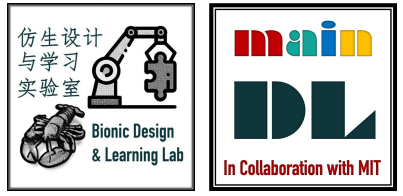
Histograms of Oriented Gradients for Human Detection

Presenter: Qirong Shen, Mingyu Zhou, Zimeng Wang, Qian Liu, Junzhe Wen



AncoraSIR.com

SUSTech
Southern University
of Science and Technology

Motivation and Main Problem

Further explaination of the title with supporting evidence

High-level description of problem being solved

- Find a robust visual objection recogition feature sets and raised the Histograms of Oriented Gradients (HOG) descriptors

- Researchers showed that the HOG descriptors is more significant than existing feature sets of human detection experimentally

- Adopting linear SVM based human detection as a test case

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Motivation and Main Problem

Further explaination of the title with supporting evidence

Why is the problem important?

Its significance towards general-purpose robot autonomy

- The new approach that HOG+SVM become the foundation of the fellowing algorithem for its success in human detection

Its potential application and societal impact of the problem

- Its high sensitivity to the outline and shape of the object appearances a good ability to recognise the human, vehicle and road marking. This is significant for the automatic drive.

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Motivation and Main Problem

Further explaination of the title with supporting evidence

The role of the AI and machine learning in tackling this problem

- Support Vector Machine (SVM) is used in the case of human detedction. The last step is to train the classifier, use SVM to train the previously extracted image feature vector from HOG than find an optimal hyperplane as the decision function, and get the target training model.

- Good use of linear SVM as a baseline classifier make the study efficient and simple

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Limitations of Prior Work

Further explaination of the title with supporting evidence

The paper's related work is a good start, but there may be others

- There are a few relevant papers on human detection

What is the key limitations of prior work(s)?

- They use more complex achitecture with mutiple detection windows. in contrast, HOG algorithem uses a simpler architecture with a single detection window, but appears to give significantly higher performance on pedestrian images

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Implementation

Gamma Transformation

$$Y(x,y) = I(x,y)^{\gamma}$$

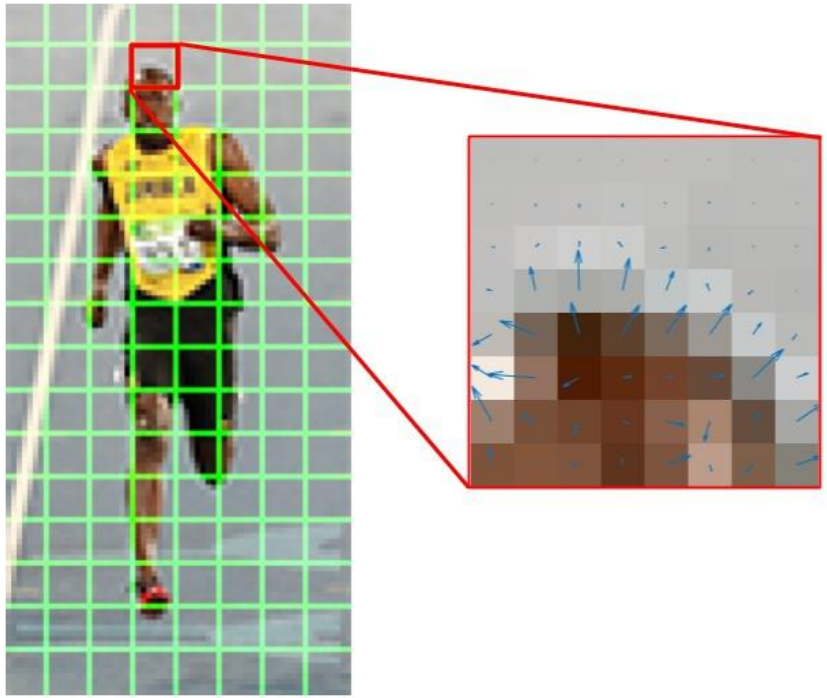
Reduce effect of difficult illumination

AncoraSIR.com

Implementation

Gradient

- Calculate the gradient and 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

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	116	17	138	110	

Gradient Direction

$$G_x(x,y) = I(x+1,y) - I(x-1,y)$$

$$G_y(x,y) = I(x,y+1) - I(x,y-1)$$

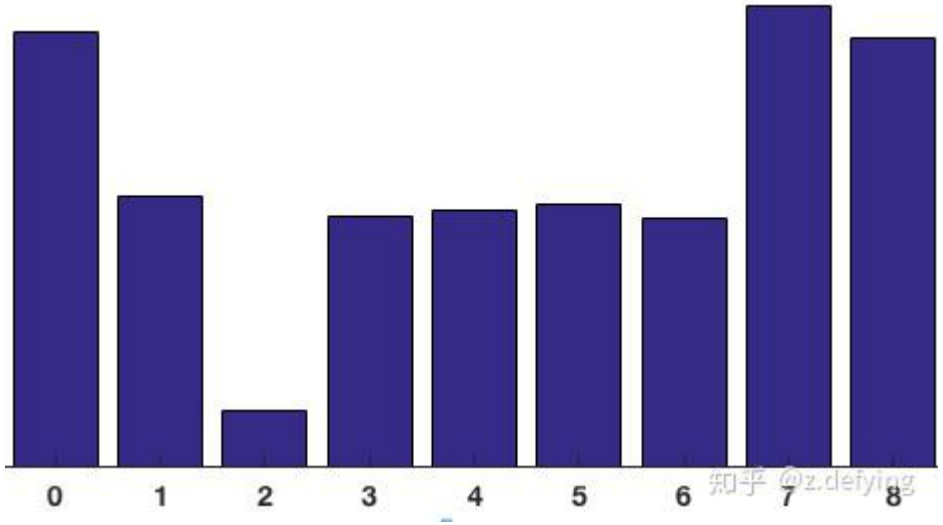
$$G(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2}$$

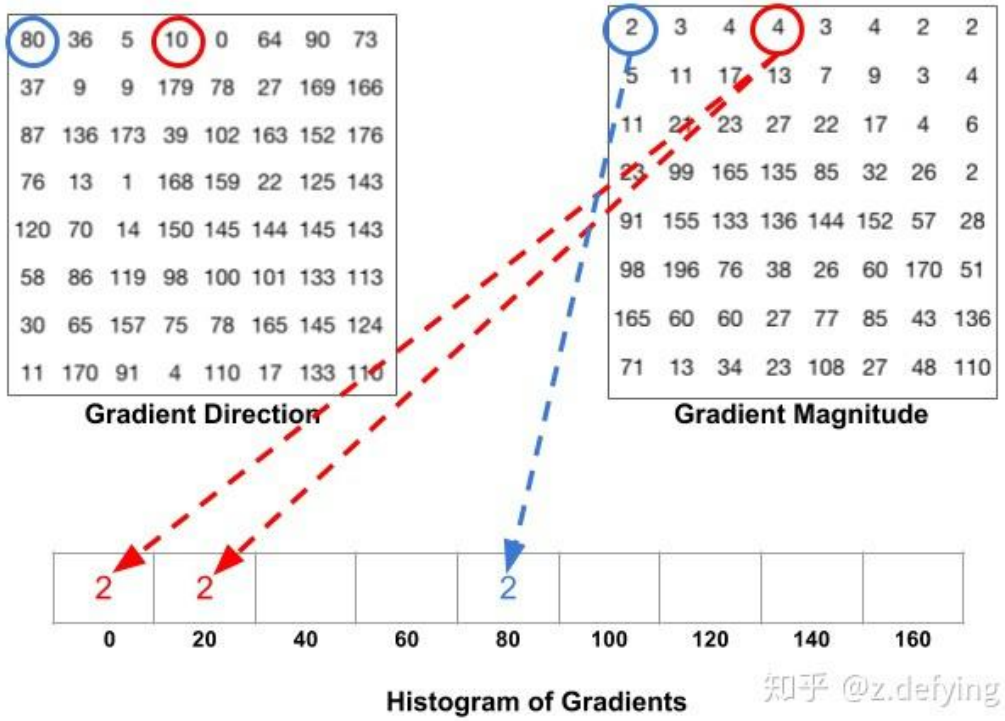
$$\theta(x,y) = arctan\left(\frac{G_y(x,y)}{G_x(x,y)}\right)$$

AncoraSIR.com

Implementation

Histogram

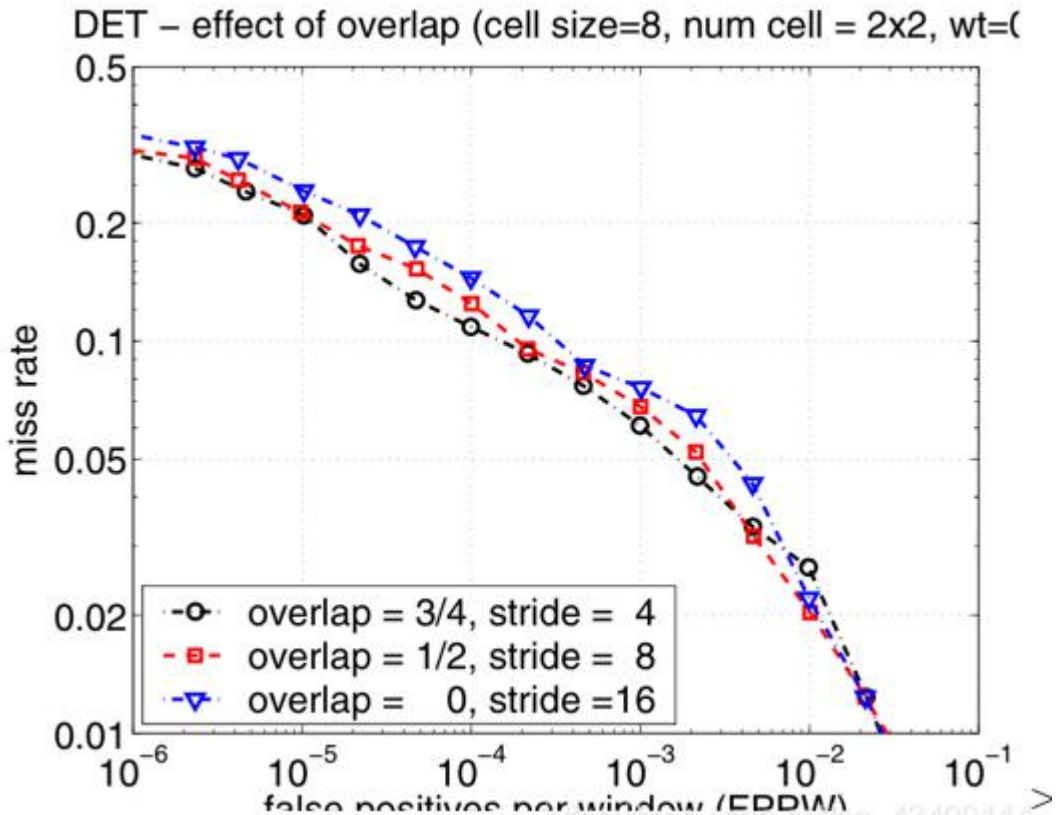




AncoraSIR.com

Implementation

Normalization

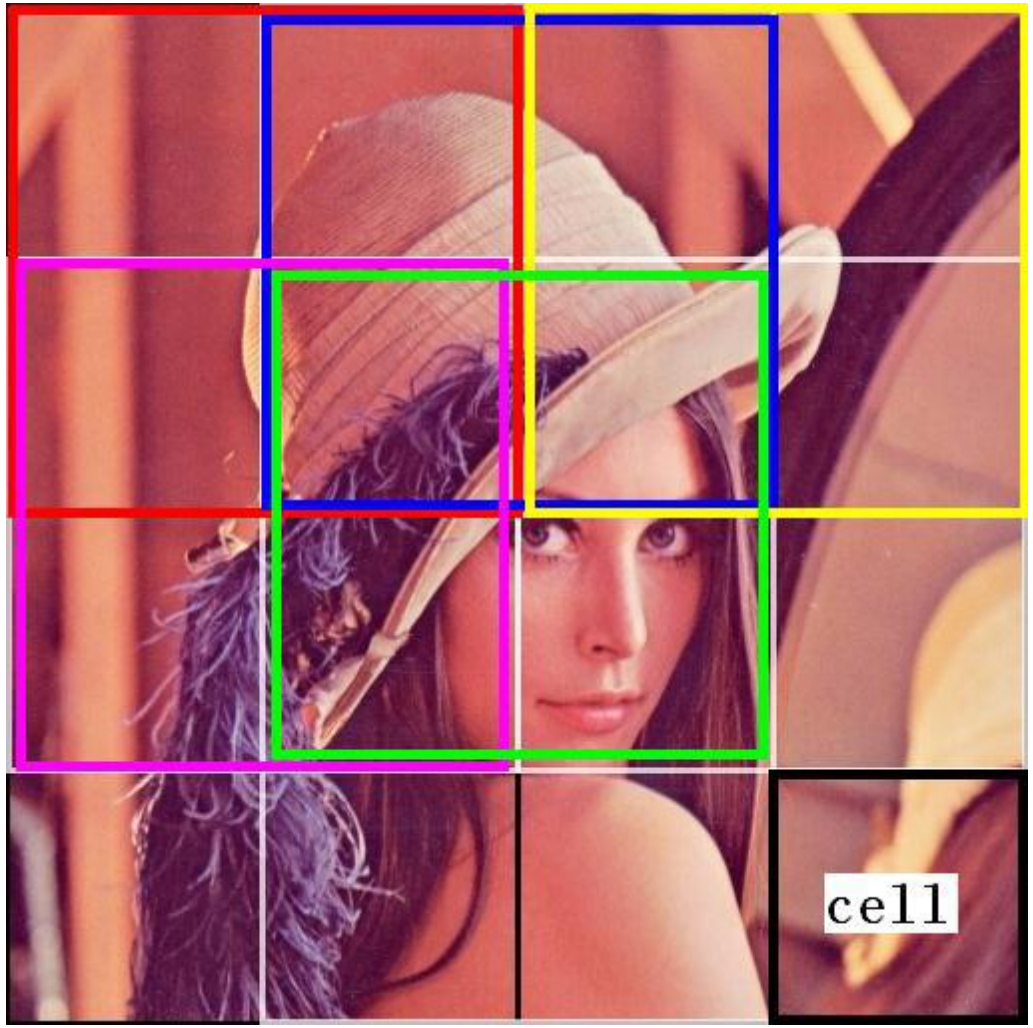


The schemes are: (a) *L2-norm*, $\mathbf{v} \rightarrow \mathbf{v}/\sqrt{\|\mathbf{v}\|_2^2 + \epsilon^2}$; (b) *L2-Hys*, L2-norm followed by clipping (limiting the maximum values of $\mathbf{v}$ to 0.2) and renormalizing, as in [12]; (c) *L1-norm*, $\mathbf{v} \rightarrow \mathbf{v}/(\|\mathbf{v}\|_1 + \epsilon)$; and (d) *L1-sqrt*, L1-norm followed by square root $\mathbf{v} \rightarrow \sqrt{\mathbf{v}/(\|\mathbf{v}\|_1 + \epsilon)}$, which amounts to treating the descriptor vectors as probability distributions and using the Bhattacharya distance between them. Fig. 4(c)

AncoraSIR.com

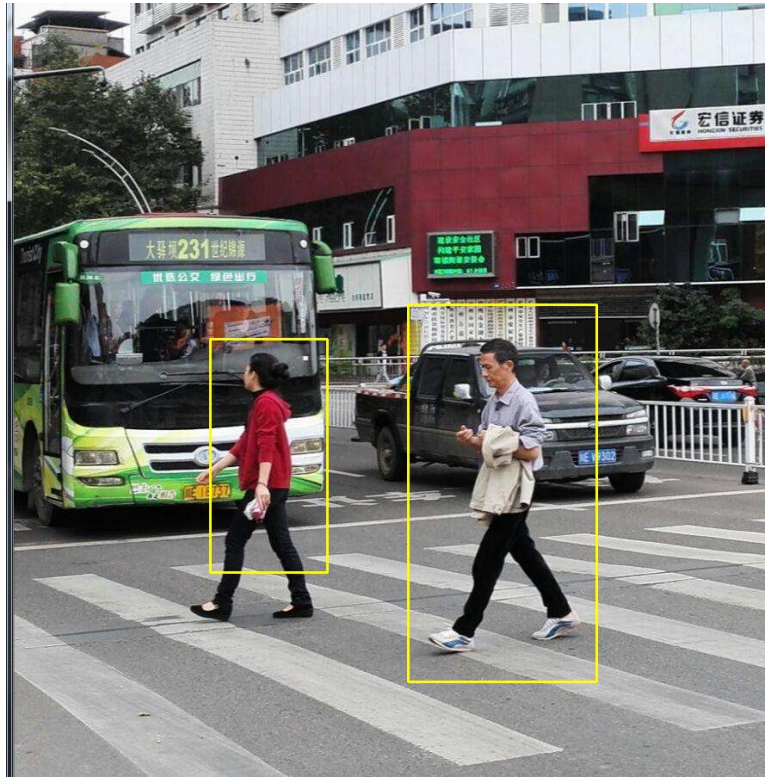
Implementation

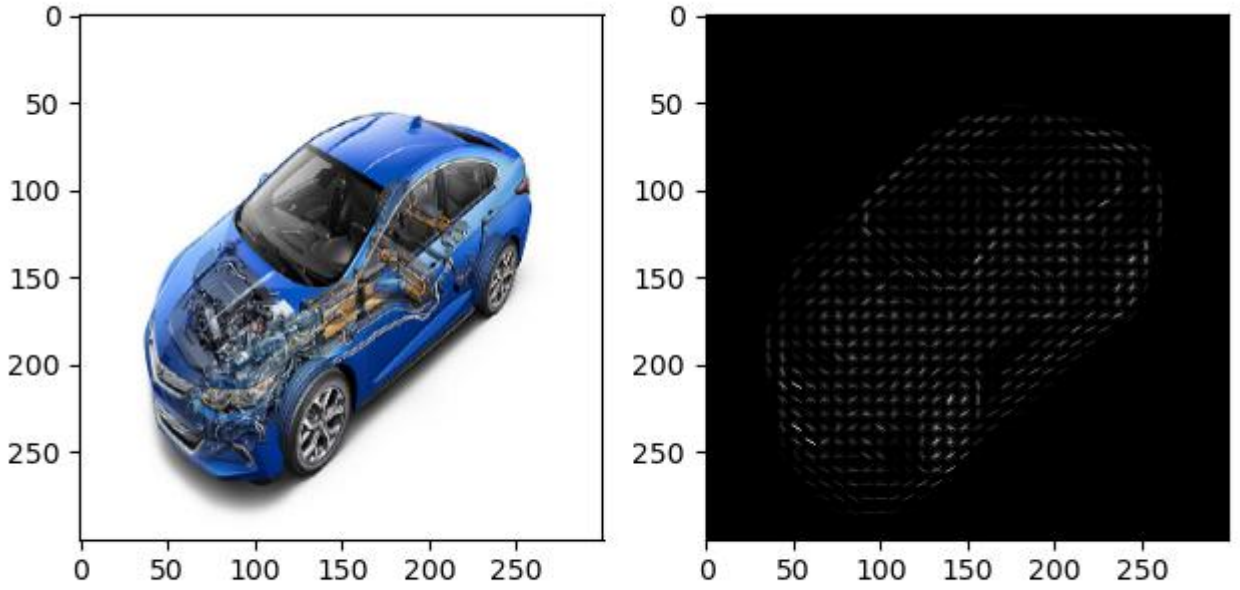
Feature Detection + SVM



AncoraSIR.com

Result





AncoraSIR.com

Experimental Setup

Further explaination of the title with supporting evidence

Datasets

A well-established MIT pedestrian database

- containing 509 training and 200 test images of pedestrians in city scenes

A new dataset‘INRIA’

- significantly more challenging
- containing 1805 images of human from a varied set of personal photos
- standing people r in any orientation and against a wide
- variety of background image including crowds.

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Experimental Setup

Further explaination of the title with supporting evidence

Methodology

Positive training sets

- 1239 of the images with their left-right reflections (2478 images in all)

Negative training sets

- A fixed set of 12180 patches sampled randomly from 1218 person-free training photos

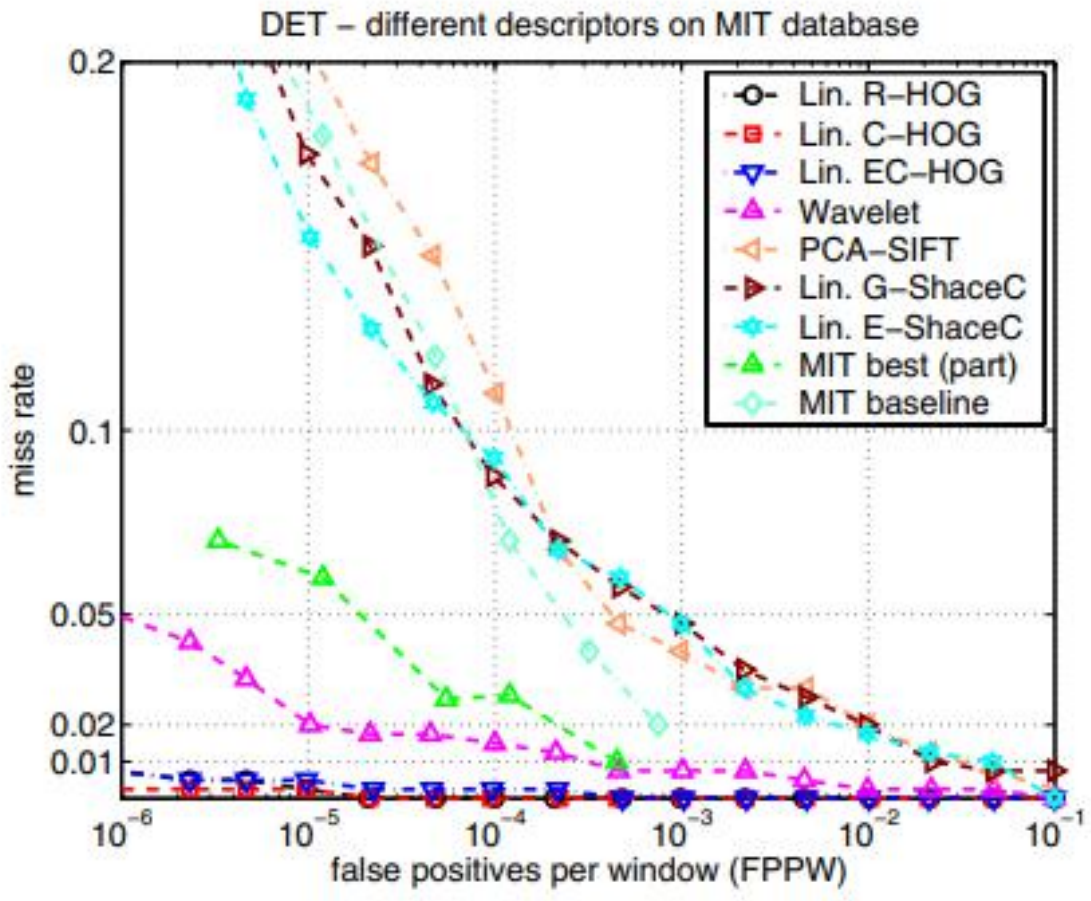
‘hard examples’+inintial 12180sets to re-traine

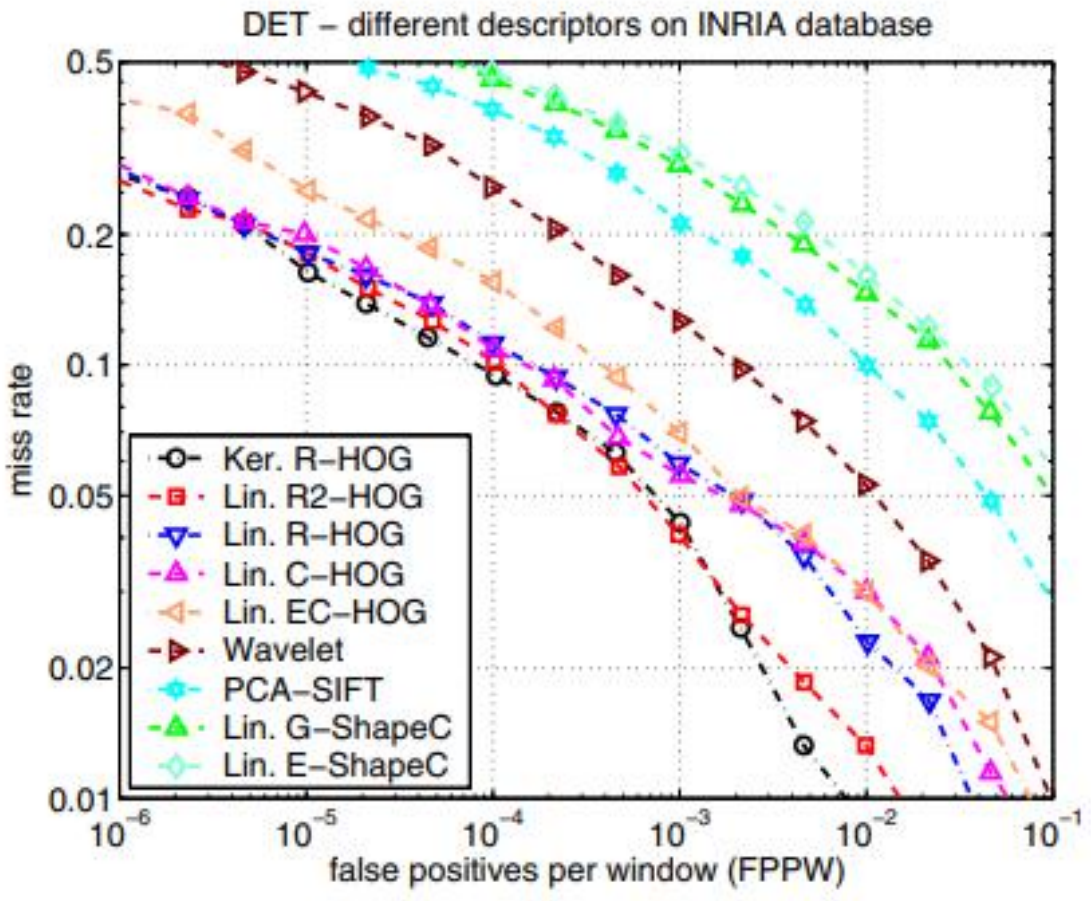
AncoraSIR.com

SUSTech
Southern University of Science and Technology

Experimental Results

Further explaination of the title with supporting evidence





AncoraSIR.com

Discussion of Results

Further explaination of the title with supporting evidence

HOG-based detectors showed near-perfect oseparation on the MIT test set

HOG-based detectorsat gived least an order of magnitude reduction in FPPW (false positives per window) on the INRIA test set

AncoraSIR.com

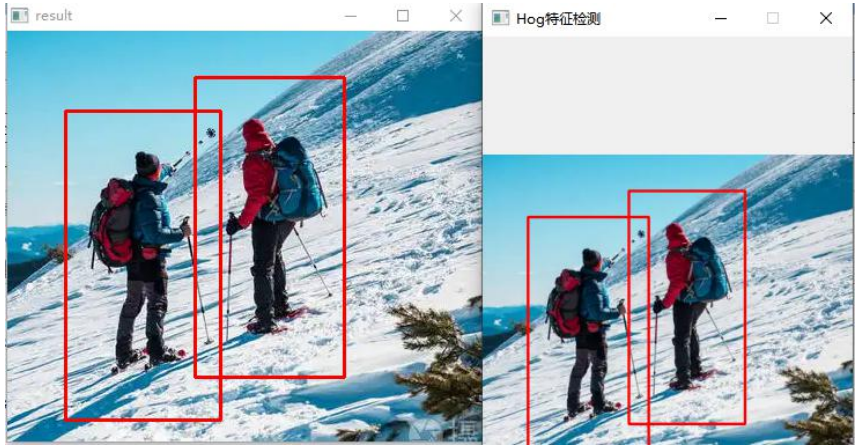
SUSTech
Southern University of Science and Technology

Future Work for Paper / Reading

Further explanation of the title with supporting evidence

- OpenCV library provides a built-in implementation of the HOG algorithm called *cv2.HOGDescriptor*

This implementation follows the same principle as the HOG algorithm proposed by Dalal and Triggs and provides an efficient and easy-to-use way to apply the HOG algorithm in pedestrian detection, face detection, and other applications



AncoraSIR.com

Extended Readings

Further explanation of the title with supporting evidence

K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN"

a. This paper introduces a novel network architecture based on the Region-based Convolutional Neural Network (R-CNN) framework that can simultaneously perform object detection and instance segmentation.

b. The significance of this paper lies in proposing an efficient and accurate method for object detection and instance segmentation, which can be widely applied in computer vision fields, including autonomous driving, medical image processing, and robotics.

c. This paper is relevant to our research as it introduces a new deep learning technique that can be combined with HOG features and SVM classifiers to improve the accuracy of human detection.

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Extended Readings

Further explanation of the title with supporting evidence

A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection"

a. This paper presents a novel object detection framework, You Only Look Once (YOLO) v4, which combines various techniques such as network architecture design, data augmentation, and post-processing to improve the speed and accuracy of object detection.

b. The significance of this paper lies in proposing an efficient and accurate method for object detection, which can be applied in fields such as autonomous driving, security monitoring, and robotics.

c. This paper is relevant to our research as it provides a new object detection technique that can replace the HOG features and SVM classifiers approach to improve the accuracy and speed of human detection.

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Summary

Further explaination of the title with supporting evidence

Main Problem

The article proposes a method for human detection using histograms of oriented gradients (HOG) and support vector machines (SVMs).

Innovation of methods

(a) the use of local image gradients to encode human appearance

(b) the use of SVMs for efficient and accurate classification of image patches

Apply

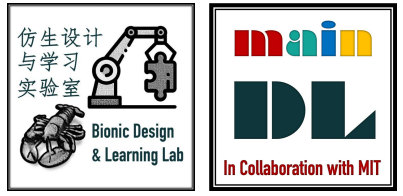
HOG-based method achieves state-of-the-art performance in human detection

AncoraSIR.com

SUSTech
Southern University of Science and Technology

Thanks for listening

Name:Qirong Shen, Mingyu Zhou, Zimeng Wang, Qian Liu, Junzhe Wen
Affiliation:SUSTCH
Supervisor：Chaoyang Son
Contact:12010311@mail.sustech.edu.cn



AncoraSIR.com

SUSTech
Southern University
of Science and Technology