img})$$

where $\text{size}(r_i)$, $\text{size}(r_j)$ and $\text{size}(\text{img})$ are the sizes of regions r_i , r_j and image respectively in pixels

- **Tính tương tự Diện vào:** Tính tương tự diện đo lường mức độ phù hợp của hai vùng với nhau. Nếu hai vùng vừa khít với nhau (Ví dụ: một vùng có mặt ở vùng khác) thì chúng sẽ được hợp nhất, nếu hai vùng thậm chí không chạm vào nhau thì chúng không nên được hợp nhất.

$$S_{\text{fill}}(r_i, r_j) = 1 - (\text{size}(\text{BB}_{ij}) - \text{size}(r_i) - \text{size}(r_j)) \div \text{size}(\text{img})$$

size(BB_{ij}) is the size of bounding box around i and j

Bây giờ, bốn điểm tương đồng ở trên kết hợp lại để tạo thành điểm tương đồng cuối cùng.

$$S_{(r_i, r_j)} = a_1 * S_{\text{color}}(r_i, r_j) + a_2 * S_{\text{texture}}(r_i, r_j) + a_3 * S_{\text{size}}(r_i, r_j) + a_4 * S_{\text{fill}}(r_i, r_j)$$

where a_i is either 0 or 1 depending upon we consider this similarity or not.

Kết quả :

Để đo lường hiệu quả của phương pháp này. Bài viết mô tả một tham số đánh giá được gọi là MABO (Chống chéo trung bình tốt nhất trung bình). Có hai phiên bản tìm kiếm chọn lọc là **Nhanh** và **Chất lượng**. Sự khác biệt giữa chúng là Chất lượng tạo ra nhiều hộp giới hạn hơn Nhanh và do đó mất nhiều thời gian hơn để tính toán nhưng có khả năng thu hồi cao hơn và ABO (Chống chéo trung bình tốt nhất) và MABO (Chống chéo trung bình tốt nhất trung bình). Chúng tôi tính toán ABO như sau.

$$ABO = \frac{1}{|G^c|} \sum_{l \in L} \max \text{Overlap}(g_i^c, l_j).$$

Như chúng ta có thể nhận thấy rằng khi tất cả những điểm tương đồng được sử dụng kết hợp, Nó mang lại cho chúng ta MABO tốt nhất. Tuy nhiên, cũng có thể kết luận rằng RGB không phải là bảng màu tốt nhất để sử dụng trong phương pháp này. HSV, Lab và rgI đều hoạt động tốt hơn RGB, điều này là do chúng không nhạy cảm với những thay đổi về bóng và độ sáng.

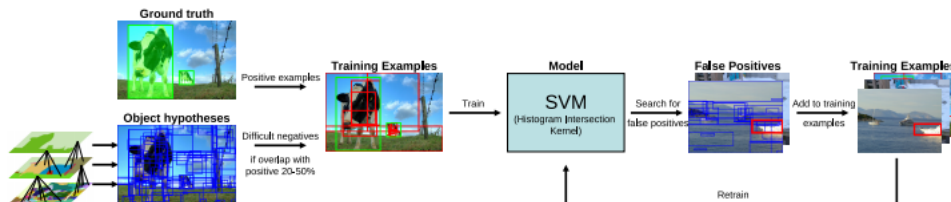
Nhưng khi chúng ta đa dạng hóa và kết hợp những điểm tương đồng khác nhau này, cách phối màu và giá trị ngưỡng (k),

Version	Diversification Strategies	MABO	# win	# strategies	time (s)
Single Strategy	HSV C+T+S+F $k = 100$	0.693	362	1	0.71
Selective Search Fast	HSV, Lab C+T+S+F, T+S+F $k = 50, 100$	0.799	2147	8	3.79
Selective Search Quality	HSV, Lab, rgI, H, I C+T+S+F, T+S+F, F, S $k = 50, 100, 150, 300$	0.878	10,108	80	17.15

Trong bài tìm kiếm chọn lọc, nó áp dụng phương pháp tham lam dựa trên MABO trên các chiến lược khác nhau để có được kết quả trên. Có thể nói rằng phương pháp kết hợp các chiến lược khác nhau này mặc dù mang lại MABO tốt hơn nhưng thời gian chạy cũng tăng lên đáng kể.

Tìm kiếm có chọn lọc trong nhận dạng đối tượng:

Trong bài viết tìm kiếm chọn lọc, các tác giả sử dụng thuật toán này để phát hiện đối tượng và huấn luyện một mô hình bằng cách đưa ra các ví dụ về sự thật cơ bản và giả thuyết mẫu trùng lặp 20-50% với sự thật cơ bản (như ví dụ tiêu cực) vào bộ phân loại SVM và huấn luyện nó để xác định dương tính giả. Kiến trúc của mô hình được sử dụng dưới đây.



Kiến trúc nhận dạng đối tượng (Nguồn: Bài báo tìm kiếm chọn lọc)

Kết quả được tạo ra trên bộ thử nghiệm VOC 2007 là,

method	recall	MABO	# windows
Arbelaez <i>et al.</i> [3]	0.752	0.649 ± 0.193	418
Alexe <i>et al.</i> [2]	0.944	0.694 ± 0.111	1,853
Harzallah <i>et al.</i> [16]	0.830	-	200 per class
Carreira and Sminchisescu [4]	0.879	0.770 ± 0.084	517
Endres and Hoiem [9]	0.912	0.791 ± 0.082	790
Felzenszwalb <i>et al.</i> [12]	0.933	0.829 ± 0.052	100,352 per class
Vedaldi <i>et al.</i> [34]	0.940	-	10,000 per class
Single Strategy	0.840	0.690 ± 0.171	289
Selective search "Fast"	0.980	0.804 ± 0.046	2,134
Selective search "Quality"	0.991	0.879 ± 0.039	10,097

Như chúng ta có thể thấy rằng nó tạo ra mức thu hồi rất cao và MABO tốt nhất trên Bộ thử nghiệm VOC 2007 và nó

Các ứng dụng :

Tìm kiếm chọn lọc được sử dụng rộng rãi trong kiến trúc tiên tiến ban đầu như R-CNN, Fast R-CNN, v.v. Tuy nhiên, do số lượng cửa sổ mà nó xử lý nên phải mất từ 1,8 đến 3,7 giây (Tìm kiếm chọn lọc nhanh) để tạo đề xuất vùng không đủ tốt cho hệ thống phát hiện đối tượng theo thời gian thực.

Thẩm quyền giải quyết :

- [Bài viết Tìm kiếm có chọn lọc \(Tìm kiếm có chọn lọc để phát hiện đối tượng\)](#).
- [Trang trình bày thị giác máy tính Stanford](#)

"Khóa học này có cấu trúc rất tốt và dễ học. Bất kỳ ai chưa có kinh nghiệm về khoa học dữ liệu, python hoặc ML đều có thể học từ khóa này. Khóa học này khiến mọi thứ trở nên dễ dàng đến mức bất kỳ ai cũng có thể tự học. Nó giúp tôi rất nhiều. Cảm ơn vì tạo ra một khóa học tuyệt vời như vậy."- **Ayushi Jain | Được đặt tại Microsoft**

Now, đây là cơ hội để bạn mở ra các cơ hội việc làm có thu nhập cao với tư cách là [Nhà khoa học dữ liệu](#) ! Tham gia [Chương trình khoa học dữ liệu và học máy hoàn chỉnh](#) của chúng tôi và có được trải nghiệm học tập 360 độ do các chuyên gia trong ngành cố vấn.

Thực hành với hơn 40 Dự án Công nghiệp, các buổi giải quyết nghi ngờ thường xuyên và hơn thế nữa. [Hãy đăng ký tham gia Chương trình ngay hôm nay!](#)

P paw... + Theo

4

Bài viết tiếp theo

Phát hiện đối tượng bằng yolov8

Bài đọc tương tự

R-CNN so với R-CNN nhanh so với R-CNN nhanh hơn | ML

R-CNN: R-CNN được đề xuất bởi Ross Girshick et al. vào năm 2014 để giải quyết vấn đề định vị đối tượng hiệu quả trong phát hiện đối tượng. Các phương pháp trước đây sử dụng cái gọi là Tìm kiếm toàn diện, sử dụng các cửa sổ trượt có tỷ lệ...

đọc 6 phút

Phát hiện bệnh viêm phổi bằng CNN trong Python

Trong bài viết này, chúng ta sẽ tìm hiểu cách xây dựng bộ phân loại bằng Mạng thần kinh chuyển đổi đơn giản, mạng này có thể phân loại hình ảnh chụp X quang của bệnh nhân để phát hiện xem bệnh nhân đó là Bình thường hay bị ảnh hưởng...

đọc 10 phút

Phát hiện đối tượng bằng Bộ biến áp phát hiện (DETR) của Facebook

Facebook vừa phát hành Mô hình phát hiện đối tượng hiện đại vào ngày 27 tháng 5 năm 2020. Họ gọi nó là DERT là viết tắt của Máy biến áp phát hiện vì nó sử dụng máy biến áp để phát hiện vật thể. Đây là lần đầu tiên máy biến áp đó được s...

đọc 7 phút

Phát hiện đối tượng so với nhận dạng đối tượng và phân đoạn hình ảnh

So sánh các thuật toán phát hiện bất thường để phát hiện ngoại lệ trên bộ dữ liệu đồ chơi trong Scikit Learn

Sự bất thường là sự sai lệch so với hành vi hoặc khuôn mẫu được mong đợi hoặc bình thường. Phát hiện bất thường là quá trình xác định các mẫu hoặc hành vi bất thường này trong dữ liệu. Điều này quan trọng vì các điểm bất thường có thể chỉ...

đọc 9 phút

Xem thêm bài viết

Thẻ bài viết : [trần](#) [Học máy](#)

Thẻ thực hành: [Học máy](#) [trần](#)

Văn bản gốc

Anomaly is a deviat pattern. Anomaly de patterns or behavior indicate important e equipment failure, o variety

Đánh giá bản dịch này
Ý kiến phản hồi của bạn sẽ
Dịch



A-143, Tầng 9, Tháp Công ty Sovereign,
Khu vực-136, Noida, Uttar Pradesh -
201305



Công ty

Về chúng tôi
Hợp pháp
Nghề nghiệp
Trong phương tiện truyền thông
Liên hệ chúng tôi
Quảng cáo với chúng tôi
Giải pháp Doanh nghiệp GFG
Chương trình đào tạo vị trí

Ngôn ngữ

Python
Java
C++
PHP
Gò Lang
SQL
Ngôn ngữ R
Hướng dẫn về Android

Khám phá

Thử thách tuyển dụng Job-A-Thon
Hack-A-Thon
Cuộc thi hàng tuần của GfG
Lớp học ngoại tuyến (Delhi/NCR)
DSA trong JAVA/C++
Thiết kế hệ thống tổng thể
CP chủ
Video của GeeksforGeeks
Cộng đồng đam mê

DSA

Cấu trúc dữ liệu
Thuật toán
DSA cho người mới bắt đầu
Các vấn đề DSA cơ bản
Lộ trình DSA
Câu hỏi phỏng vấn DSA
Lập trình cạnh tranh

Chúng tôi sử dụng cookie để đảm bảo bạn có trải nghiệm duyệt web tốt nhất trên trang web của chúng tôi. Bằng cách sử dụng trang web của chúng tôi, bạn xác nhận rằng bạn đã đọc và hiểu [Chính sách cookie](#) và [Chính sách quyền riêng tư của chúng tôi](#)

Hướng dẫn học máy

Toán học ML

Hướng dẫn trực quan hóa dữ liệu

Hướng dẫn về cấu trúc

Hướng dẫn về NumPy

Hướng dẫn NLP

Hướng dẫn học sâu

Hướng dẫn Python

Ví dụ lập trình Python

Hướng dẫn Django

Dự án Python

Python Tkinter

Rút trích nội dung trang web

Hướng dẫn OpenCV

Câu hỏi phỏng vấn Python

DevOps

Git

AWS

Docker

Kubernetes

Azure

GCP

Lộ trình DevOps

Những môn học ở trường

toán học

Vật lý

Hoá học

Sinh vật học

Khoa học xã hội

Ngữ pháp tiếng Anh

Cơ sở dữ liệu

SQL

MYSQL

PostgreSQL

PL/SQL

MongoDB

Kỳ thi cạnh tranh

JEE nâng cao

Mạng UGC

UPSC

SSC CGL

SBI PO

Thư ký SBI

IBPS PO

Thư ký IBPS

Công cụ trực tuyến miễn phí

Kiểm tra đánh máy

Biên tập hình ảnh

Trình định dạng mã

Bộ chuyển đổi mã

Chuyển đổi tiền tệ

Trình tạo số ngẫu nhiên

Trình tạo mật khẩu ngẫu nhiên

JavaScript

Bản đánh máy

ReactJS

Tiếp theoJS

NodeJ

Khởi động

CSS theo gió

Khoa học máy tính

GATE CS Ghi chú

Các hệ điều hành

Mạng máy tính

Hệ thống quản lý cơ sở dữ liệu

Kỹ thuật phần mềm

Thiết kế logic kỹ thuật số

Toán kỹ thuật

Thiết kế hệ thống

Thiết kế cấp cao

Thiết kế cấp thấp

Sơ đồ UML

Hướng dẫn phỏng vấn

Mẫu thiết kế

OOAD

Chương trình đào tạo về thiết kế hệ thống

Câu hỏi phỏng vấn

thương mại

Kế toán

Nghiên cứu kinh doanh

Kinh tế học

Sự quản lý

Quản lý nhân sự

Tài chính

Thuế thu nhập

Góc chuẩn bị

Quy trình tuyển dụng thông minh của công ty

Mẫu sơ yếu lý lịch

Chuẩn bị năng khiếu

Câu đố

Sự chuẩn bị khôn ngoan của công ty

Các công ty

Trường Cao đẳng

Hướng dẫn khác

Phát triển phần mềm

Kiểm thử phần mềm

Quản lý sản phẩm

Quản lý dự án

Linux

Excel

Tất cả các bảng cheat

Viết và kiểm tiền

Viết một tiêu đề

Cải thiện một bài viết

Chọn chủ đề để viết

Chia sẻ kinh nghiệm của bạn

Thực tập