

Báo cáo Template matching

1: Bài toán finding & counting:

- Bài toán finding là bài toán tìm trong ảnh gốc những đồ vật có sẵn bên ngoài, còn bài toán counting là đếm số lượng một vật thể nào đó xuất hiện trong ảnh
- Điểm chung của hai bài toán này là ta đều có thể đưa về dạng template matching. Với bài toán finding thì ảnh template đã được cho sẵn, còn bài toán counting thì khó hơn hẳn, ta phải tự tìm một template nào đó đủ tốt.

2: Thuật toán Template matching:

Với sự hỗ trợ của thư viện OpenCV, ta có thể sử dụng hàm `cv2.matchTemplate()` để thực hiện phép template matching. Bản chất của phép toán này là sử dụng template làm cửa sổ trượt để tính toán mức độ sai khác giữa ảnh template và ảnh trong cửa sổ. Trong bài toán này, tôi sử dụng tham số `cv2.TM_CCOEFF_NORMED`, tương đương với hàm:

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

Trong đó, $R(x, y)$ là độ tương quan của template và ảnh ở tọa độ (x, y) . Độ tương quan này được tính bởi việc lấy tổng của tích giữa các giá trị trong template T' và vùng tương ứng của ảnh I' , sau đó chia chuẩn hóa cho Norm-2 của chúng để đưa kết quả về giá trị từ -1 đến 1.

3: Thuật toán Non-Maximum Suppression (NMS):

Non-Maximum Suppression (NMS) là một phương pháp phổ biến được sử dụng sau khi thực hiện template matching để loại bỏ các khung bao (bounding boxes) không cần thiết hoặc bị lặp lại. Sau quá trình template matching, có thể xuất hiện nhiều vùng phù hợp, bao gồm các khung bao có kích thước khác nhau và chồng chéo lên nhau. NMS có nhiệm vụ giữ lại các khung bao có giá trị cao nhất và loại bỏ các khung trùng lặp không đáng kể.

Các bước thực hiện NMS:

- **Bước 1:** Xếp hạng các khung bao theo điểm số tương đồng (score) từ cao xuống thấp.
- **Bước 2:** Với mỗi khung có điểm cao nhất, loại bỏ những khung bao khác có mức độ chồng chéo (Intersection over Union - IoU) vượt quá ngưỡng đã xác định.
- **Bước 3:** Lặp lại quá trình này cho đến khi không còn khung nào cần loại bỏ.

4: Một số kết quả:

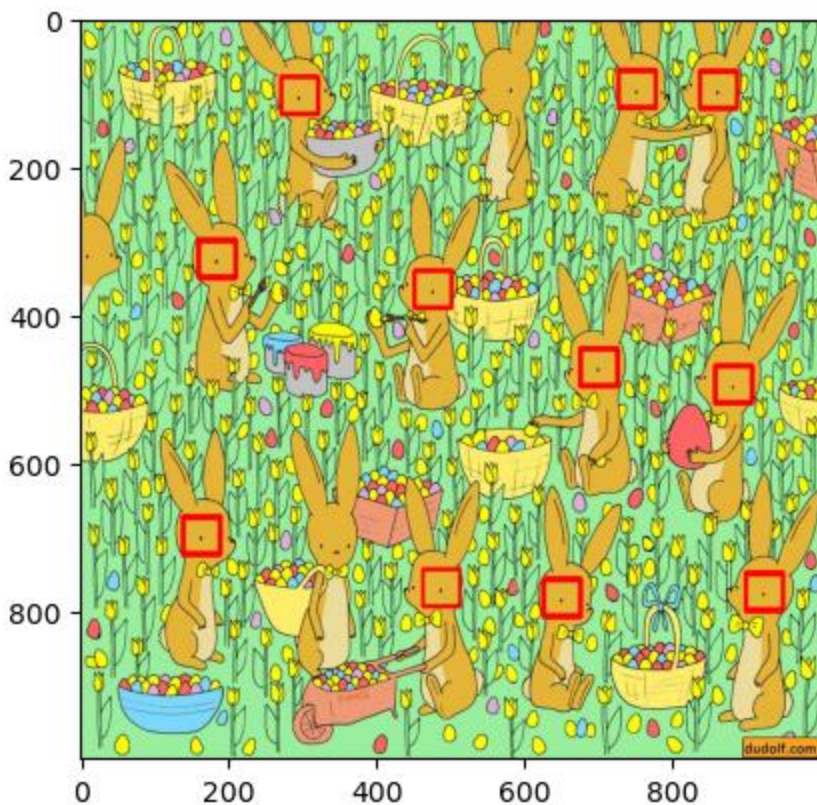
Do phép Template matching này là phương pháp khớp theo từng pixel, nên các đặc trưng của template sẽ rất khó để match chính xác nếu template scale khác (hay có biến đổi affine) so với vật thể trong ảnh. Do đó, để tìm kiếm 1 template hoàn hảo là một việc khó khăn.

Ngoài ra, việc đưa ảnh về gray scale, hoặc đưa thẳng về đen trắng (binary) đôi lúc sẽ giúp đơn giản hóa bài toán nếu như đặc trưng về màu là không cần thiết để xét đến.

Trong bài toán Finding, các vật thể template đã được cho trước, vì thế nên ta sẽ thực hiện một bước scale lại template theo nhiều tỉ lệ khác nhau để tìm kiếm được nó trong ảnh. Sau đây là kết quả sau khi đã match và tìm các vật thể có trong ảnh.



Trong bài toán Counting thì việc khó khăn lại nằm ở việc tìm ra được một template hợp lý để có thể ghép được với tất cả các vật thể và đếm. Việc này khiến ta phải tìm các đặc trưng thật sự đơn giản mà cái nào cũng có, nhưng đôi khi đơn giản quá thì lại khiến việc tìm kiếm gặp nhiều False Positive bounding box. Sau đây là ví dụ một ảnh của bài toán đếm thỏ.



Có thể thấy việc chọn mắt làm template phần nào hiệu quả khi quét được khá nhiều thỏ nhìn ngang, tuy nhiên lại không quét được những con thỏ nhìn thẳng. Điều này cho thấy hạn chế rất lớn của phương pháp Template matching.