

Computer Vision

Histogram of Oriented Gradient

trong bài toán phát hiện người đi bộ

Lê Hoàng Ngân
Nguyễn Đức Thắng

Hanoi University of Science and Technology

10/2020

Nội dung

- 1 Kiến thức cơ sở
- 2 Histogram of Oriented Gradient cho bài toán phát hiện người đi bộ

Kiến thức cơ sở

- 1 Support Vector Machine - SVM
- 2 Histogram of Oriented Gradient - HOG

Support Vector Machine - SVM

Khoảng cách từ một điểm đến một siêu phẳng

Không gian 2 chiều: khoảng cách từ một điểm có tọa độ (x_0, y_0) tới đường thẳng $w_1x + w_2y + b = 0$:

$$\frac{|w_1x_0 + w_2y_0 + b|}{\sqrt{w_1^2 + w_2^2}}$$

Không gian 3 chiều, khoảng cách từ một điểm có tọa độ (x_0, y_0, z_0) tới một mặt phẳng $w_1x + w_2y + w_3z + b = 0$:

$$\frac{|w_1x_0 + w_2y_0 + w_3z_0 + b|}{\sqrt{w_1^2 + w_2^2 + w_3^2}}$$

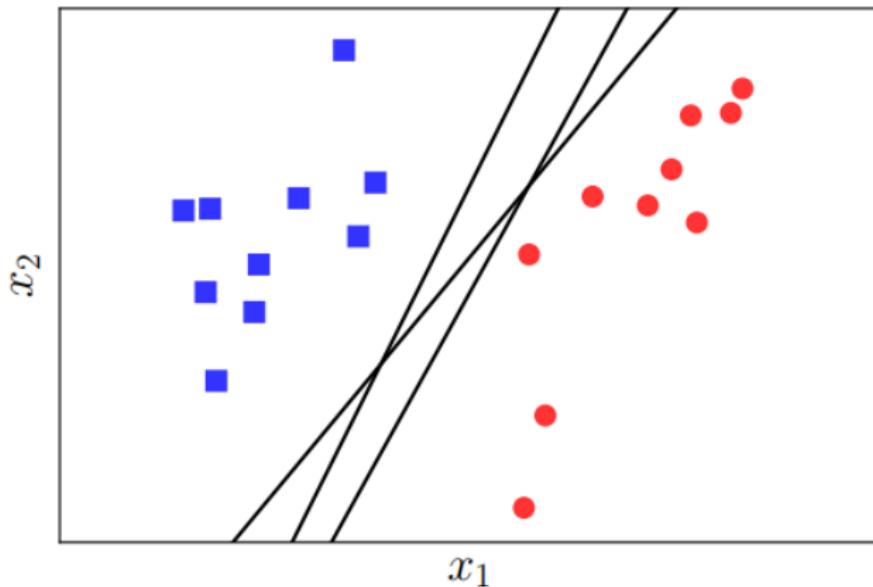
Support Vector Machine - SVM

Không gian d chiều: khoảng cách từ một điểm (vector) có tọa độ $(x_{10}, x_{20}, \dots, x_{d0})$ tới siêu phẳng $w_1x_1 + w_2x_2 + \dots + w_dx_d + b = 0$:

$$\frac{|w_1x_{10} + w_2x_{20} + \dots + w_dx_{d0} + b|}{\sqrt{w_1^2 + w_2^2 + w_3^2 + \dots + w_d^2}} = \frac{|w^T x_0 + b|}{\|w\|_2}$$

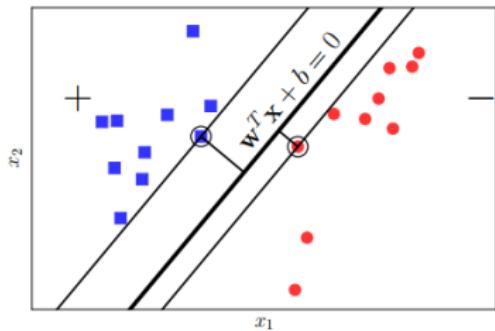
Với $x_0 = [x_{10}, x_{20}, \dots, x_{d0}]^T$, $w = [w_1, w_2, \dots + w_d]^T$

Support Vector Machine - SVM

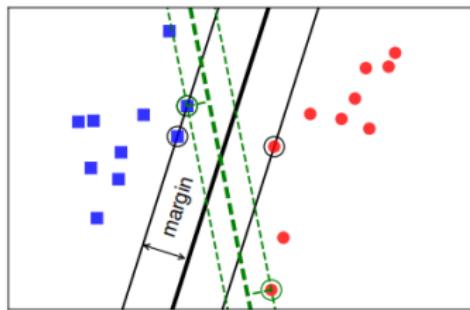


Hình: Hai lớp dữ liệu đỏ và xanh là linearly separable. Có vô số các đường thẳng có thể phân tách chính xác hai lớp dữ liệu này

Support Vector Machine - SVM



(a)



(b)

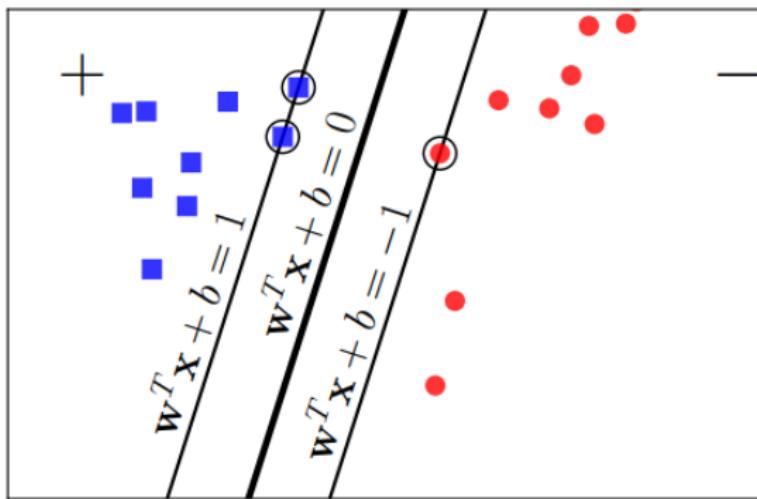
Hình: Ý tưởng của SVM: Margin của một lớp được định nghĩa là khoảng cách từ các điểm gần nhất của lớp đó tới mặt phân chia. Margin của hai lớp phải bằng nhau và lớn nhất có thể.

Support Vector Machine - SVM

- ▶ Các cặp dữ liệu trong tập huấn luyện là $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$
- ▶ Vector $x_i \in \mathbb{R}^d$ thể hiện đầu vào của một điểm dữ liệu
- ▶ y_i là nhãn của điểm dữ liệu, d là số chiều của dữ liệu, N là số điểm của dữ liệu.

Giả sử nhãn của mỗi điểm dữ liệu được xác định bởi $y_i = 1$ hoặc $y_i = -1$.

Support Vector Machine - SVM



Hình: Giả sử mặt phân chia có phương trình $w^T x + b = 0$. Không mất tính tổng quát, bằng cách nhân các hệ số w, b với các hằng số phù hợp, ta có thể giả sử rằng điểm gần nhất của lớp màu xanh tới mặt này thỏa mãn $w^T x + b = 1$. Khi đó, điểm gần nhất của lớp đỏ thỏa mãn $w^T x + b = -1$

Support Vector Machine - SVM

Với cặp dữ liệu (x_n, y_n) bất kỳ, khoảng cách từ điểm đó tới mặt phân chia:

$$\frac{y_n(w^T x_n + b)}{\|w\|_2}$$

Với mặt phân chia này, margin được tính là khoảng cách gần nhất từ một điểm (trong cả hai lớp, vì cuối cùng margin của cả hai lớp sẽ như nhau) tới mặt đó:

$$margin = \min_n \frac{y_n(w^T)}{\|w\|_2}$$

Support Vector Machine - SVM

Bài toán tối ưu của SVM:

$$(w, b) =_{w,b} \left\{ \min_n \frac{y_n(w^T)}{\|w\|_2} \right\} =_{w,b} \left\{ \frac{1}{\|w\|_2} \min_n y_n(w^T x_n + b) \right\} \quad (1)$$

Bài toán tối ưu có ràng buộc có dạng:

$$(w, b) =_{w,b} \frac{1}{\|w\|_2} \quad (2)$$

$$\forall k : 1 - y_n(w^T x_n + b) \leq 0 \quad \forall n = 1, 2, \dots, N$$

Sử dụng hàm `sklearn.svm.SVC` của thư viện `scikit-learn` để giải quyết bài toán.

Histogram of Oriented Gradient - HOG

- ▶ một feature descriptor, dùng để phát hiện một đối tượng.
- ▶ Các khái niệm về HOG được nêu ra từ năm 1986, năm 2005, HOG được sử dụng rộng rãi sau khi Navneet Dalal và Bill Triggs công bố những bổ sung.
- ▶ Mục đích: trừu tượng hóa đối tượng bằng cách trích xuất ra những đặc trưng của đối tượng và bỏ đi những thông tin không hữu ích.

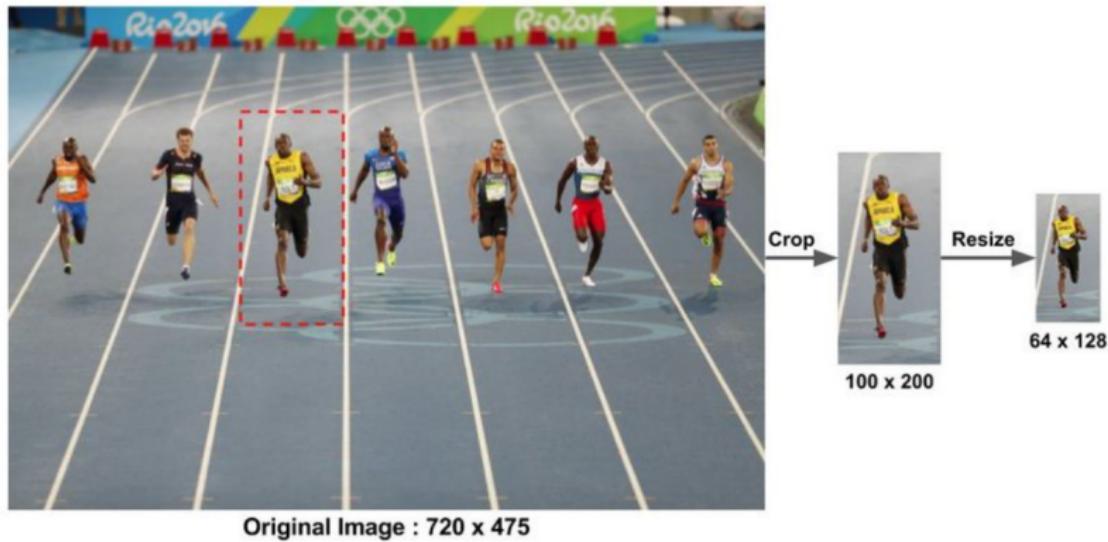
Histogram of Oriented Gradient - HOG

5 bước cơ bản để xây dựng một vector HOG cho hình ảnh:

- ▶ Bước 1: Tiền xử lý
- ▶ Bước 2: Tính gradient
- ▶ Bước 3: Tính vector đặc trưng cho từng ô (cell)
- ▶ Bước 4: Chuẩn hóa block
- ▶ Bước 5: Tính toán vector HOG

Histogram of Oriented Gradient - HOG

Tiền xử lý



Histogram of Oriented Gradient - HOG

Tính Gradient

$$D_x = [-1, 0, 1]$$

$$D_y = [1, 0, -1]^T$$

-1	0	1
----	---	---

-1
0
1

Gọi ảnh đầu vào là I , 2 ảnh đạo hàm riêng theo Ox, Oy :

$$I_x = I * D_x$$

$$I_y = I * D_y$$

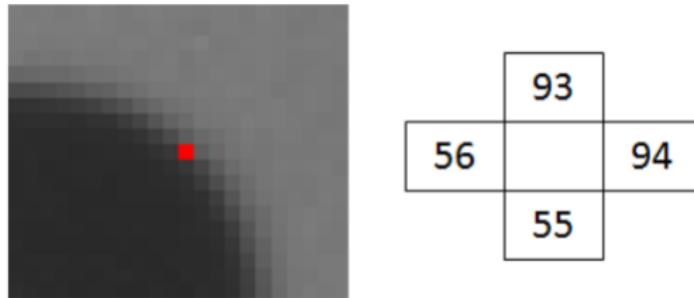
Histogram of Oriented Gradient - HOG

Gradient bao gồm hai thành phần là: cường độ (*Gradient Magnitude*), và hướng (*Gradient Direction*):

$$\text{Cường độ : } |G| = \sqrt{I_x^2 + I_y^2}$$

$$\text{Hướng : } \theta = \arctan \frac{I_y}{I_x}$$

Histogram of Oriented Gradient - HOG



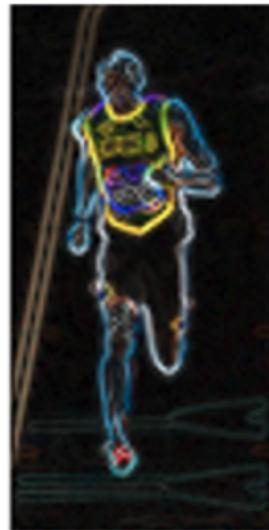
$$I_x = I * D_x = [56, x, 94] * [-1, 0, 1]^T = [38]$$

$$I_y = I * D_y = [93, y, 55]^T * [1, 0, -1] = [38]$$

Cường độ: $|G| = \sqrt{I_x^2 + I_y^2} = \sqrt{38^2 + 38^2} \approx 53.74$

Hướng: $\theta = \arctan \frac{I_y}{I_x} \approx 0.79$

Histogram of Oriented Gradient - HOG



Left : Absolute value of x-gradient. Center : Absolute value of y-gradient.
Right : Magnitude of gradient.

Histogram of Oriented Gradient - HOG

Tính vector đặc trưng cho từng cell

Chia hình ảnh thành các block, mỗi block lại chia đều thành các cell. Xác định được số block:

$$n_{block,image} = \left(\frac{W_{image} - W_{block} * W_{cell}}{W_{cell}} + 1 \right) * \left(\frac{H_{image} - H_{block} * H_{cell}}{H_{cell}} + 1 \right)$$

Tính toán vector đặc trưng cho từng cell:

- ▶ Chia không gian hướng thành p bin (số chiều vector đặc trưng của ô).
- ▶ Rời rạc hóa góc hướng nghiêng tại mỗi điểm ảnh vào trong các bin.

Histogram of Oriented Gradient - HOG

Giả sử hướng góc nghiêng tại pixel ở vị trí (x, y) có độ lớn là $\alpha(x, y)$.

- ▶ Trường hợp rời rạc hóa unsigned - HOG với $p = 9$:

$$B(x, y) = \text{round}\left(\frac{p * \alpha(x, y)}{\pi}\right) \mod p$$

- ▶ Trường hợp rời rạc hóa signed-HOG với $p = 18$:

$$B(x, y) = \text{round}\left(\frac{p * \alpha(x, y)}{2\pi}\right) \mod p$$

Histogram of Oriented Gradient - HOG

Số chiều vector đặc trưng block được tính theo công thức:

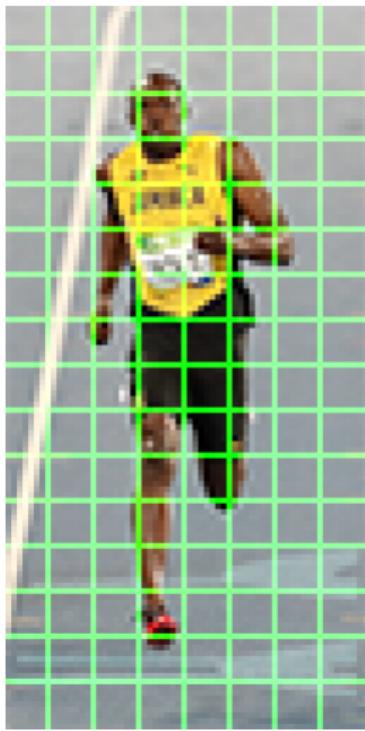
$$size_{block} = n * size_{cell}$$

Trong đó:

n là số cell trong block

$size_{cell}$ là số chiều của vector đặc trưng cell ($size_{cell} = 9$ nếu sử dụng unsigned-HOG và $size_{cell} = 18$ nếu sử dụng signed-HOG)

Histogram of Oriented Gradient - HOG

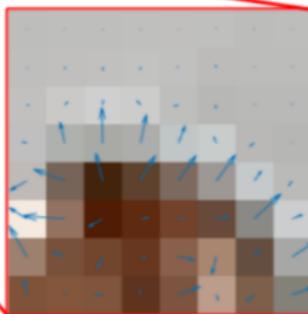
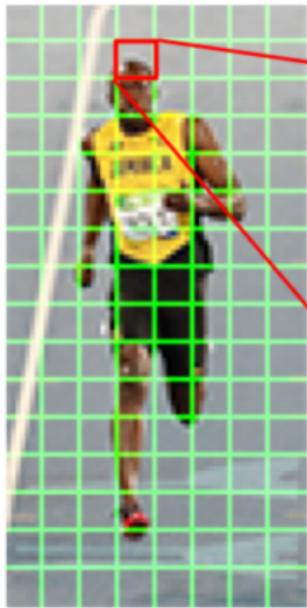


Histogram of Oriented Gradient - HOG

Không gian hướng 9 bin, trường hợp unsigned-HOG. Hướng gradient sẽ chạy trong khoảng từ 0° đến 180° , trung bình 20° mỗi bin:

- ▶ Hướng $0^\circ - 20^\circ$: vote hướng thuộc đoạn này vào bin 0.
- ▶ Hướng $20^\circ - 40^\circ$: vote hướng thuộc đoạn này vào bin 1.
- ▶ Hướng $40^\circ - 60^\circ$: vote hướng thuộc đoạn này vào bin 2.
- ▶ Hướng $60^\circ - 80^\circ$: vote hướng thuộc đoạn này vào bin 3.
- ▶ Hướng $80^\circ - 100^\circ$: vote hướng thuộc đoạn này vào bin 4.
- ▶ Hướng $100^\circ - 120^\circ$: vote hướng thuộc đoạn này vào bin 5.
- ▶ Hướng $120^\circ - 140^\circ$: vote hướng thuộc đoạn này vào bin 6.
- ▶ Hướng $140^\circ - 160^\circ$: vote hướng thuộc đoạn này vào bin 7.
- ▶ Hướng $160^\circ - 180^\circ$: vote hướng thuộc đoạn này vào bin 8.

Histogram of Oriented Gradient - HOG



2	3	4	4	3	4	2	2
5	11	17	13	7	9	3	4
11	21	23	27	22	17	4	6
23	99	165	135	85	32	26	2
91	155	133	136	144	152	57	28
98	196	76	38	26	60	170	51
165	60	60	27	77	85	43	136
71	13	34	23	108	27	48	110

Gradient Magnitude

80	36	5	10	0	64	90	73
37	9	9	179	78	27	169	166
87	136	173	39	102	163	152	176
76	13	1	168	159	22	125	143
120	70	14	150	145	144	145	143
58	86	119	98	100	101	133	113
30	65	157	75	78	165	145	124
11	170	91	4	110	17	133	110

Gradient Direction

Center : The RGB patch and gradients represented using arrows. Right : The gradients in the same patch represented as numbers

Histogram of Oriented Gradient - HOG

80	36	5	10	0	64	90	73
37	9	9	179	78	27	169	166
87	136	173	39	102	163	152	176
76	13	1	168	159	22	125	143
120	70	14	150	145	144	145	143
58	86	119	98	100	101	133	113
30	65	157	75	78	165	145	124
11	170	91	4	110	17	133	110

Gradient Direction

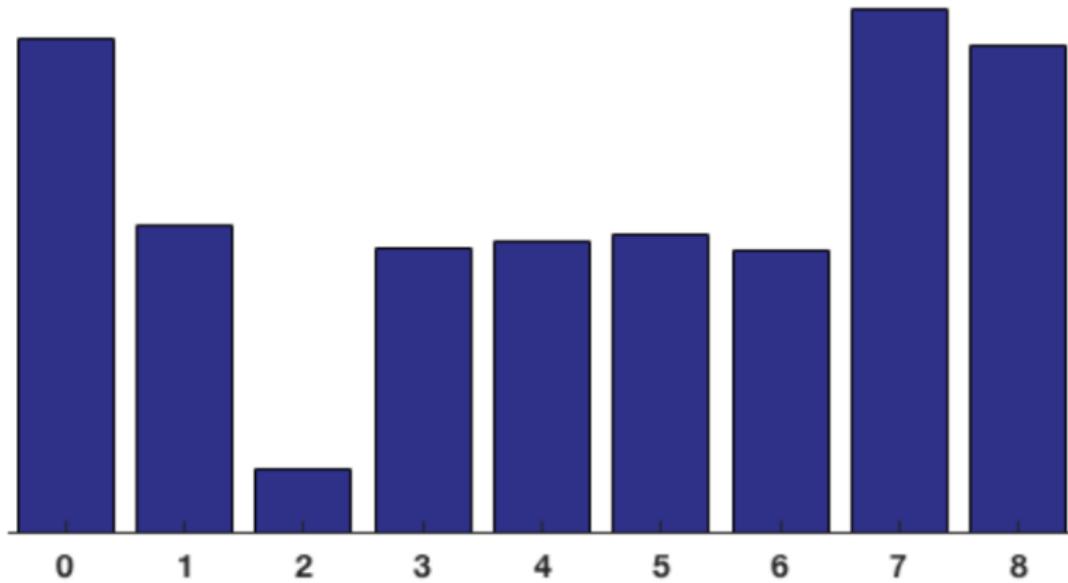
2	3	4	4	3	4	2	2
5	11	17	13	7	9	3	4
11	21	23	27	22	17	4	6
23	99	165	135	85	32	26	2
91	155	133	136	144	152	57	28
98	196	76	38	26	60	170	51
165	60	60	27	77	85	43	136
71	13	34	23	108	27	48	110

Gradient Magnitude



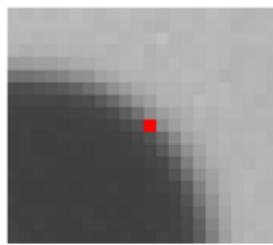
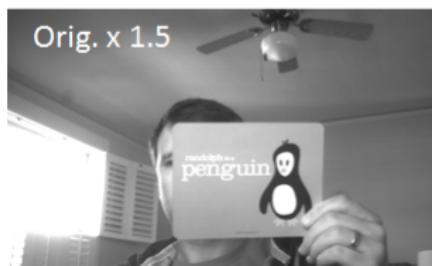
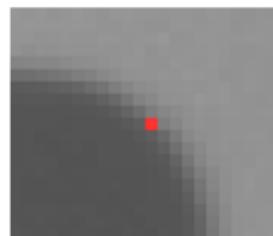
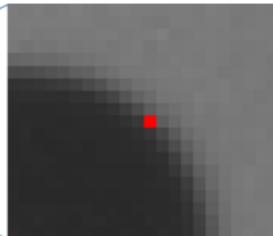
Histogram of Gradients

Histogram of Oriented Gradient - HOG

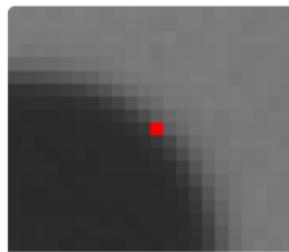


Histogram of Oriented Gradient - HOG

Chuẩn hóa Block



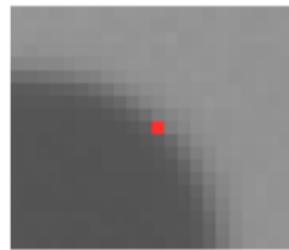
Histogram of Oriented Gradient - HOG



	93	
56		94
	55	

$$\nabla f = \begin{bmatrix} 38 \\ 38 \end{bmatrix}$$

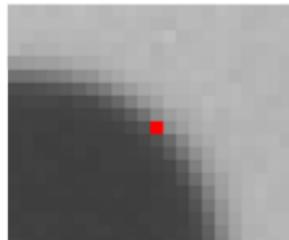
$$|\nabla f| = \sqrt{(38)^2 + (38)^2} = 53.74$$



	143	
106		144
	105	

$$\nabla f = \begin{bmatrix} 38 \\ 38 \end{bmatrix}$$

$$|\nabla f| = \sqrt{(38)^2 + (38)^2} = 53.74$$



	140	
84		141
	83	

$$\nabla f = \begin{bmatrix} 57 \\ 57 \end{bmatrix}$$

$$|\nabla f| = \sqrt{(57)^2 + (57)^2} = 80.61$$

Histogram of Oriented Gradient - HOG

Các giá trị chuẩn hóa có thể tính bằng một trong các công thức sau:

$$\text{L2-norm: } f = \frac{v}{\sqrt{\|v_2\|^2 + e^2}}$$

$$\text{L1-norm: } f = \frac{v}{\|v_1\| + e}$$

$$\text{L1-sqrt: } f = \sqrt{\frac{v}{\|v_1\| + e}}$$

Thực nghiệm

Dữ liệu:

- ▶ Penn Fudan
- ▶ INRIA Person

Code:

[https:](https://github.com/nducthang/Pedestrians-Detections-HOG)

//github.com/nducthang/Pedestrians-Detections-HOG

Thực nghiệm

Sử dụng 2 bộ dữ liệu đã giới thiệu để tách các mẫu thành positive (người đi bộ) và negative (không phải người đi bộ). Sau đó trích rút đặc trưng HOG tương ứng với các mẫu từ dữ liệu, đầu ra với mỗi ảnh của HOG là một vector 1D có 3780 phần tử. Với các mẫu là người đi bộ, đặt là nhãn 1, và ngược lại đặt là 0. Sau đó vector đặc trưng và vector nhãn được đưa qua mô hình SVM để huấn luyện phân lớp, file weight của mô hình được lưu lại phục vụ cho việc kiểm thử.

Thực nghiệm

Kiểm thử:

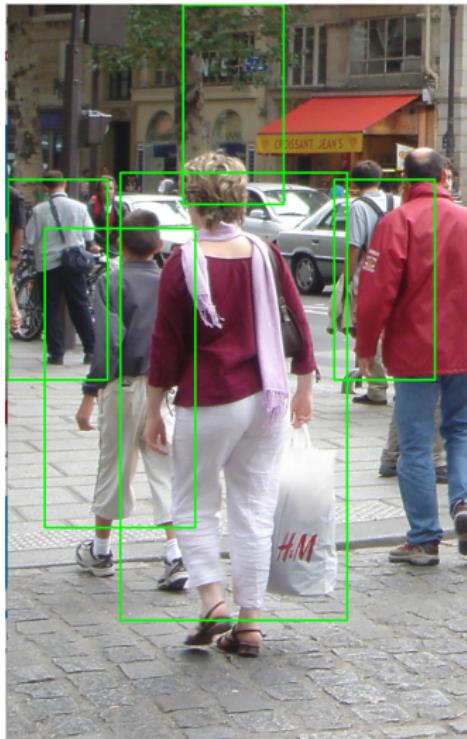
1. Load model đã huấn luyện cho dự đoán
2. Dùng sliding window để lấy ra các box với kích cỡ khác nhau của ảnh.
3. Các box được đưa qua HOG để biến đổi thành vector đặc trưng. Các vector đặc trưng này sau đó được đưa qua mô hình đã huấn luyện cho phân lớp để dự đoán. Nếu xác suất dự đoán người đi bộ của vector đó lớn hơn ngưỡng đặt ra (Threshold) thì box đó được lưu giữ lại. Mỗi box lưu toạ độ góc trên trái nhất và góc dưới phải nhất.
4. Sử dụng Non-maximum suppression để loại bỏ các box mà chồng lấn nhau lớn hơn ngưỡng đặt ra (overlapThreshold).
5. Vẽ các box còn lại lên ảnh và hiển thị kết quả.

Thực nghiệm



Hình: Kết quả dự đoán của mô hình cho 1 mẫu ảnh

Thực nghiệm



Hình: Kết quả dự đoán của mô hình cho 1 mẫu ảnh

Thực nghiệm

Ưu điểm:

- ▶ Thuật toán đơn giản, dễ triển khai.
- ▶ Nhận dạng được người đi bộ ở mức khá.

Nhược điểm:

- ▶ Vẫn sai sót ở nhiều mẫu ảnh,
- ▶ Chương trình chạy chậm, chưa thích hợp với các dự án thực tế đòi hỏi thời gian thực.