

Finding/Counting entities - Image Processing

Bài toán

Phương pháp Finding

Template matching

Hàm Square Difference Normalized kết hợp Masking

Rescale ảnh

Phương pháp Counting

Tìm điểm chung

Biến hóa ảnh/template theo mong muốn

Kết quả

Finding

Ảnh 1: Tìm được 13 trên 15 đồ vật

Ảnh 2: Tìm được 12 trên 12 đồ vật

Counting

Ảnh mèo: Tìm được ~56 trên ? mèo

Ảnh ủng màu đỏ: Tìm được 8/9 đôi ủng

Ảnh chuột: Tìm được 9/10 con chuột

Ảnh thỏ 1: Tìm được 14 trên 14 con thỏ

Ảnh thỏ 2: Tìm được 29 trên 29 con thỏ

Phân tích kết quả

Bài toán

Đầu vào: Ảnh chứa hàng loạt vật thể, bao gồm nhiều nhiễu.

Yêu cầu bài toán: Tìm ra các vật thể được yêu cầu đang ẩn bên trong ảnh bằng các phương pháp xử lý ảnh đã được học.

Đầu ra: Tìm ra/Khoanh vùng/Đếm số vật thể được yêu cầu.

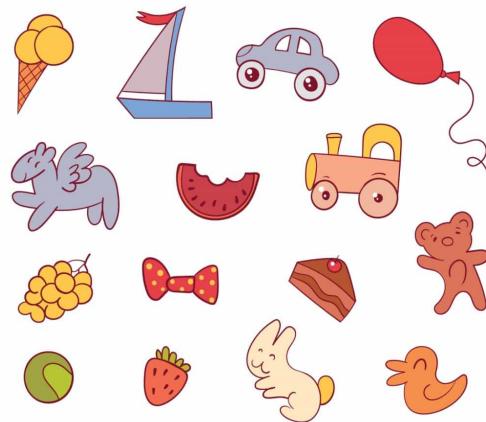
Phương pháp Finding

Template matching

Một đặc điểm nhận diện được ở ảnh và các template của vật cần tìm là chúng khác kích cỡ, nhưng có cùng hướng. Trong trường hợp này, ta có thể sử dụng phương pháp Template matching và tính toán trên pixel mà không sợ xảy ra sai sót khi ảnh bị xoay hay biến dạng.



FIND
15
HIDDEN
OBJECTS
IN THE
PICTURE



Hàm Square Difference Normalized kết hợp Masking

Khi so sánh độ tương đồng giữa 2 template, sử dụng công thức SQDIFF_NORMED kết hợp Masking cho template là 1 hướng đi hợp lý, do với cách này, ta có thể hiểu một cách rất trực giác rằng việc so khớp đang diễn ra bằng cách tính độ chênh lệch về độ đậm nhạt của template. Đồng thời với masking, ta đảm bảo các pixel không liên quan đến template không bị đưa vào công thức so sánh gây nhiễu

cv2.TM_SQDIFF_NORMED

$$R(x, y) = \frac{\sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2}{\sqrt{\sum_{x', y'} T(x', y')^2 \cdot \sum_{x', y'} I(x + x', y + y')^2}}$$

Rescale ảnh

Phương pháp template matching đặt nặng vấn đề rằng ảnh và template cần có cùng kích cỡ, do chỉ chênh 1 lượng nhỏ có thể dẫn đến sai lệch lớn trong tính toán trên từng pixel.

Phương pháp Counting

Tìm điểm chung

Để tìm ra số lượng vật thể bằng template matching, ta tìm ra điểm chung mà tất cả vật thể trong ảnh đều có, từ đó trích xuất để có thể sử dụng để truy ra mọi vật thể. Ví dụ có thể kể đến:

- Miệng của những con mèo đều màu vàng và dạng kim cương.
- Cả 2 ảnh các con thỏ đều giống nhau ở mắt, đều là chấm đen.
- Các con chuột đều có mảng lông màu xám
- Các đôi ủng đỏ đều lộ ra ít nhất 1 chiếc ủng

Biến hóa ảnh/template theo mong muốn

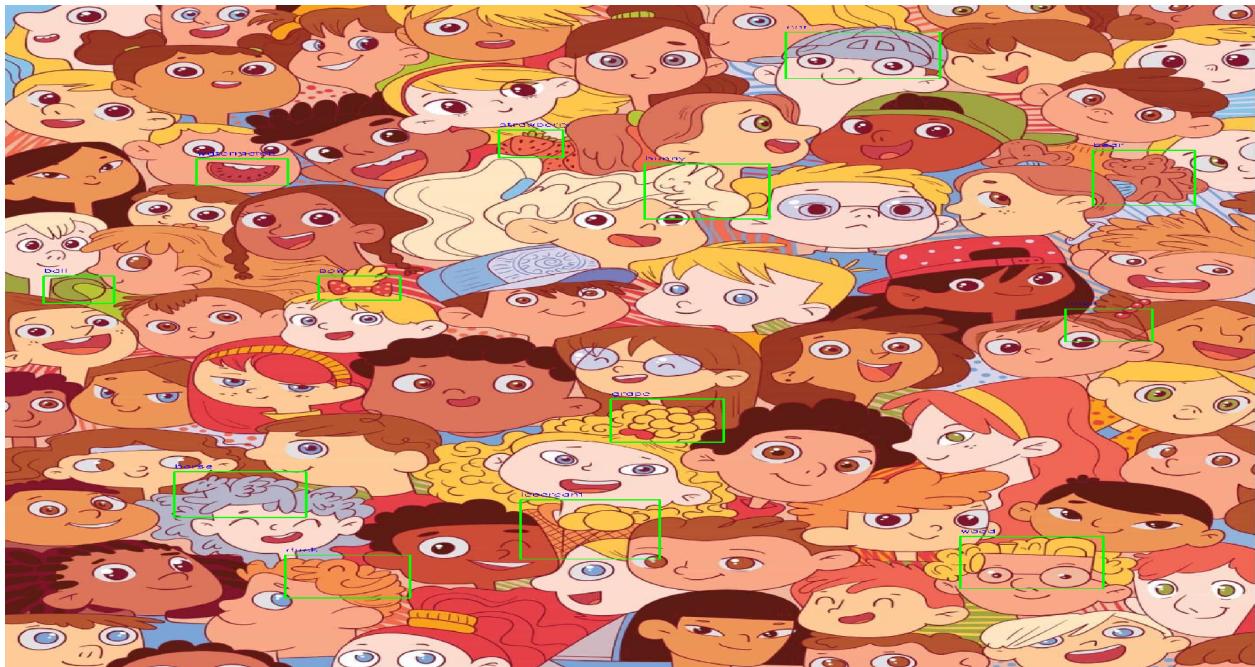
Để đạt kết quả tốt nhất, ta không chỉ trích xuất đặc trưng, mà còn có thể biến hóa ảnh tùy ý để phục vụ cho việc so khớp template. Ví dụ như:

- Đổi màu lông thỏ để dễ so sánh với template là mắt thỏ.
- Chỉnh sửa màu template miệng của mèo để dễ generalize giữa những con mèo với nhau.

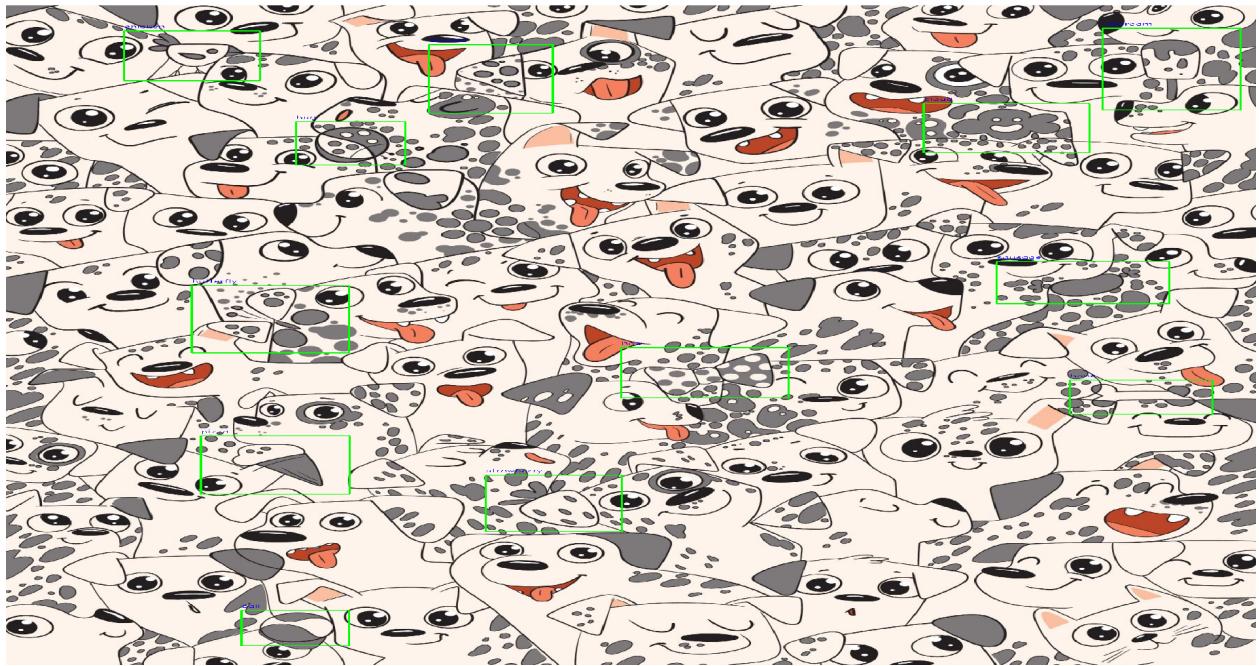
Kết quả

Finding

Ảnh 1: Tìm được 13 trên 15 đồ vật

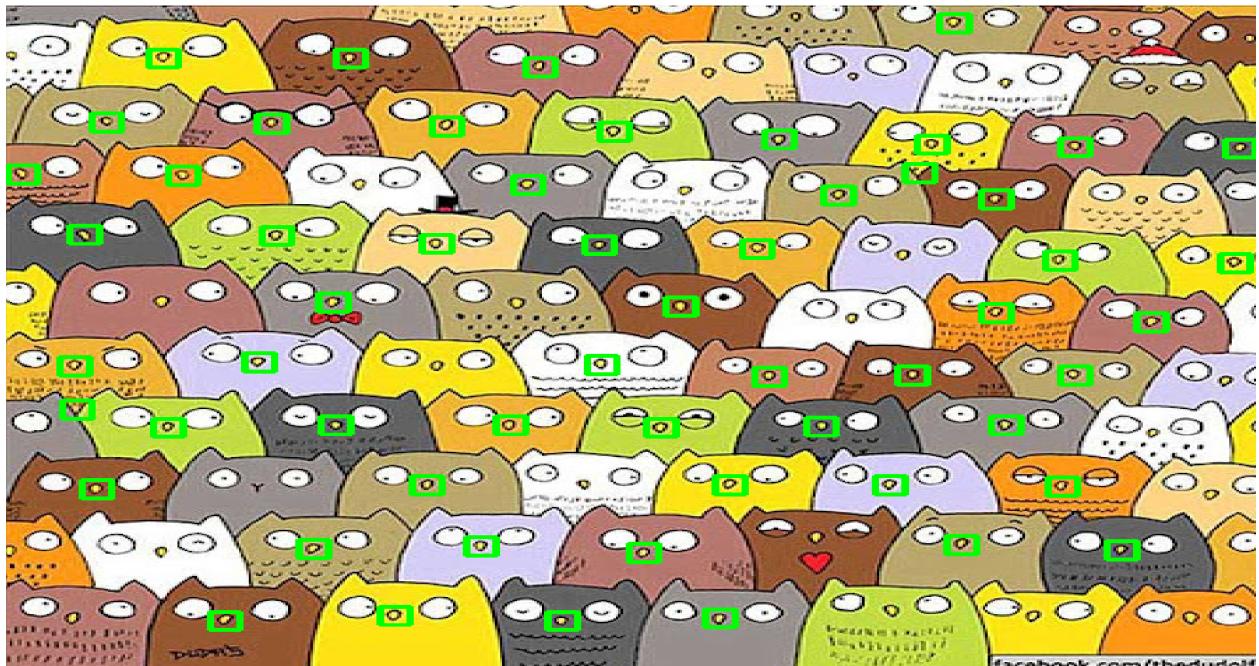


Ảnh 2: Tìm được 12 trên 12 đồ vật



Counting

Ảnh mèo: Tìm được ~56 trên ? mèo



Ảnh ủng màu đỏ: Tìm được 8/9 đôi ủng



Ảnh chuột: Tìm được 9/10 con chuột

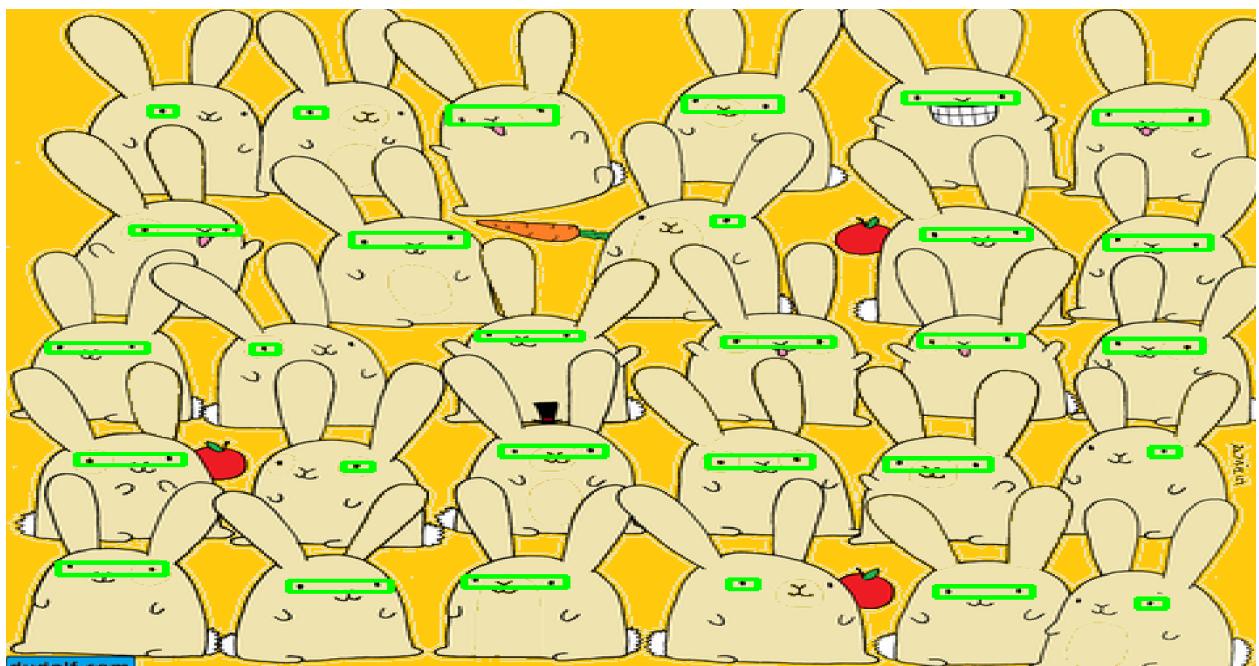
Can you find the dice?



Ảnh thò 1: Tìm được 14 trên 14 con thỏ



Ảnh thò 2: Tìm được 29 trên 29 con thò



Phân tích kết quả

- Ở phần Finding, việc scale lại template để cùng cỡ với vật thể ẩn trong ảnh là 1 trở ngại. Tuy nhiên, dễ thấy bằng mắt thường rằng phần lớn vật thể trong ảnh có kích cỡ nằm trong khoảng 60~70% template. Bằng thông tin đó, ta đơn giản hóa quá trình bằng việc scale ảnh thủ công, tránh vét cạn các trường hợp scale ảnh gây mất thời gian.
- Ở phần Counting, việc đặt ra ngưỡng (threshold) để coi 1 phần ảnh là trùng với template còn thiếu sự ổn định, yêu cầu sự “đoán mò” threshold tối ưu.
- Trong 1 số trường hợp, việc sử dụng template matching với màu RGB thay vì Grayscale đem lại lợi ích lớn, điển hình như trong bài toán Counting số chuột, nhờ đặc điểm đơn giản là lông chuột có màu xám, nổi bật với màu vàng của phô mai.