

THỊ GIÁC MÁY TÍNH

Bài 7: Phát hiện đối tượng

Lê Thị Lan, Nguyễn Văn Tới

AI Academy Vietnam

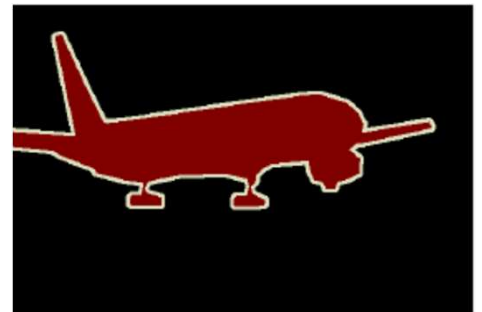
Nội dung buổi học

- Giới thiệu về phát hiện đối tượng
- Nguyên lý của phát hiện đối tượng
- Một số phương pháp phát hiện đối tượng
- Phát hiện đối tượng dựa trên Haar-like và Adaboost
- Giới thiệu bài toán phát hiện tay ứng dụng trong các hệ thống điều khiển và tương tác người-robot

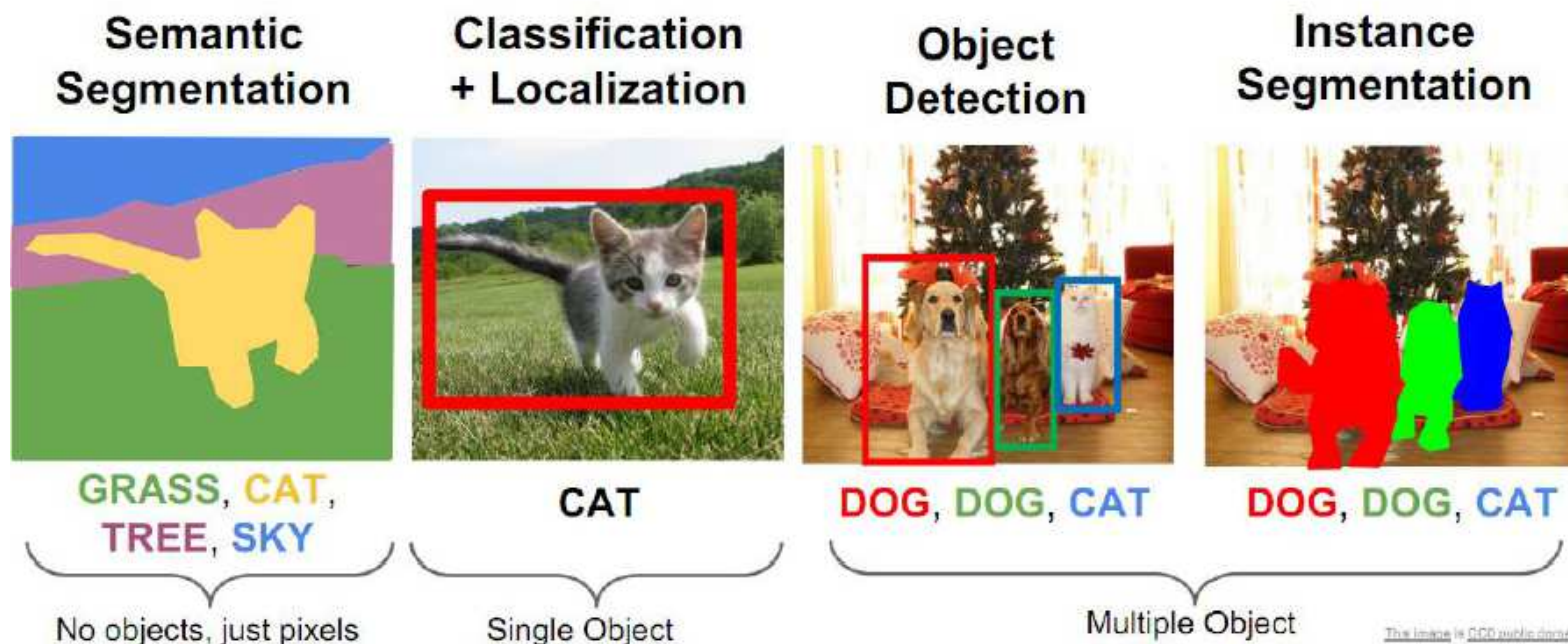
Giới thiệu về phát hiện đối tượng

Giới thiệu về phát hiện đối tượng

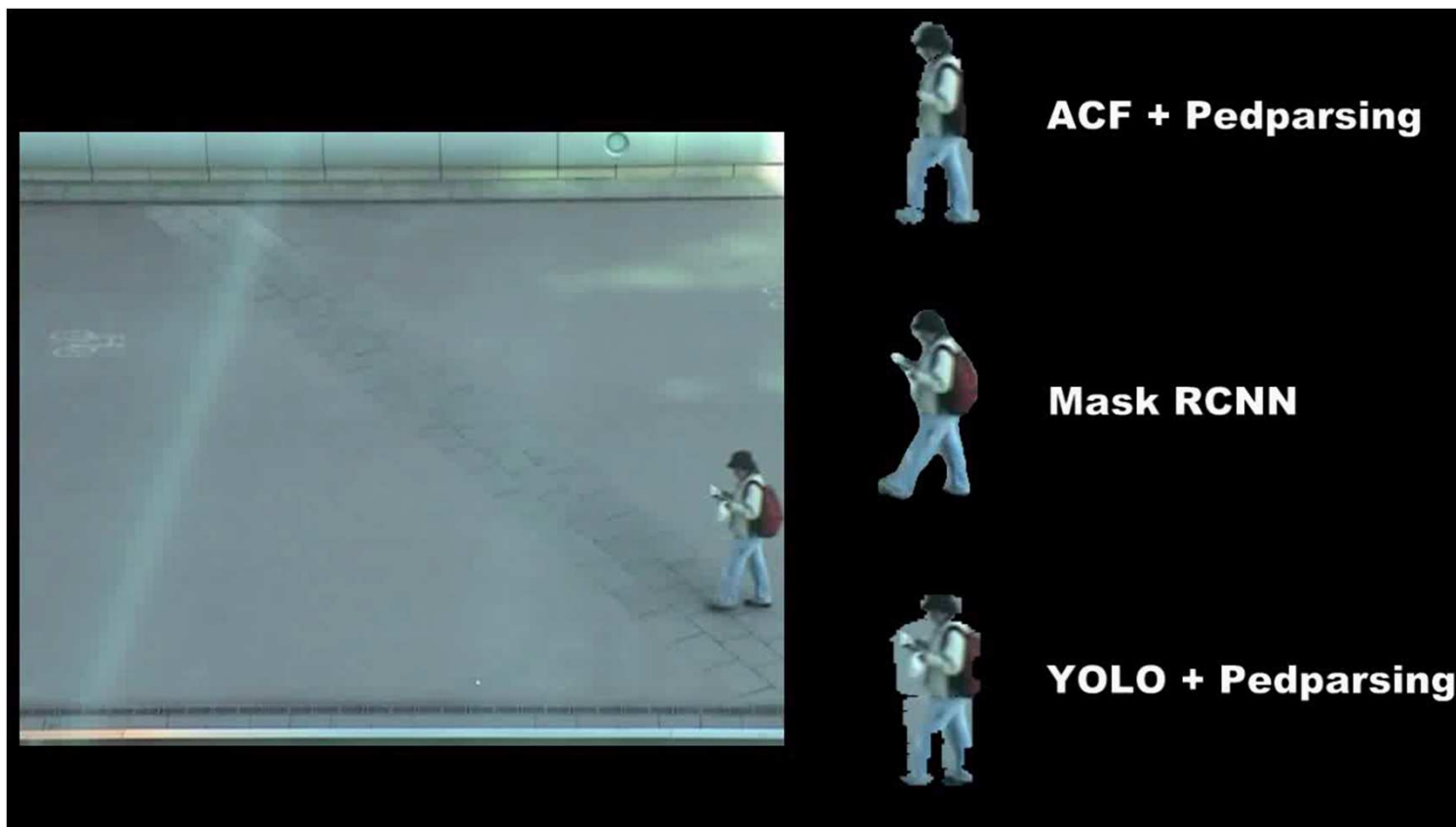
- Phân loại ảnh (Image classification)
 - Xác định sự tồn tại của đối tượng quan tâm trong ảnh.
- Phát hiện/định vị đối tượng (Object detection)
 - Xác định sự tồn tại và vị trí của đối tượng quan tâm
- Phân vùng đối tượng mức điểm ảnh
(Object pixel-level segmentation)
 - Xác định các điểm ảnh thuộc vào đối tượng quan tâm



Giới thiệu về phát hiện đối tượng

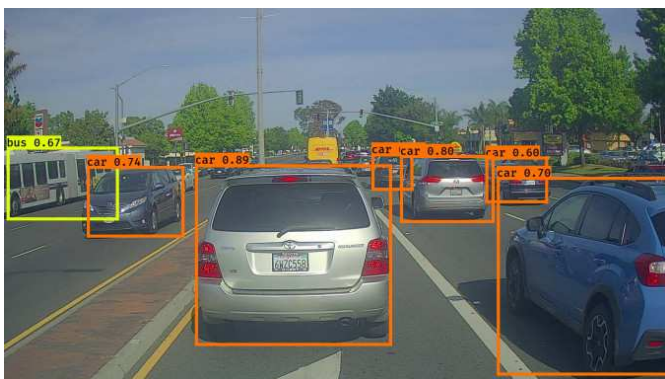


Giới thiệu về phát hiện đối tượng

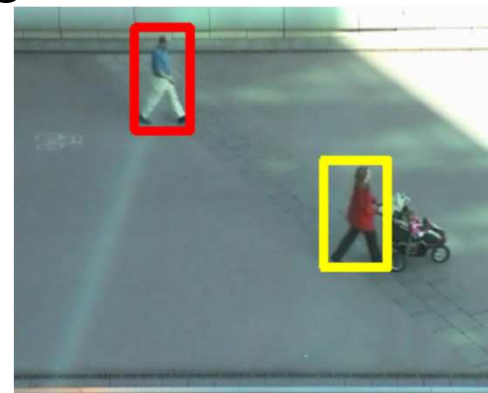


Giới thiệu về phát hiện đối tượng

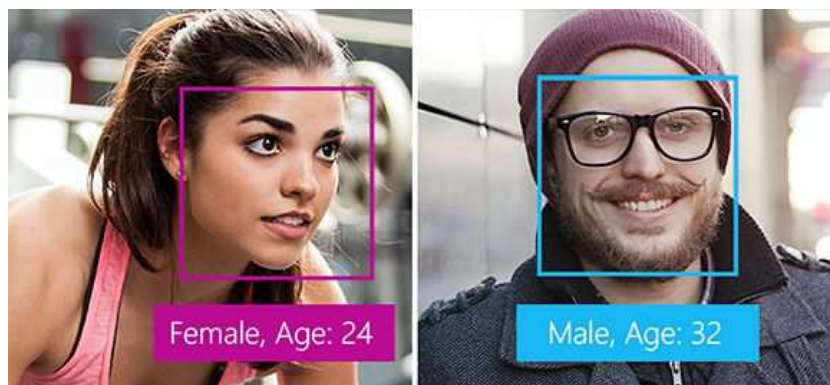
- Ứng dụng của bài toán phát hiện đối tượng



Phát hiện xe trong các hệ thống xe không người lái, các hệ thống trợ giúp



Phát hiện người trong hệ thống camera giám sát → phân tích hành vi, nhận dạng hoạt động, đánh giá hành trình....

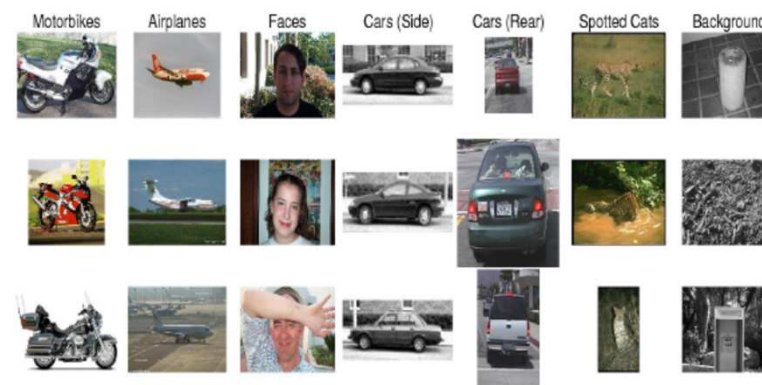


Phát hiện mặt người trong hệ thống kiểm soát vào ra, nhận dạng cảm xúc...

Giới thiệu về phát hiện đối tượng

- Các khó khăn và thách thức của bài toán phát hiện đối tượng

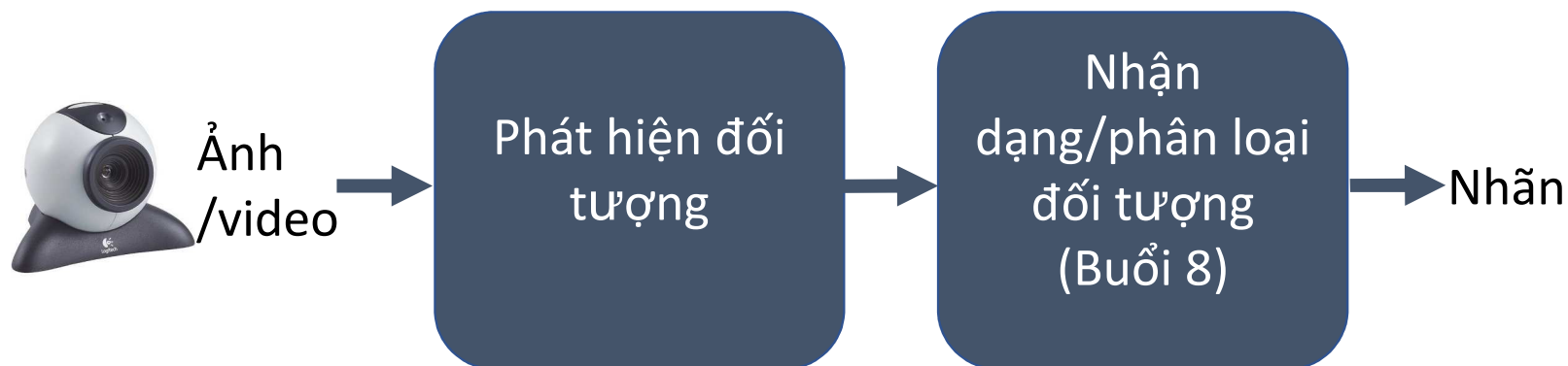
- Số lớp lớn
- Che khuất
- Điều kiện chiếu sáng
- Sự đa dạng về hình dáng, thể hiện của các đối tượng trong cùng 1 lớp (intra-class variation)
- Nền phức tạp
- Sự giống nhau của các đối tượng ở các lớp khác nhau (inter-class similarity)



Một số lớp đối tượng trong CSDL Caltech

Giới thiệu về phát hiện đối tượng


- Các thành phần chính của hệ thống phát hiện và nhận dạng đối tượng

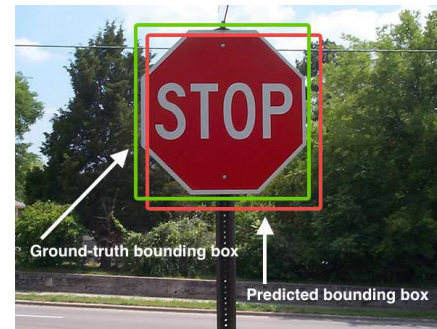


Hand detection and hand posture recognition

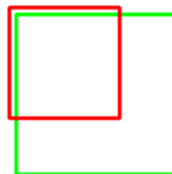
Giới thiệu về phát hiện đối tượng

- Các độ đo đánh giá phát hiện đối tượng
 - Chỉ số Jaccard (Jaccard index)

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$


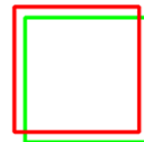


IoU: 0.4034



Poor

IoU: 0.7330



Good

IoU: 0.9264



Excellent

Giới thiệu về phát hiện đối tượng

- Ma trận nhầm lẫn (confusion matrix)

		Predicted Label	
		Positive	Negative
Known Label	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

		Predicted Label	
		Positive	Negative
Known Label	Positive	Hit	Miss
	Negative	False Alarm	Correct Rejection

Giới thiệu về phát hiện đối tượng

• Các độ đo

Measure	Formula	Intuitive Meaning
Precision	$TP / (TP + FP)$	The percentage of positive predictions that are correct.
Recall / Sensitivity	$TP / (TP + FN)$	The percentage of positive labeled instances that were predicted as positive.
Specificity	$TN / (TN + FP)$	The percentage of negative labeled instances that were predicted as negative.
Accuracy	$(TP + TN) / (TP + TN + FP + FN)$	The percentage of predictions that are correct.

- **Recall: true positive rate hit rate**
- **False positive rate or false alarm rate = $FP / (FP + TN)$**
- **Specificity = $1 - \text{False positive rate}$**
- **F-measure = $2 / (1/\text{precision} + 1/\text{recall})$**

Giới thiệu về phát hiện đối tượng

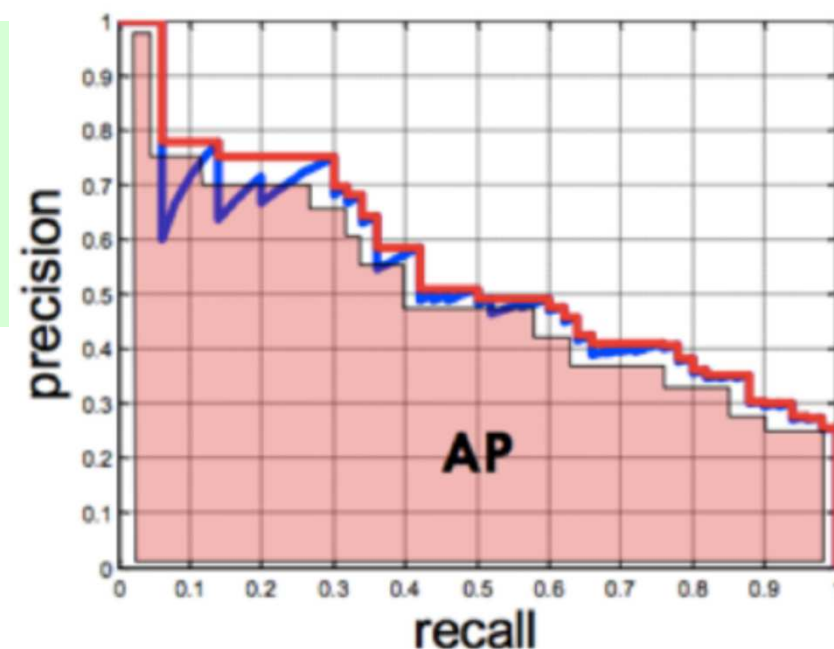
- **AP (Average Precision):** Diện tích phần bên dưới của đường cong Precision-Recall
- **mAP:** giá trị trung bình AP được tính trên các lớp

Average Precision (AP):

AP % AP at IoU=.50:.05:.95 (primary challenge metric)
 $AP^{IoU=.50}$ % AP at IoU=.50 (PASCAL VOC metric)
 $AP^{IoU=.75}$ % AP at IoU=.75 (strict metric)

AP Across Scales:

AP^{small} % AP for small objects: $area < 32^2$
 AP^{medium} % AP for medium objects: $32^2 < area < 96^2$
 AP^{large} % AP for large objects: $area > 96^2$

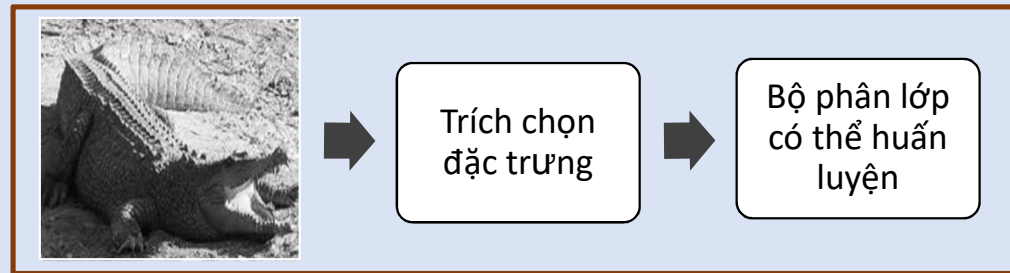


Nguyên lý phát hiện đối tượng

Nguyên lý của phát hiện đối tượng

- Các cách tiếp cận chính

Cách tiếp cận truyền thống:



Cách tiếp cận dựa trên học sâu (Buổi 13)



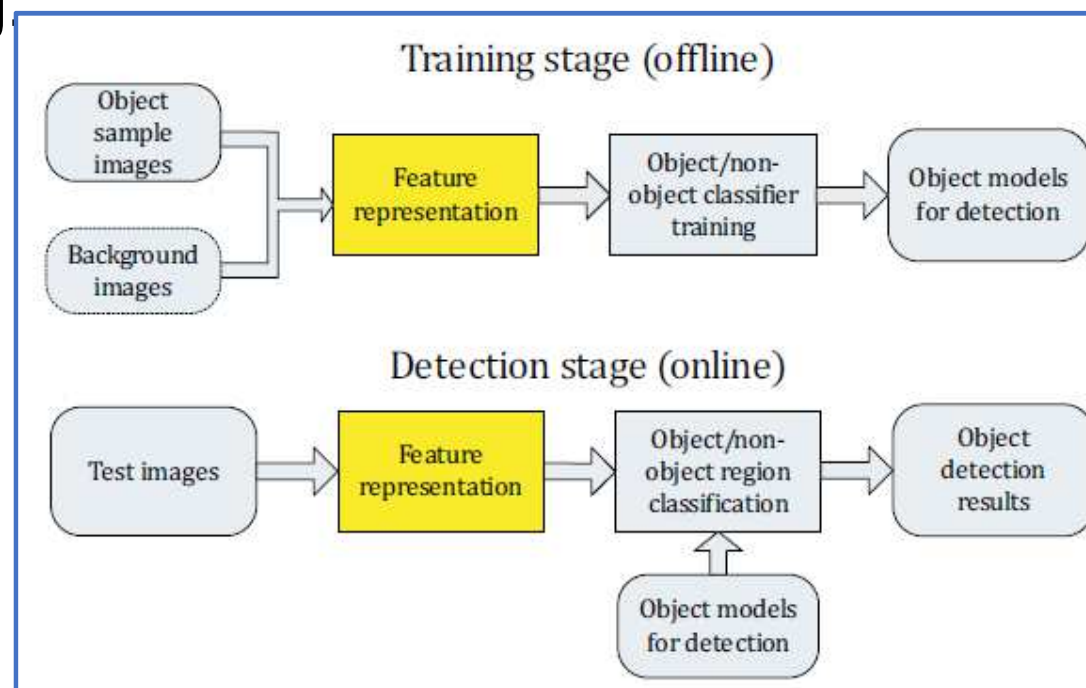
Nguyên lý của phát hiện đối tượng

- **Hai thành phần quan trọng:**

- Biểu diễn đối tượng
- Học mô hình biểu diễn

- **Hai giai đoạn chính:**

- Huấn luyện
- Nhận dạng



Yali Li, Shengjin Wang, Qi Tian, Xiaoqing Ding, Feature representation for statistical learning-based object detection: a review, PR 2015

Nguyên lý của phát hiện đối tượng

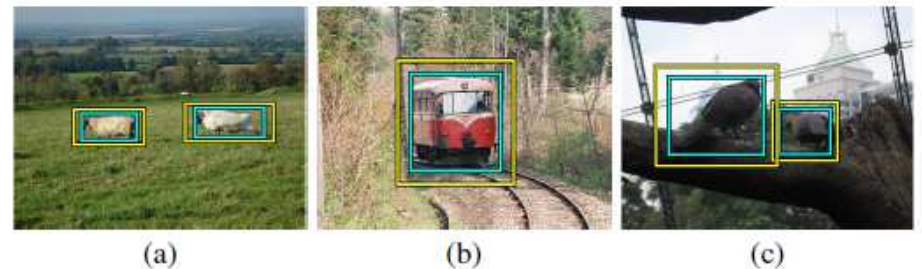
- **Nguyên lý:**

- Tìm đặc điểm (thông qua đặc trưng) **cho phép phân biệt đối tượng** quan tâm với các đối tượng khác hoặc với nền
- **Huấn luyện bộ phân lớp để học đặc điểm đó** → hàm khoảng cách
- Đối tượng quan tâm (positive samples), đối tượng khác (negative samples)
- **Phát hiện đối tượng/lớp đối tượng cụ thể:**
 - Tìm đặc trưng để biểu diễn đặc điểm của lớp đối tượng quan tâm (ví dụ HOG, SIFT,...)
- **Phát hiện đối tượng tổng quát:**
 - Bộ ước lượng Objectness: “It quantifies how likely it is for an image window to cover an object of any class. Objects are **standalone things** with a **welldefined boundary** and **center**, such as cows, cars, and telephones, as opposed to amorphous background stuff, such as sky, grass, and road”

Nguyên lý của phát hiện đối tượng

- **Phát hiện đối tượng tổng quát:**

- “It quantifies how likely it is for an image window to cover an object of any class. Objects are **standalone things** with a **welldefined boundary** and **center**, such as cows...**as opposed** to amorphous background stuff, such as sky, grass...”
- 4 dấu hiệu của đối tượng
 - **Multi-scale Saliency (MS)**
 - **Color Contrast (CC)**
 - **Edge Density (ED)**
 - **Superpixels Straddling (SS)**



Kết quả dựa trên màu sắc (CC)

Codes available: <http://groups.inf.ed.ac.uk/calvin/objectness/>

Nguyên lý của phát hiện đối tượng

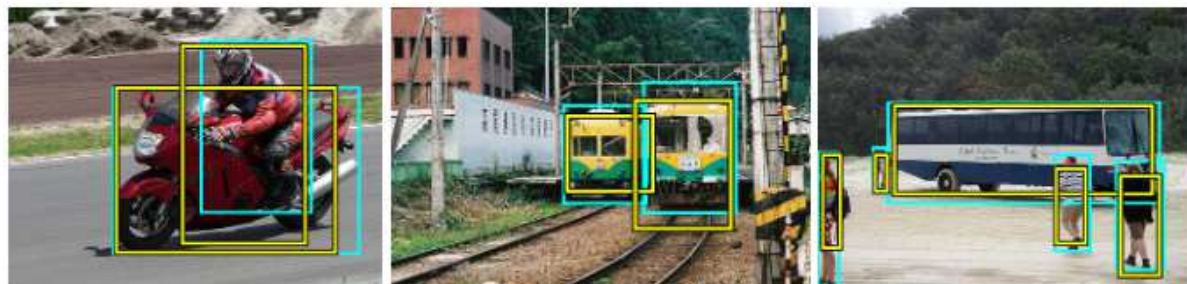
- **Phát hiện đối tượng tổng quát:**

- CC được xác định là khoảng cách của histogram màu trong không gian LAB của cửa sổ w và cửa sổ bao quanh $Surr$

$$CC(w, \theta_{CC}) = \chi^2(h(w), h(Surr(w, \theta_{CC})))$$



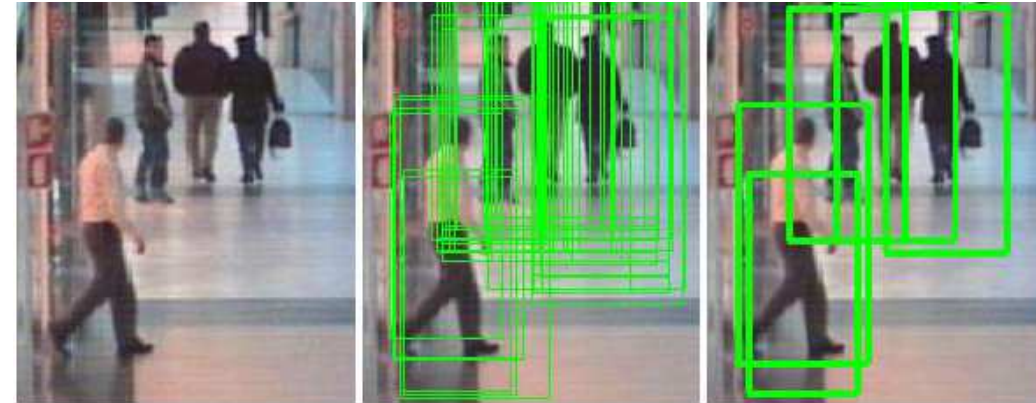
Bản đồ của objectness



Kết quả trên PASCAL VOC 2017, màu vàng: ước lượng của phương pháp, màu cyan: grountruth

Nguyên lý của phát hiện đối tượng

- Tạo ra các cửa sổ quét với các kích thước và vị trí khác nhau
 - Tại sao???
- Xác định sự tồn tại của đối tượng quan tâm trong các cửa sổ
- Tổng hợp kết quả, chiến lược NMS (Non-maxima Suppression)
 - Ở mỗi bước lặp chọn cửa sổ có độ tin cậy cao nhất
 - Kiểm tra IoU và loại tất cả các cửa sổ có $\text{IoU} \geq$ ngưỡng nào đó với cửa sổ đã chọn
 - Tại sao???



Ảnh gốc

Các cửa sổ có đối tượng quan tâm

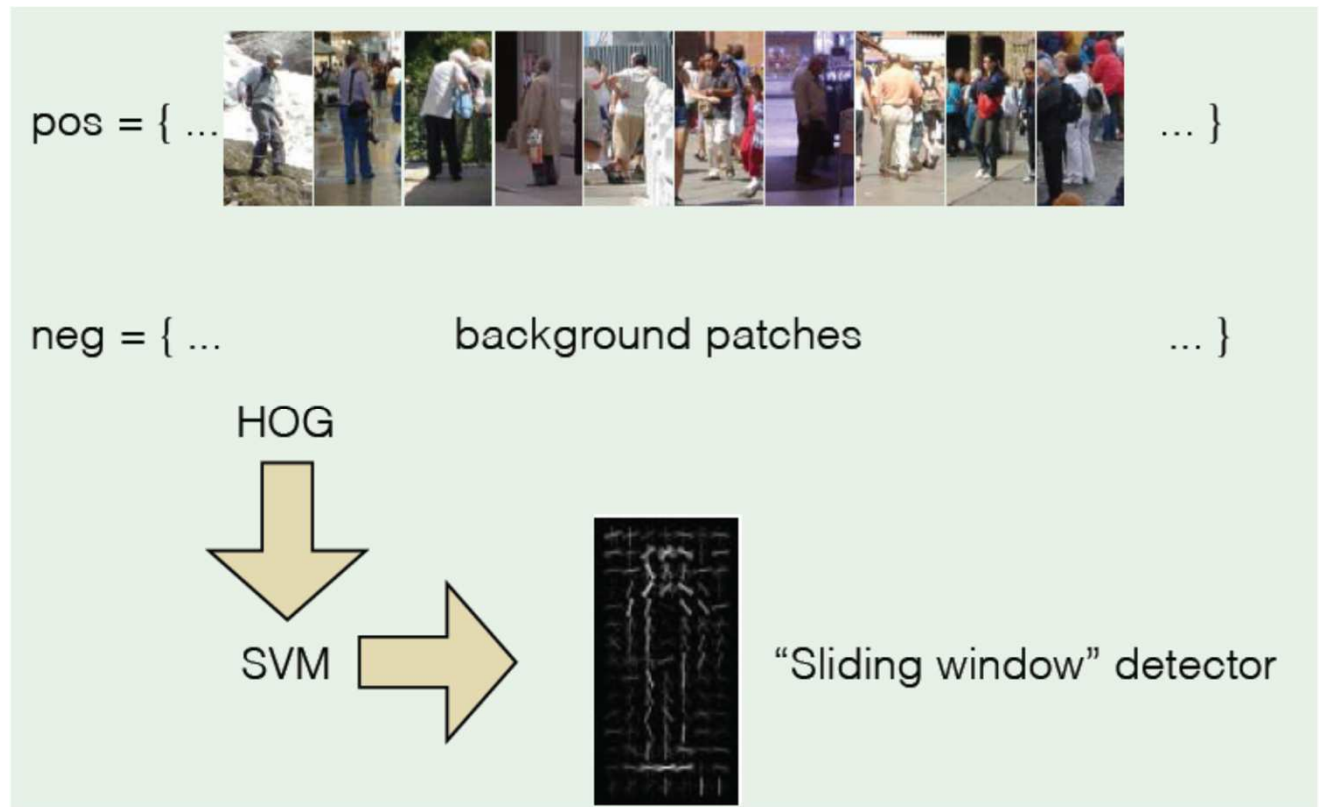
Tổng hợp kết quả

Một số phương pháp phát hiện đối tượng

Một số phương pháp phát hiện đối tượng

Histogram of Gradient (HOG)

- Histogram of Gradient

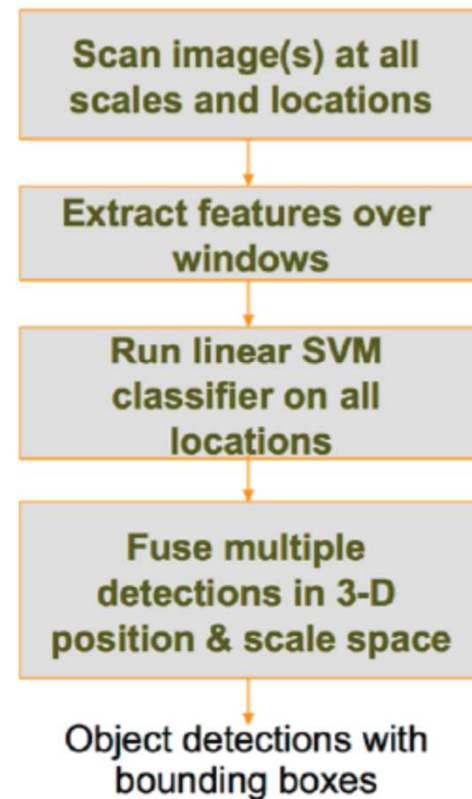


(Dalal, CVPR 2005)

Một số phương pháp phát hiện đối tượng

Histogram of Gradient (HOG)

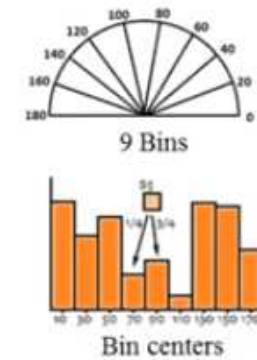
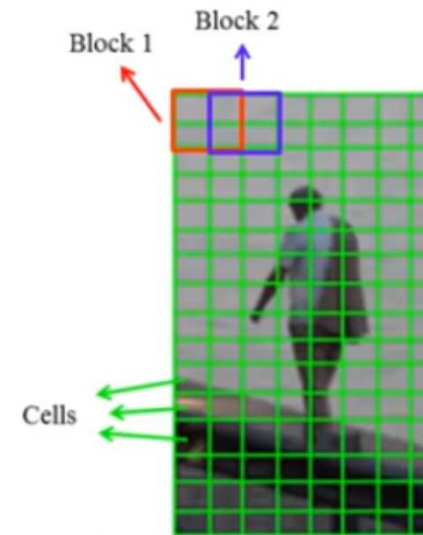
Detection Phase



Một số phương pháp phát hiện đối tượng

Histogram of Gradient (HOG)

- Chia ảnh thành các blocks có kích thước 16*16 chồng nhau 50%
- Block được chia thành 4 cells có kích thước 8*8 không chồng nhau
- Với mỗi cell, tính phân bố gradient theo hướng
- Tính đặc trưng cho mỗi blocks gồm 4 cells
- Số đặc trưng HOG của 1 ảnh:



$$\# \text{ features} = 15 \times 7 \times 9 \times 4 = 3780$$

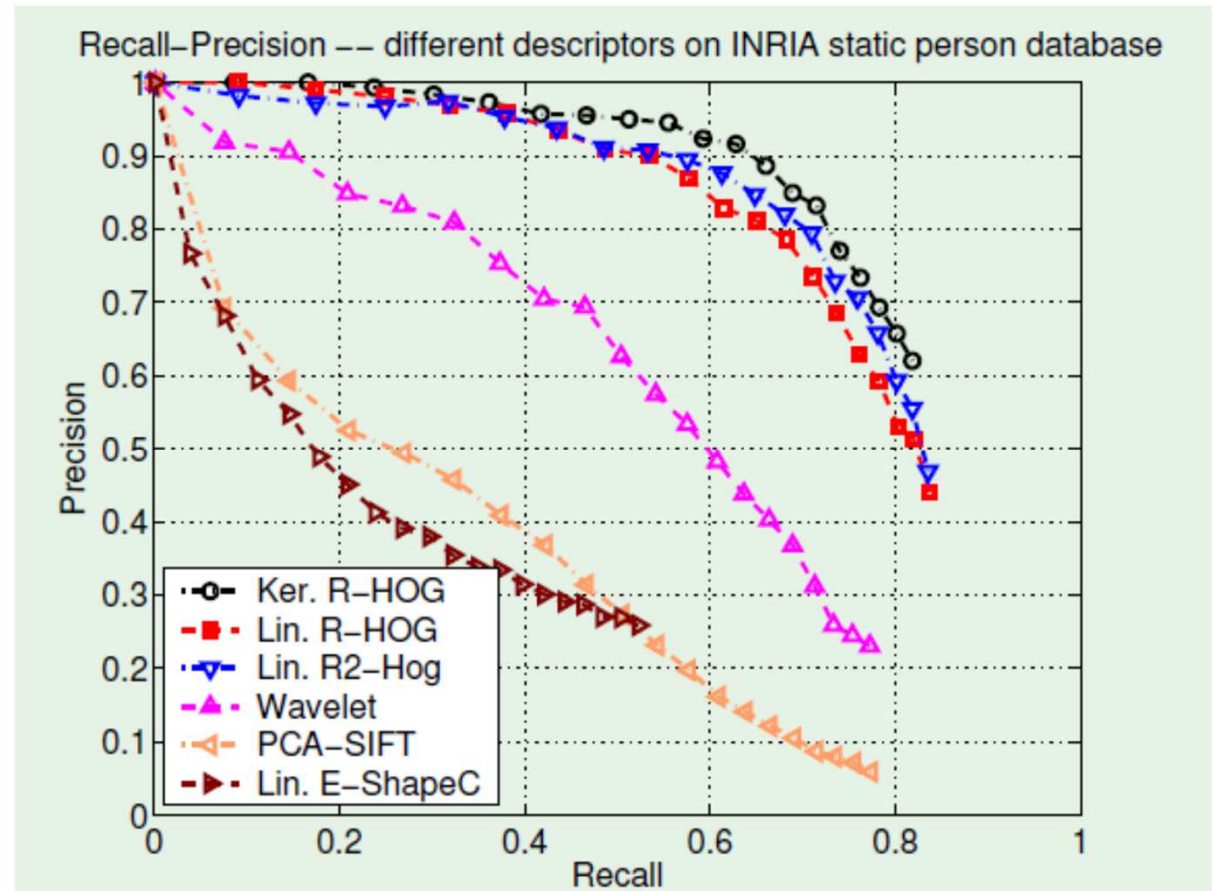
Annotations for the formula:

- 15 : # cells
- 7 : # orientations
- 9 : # normalizations by neighboring cells
- 4 : (implied from context, representing the 4 cells in a block)

Một số phương pháp phát hiện đối tượng

Histogram of Gradient (HOG)

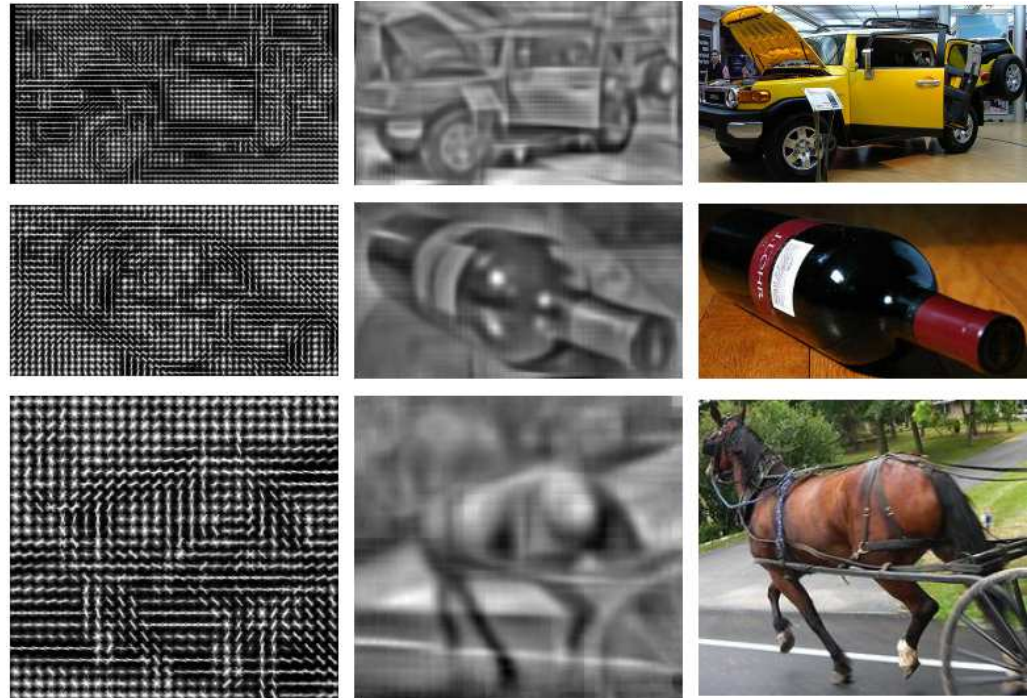
- Kết quả so sánh trên CSDL INRIA



Một số phương pháp phát hiện đối tượng

Histogram of Gradient (HOG)

- Biểu diễn HOG



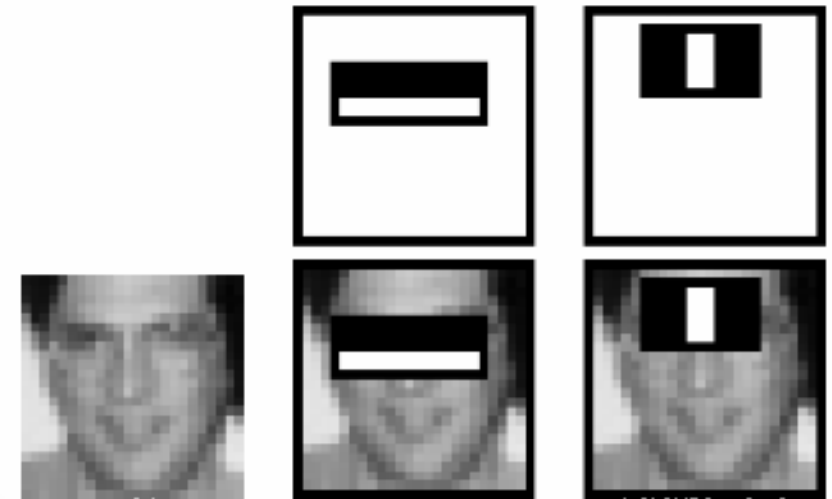
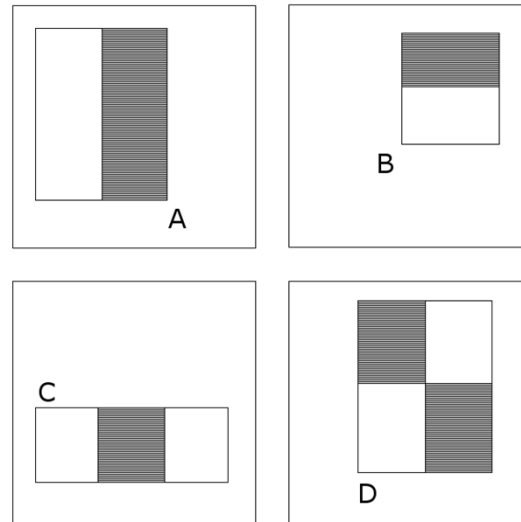
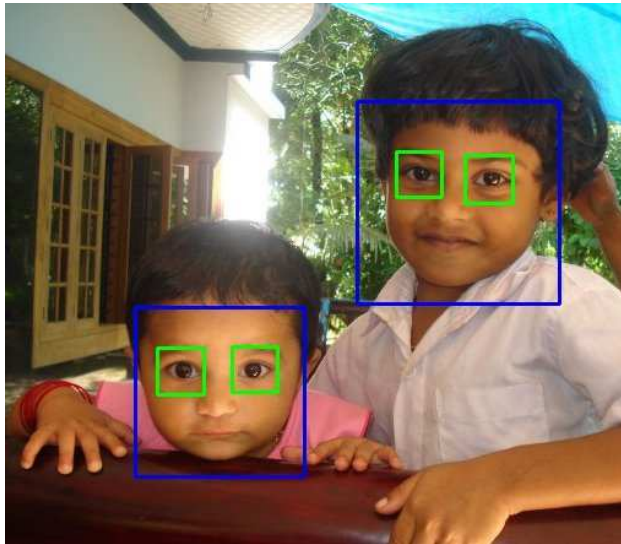
Đặc trưng HOG

Ảnh biểu diễn

Ảnh gốc

Một số phương pháp phát hiện đối tượng

Haarlike và Adaboost



Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. (Google scholar: Cited by **18120**). Searched on 11/12/2018.

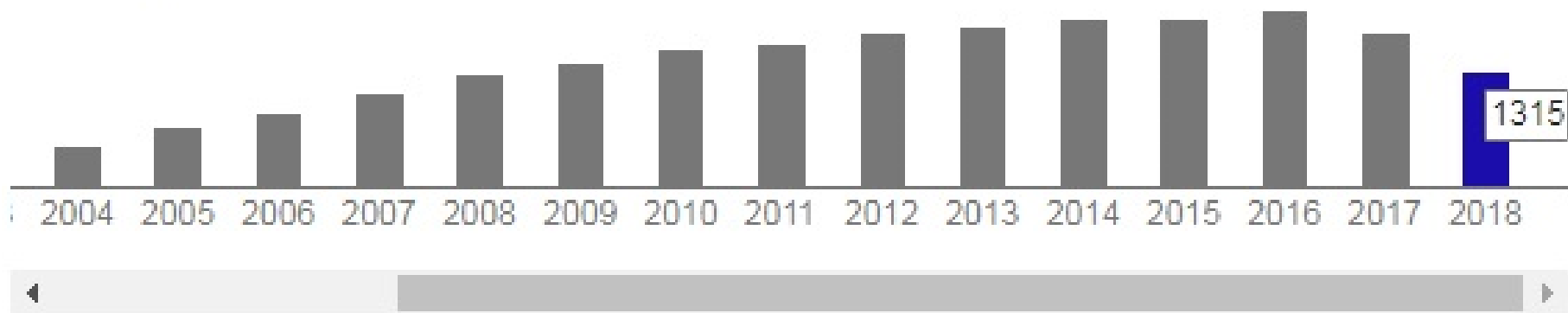
Viola, Paul, and Michael J. Jones. "Robust real-time face detection." *International journal of computer vision* 57, no. 2 (2004): 137-154. (Google scholar: Cited by **16486**)

Một số phương pháp phát hiện đối tượng

Haarlike và Adaboost

- Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. (Google scholar: Cited by **18120**). Searched on 11/12/2018.

Cited by 22769

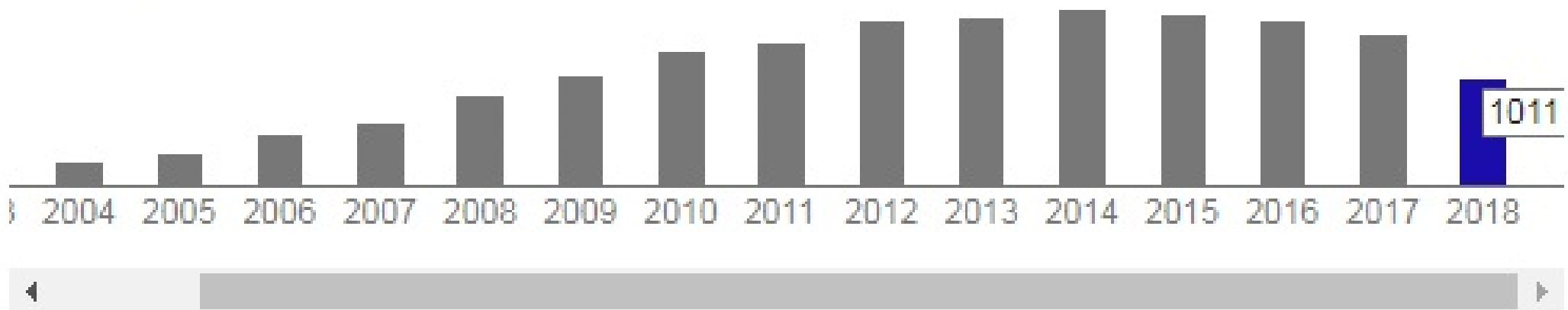


Một số phương pháp phát hiện đối tượng

Haarlike và Adaboost

- Viola, Paul, and Michael J. Jones. "Robust real-time face detection." *International journal of computer vision* 57, no. 2 (2004): 137-154. (Google scholar: Cited by **16486**).

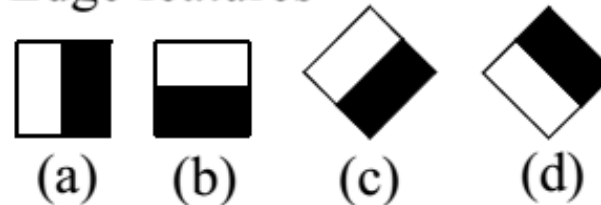
Cited by 16975



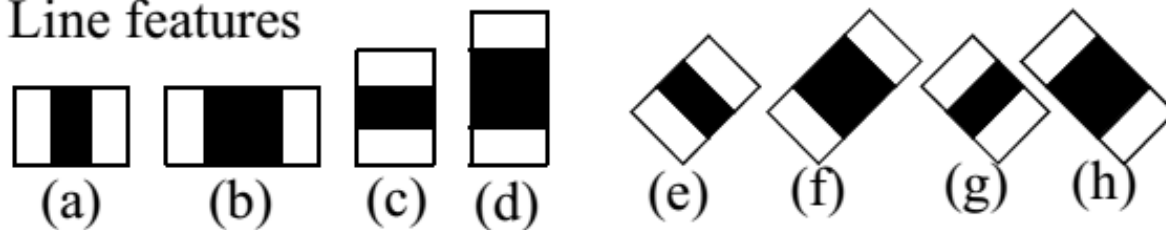
Một số phương pháp phát hiện đối tượng Haarlike và Adaboost

- Lienhart, Rainer, and Jochen Maydt. "An extended set of haar-like features for rapid object detection." *Image Processing. 2002. Proceedings. 2002 International Conference on*. Vol. 1. IEEE, 2002. (Google scholar: Cited by **3643**)

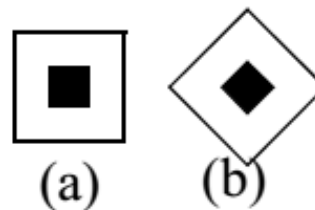
1. Edge features



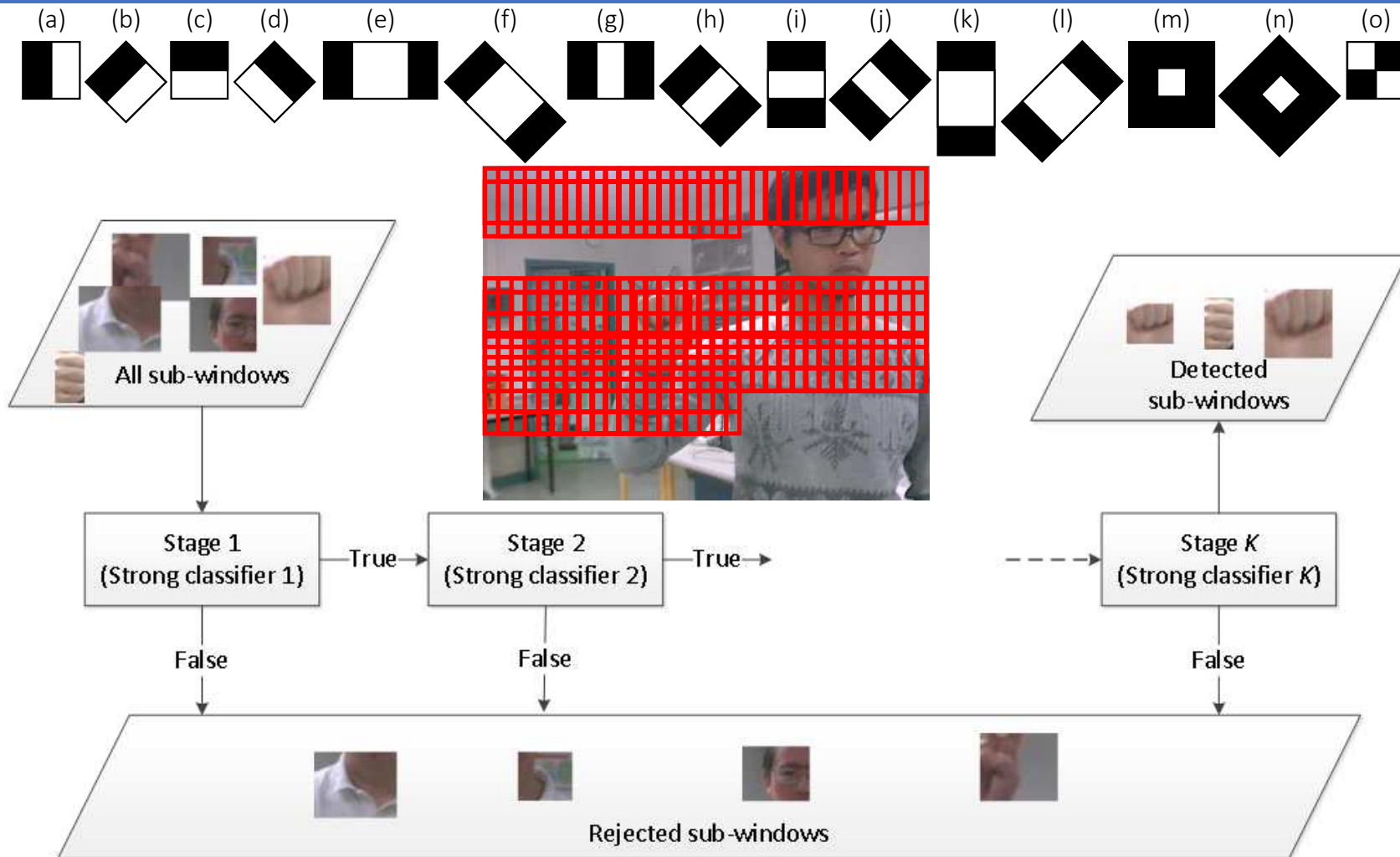
2. Line features



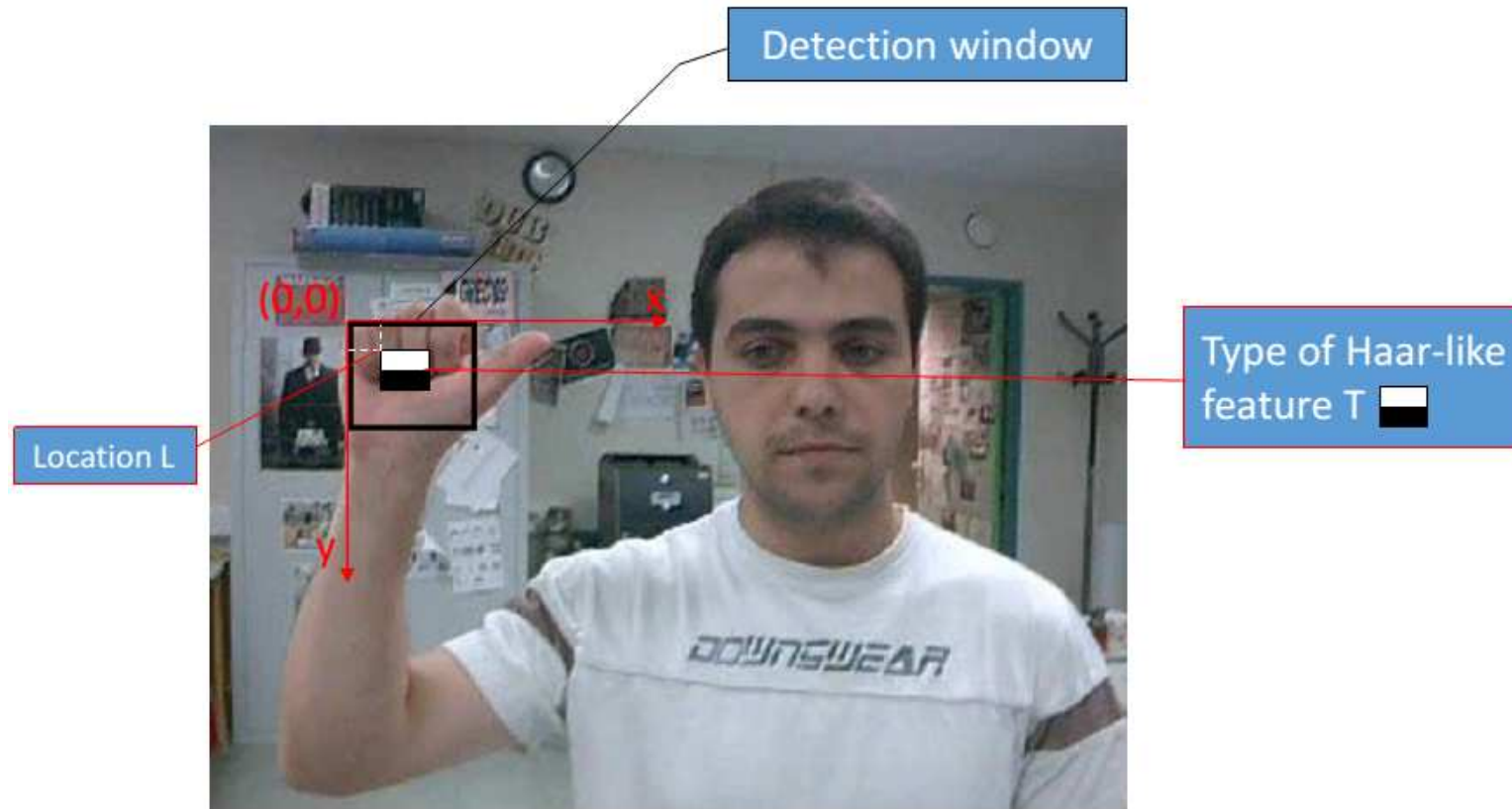
3. Center-surround features



Một số phương pháp phát hiện đối tượng Haarlike và Adaboost



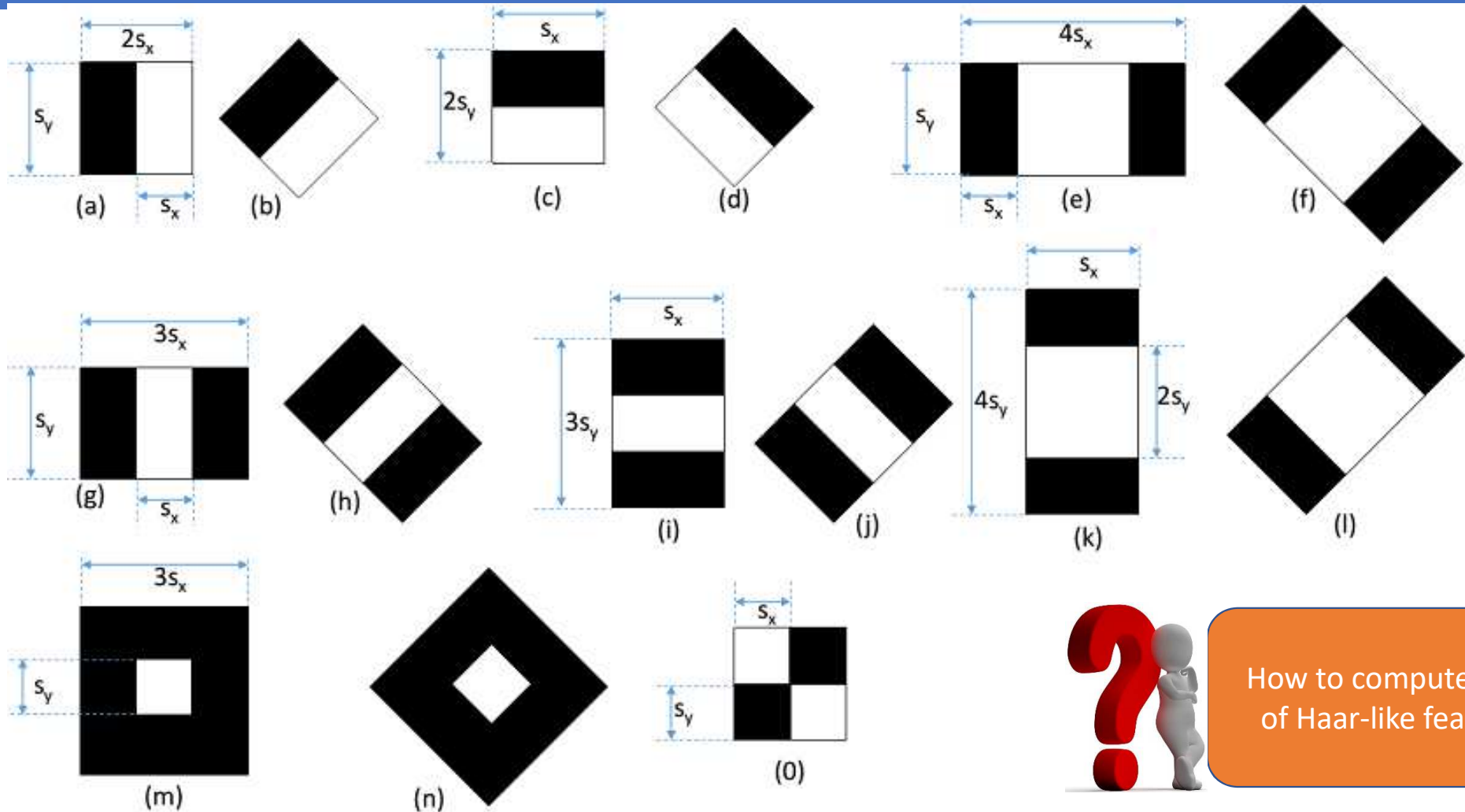
Một số phương pháp phát hiện đối tượng Haarlike và Adaboost



$$f(I) = \sum_{r \in RECT} w_r \cdot RectangleSum(r, I)$$

$$RectangleSum(r, I) = \sum_{(x,y) \in r} I(x, y)$$

Một số phương pháp phát hiện đối tượng Haarlike và Adaboost



How to compute values
of Haar-like features?

Một số phương pháp phát hiện đối tượng

Haarlike và Adaboost

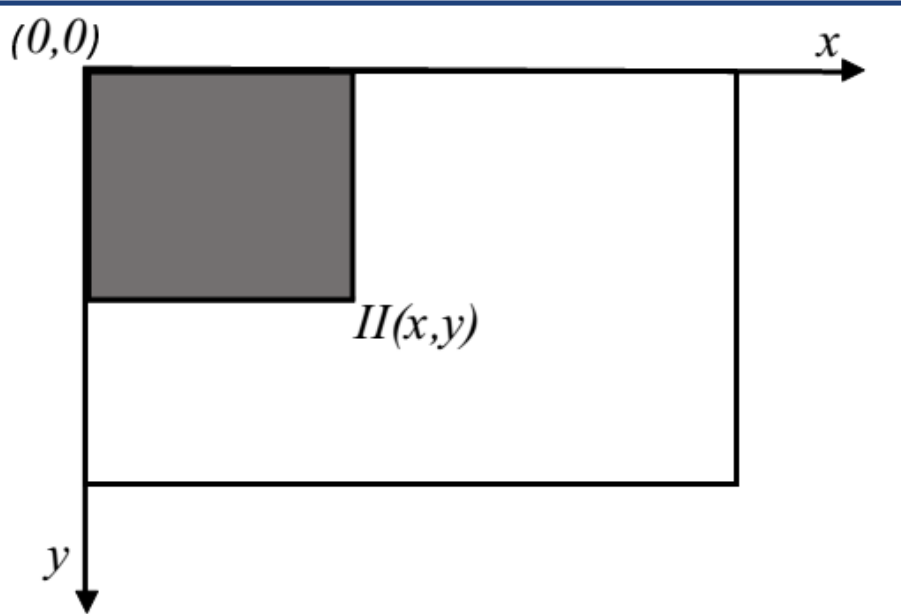


Figure 3.6: Axis-aligned integral image.

$$II(x, y) = \sum_{i=0}^x \sum_{j=0}^y I(i, j)$$

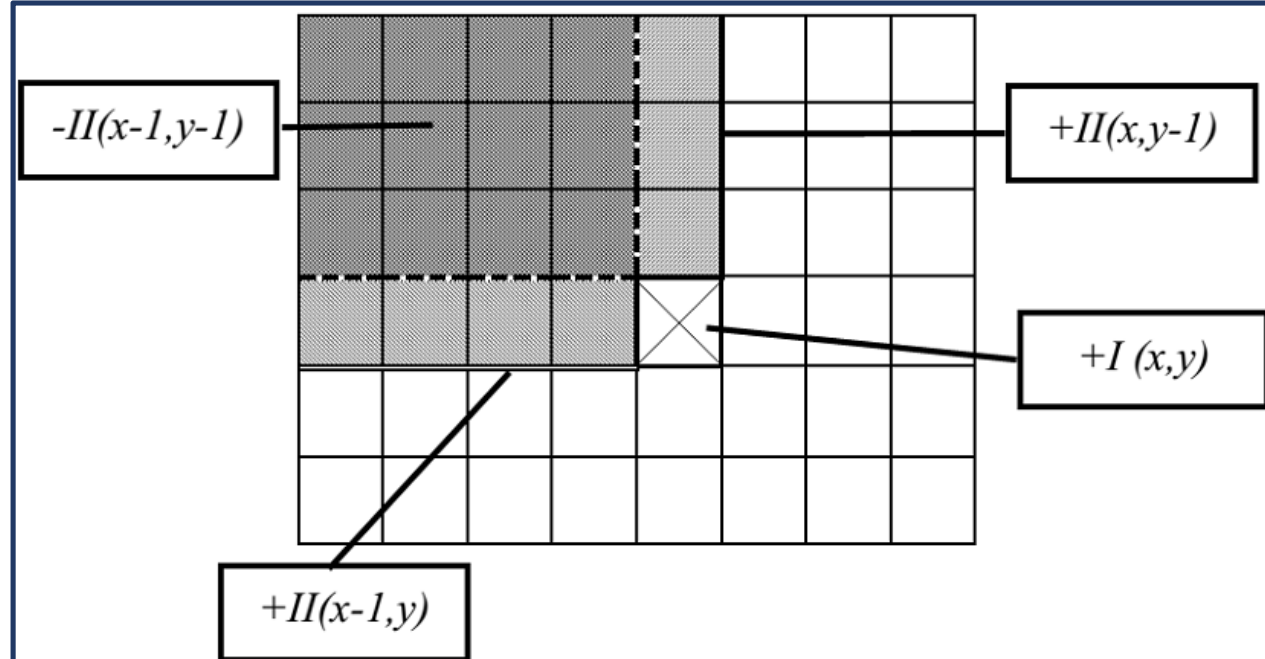
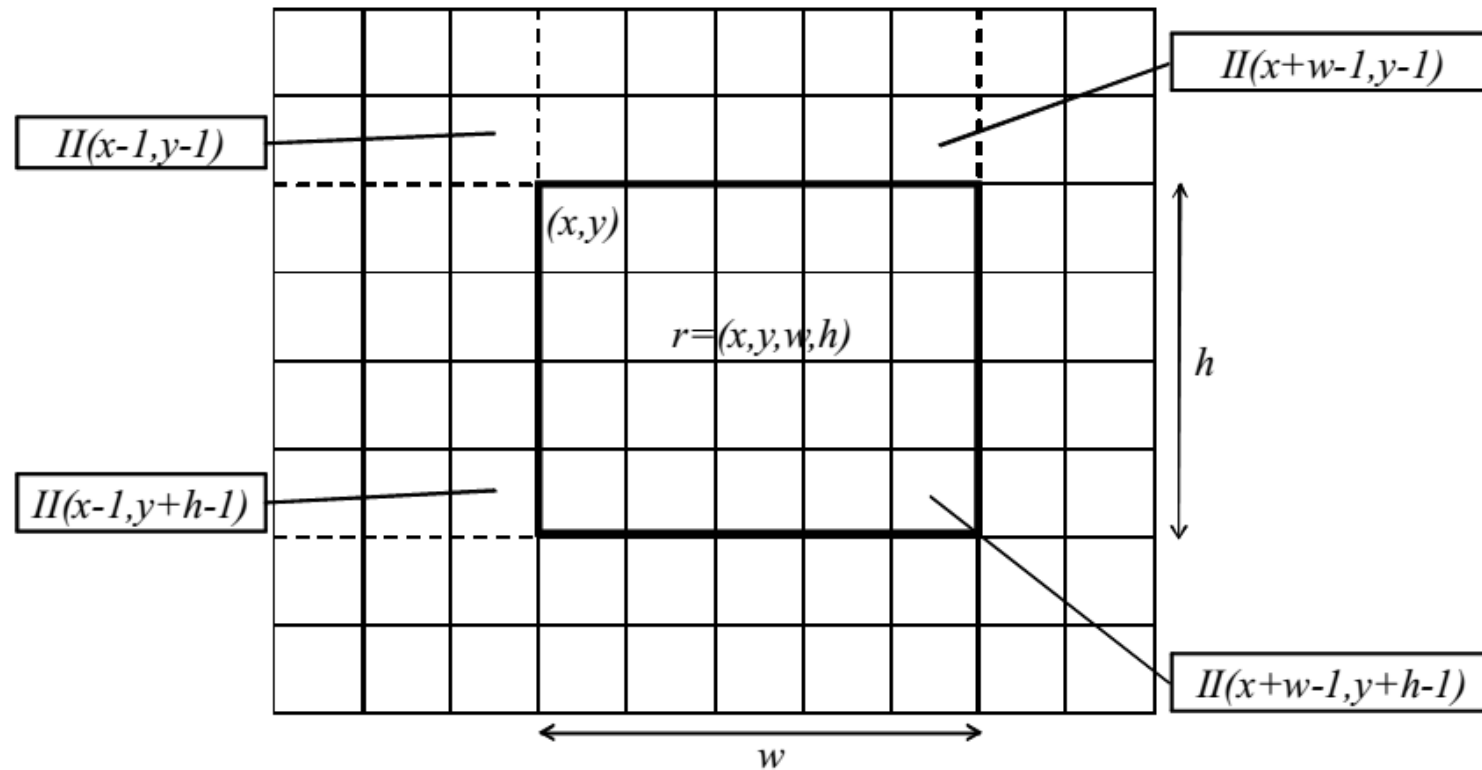


Figure 3.7: Computation of integral image.

Một số phương pháp phát hiện đối tượng Haarlike và Adaboost



$$\begin{aligned} \text{RectangleSum}(r, I) = & II(x + w - 1, y + h - 1) - II(x + w - 1, y - 1) \\ & - II(x - 1, y + h - 1) + II(x - 1, y - 1) \end{aligned}$$

Một số phương pháp phát hiện đối tượng Haarlike và Adaboost

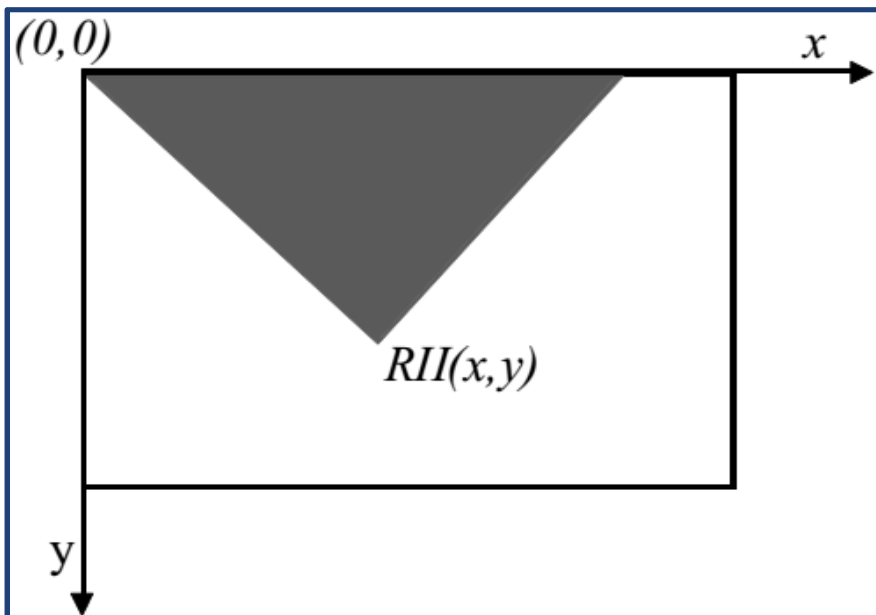


Figure 3.10: 45° rotated integral image.

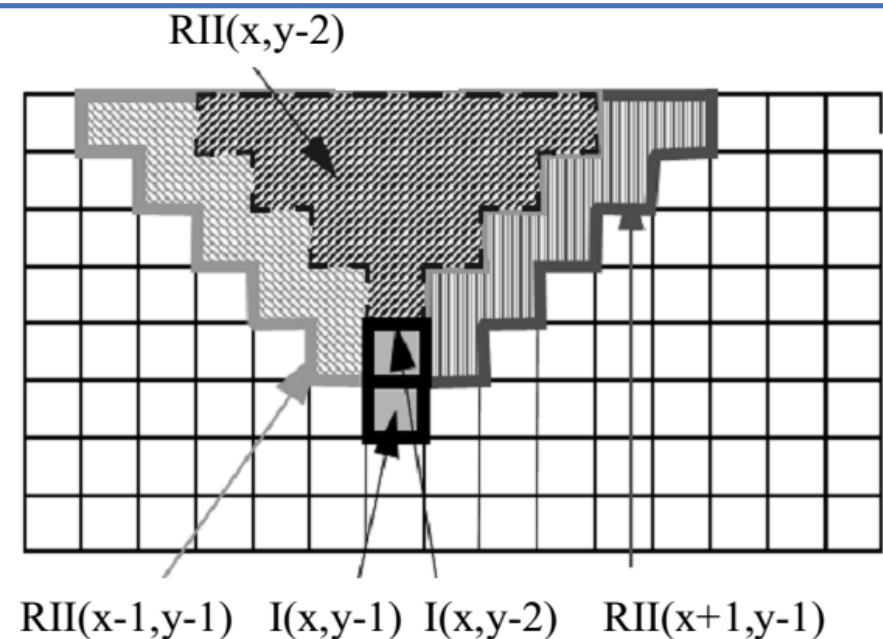
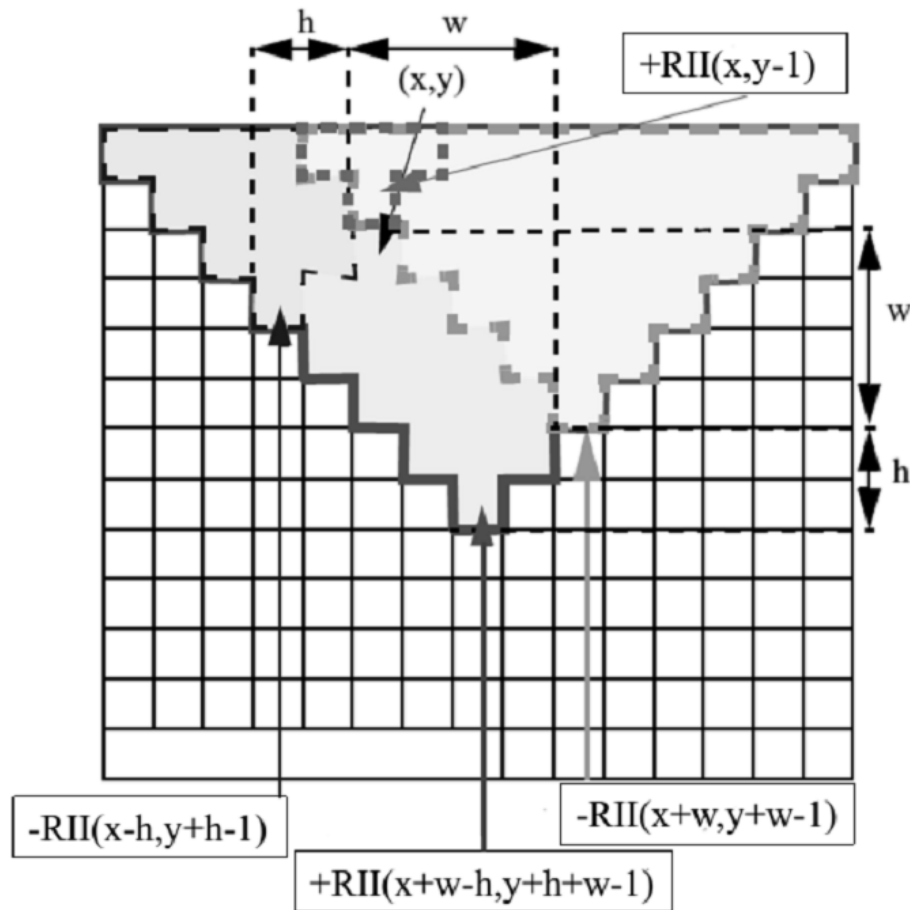


Figure 3.11: Calculation scheme for 45° rotated integral image.

$$\begin{aligned}
 RII(x, y) = & RII(x - 1, y - 1) - RII(x + 1, y - 1) - RII(x, y - 2) \\
 & + I(x, y) + I(x, y - 1)
 \end{aligned} \tag{3.11}$$

Một số phương pháp phát hiện đối tượng Haarlike và Adaboost



$RectangleSum(r, I)$

$$= RII(x + w - h, y + h + w - 1)$$

$$- RII(x + h, y + h - 1)$$

$$- RII(x + w, y + w - 1) + RII(x, y - 1)$$

More: Pham Minh Tri

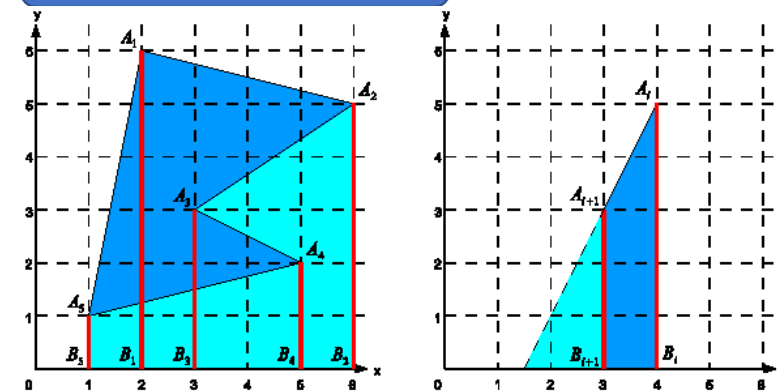
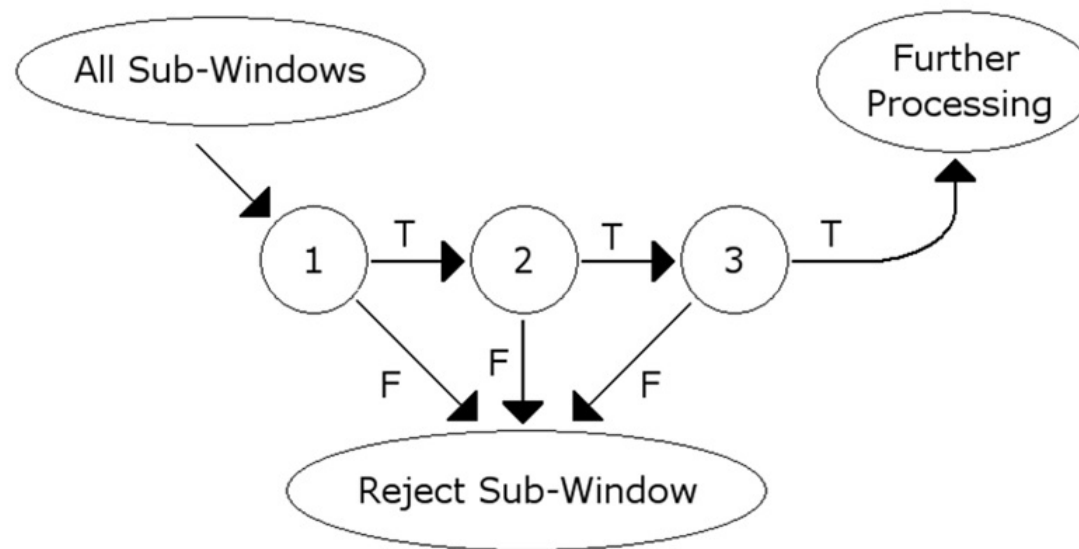


Figure 3.12: Calculation scheme for sum of 45° rotated rectangle.

Một số phương pháp phát hiện đối tượng Haarlike và Adaboost

- Boosting
- AdaBoost
- Cascade of AdaBoost classifiers



Một số phương pháp phát hiện đối tượng Haarlike và Adaboost

Algorithm 1: Boosting algorithm [84]

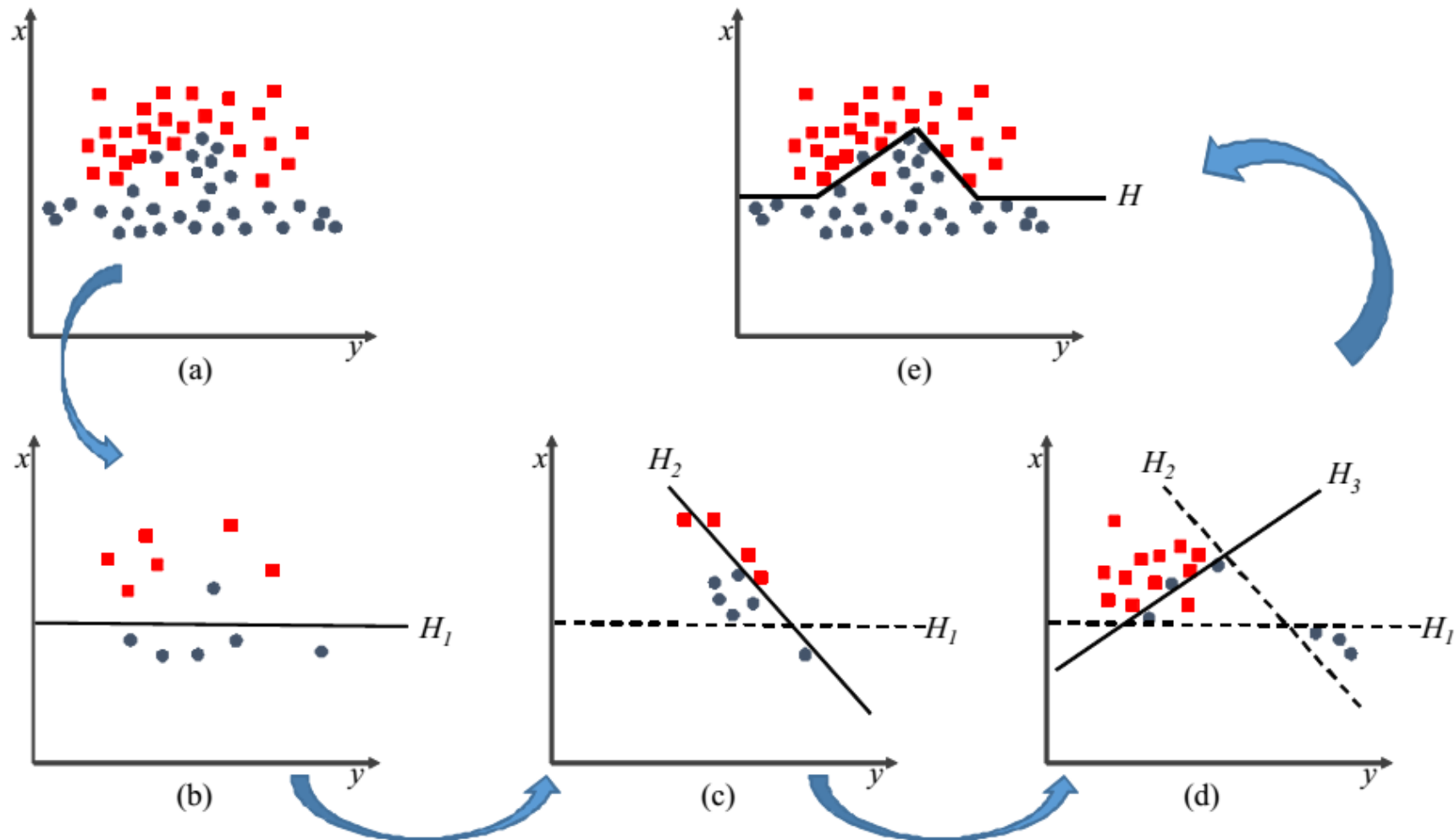
input : N examples $(x_1, y_1), \dots, (x_N, y_N)$ with $x \in \mathbb{R}^k, y_i \in \{-1, 1\}$

output: $H(x)$, a classifier suited for the training set

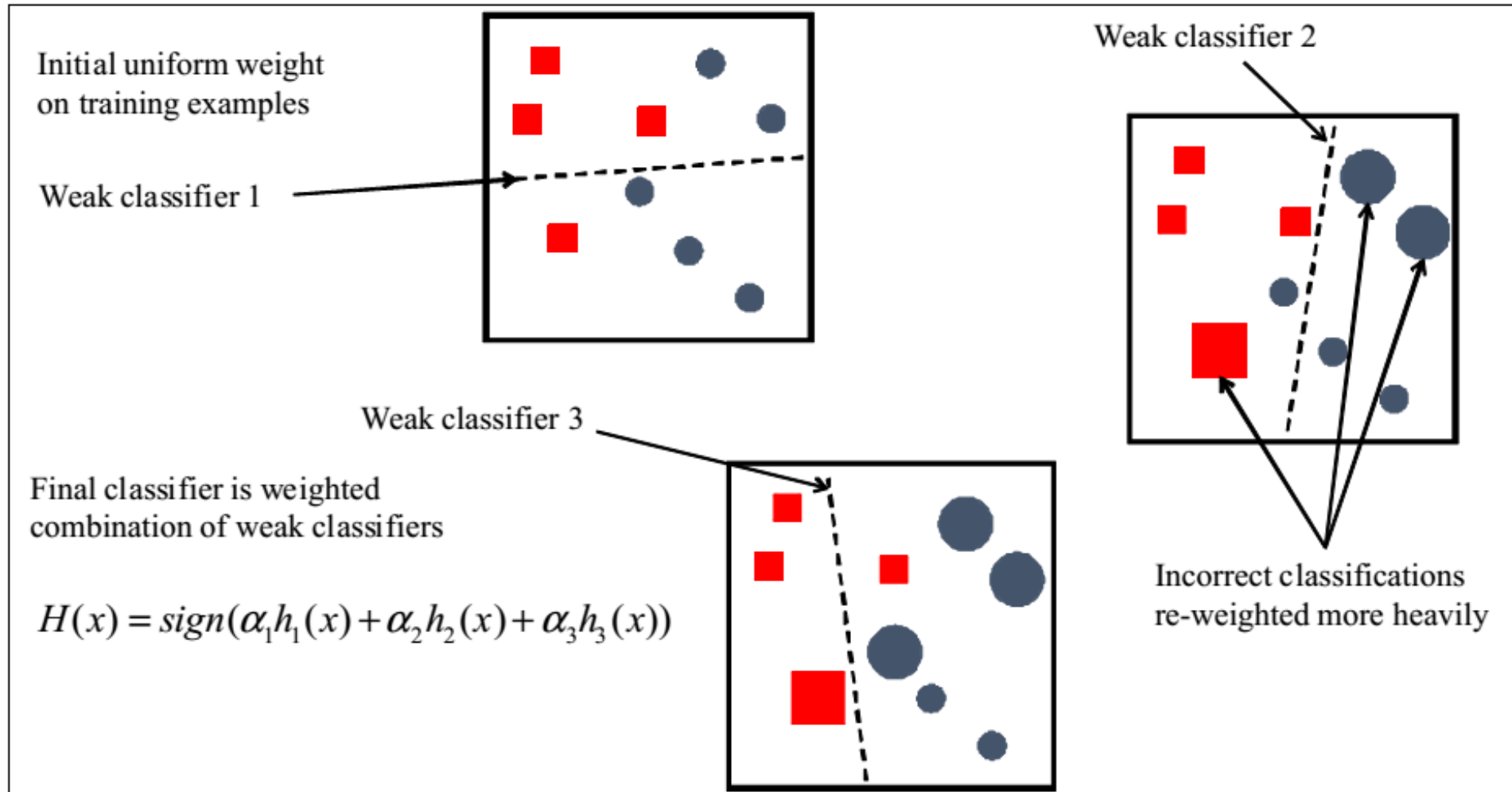
- 1 Randomly select, without replacement, $L_1 < N$ samples from Z to obtain Z_1 ; train weak learner H_1 on it.
- 2 Select $L_2 < N$ samples from Z with half of the samples misclassified by H_1 to obtain Z_2 ; train weak learner H_2 on it.
- 3 Select all samples from Z that H_1 and H_2 disagree on; train weak learner H_3 .
- 4 Produce final classifier as a vote of weak learners $H(x) = \text{sign} \left(\sum_{n=1}^3 H_n(x) \right)$.
- 5 **return** $H(x)$

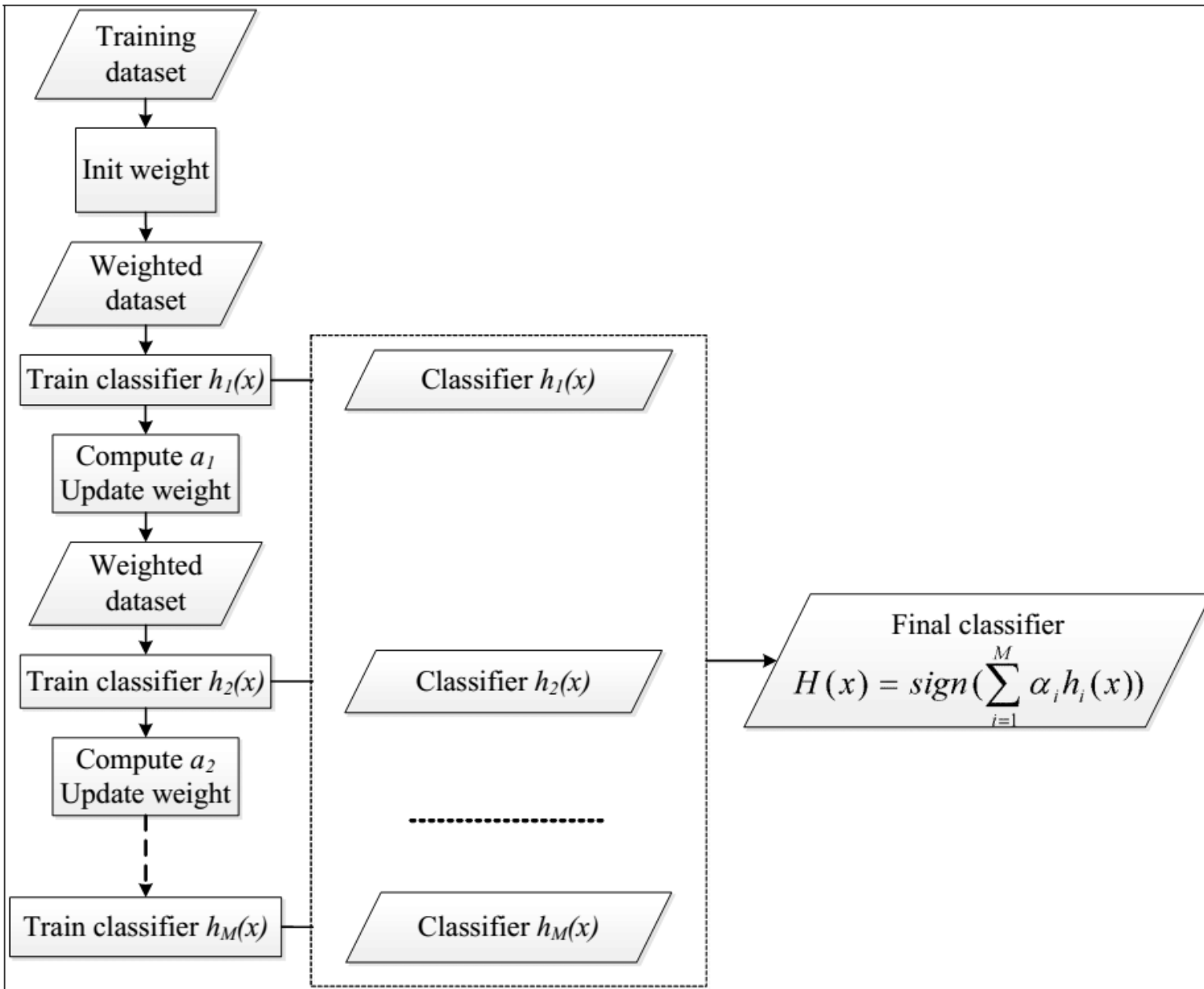
[84] Robert E Schapire. The strength of weak learnability. *Machine learning*, 5(2):197–227, 1990.

Một số phương pháp phát hiện đối tượng Haarlike và Adaboost



Một số phương pháp phát hiện đối tượng Haarlike và Adaboost





Freund, Yoav, and Robert E. Schapire. "A decision-theoretic generalization of on-line learning and an application to boosting." *Journal of computer and system sciences* 55.1 (1997): 119-139.

Freund, Yoav, and Robert E. Schapire. "Experiments with a new boosting algorithm." *lcm1*. Vol. 96. 1996.

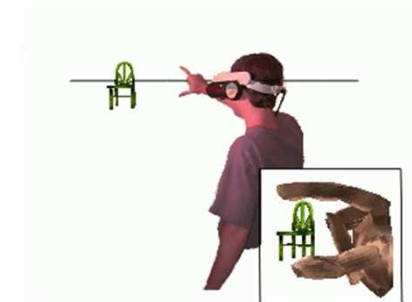
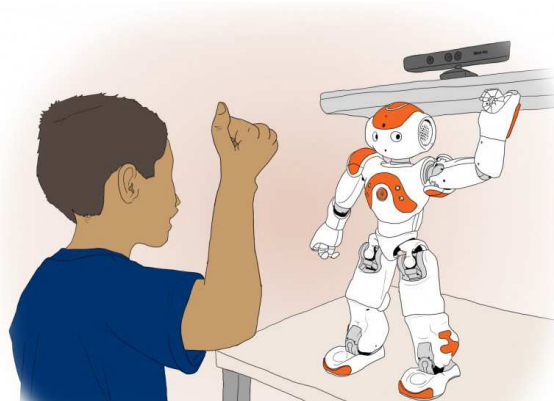
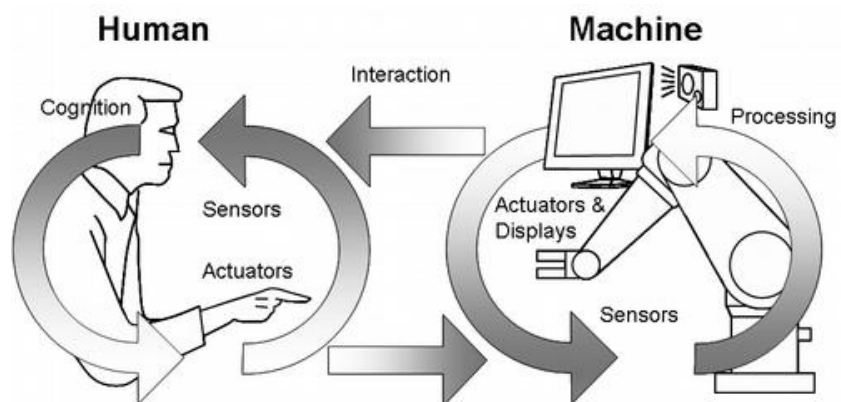
Thực hành

Các bài thực hành

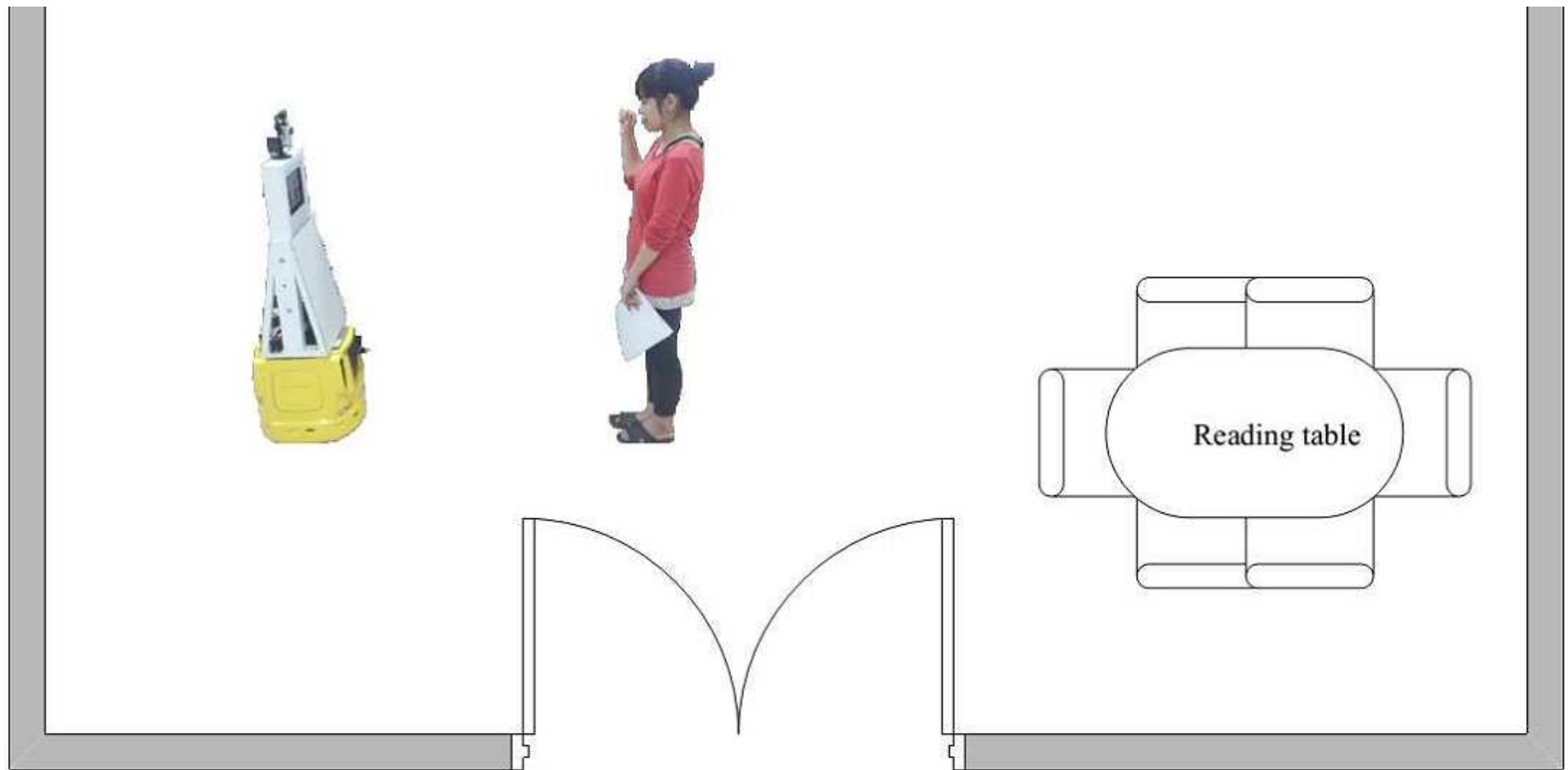
- Phát hiện tay trong hệ thống tương tác người-máy → thực hành tại lớp
- Phát hiện mặt người trong các hệ thống giám sát, kiểm soát vào ra → bài làm về nhà
- Phát hiện phương tiện giao thông trong các hệ thống giao thông thông minh

Giới thiệu bài toán phát hiện tay








- Các ứng dụng



Giới thiệu bài toán phát hiện tay



Giới thiệu bài toán phát hiện tay

Interface Controlling				Consultation		
Back	Next	End	Next Page	Open User Info	Open Book Info	Read Book Summary
						

Giới thiệu bài toán phát hiện tay



Thực hành

- Chuẩn bị dữ liệu
- Huấn luyện mô hình
- Sử dụng mô hình đã huấn luyện để phát hiện bàn tay
- Đánh giá hiệu năng của mô hình đã huấn luyện

Tài liệu tham khảo

- [1] Andrew Zisserman, Object Recognition and Localization, ICVSS 2008.
- (Dalal, 2005) Dalal, N. and Triggs, B, Histograms of Oriented Gradients for Human Detection, Conference on Computer Vision and Pattern Recognition (CVPR'05)

Bài tập làm thêm

- Thử nghiệm và so sánh hiệu quả của Haar và HOG trên cùng cơ sở dữ liệu
- Đọc thêm:
 - Salient Object Detection by Composition, ICCV 2011
- Xem thêm:
 - Bài giảng về HOG: <https://www.youtube.com/watch?v=0Zib1YEE4LU>

THỊ GIÁC MÁY TÍNH

AI Academy Vietnam

CẢM ƠN!