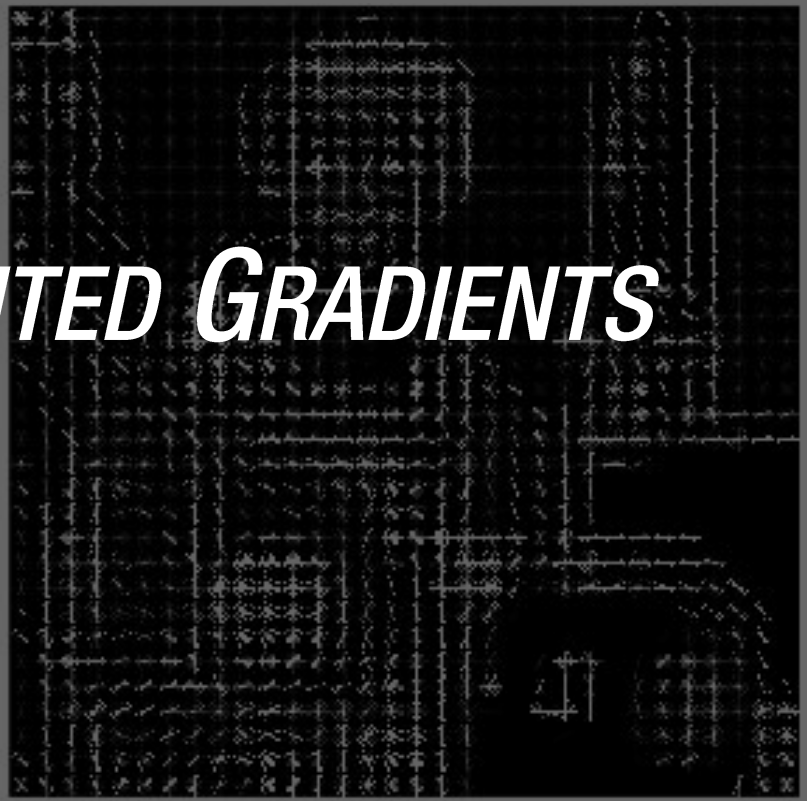
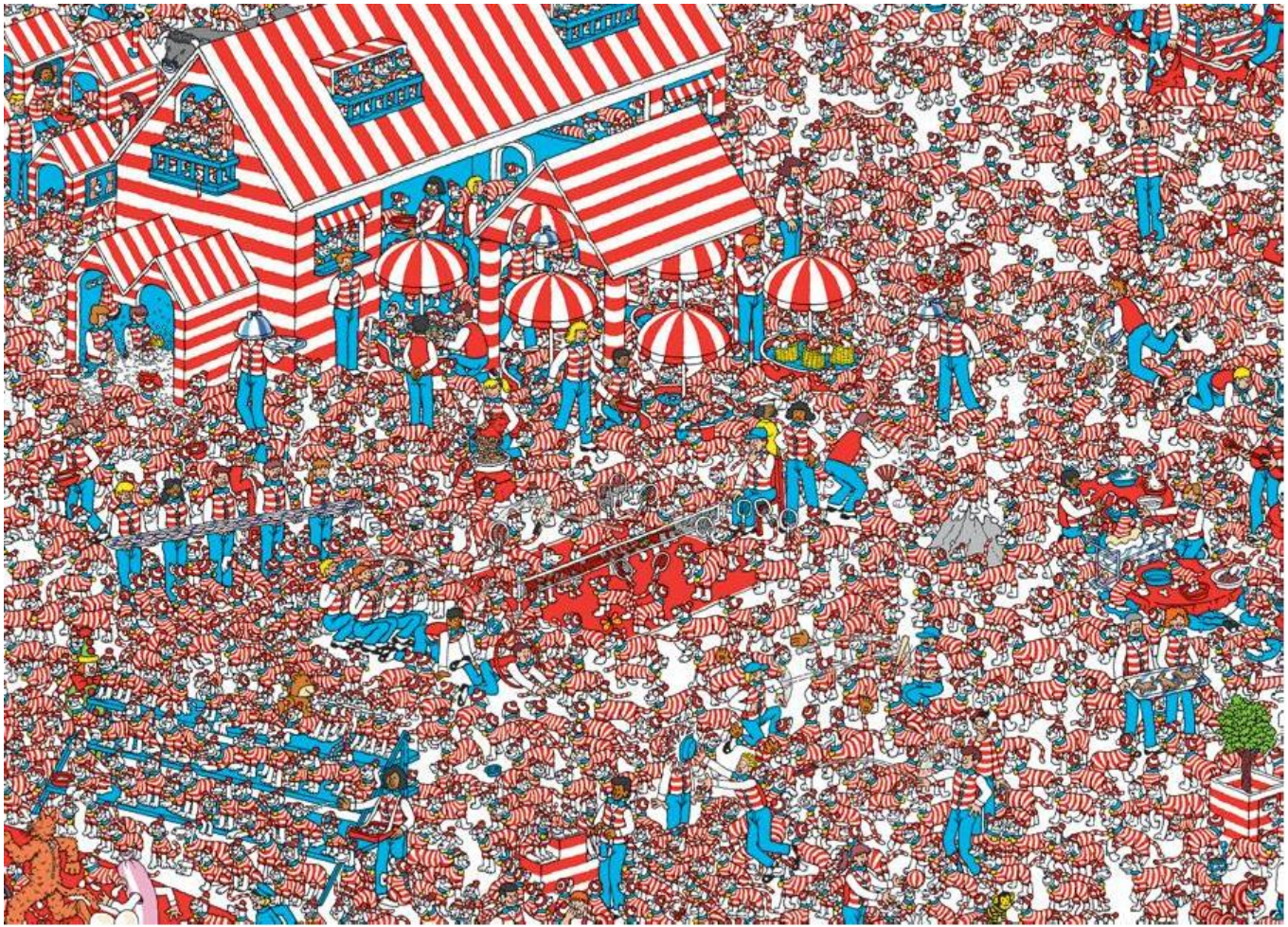
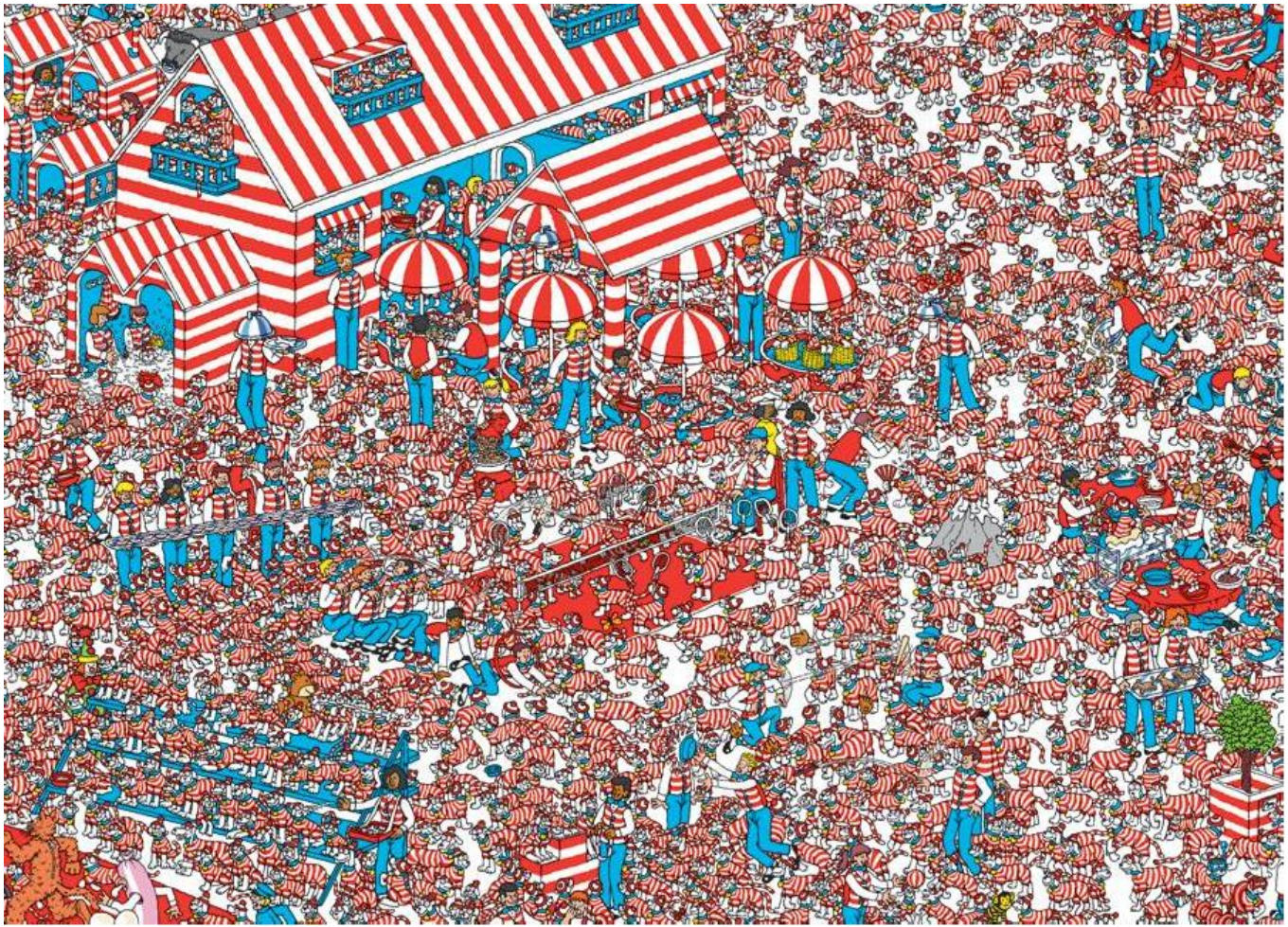
Histogram of Oriented Gradients

Hyun Soo Park





20	35	21
53	22	23
22	34	32
22	56	55
45	11	13



20	35	21
53	22	23
22	34	32
22	56	55
45	11	13

$$J = \begin{bmatrix} 20 \\ 53 \\ 22 \\ 22 \\ 45 \\ \cdot \\ \cdot \\ 55 \\ 13 \end{bmatrix}$$

11	13	34
34	24	21
64	55	64
68	45	25
80	31	45

$$I = \begin{bmatrix} 11 \\ 34 \\ 64 \\ 68 \\ 80 \\ \cdot \\ \cdot \\ 25 \\ 45 \end{bmatrix}$$

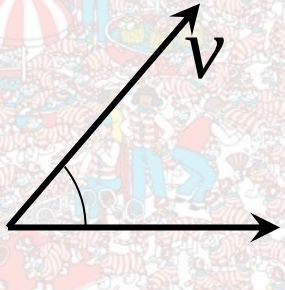
20	35	21
53	22	23
22	34	32
22	56	55
45	11	13

$$J = \begin{bmatrix} 20 \\ 53 \\ 22 \\ 22 \\ 45 \\ \cdot \\ \cdot \\ 55 \\ 13 \end{bmatrix}$$

11	13	34
34	24	21
64	55	64
68	45	25
80	31	45

Correlation == Dot Product

$$I = \begin{bmatrix} 11 \\ 34 \\ 64 \\ 68 \\ 80 \\ \cdot \\ \cdot \\ 25 \\ 45 \end{bmatrix}$$



$$\frac{u \cdot v}{\|u\|\|v\|} = \cos\theta$$

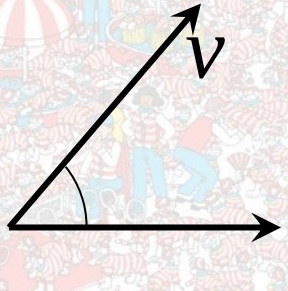
20	35	21
53	22	23
22	34	32
22	56	55
45	11	13

$$J = \begin{bmatrix} 20 \\ 53 \\ 22 \\ 22 \\ 45 \\ \cdot \\ \cdot \\ 55 \\ 13 \end{bmatrix}$$

11	13	34
34	24	21
64	55	64
68	45	25
80	31	45

CORRELATION == DOT PRODUCT

$$I = \begin{bmatrix} 11 \\ 34 \\ 64 \\ 68 \\ 80 \\ \cdot \\ \cdot \\ 25 \\ 45 \end{bmatrix}$$



$$\frac{u \cdot v}{\|u\|\|v\|} = \cos\theta$$

20	35	21
53	22	23
22	34	32
22	56	55
45	11	13

$$J = \begin{bmatrix} 20 \\ 53 \\ 22 \\ 22 \\ 45 \\ \cdot \\ \cdot \\ 55 \\ 13 \end{bmatrix}$$

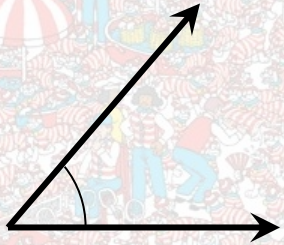
11	13	34
34	24	21
64	55	64
68	45	25
80	31	45

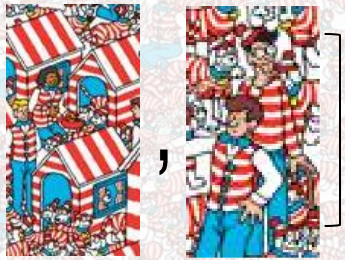
CORRELATION == DOT PRODUCT

20	35	21
53	22	23
22	34	32
22	56	55
45	11	13

$I =$

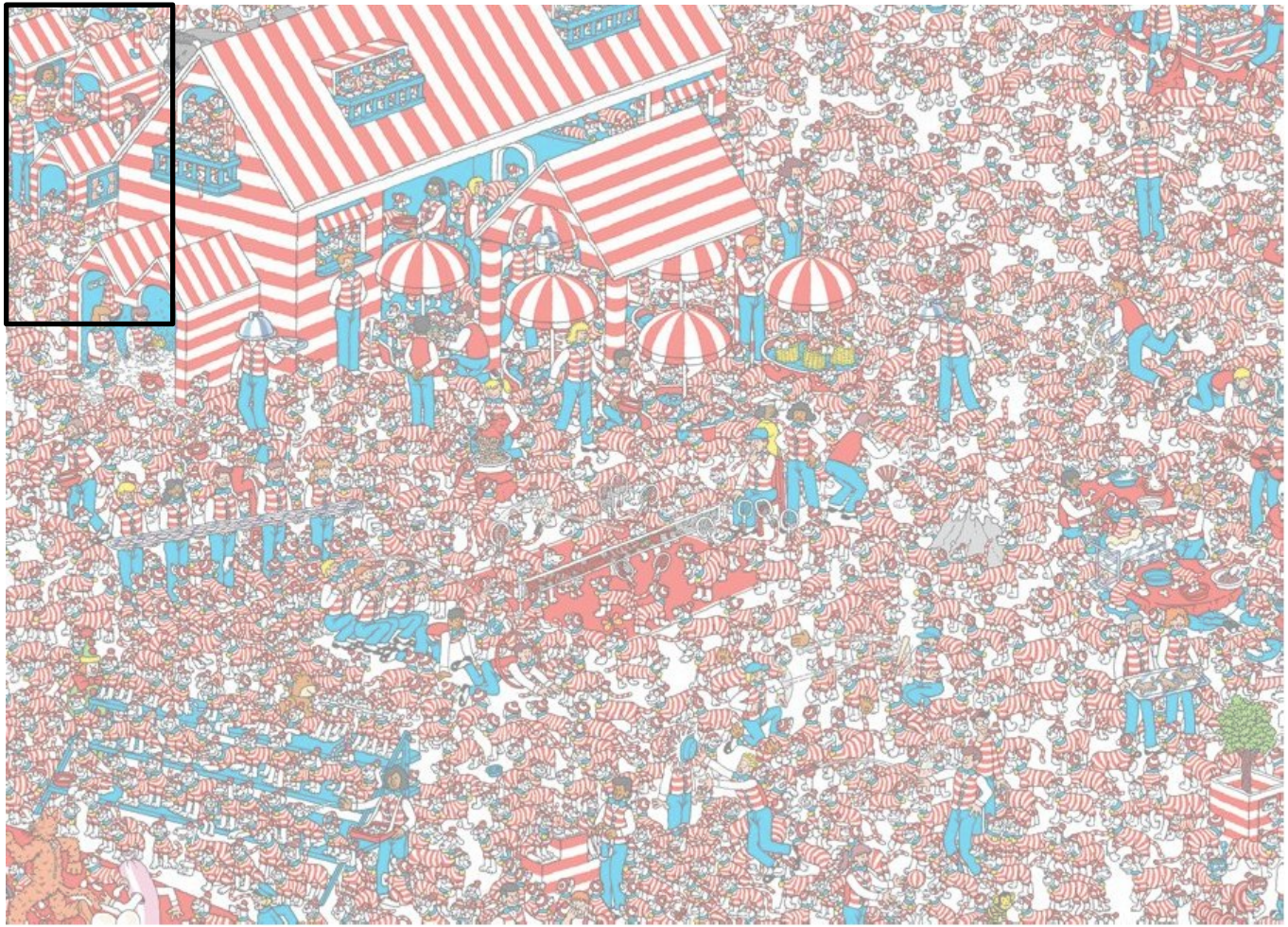
11
34
64
68
80
.
.
25
45

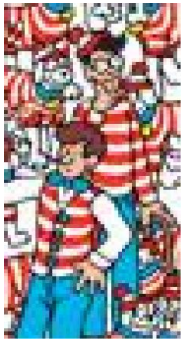


$Corr$ 

$J =$

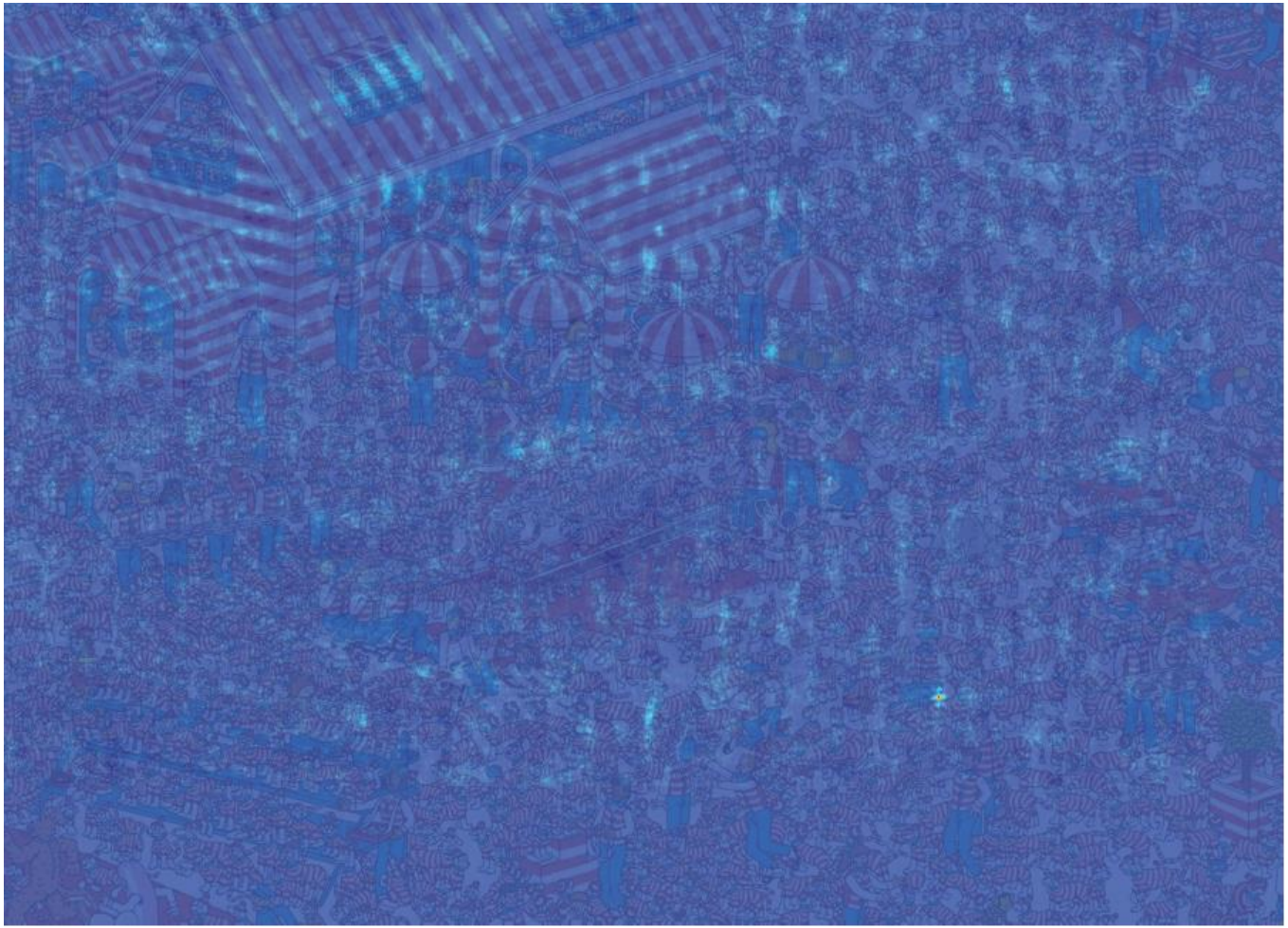
20
53
22
22
45
.
.
55
13

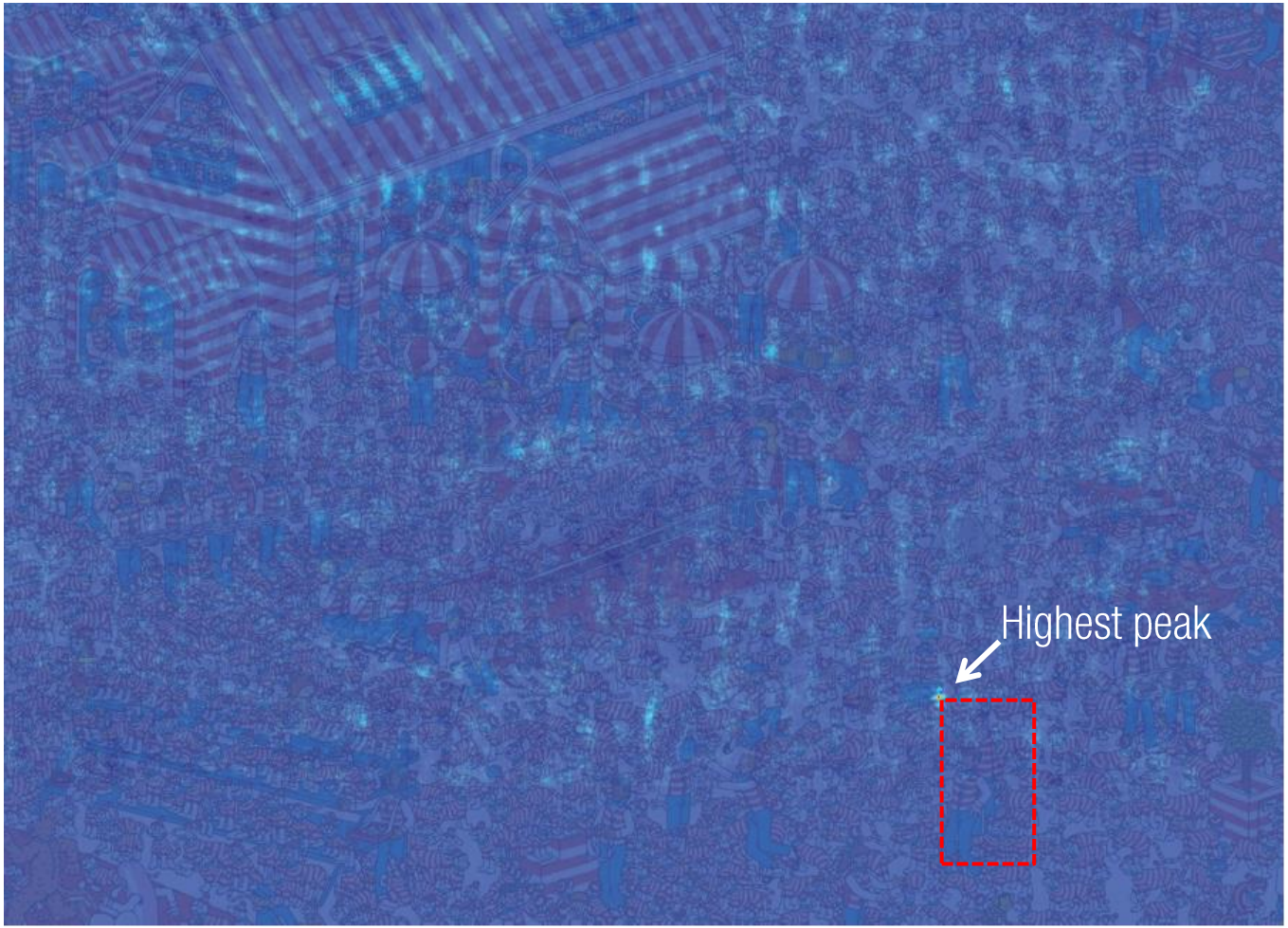


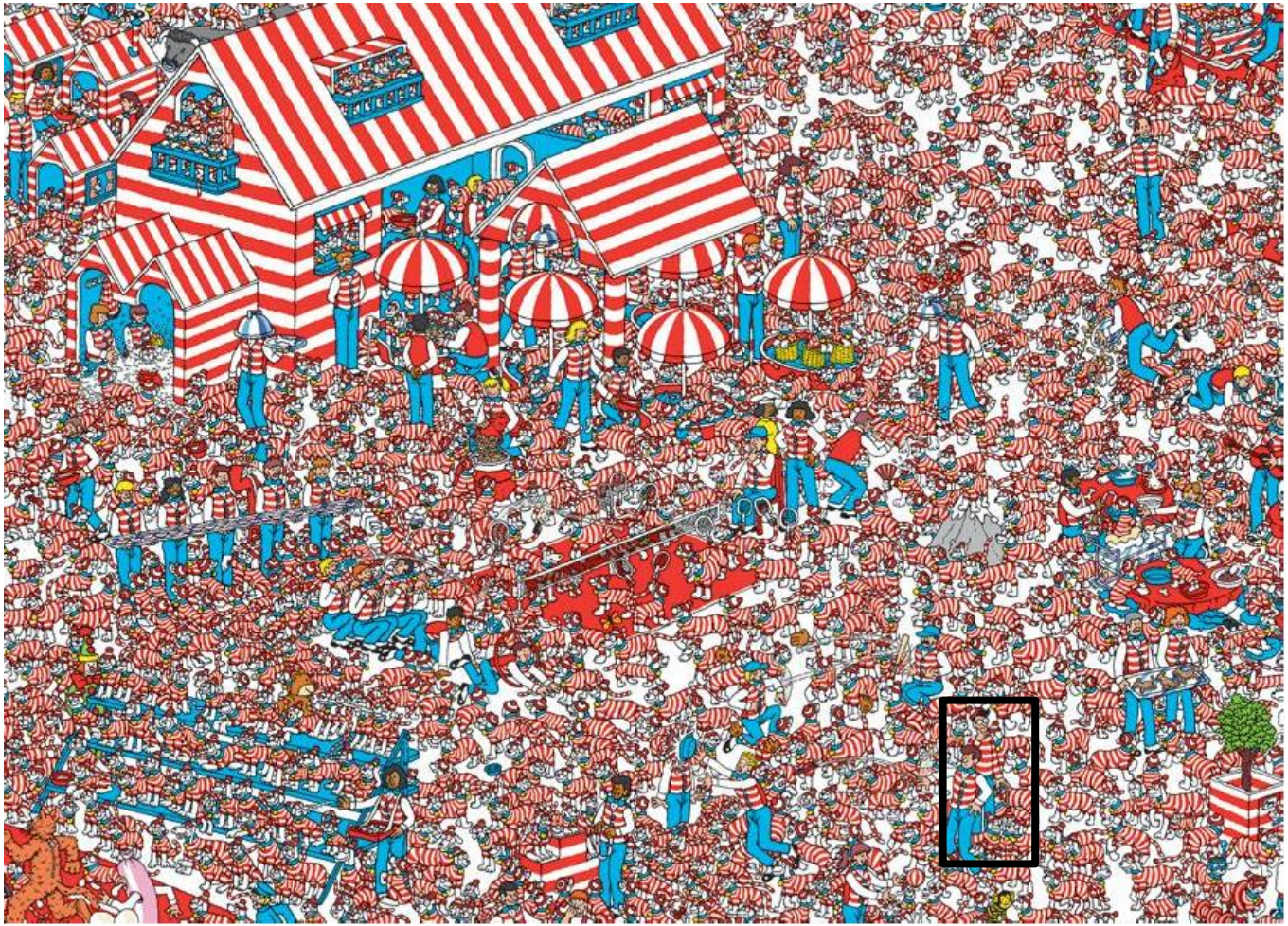


20	35	21
53	22	23
22	34	32
22	56	55
45	11	13

$$J = \begin{bmatrix} 20 \\ 53 \\ 22 \\ 22 \\ 45 \\ \cdot \\ \cdot \\ 55 \\ 13 \end{bmatrix}$$




Highest peak

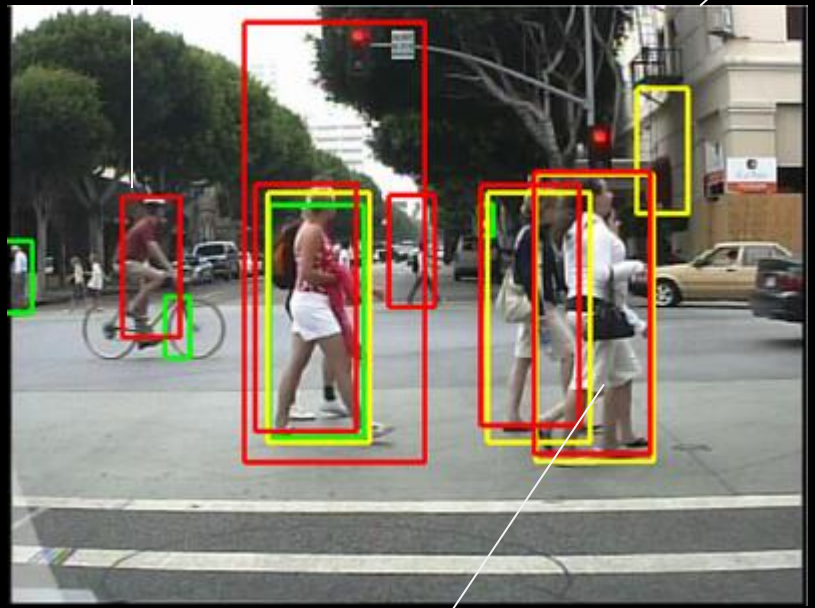


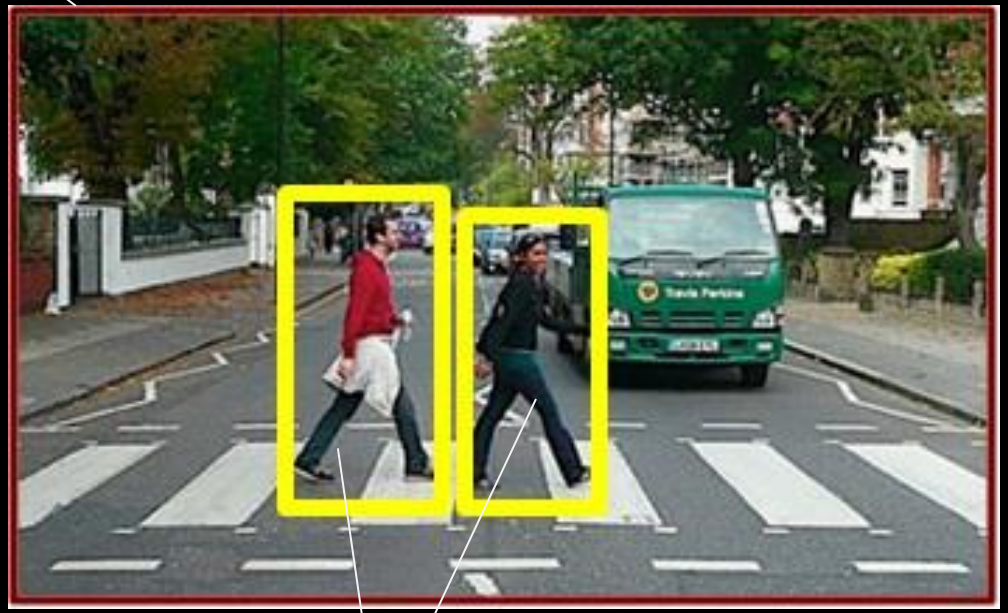
Assumptions:

+ Exactly same shape (no occlusion, no illumination change)
+ Exactly same size
+ Exactly same orientation
+ Exactly same viewpoint

Different illumination

Views



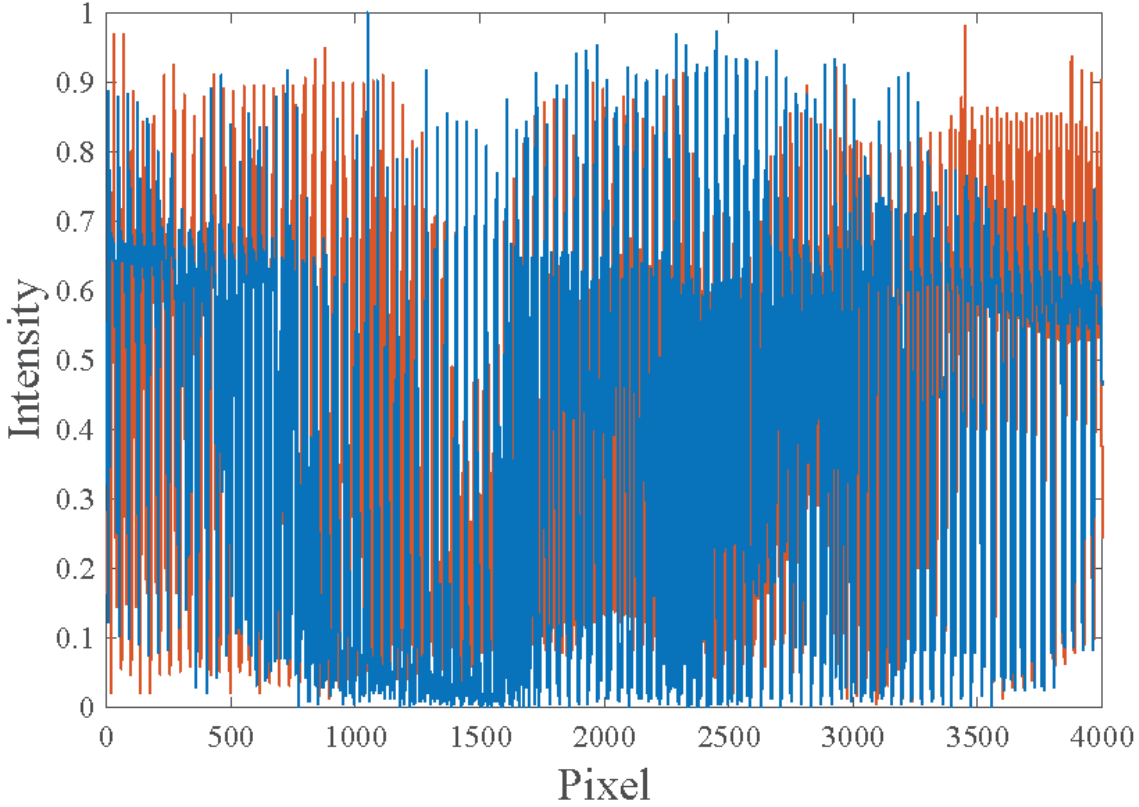


Occlussion

Various poses

Pixel Intensity Correlation

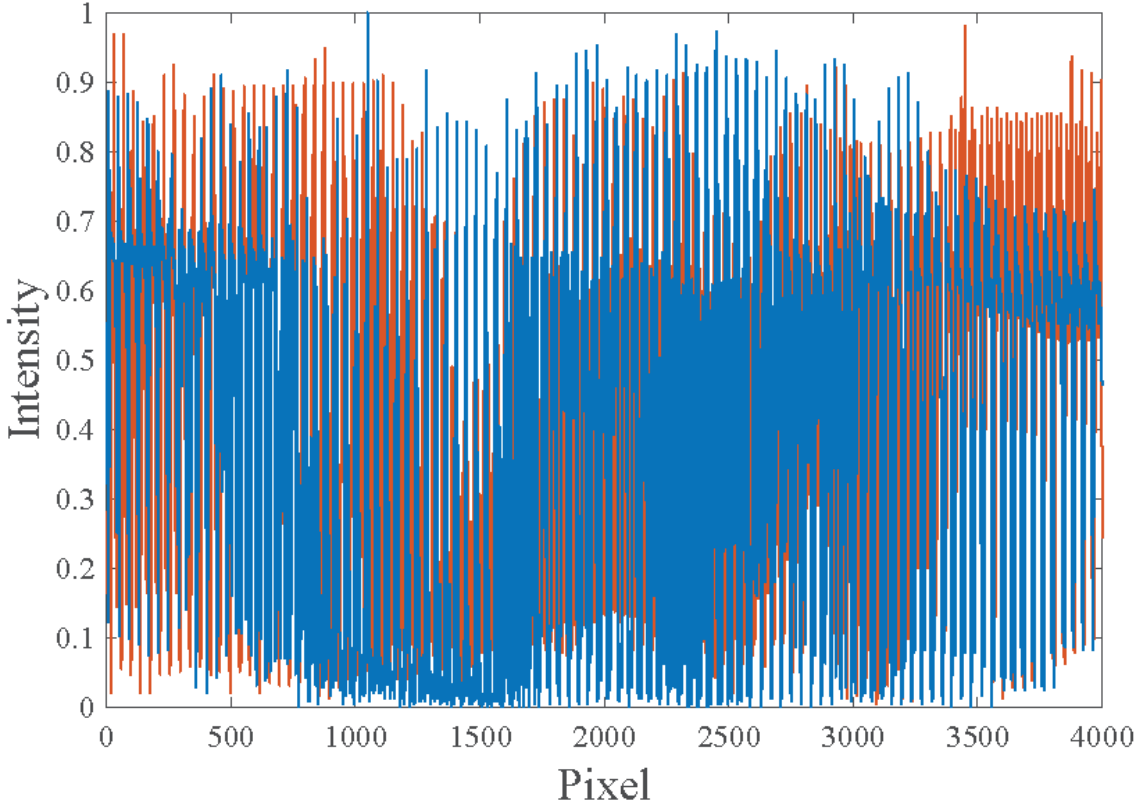
Corr = 0.24



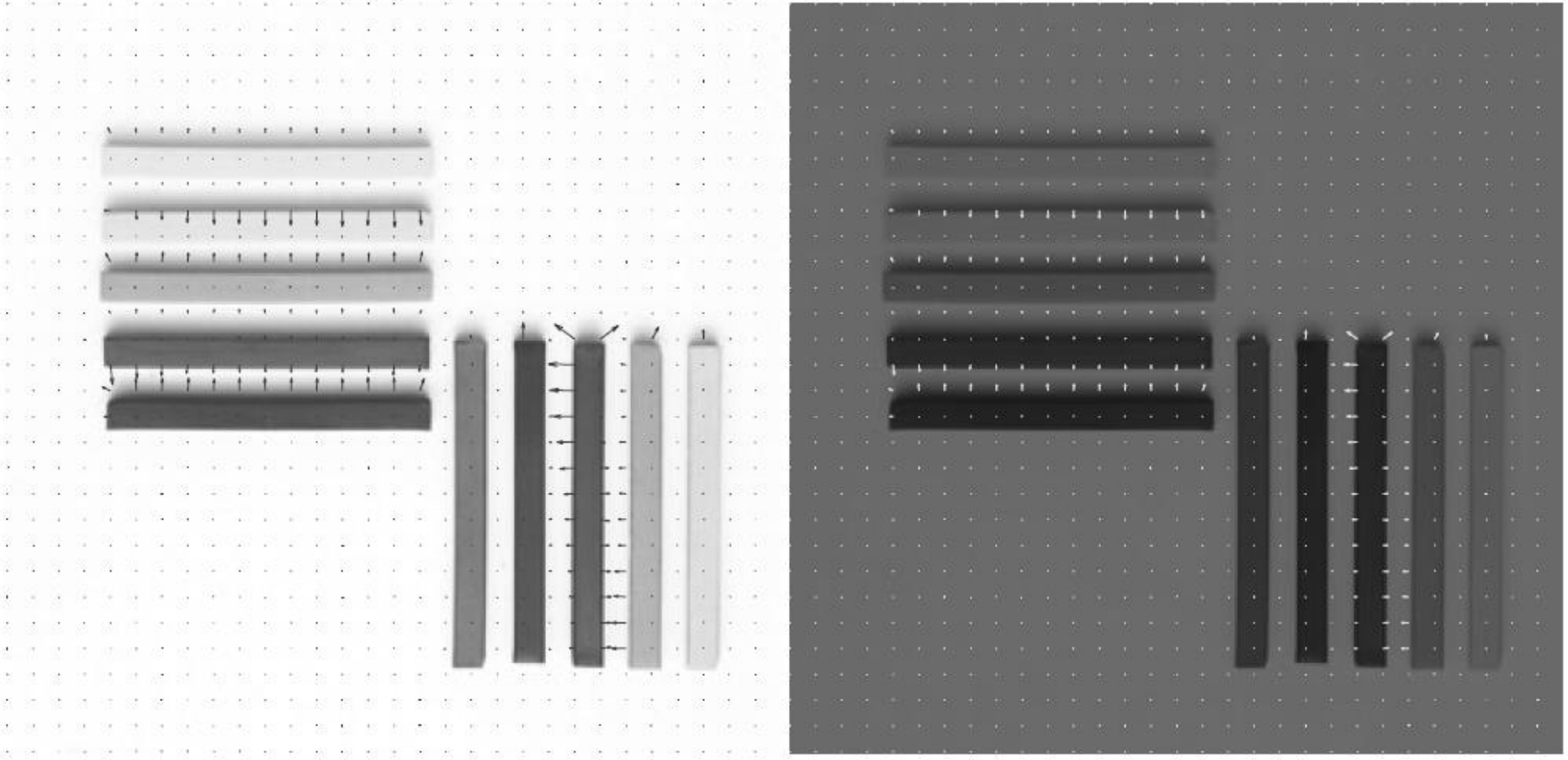
Pixel Intensity Correlation

Corr = 0.24

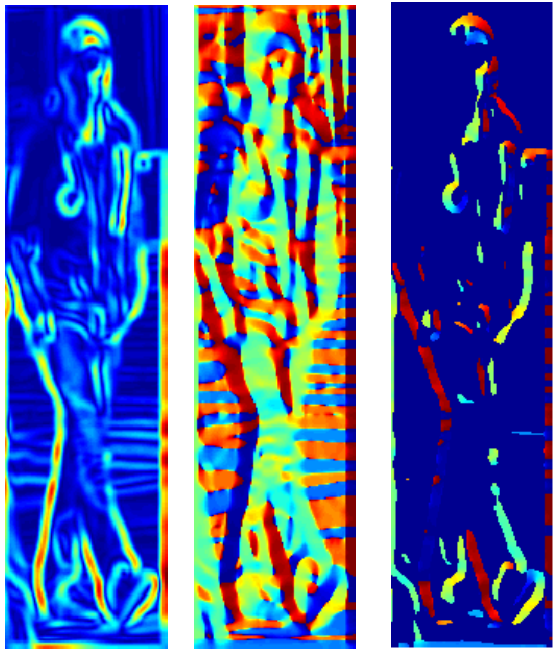
Corr = 0.31



Recall: Illumination Invariant Gradient

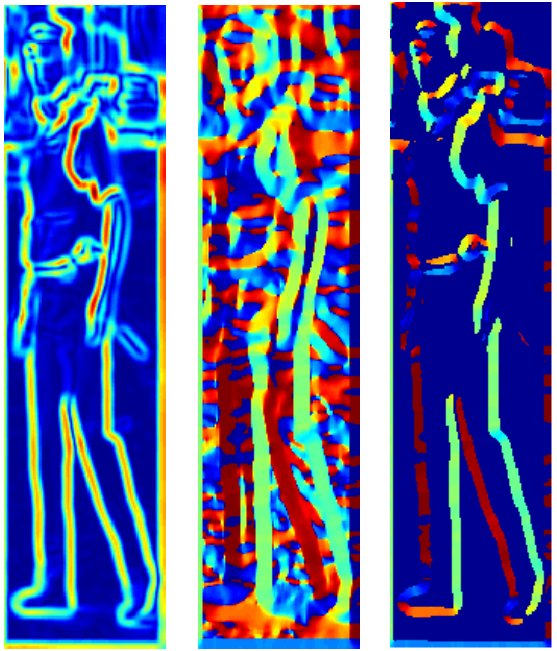


GRADIENT



$\|\nabla I\|$ $\measuredangle \nabla I$ $\measuredangle \nabla I$

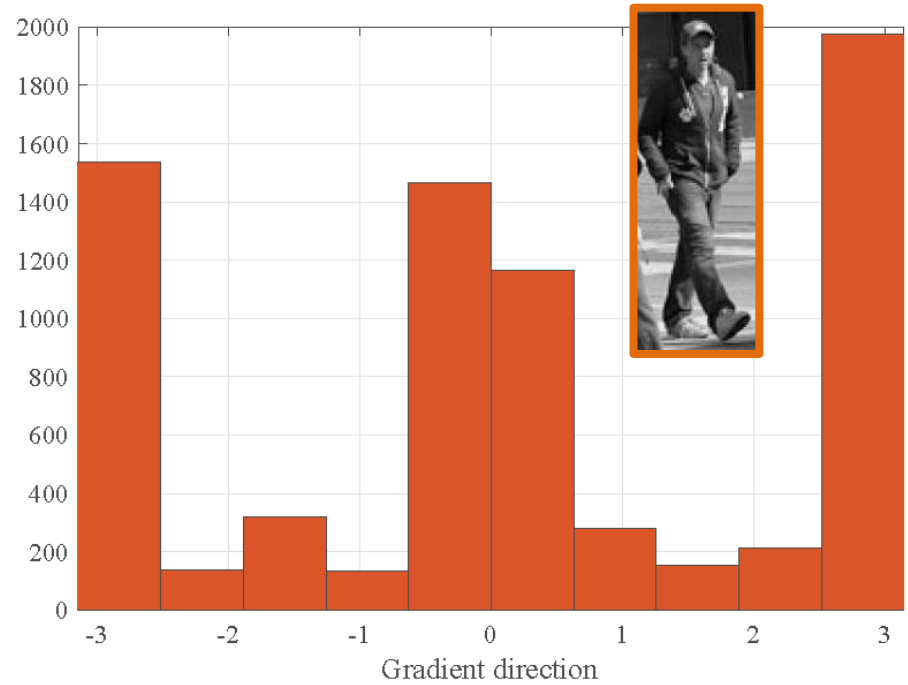
with mag thr.

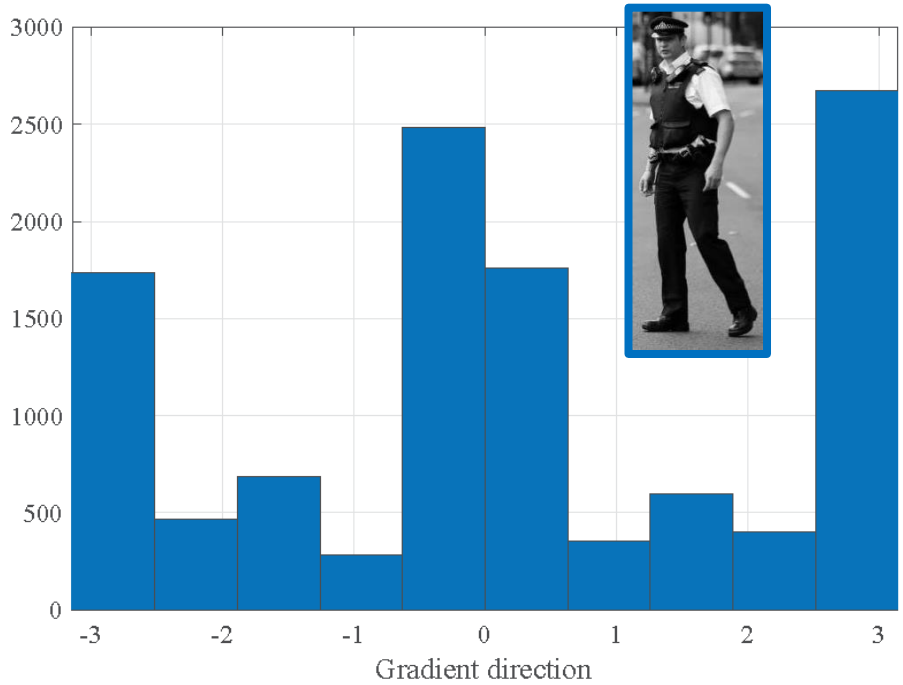


$\|\nabla J\|$ $\measuredangle \nabla J$ $\measuredangle \nabla J$

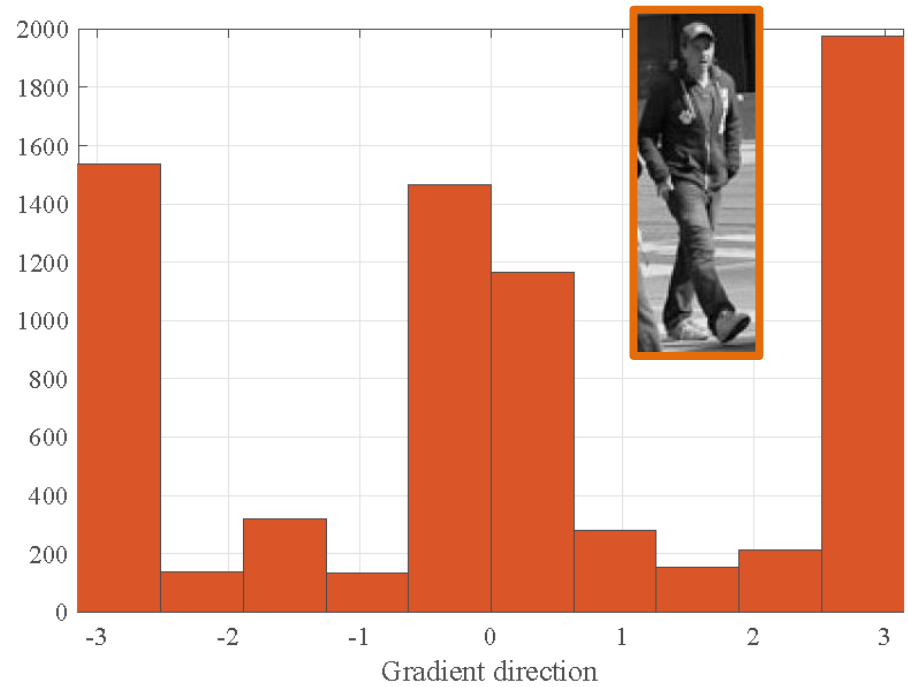
with mag thr.

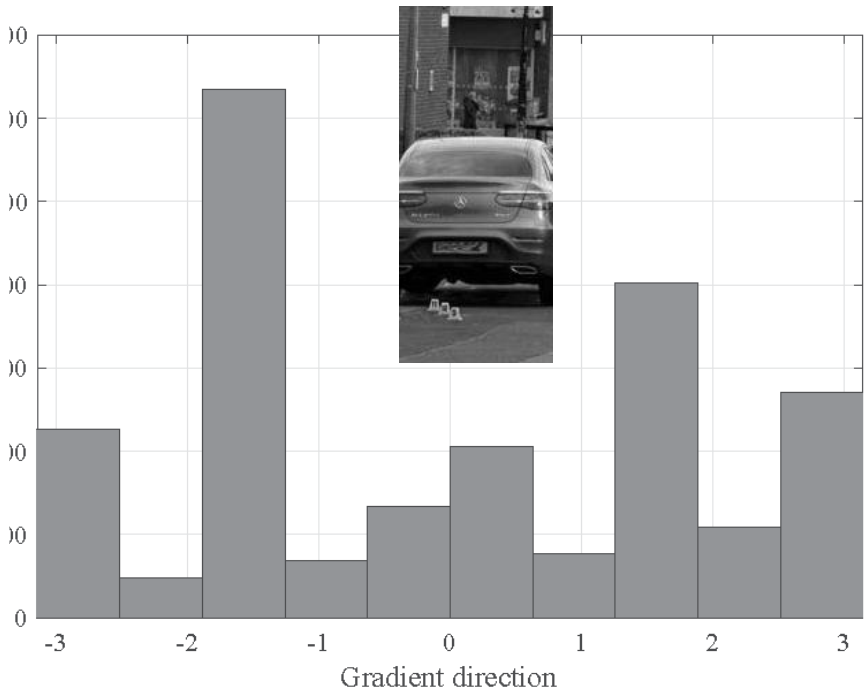
Gradient Distribution (Histogram of Grad. Direction)



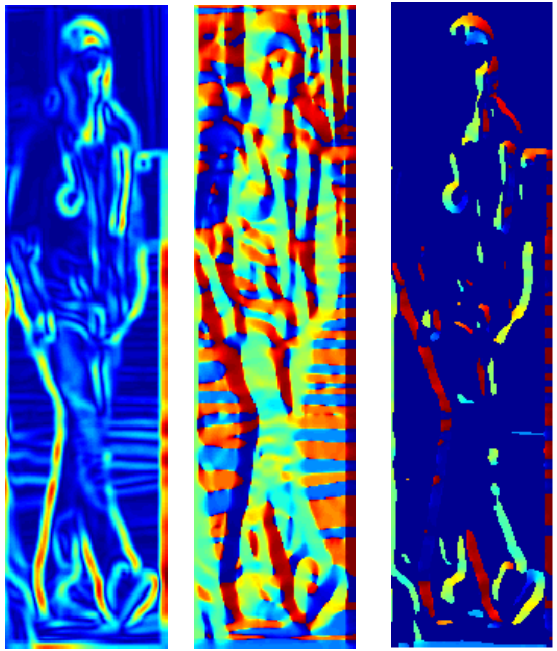


GRADIENT DISTRIBUTION (HISTOGRAM OF GRAD. DIRECTION)



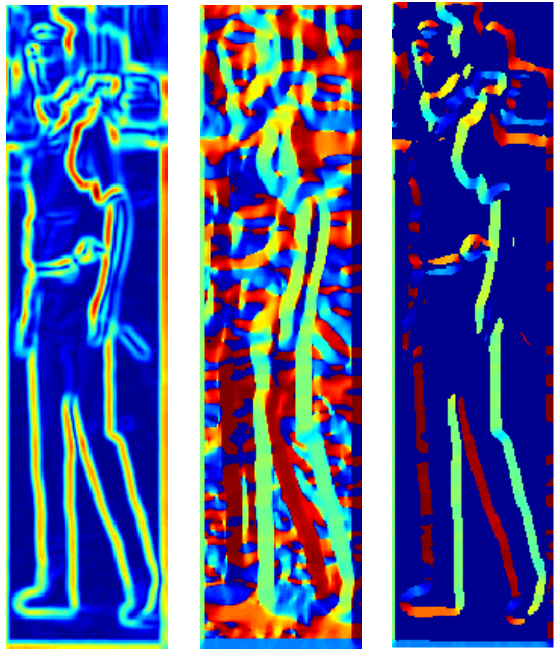


GLOBAL GRADIENTS



$\|\nabla I\|$ $\measuredangle \nabla I$ $\measuredangle \nabla I$

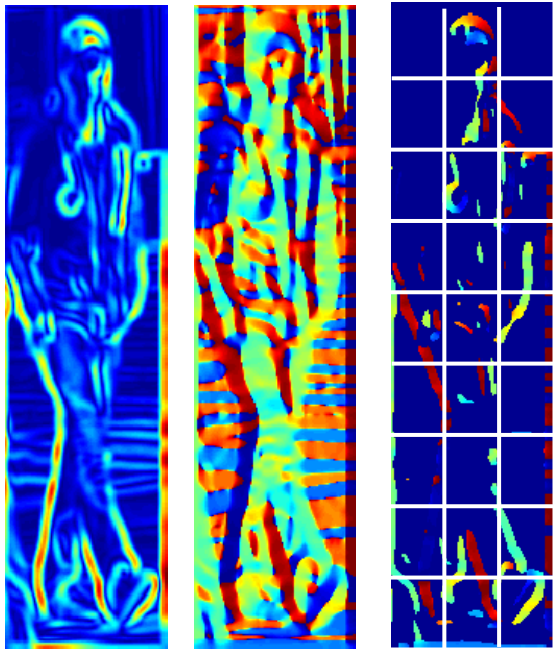
with mag thr.



$\|\nabla J\|$ $\measuredangle \nabla J$ $\measuredangle \nabla J$

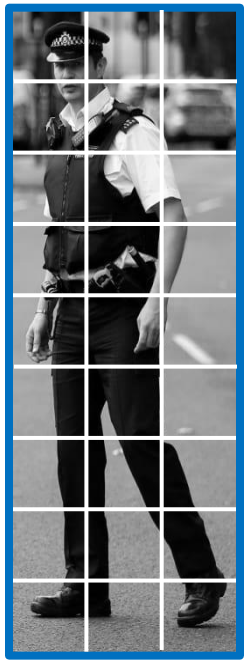
with mag thr.

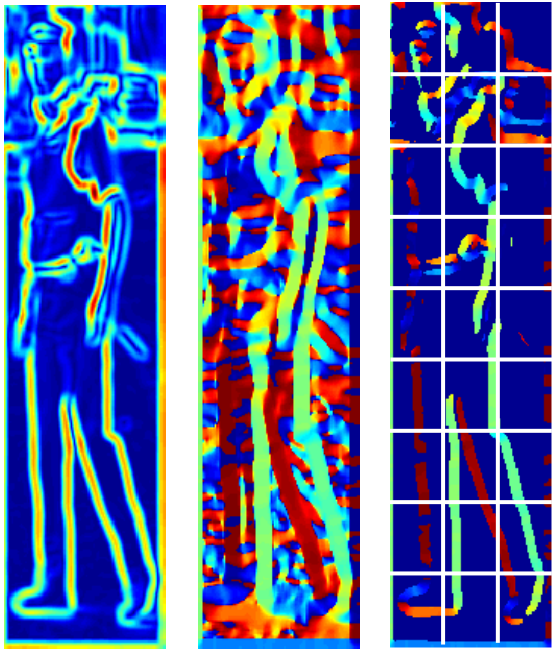
LOCAL GRADIENTS



$\|\nabla I\|$ $\measuredangle \nabla I$ $\measuredangle \nabla I$

with mag thr.

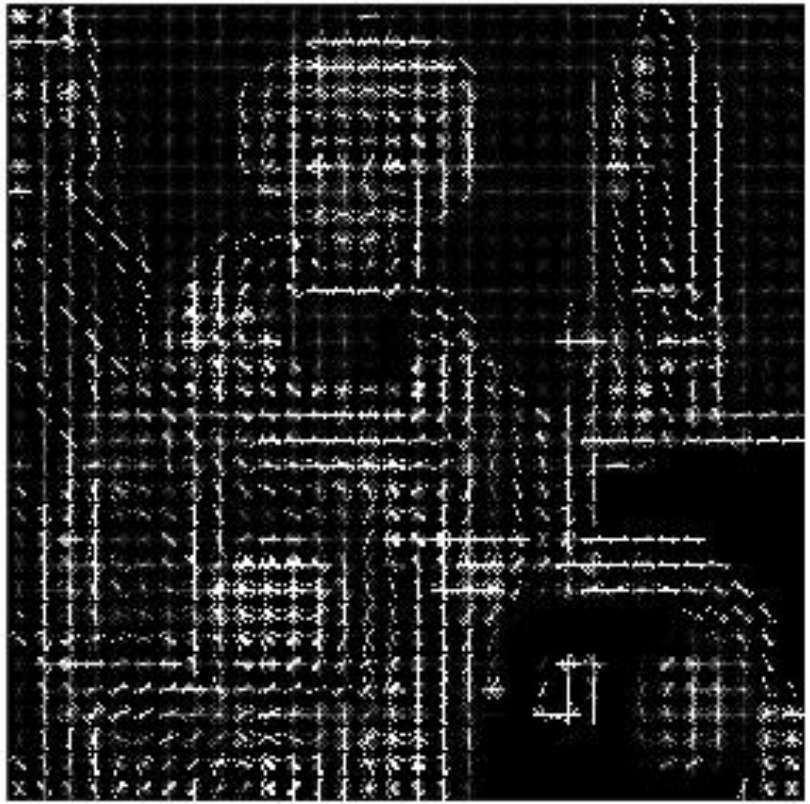




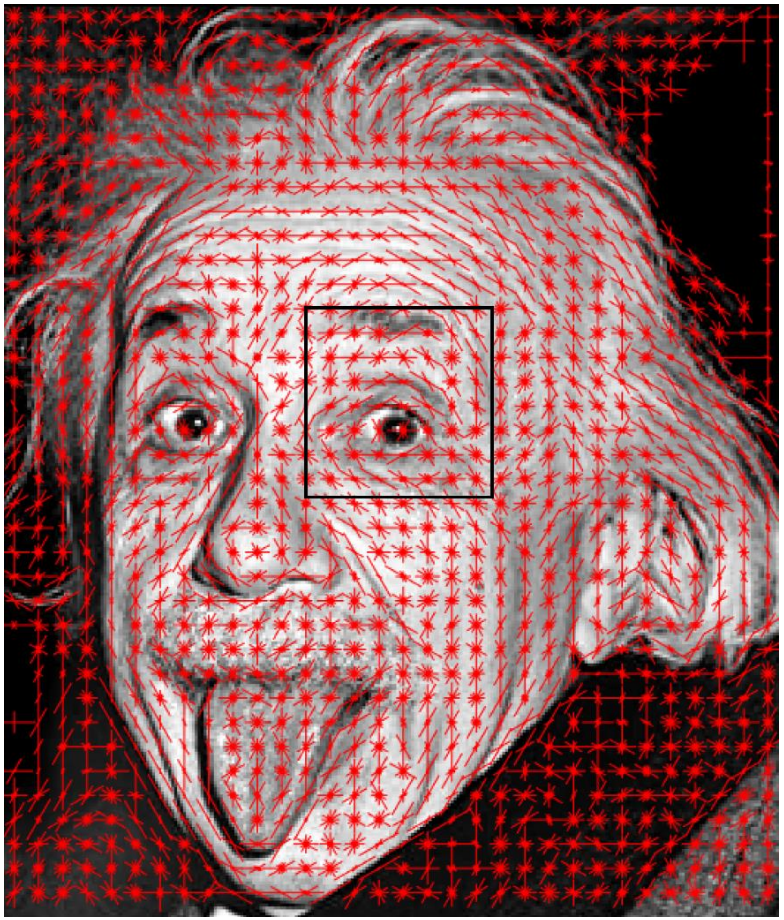
$\|\nabla J\|$ $\measuredangle \nabla J$ $\measuredangle \nabla J$

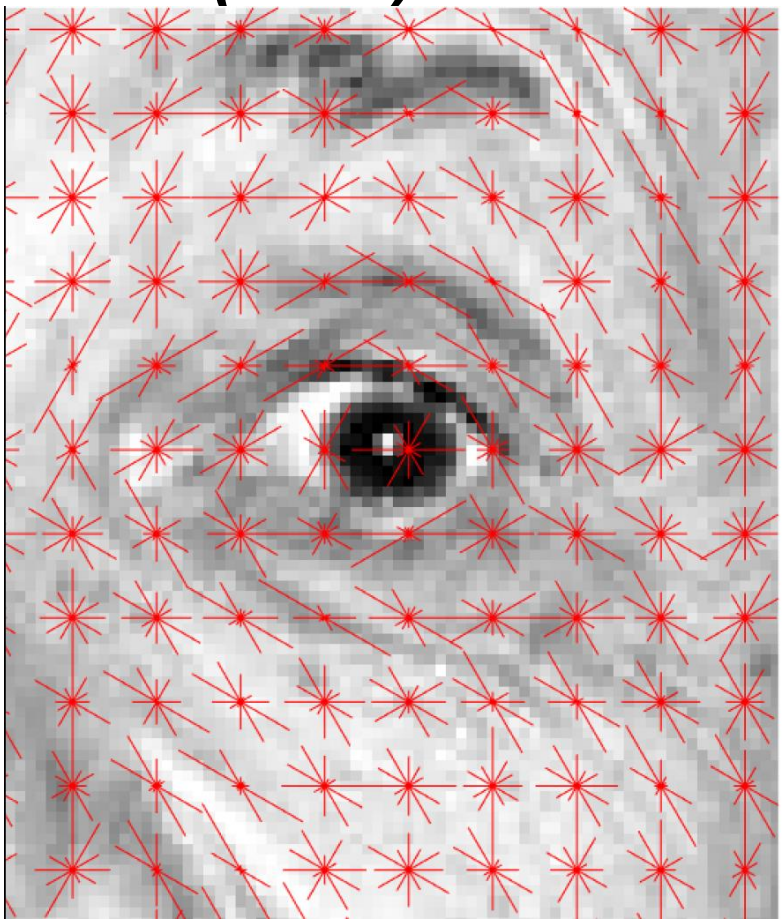
with mag thr.

Histogram of Oriented Gradients (HOG)

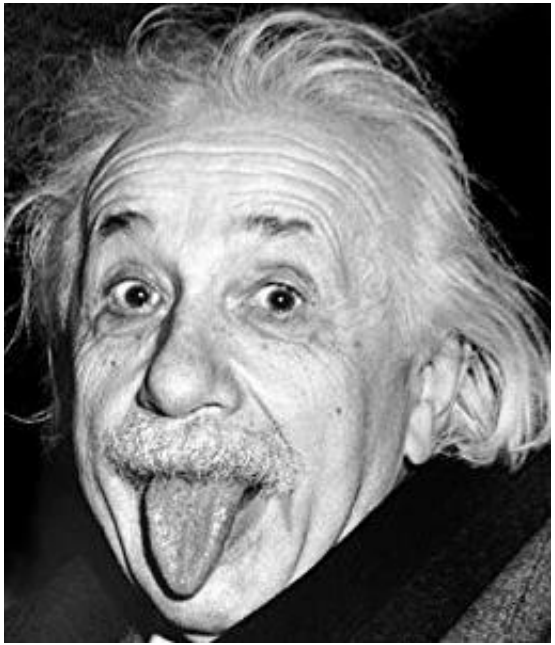


Histogram of Oriented Gradients (HOG)

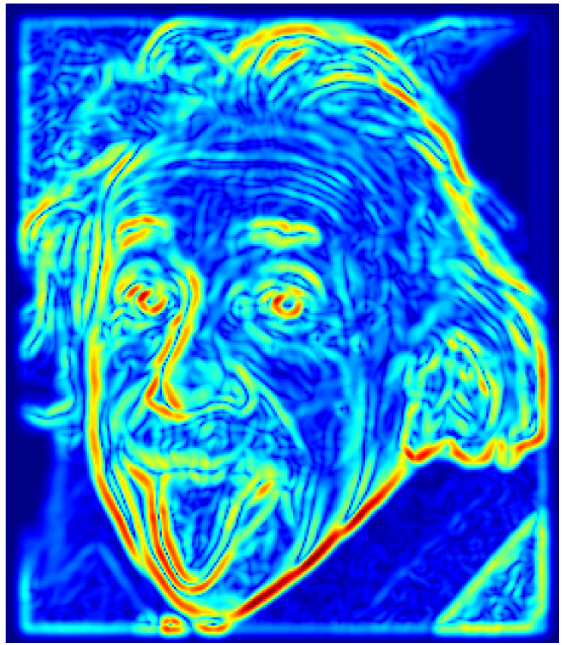




Gradient Image



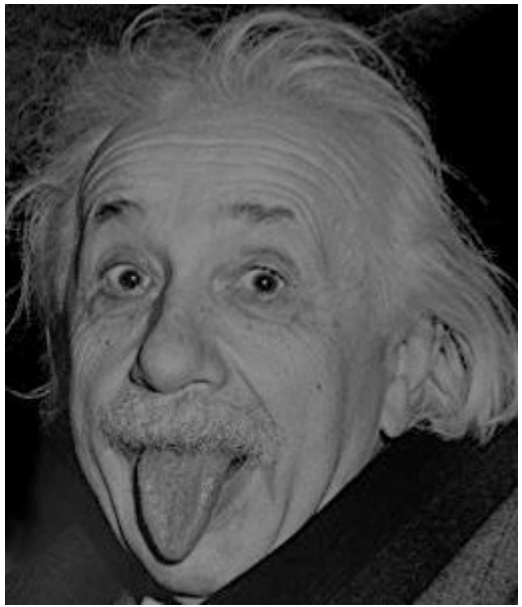
I



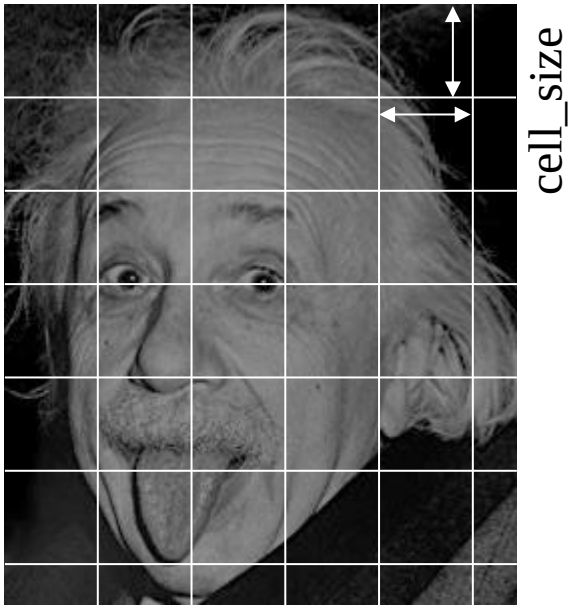
$\|\nabla I\|$

$\measuredangle \nabla I$

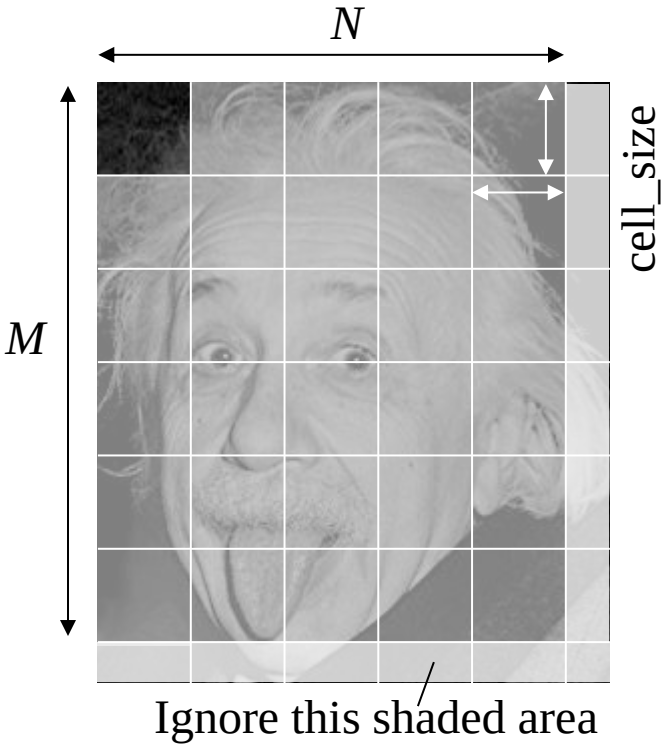
Orientation Binning



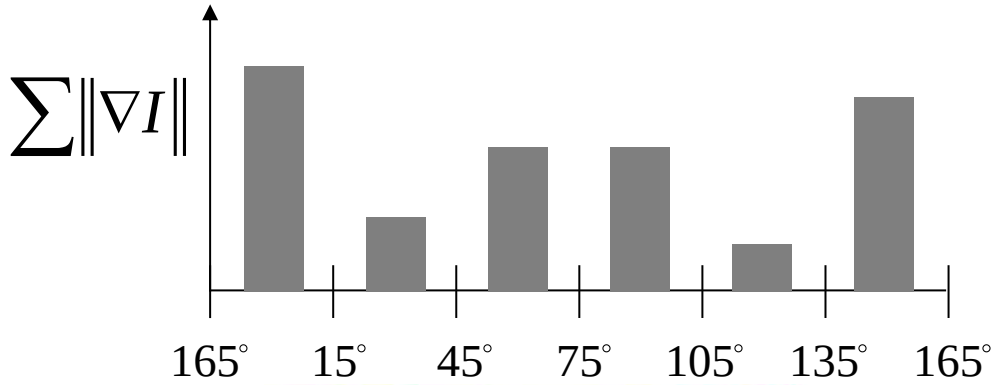
Orientation Binning

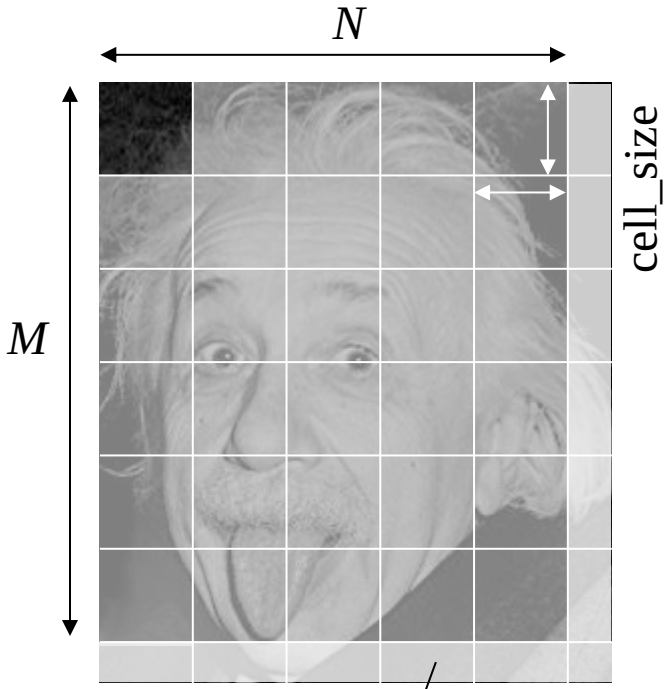


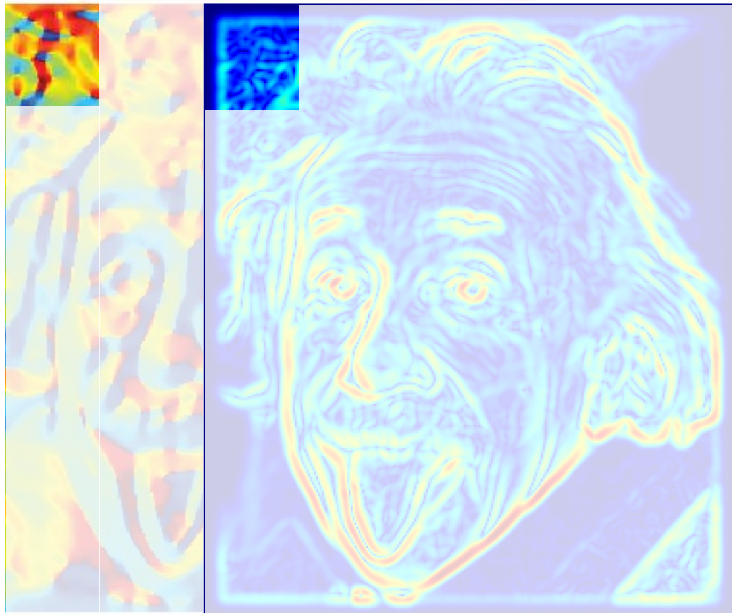
ORIENTATION BINNING



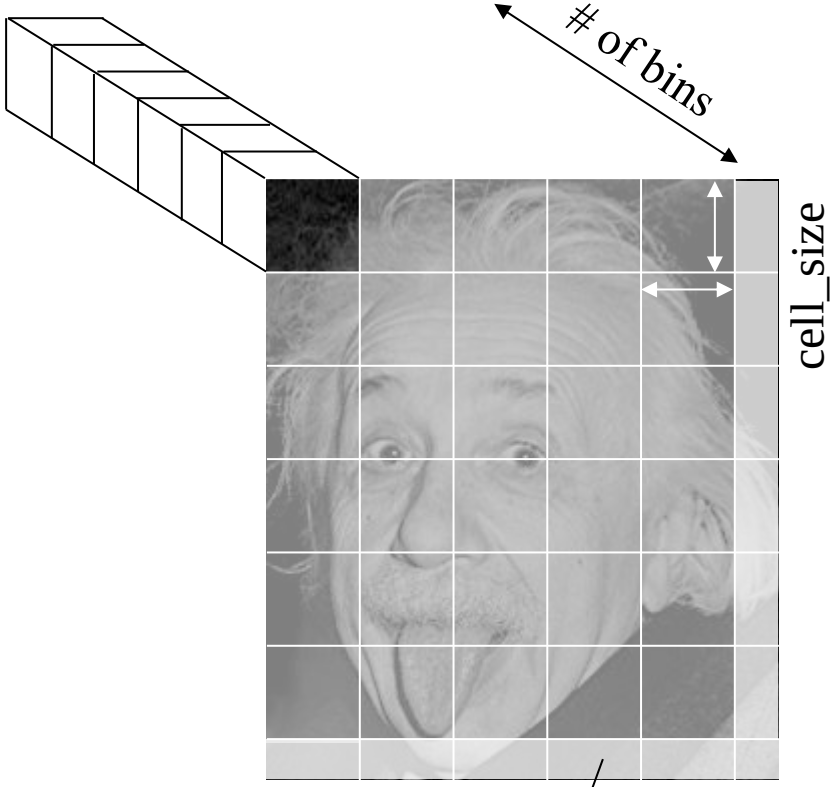
ORIENTATION BINNING

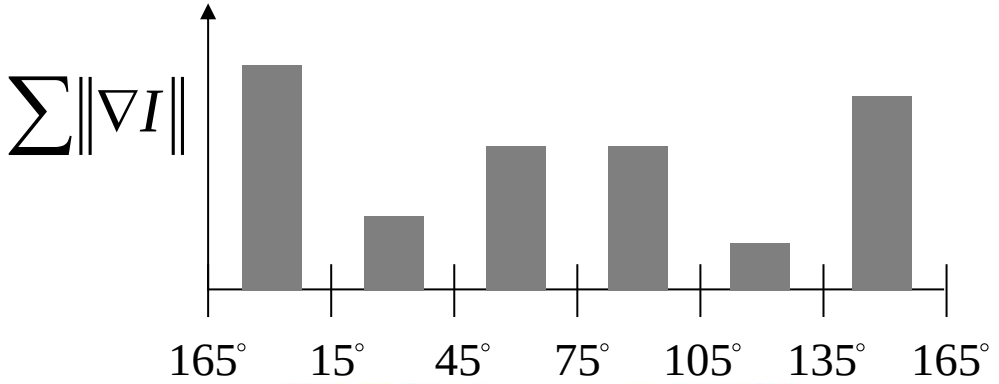


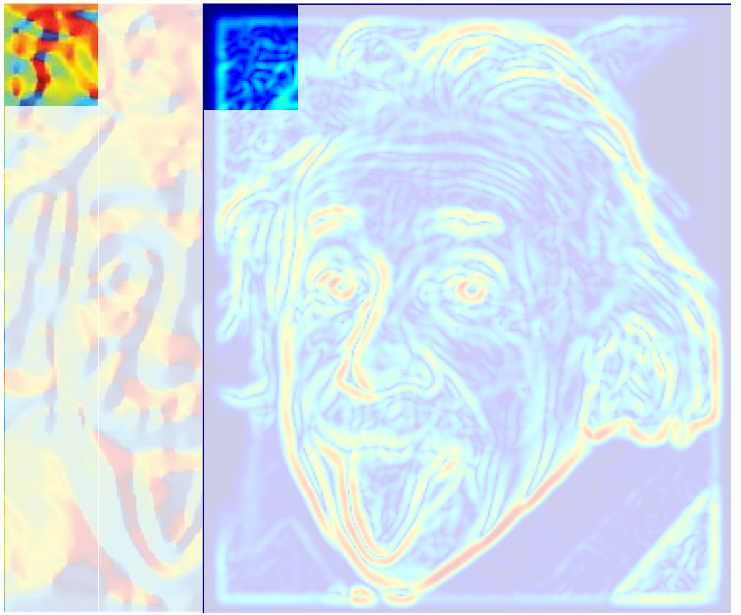




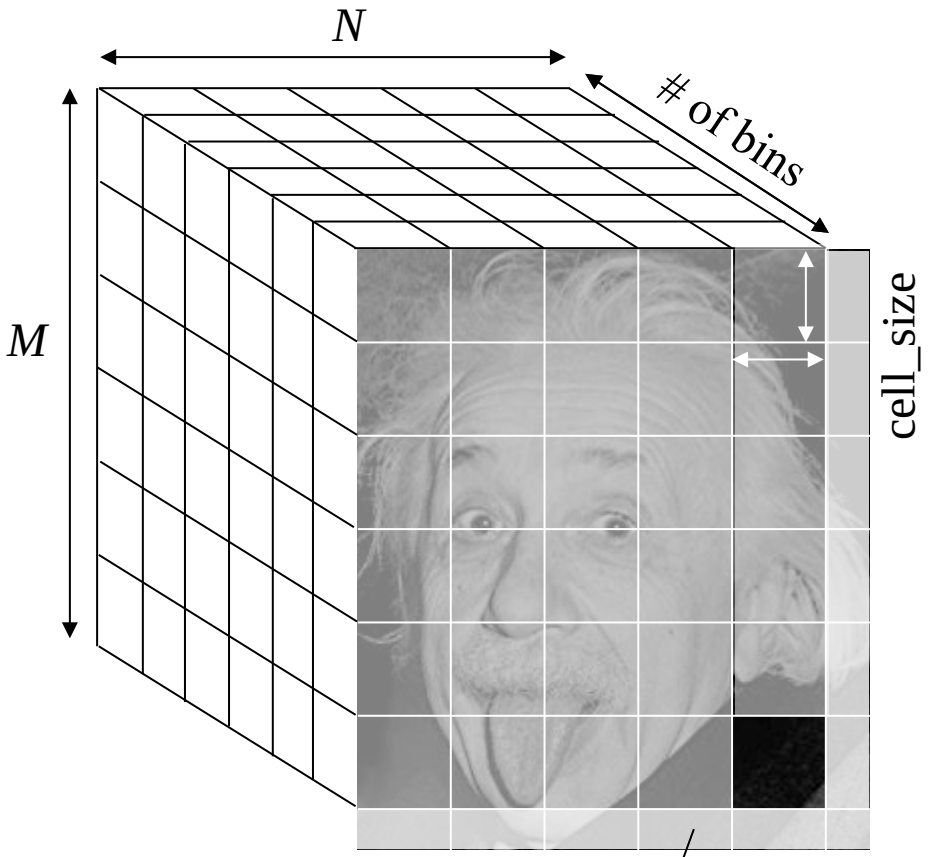
ORIENTATION BINNING

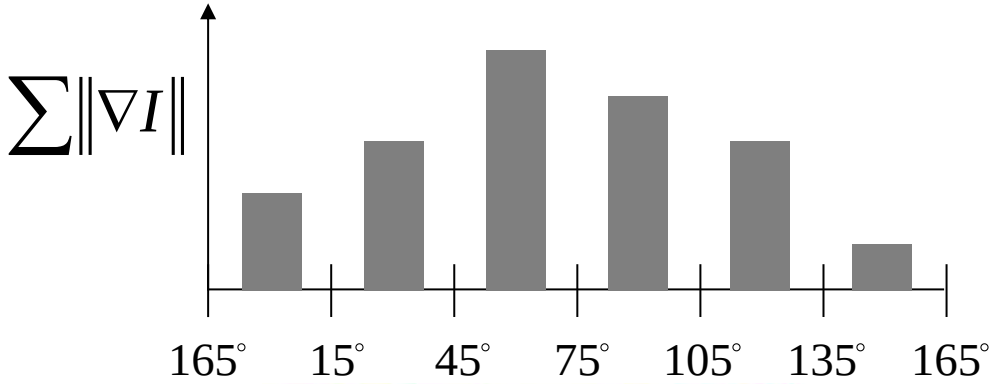


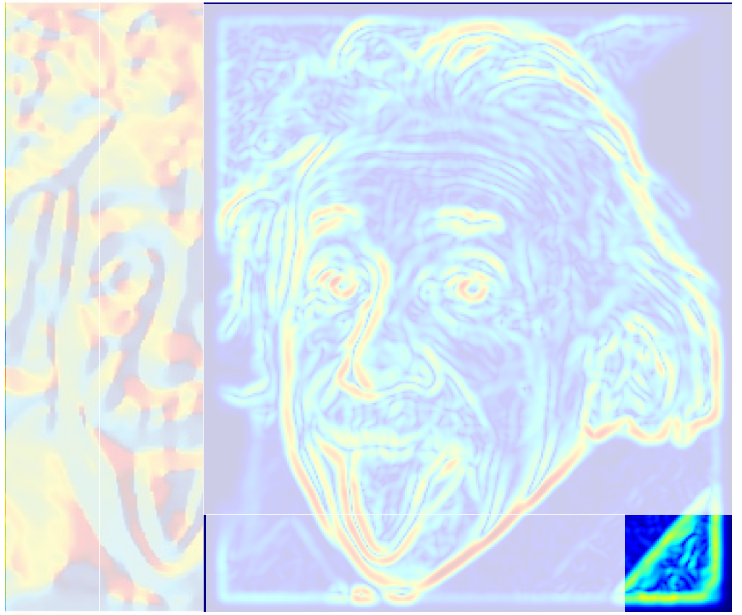




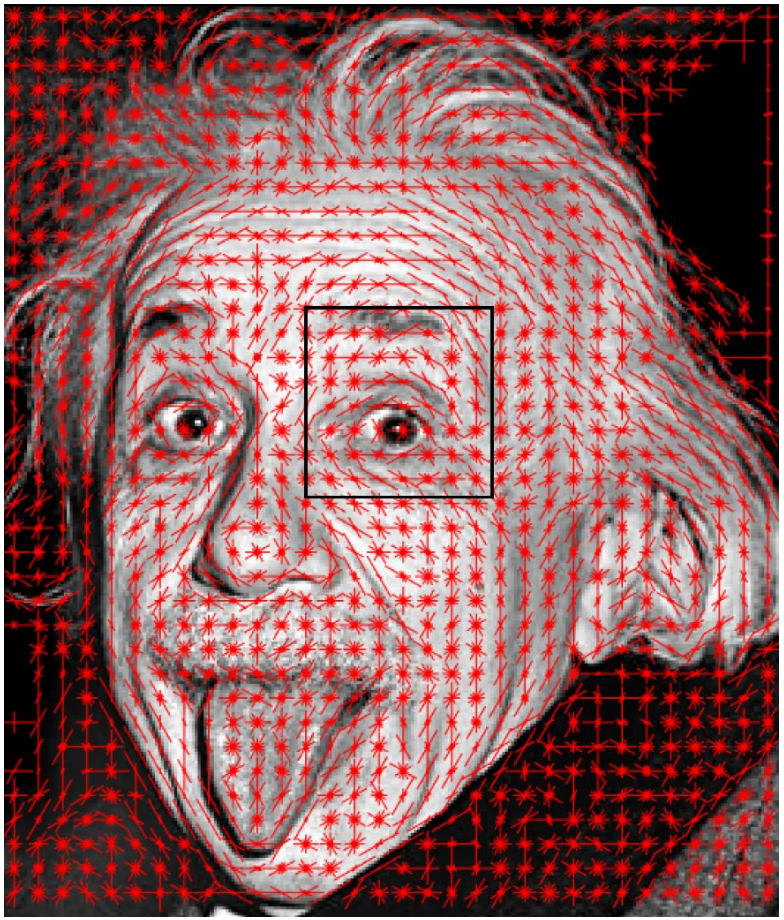
ORIENTATION BINNING

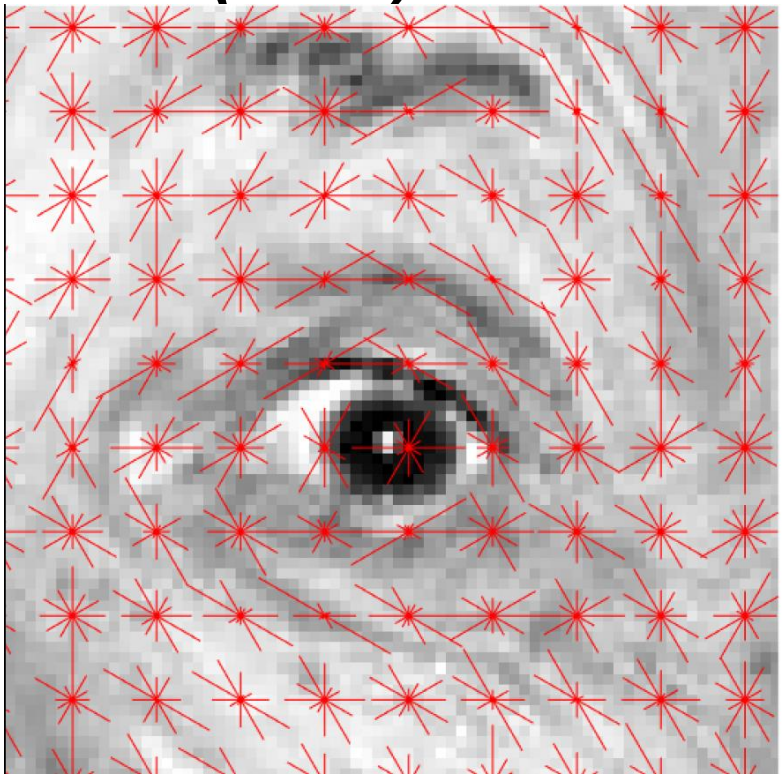






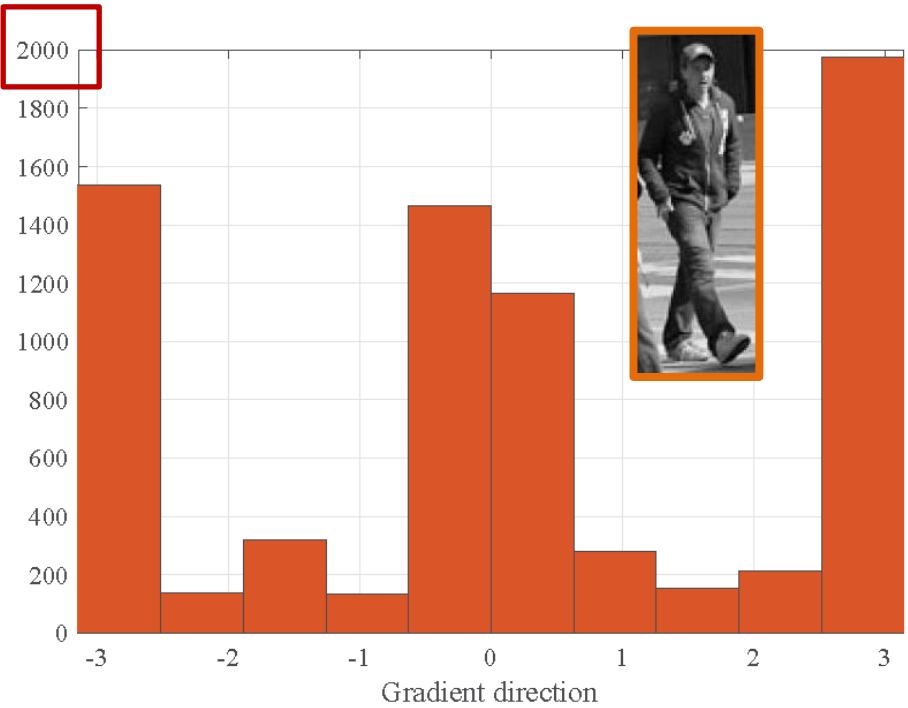
Histogram of Oriented Gradients (HOG)

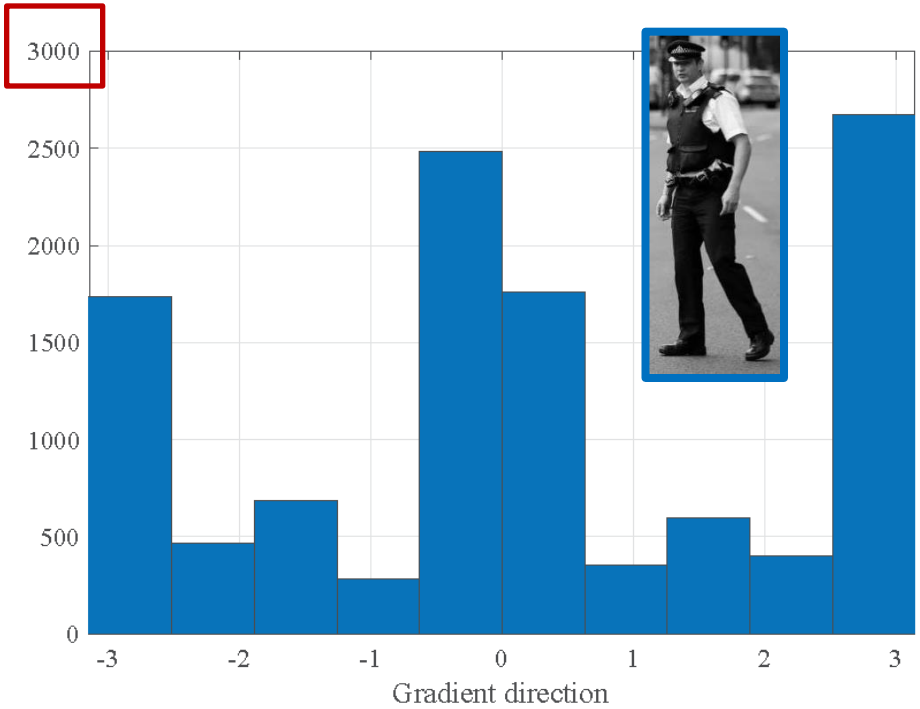




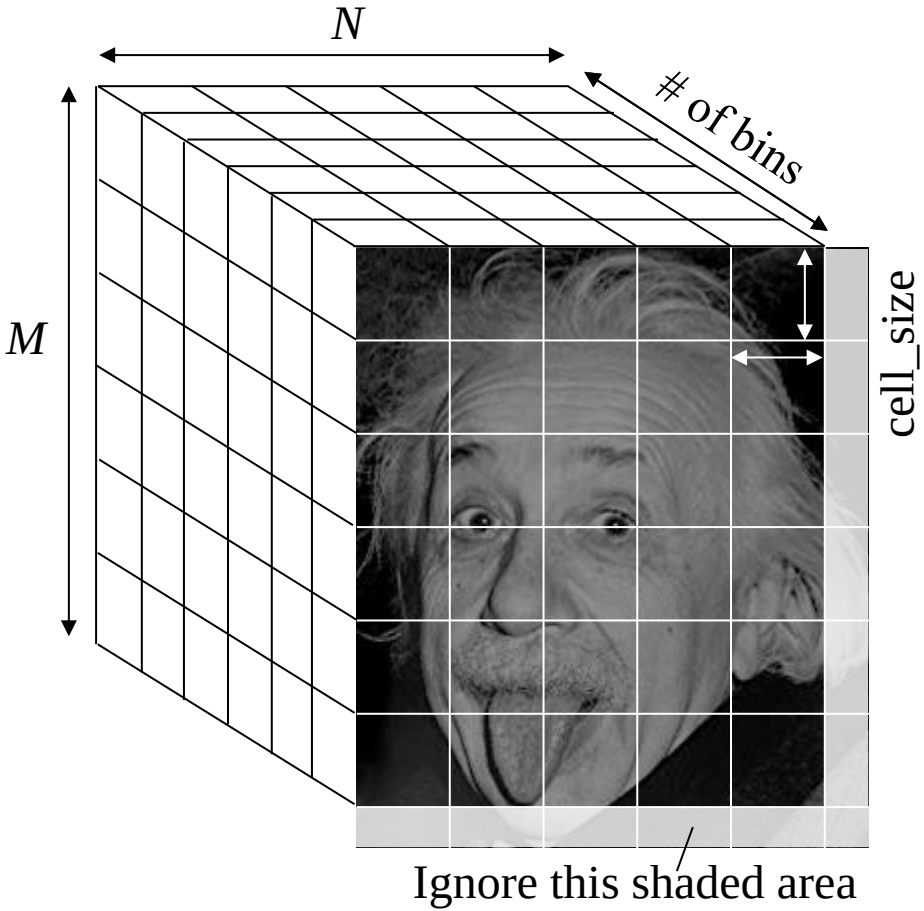
For visualization purpose, line perpendicular to the gradient directions are drawn.

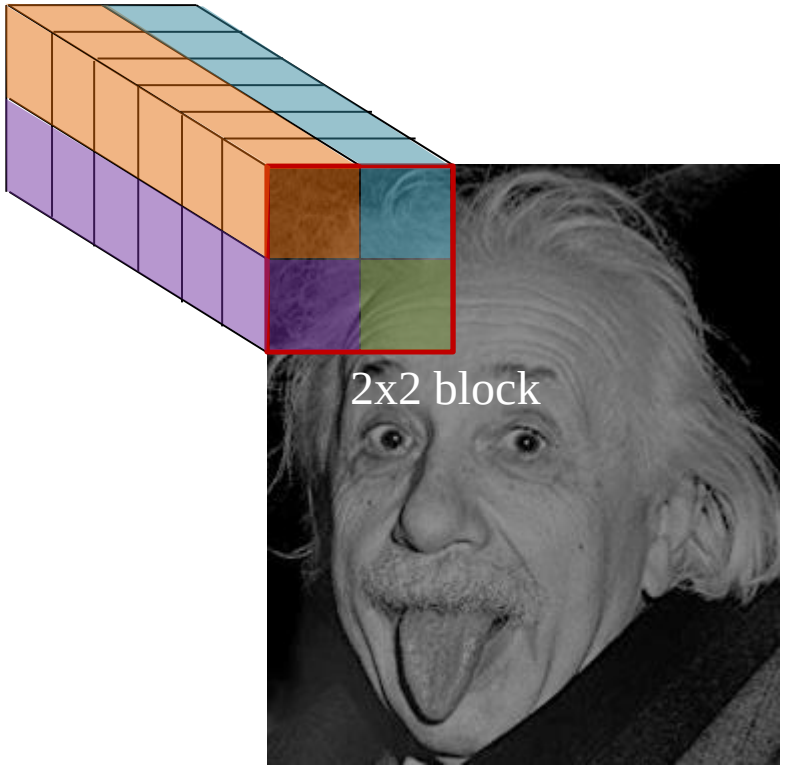
RECALL: SCALE



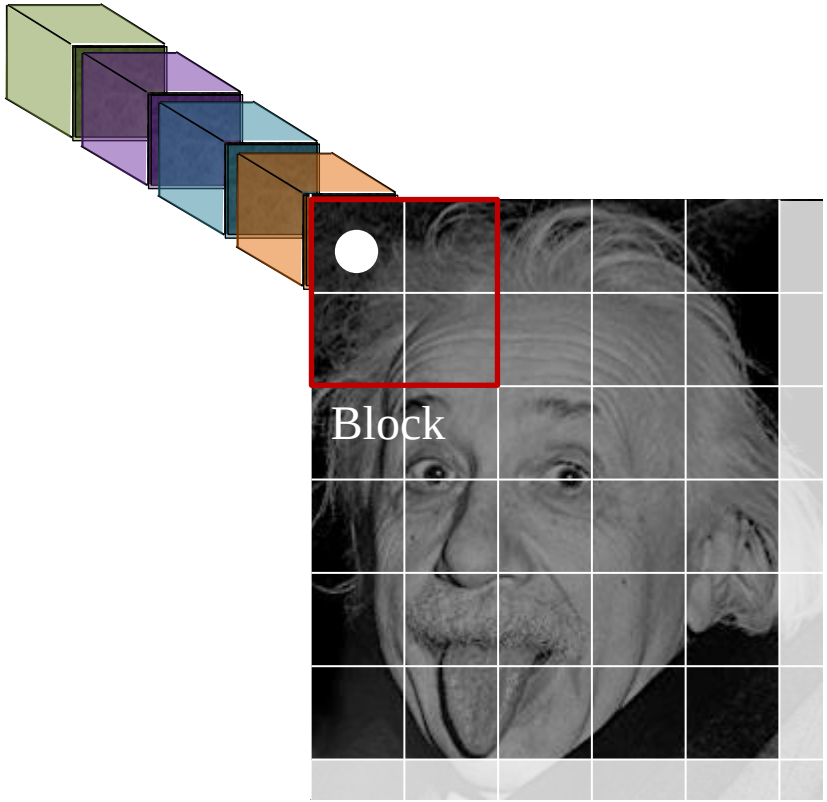


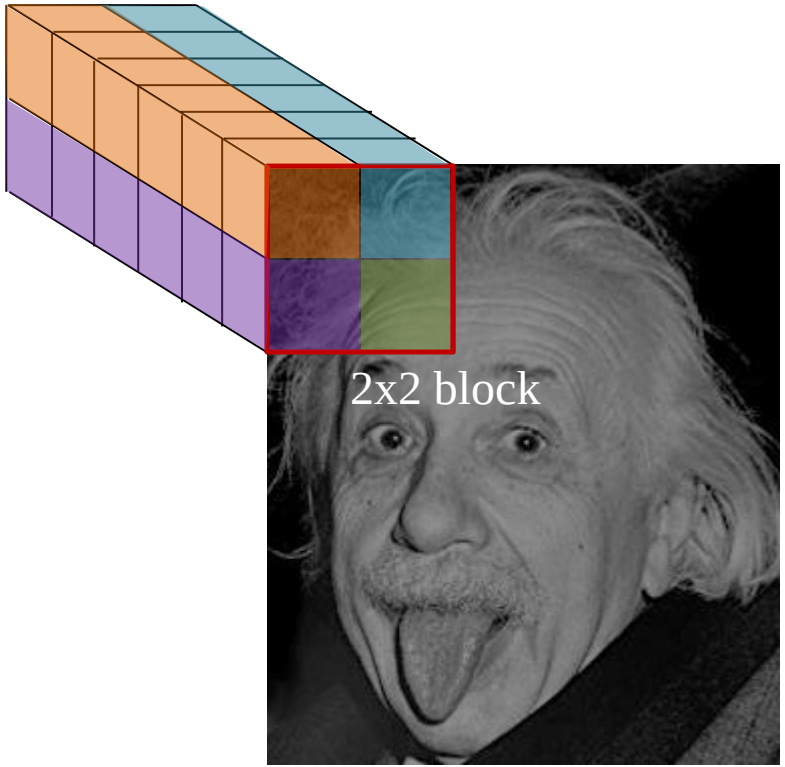
BLOCK NORMALIZATION



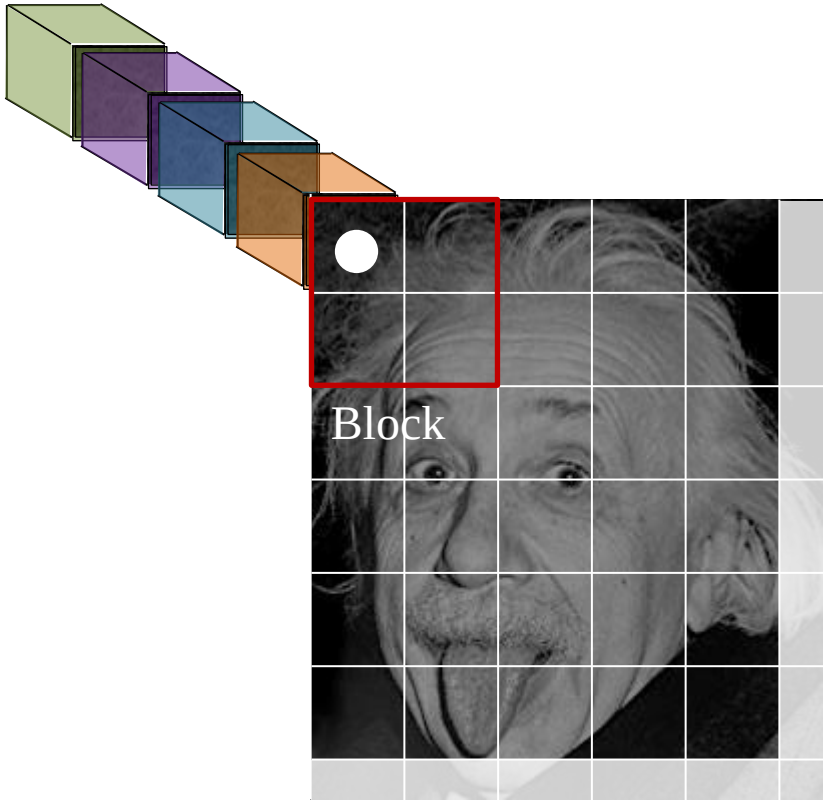


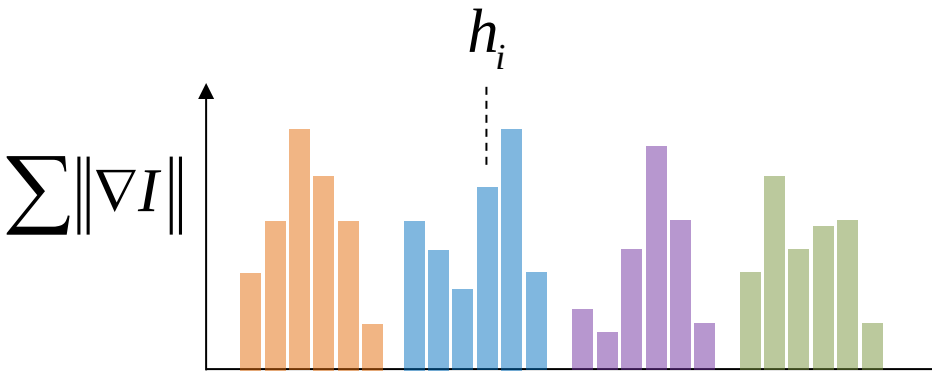
Block Normalization





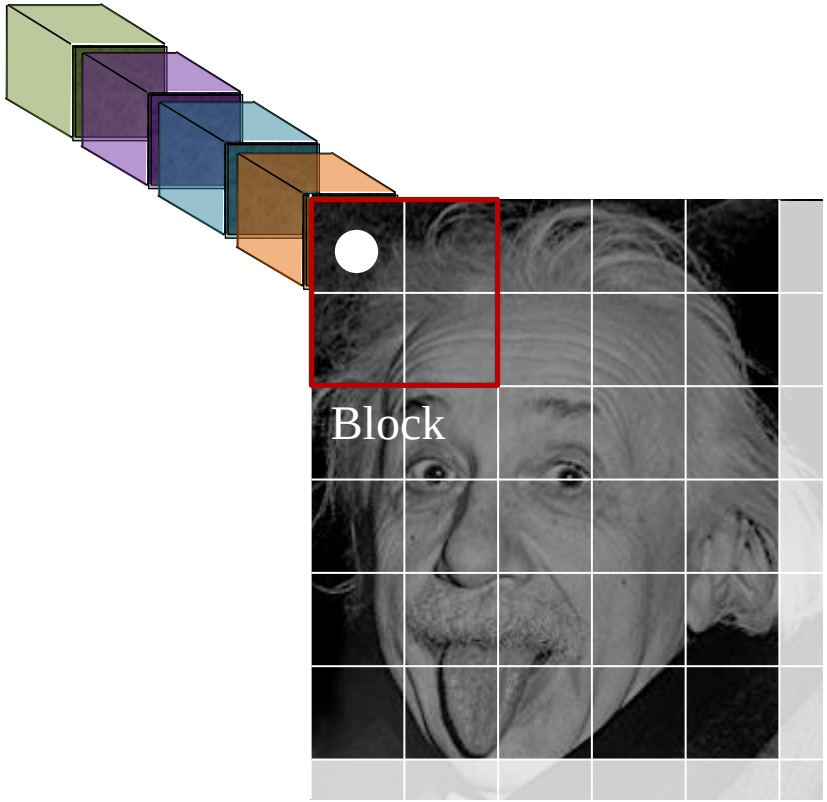
Block Normalization

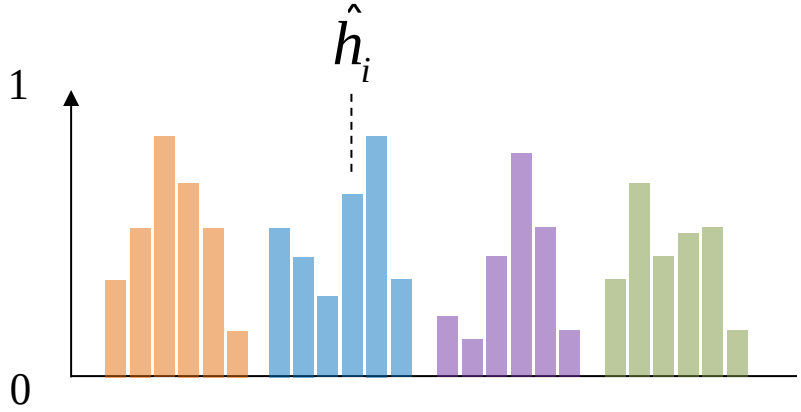




$$\hat{h}_i = \frac{h_i}{\sqrt{\sum_i h_i^2}}$$

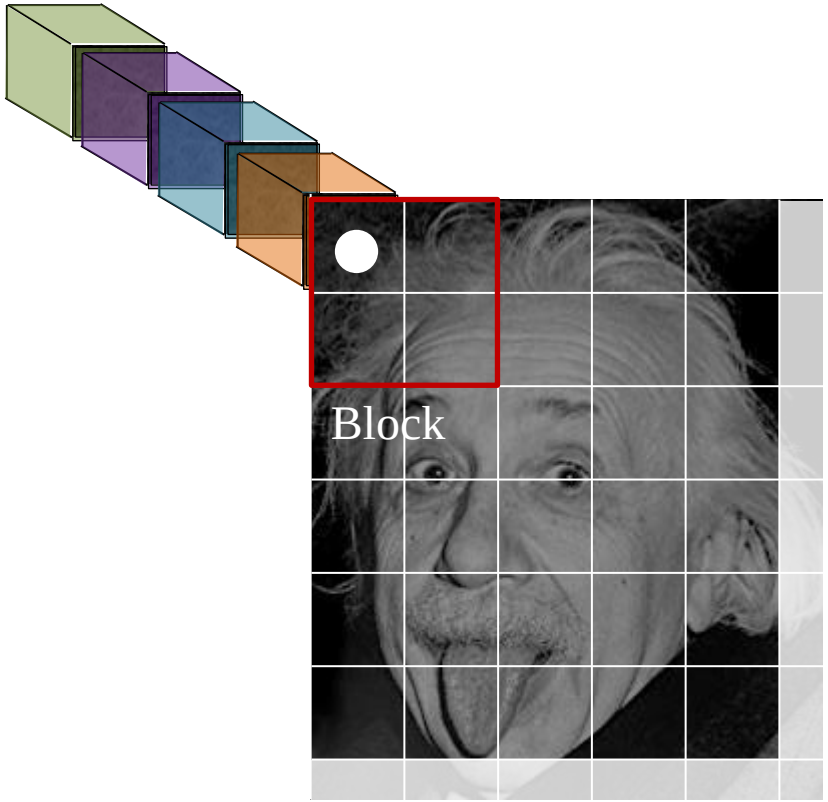
Block Normalization

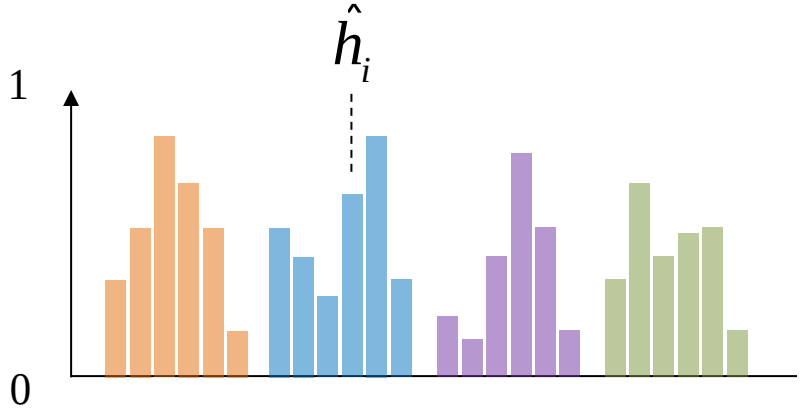




$$\hat{h}_i = \frac{h_i}{\sqrt{\sum_i h_i^2}}$$

Block Normalization

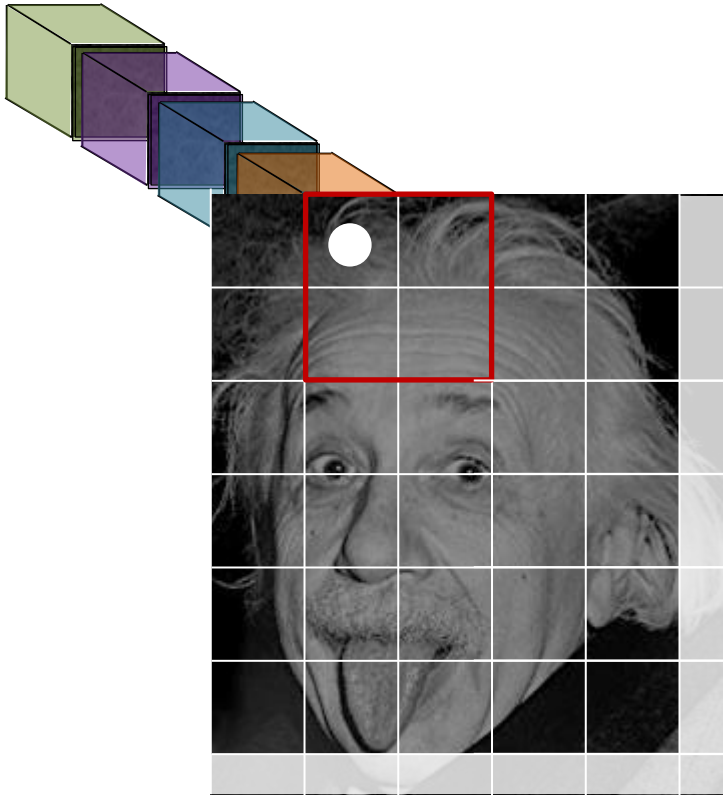


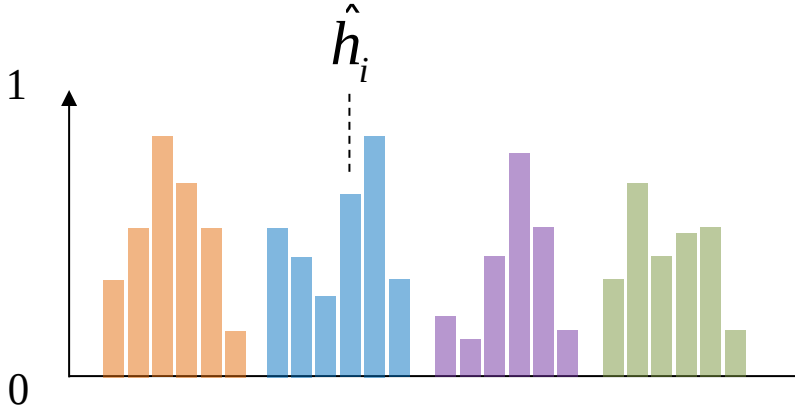


$$\hat{h}_i = \frac{h_i}{\sqrt{\sum_i h_i^2 + e^2}}$$

Preventing division by zero

Block Normalization

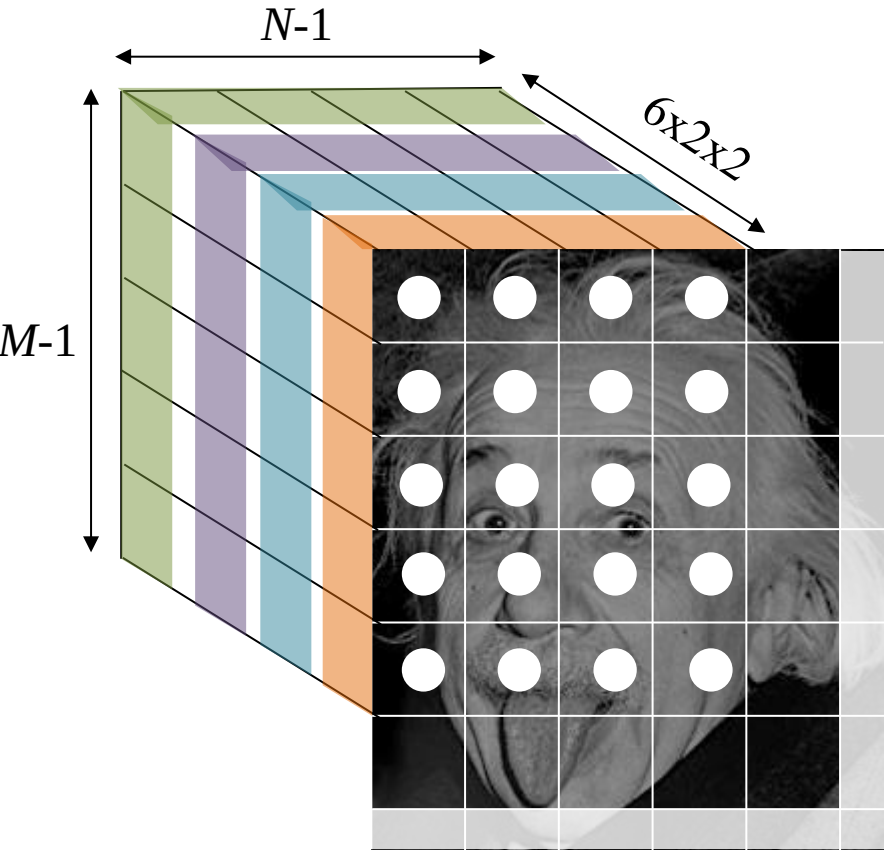


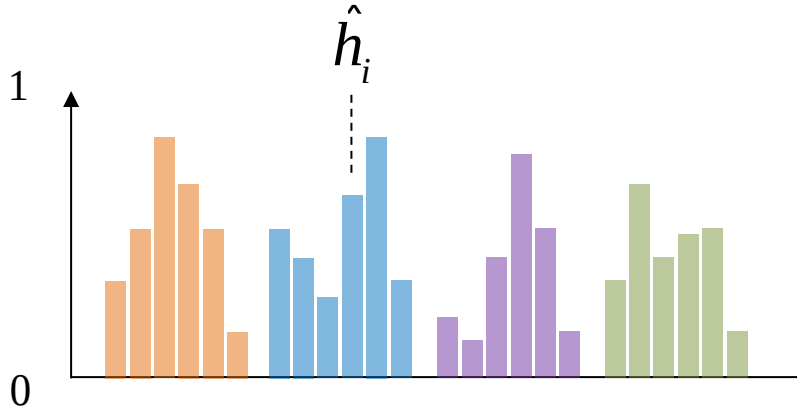


$$\hat{h}_i = \frac{h_i}{\sqrt{\sum_i h_i^2 + e^2}}$$

Preventing division by zero

BLOCK NORMALIZATION

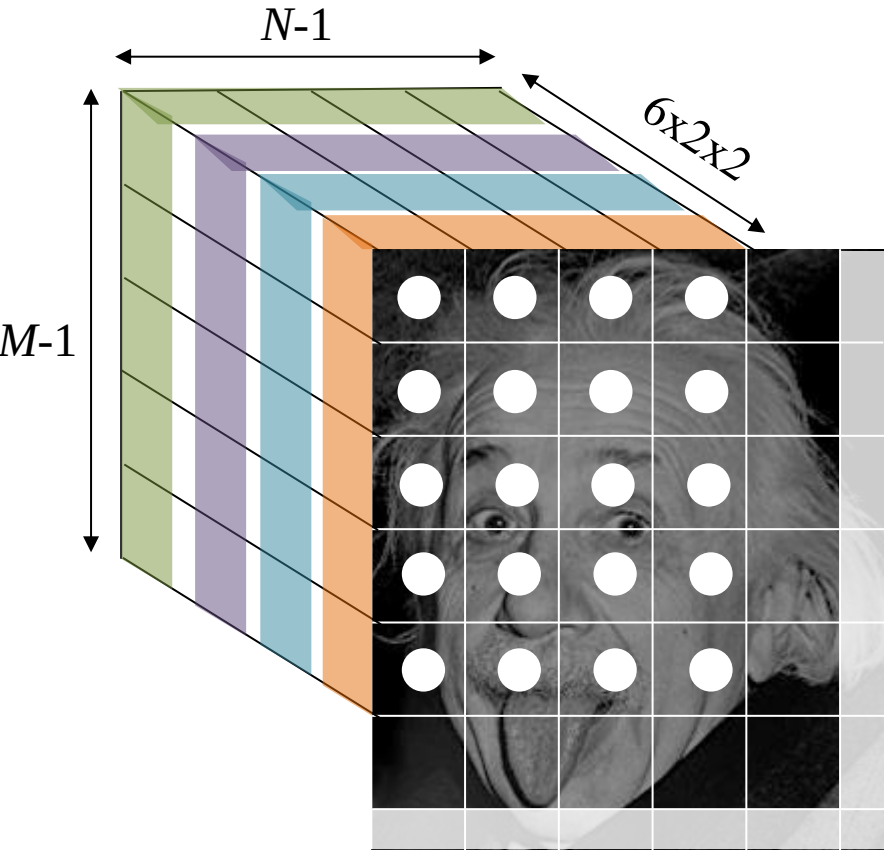




$$\hat{h}_i = \frac{h_i}{\sqrt{\sum_i h_i^2 + e^2}}$$

Preventing division by zero

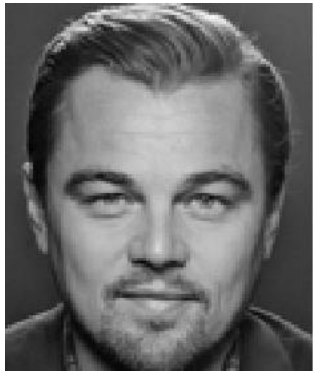
VECTORIZATION



$$\mathbf{x} = \begin{bmatrix} \\ \\ \\ \\ \vdots \\ \\ \\ \\ \end{bmatrix}$$

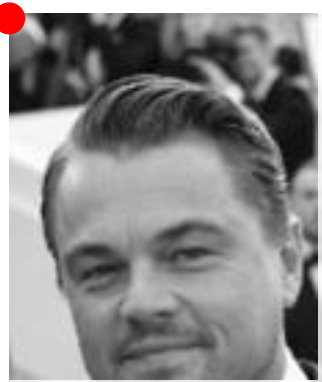
Feature descriptor of image I

Face Detection

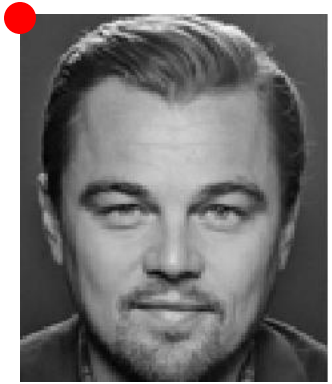


FACE DETECTION

(x,y)



u



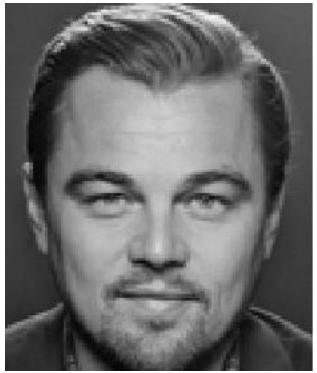
v

$$\frac{u \cdot v}{\|u\|\|v\|} = \cos\theta$$

FACE DETECTION

(x,y)

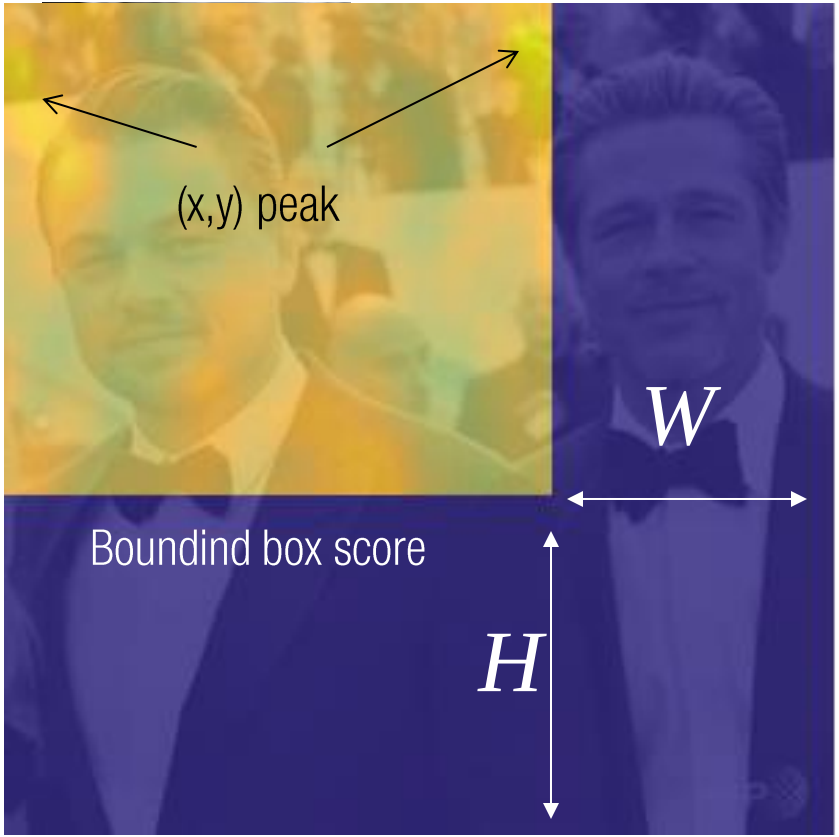
u

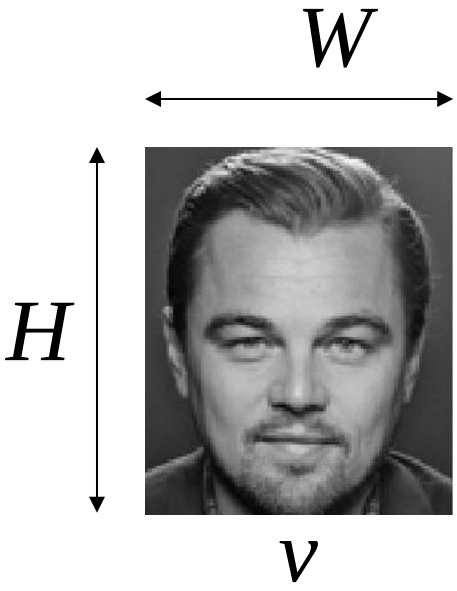


v

$$\frac{u \cdot v}{\|u\|\|v\|} = \cos\theta$$

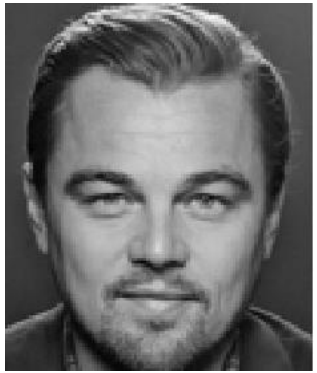
FACE DETECTION





$$\frac{u \cdot v}{\|u\|\|v\|} = \cos\theta$$

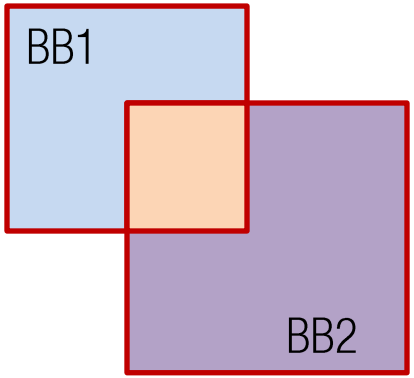
THRESHOLDING



Detected bounding box

$$\frac{u \cdot v}{\|u\|\|v\|} > \varepsilon$$

Intersection of Union

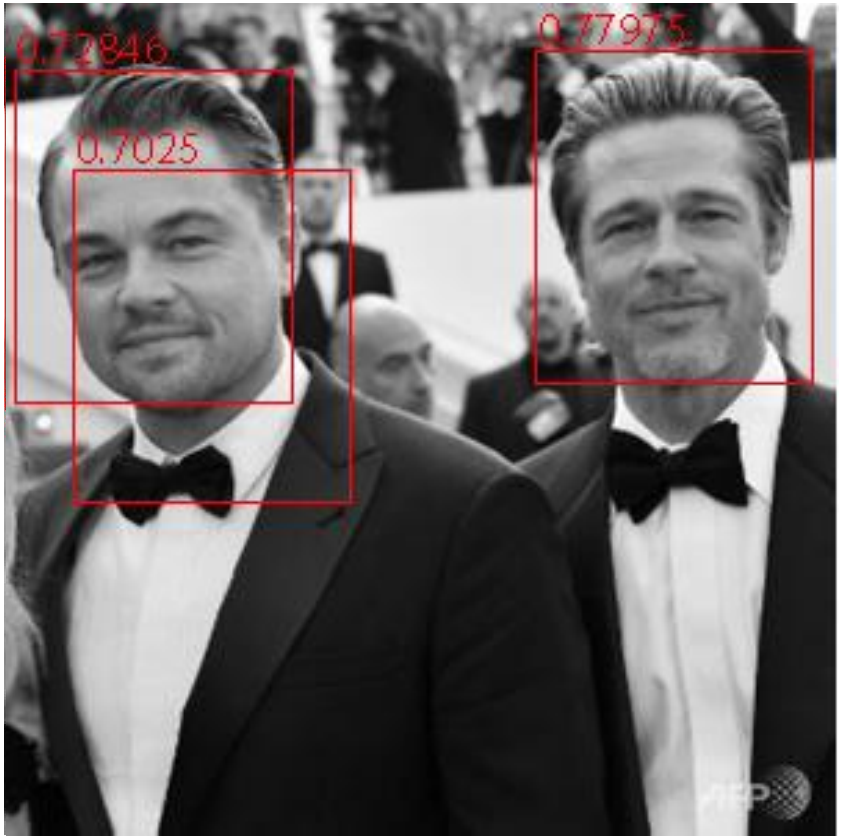


$$IoU = \frac{\square}{\square + \square + \square}$$

Non-maximum Suppression

1. Find the BB of maximum score from BB set
2. Suppress/delete all BBs in BB set that have IoU greater than 0.5
3. Break if no element in BB set
4. Go to 1

Non-maximum Suppression



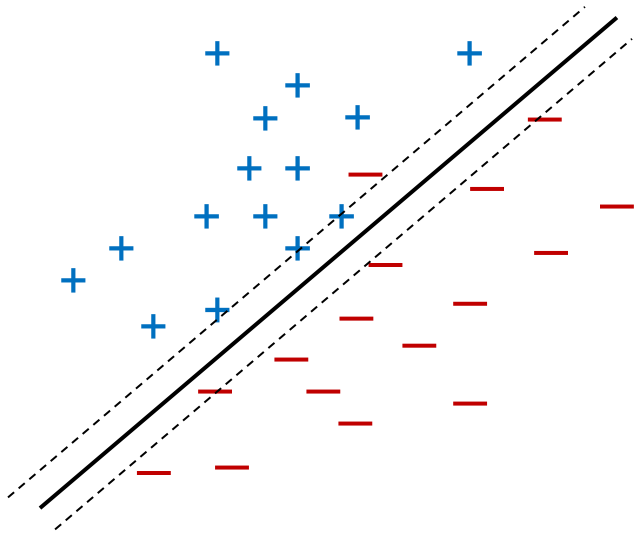
1. Find the BB of maximum score from BB set
2. Suppress/delete all BBs in BB set that have IoU greater than 0.5
3. Break if no element in BB set
4. Go to 1

OBJECT RECOGNITION WITH HOG

Positive D.

Negative D.

Random patches

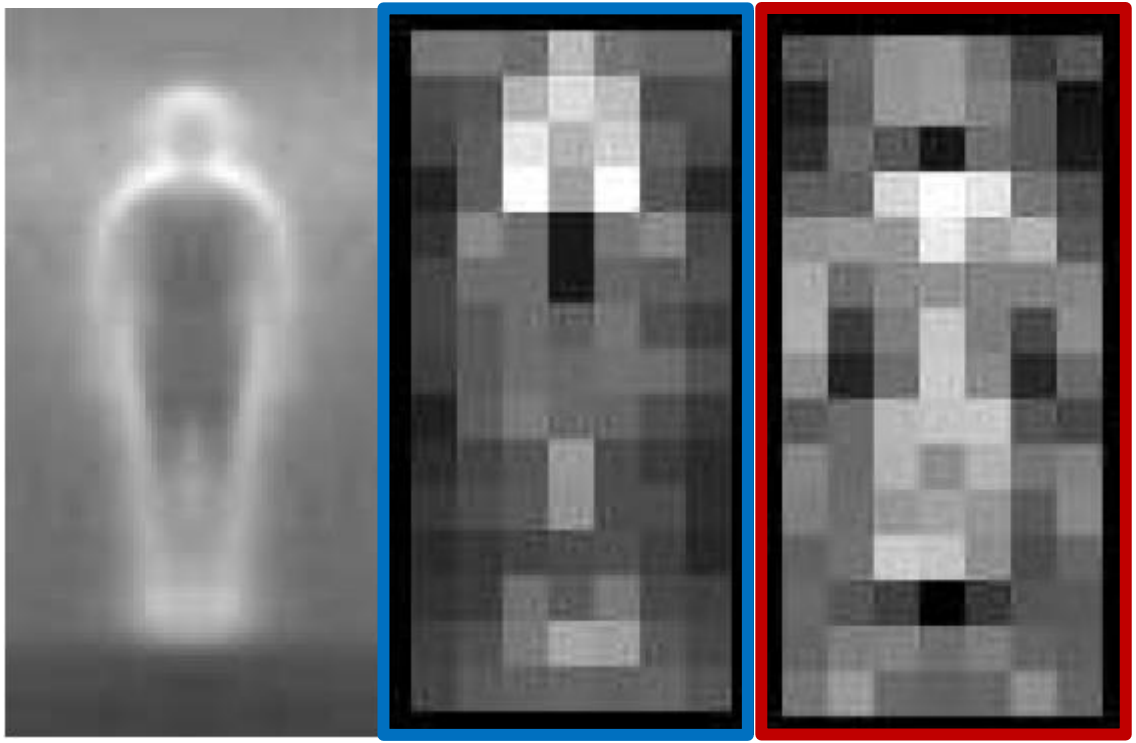


Max margin SVM classifier

$$\mathbf{x} \cdot \mathbf{w} + \mathbf{b} > 0 \qquad \text{Positive D.}$$

$$\mathbf{x} \cdot \mathbf{w} + \mathbf{b} < 0 \qquad \text{Negative D.}$$

OBJECT RECOGNITION WITH HOG

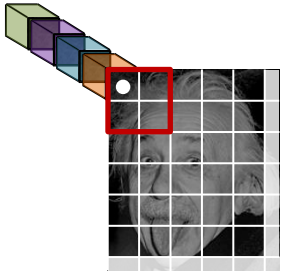


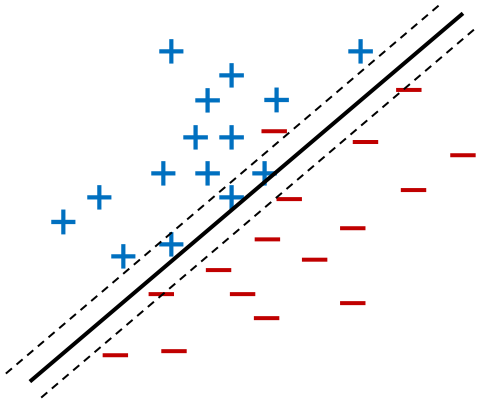
Average pedestrian

Positive

Negative

Maximum +/- SVM weight on each block

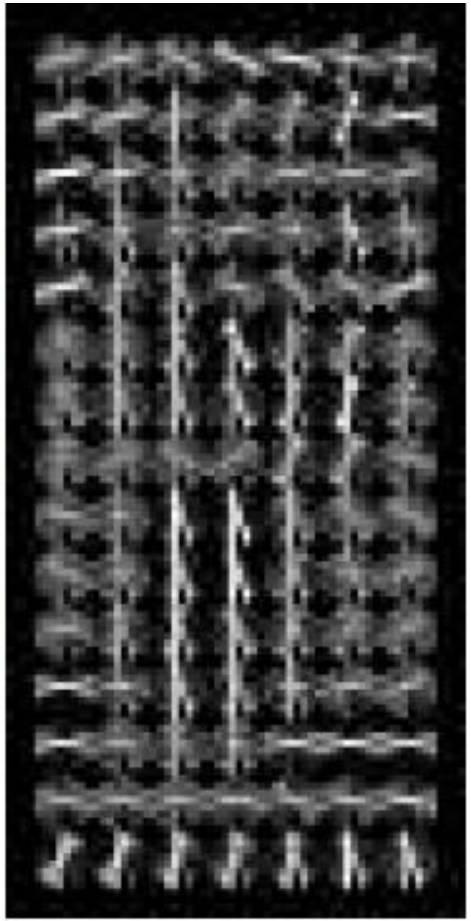




Max margin SVM classifier

$\mathbf{x} \cdot \mathbf{w} + \mathbf{b} > 0$ Positive D.

$\mathbf{x} \cdot \mathbf{w} + \mathbf{b} < 0$ Negative D.

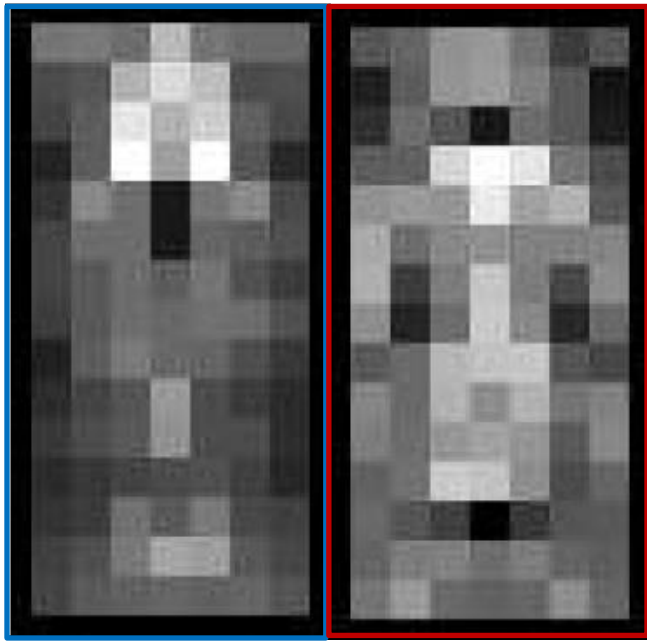


HOG

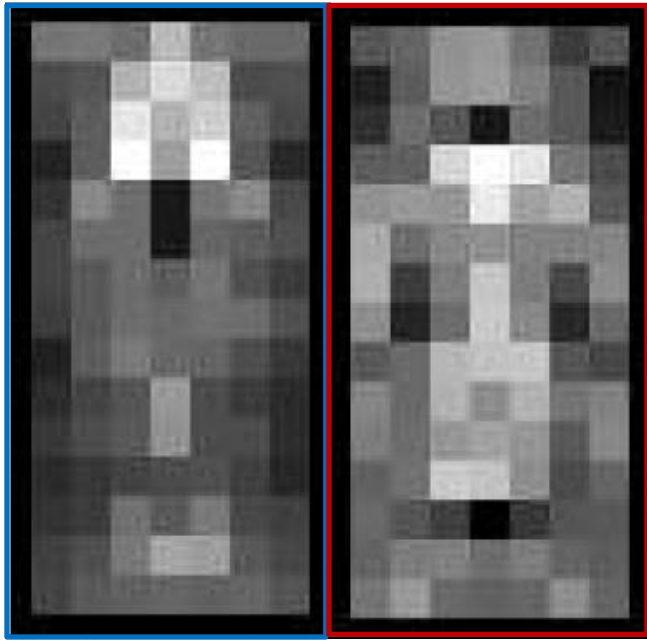
HOG weighted by pos. SVM weight

HOG weighted by neg. SVM weight

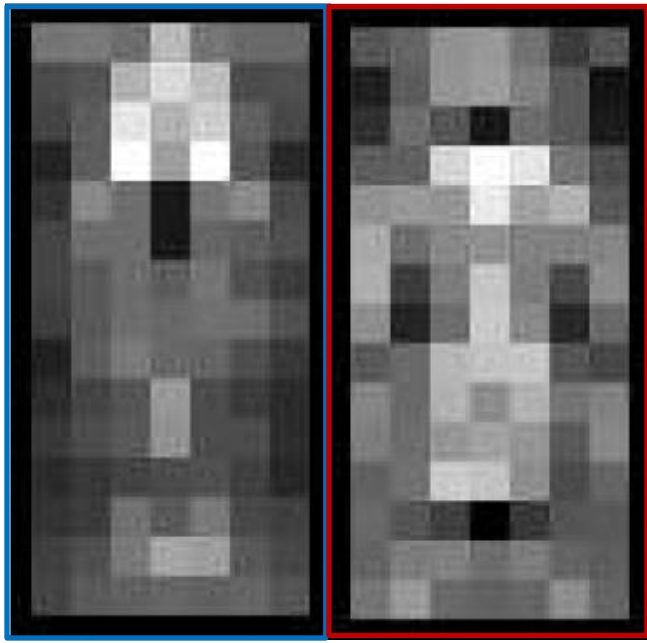
Object Recognition with HOG



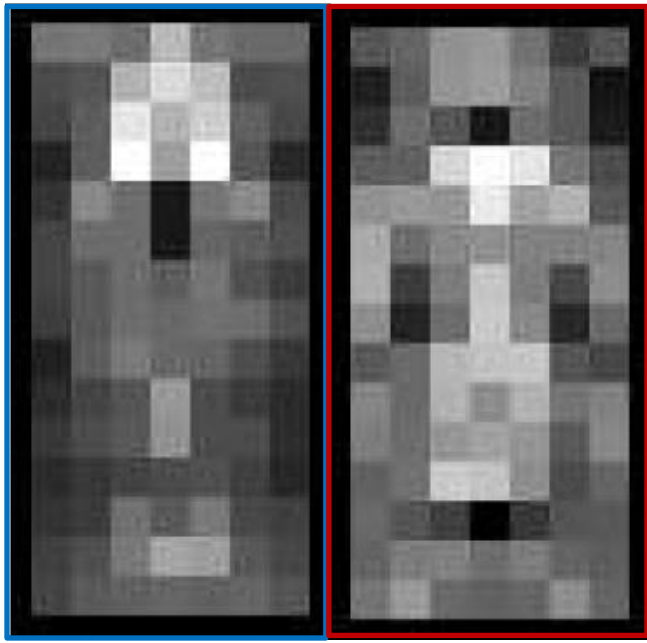
Object Recognition with HOG

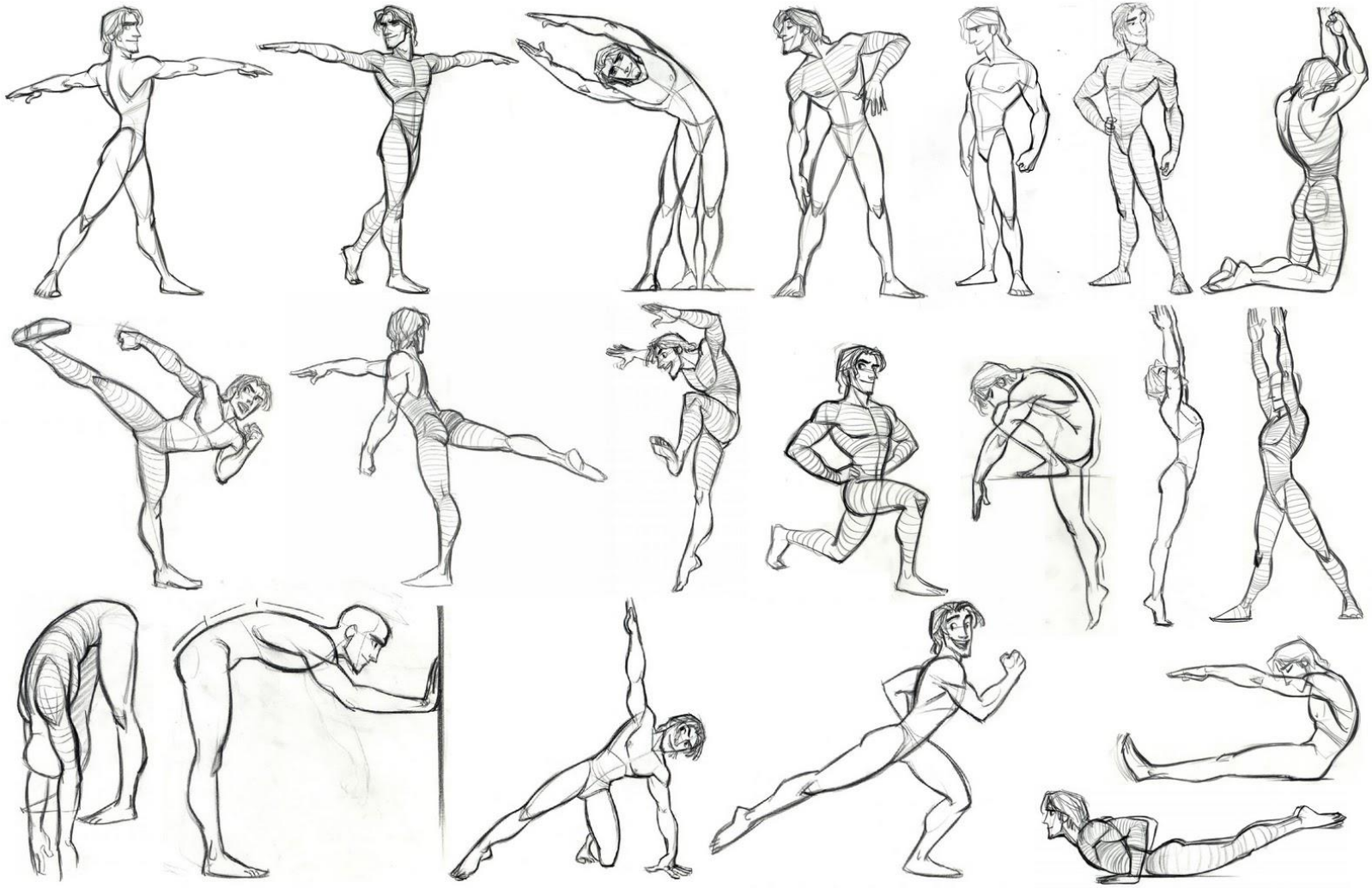


Object Recognition with HOG

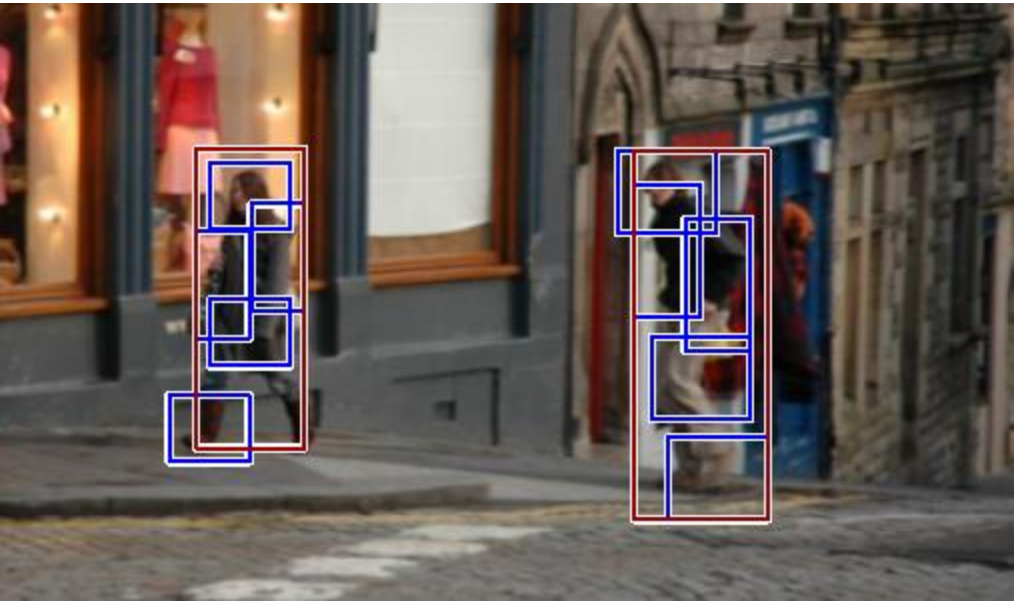


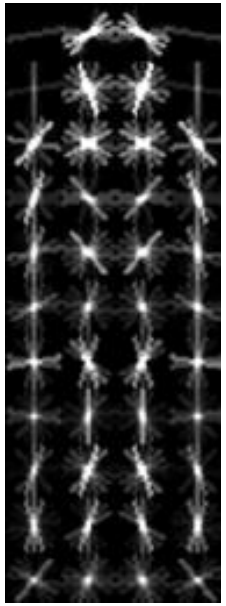
Object Recognition with HOG



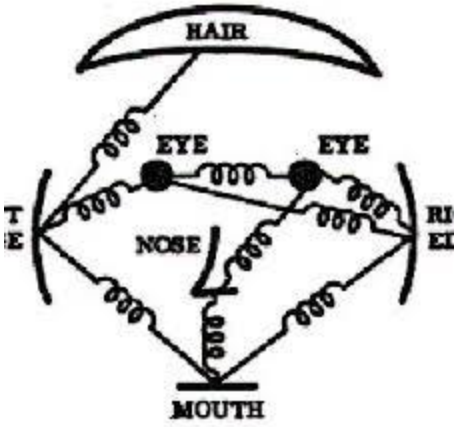


HOG Extension: Deformable Part Model (DPM)

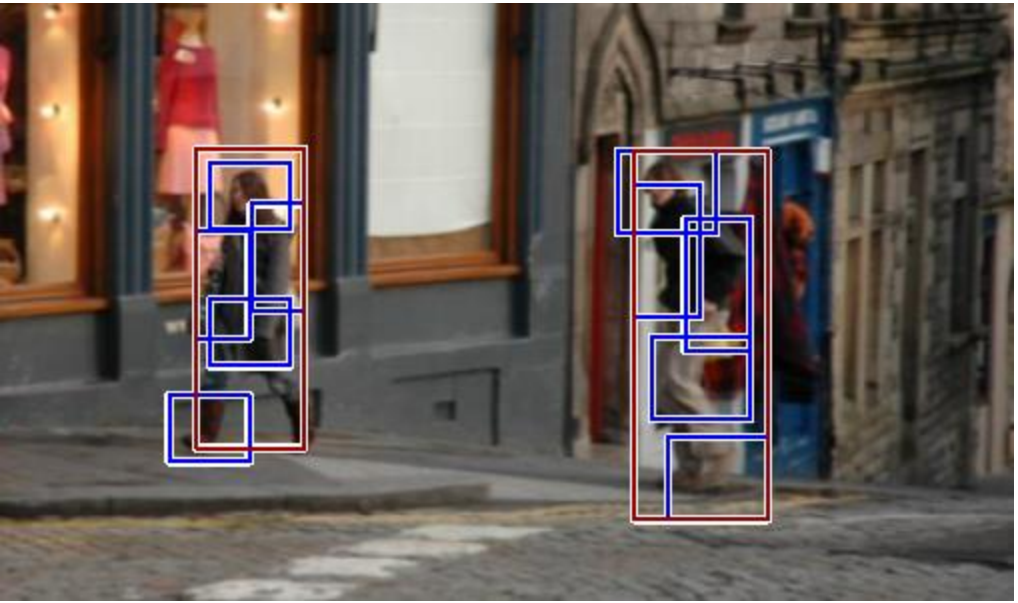


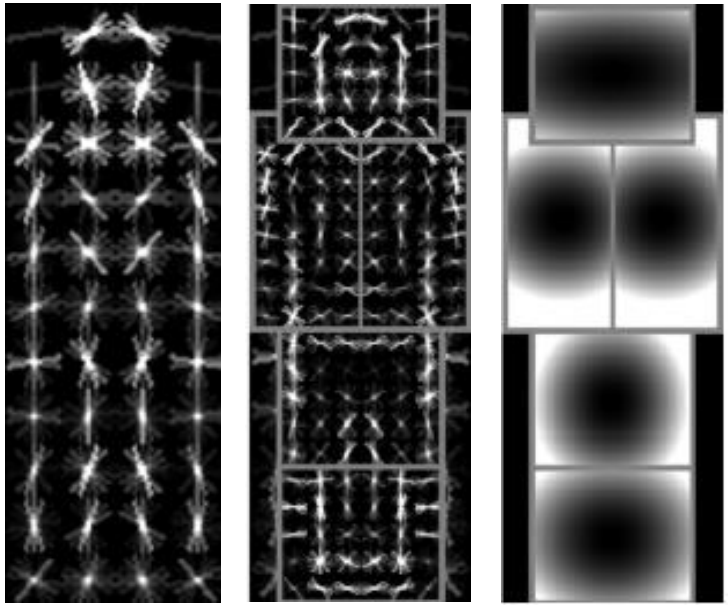


Root filter

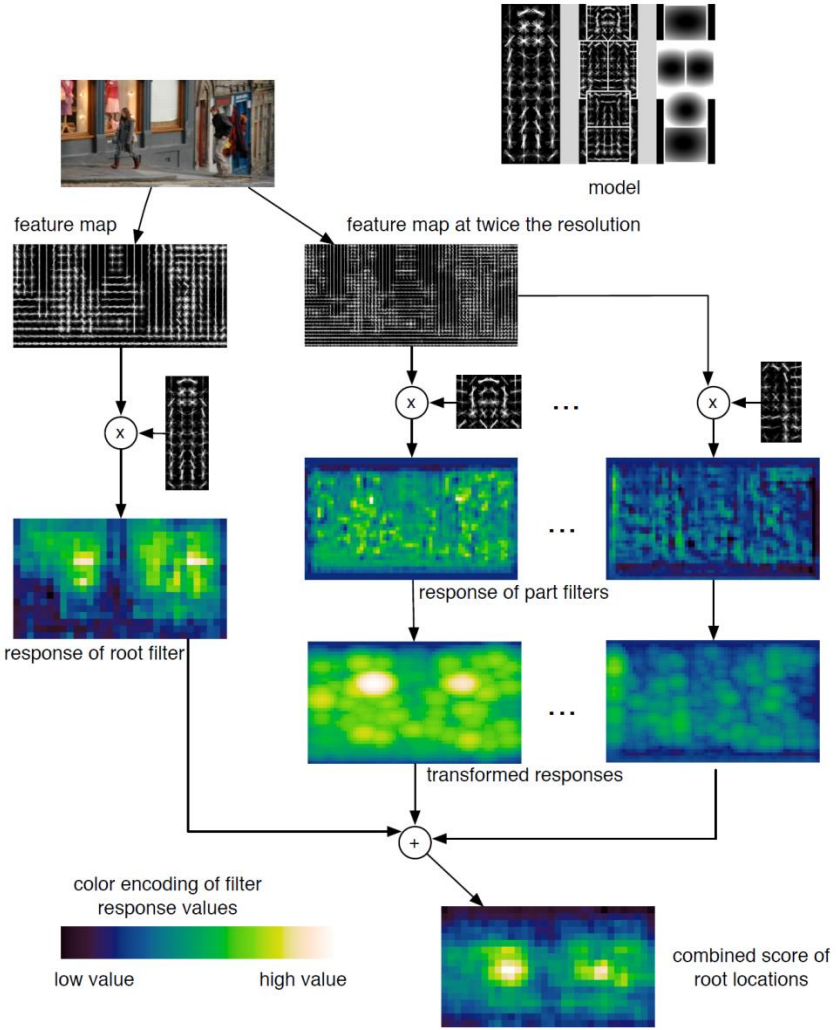


HOG Extension: Deformable Part Model (DPM)

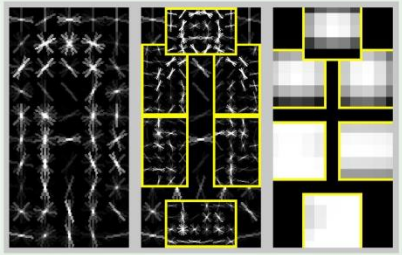




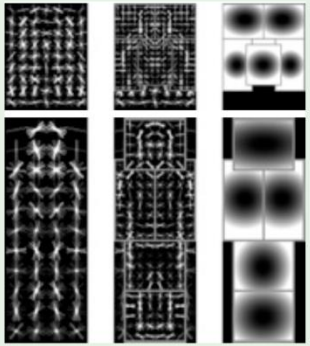
Root filter Part desc. Deformation


model
feature map
feature map at twice the resolution
x
x
x
response of part filters
response of root filter
transformed responses
+
color encoding of filter
response values
low value
high value
combined score of
root locations

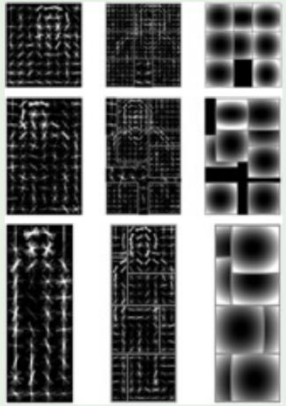
[DT'05]
AP 0.12



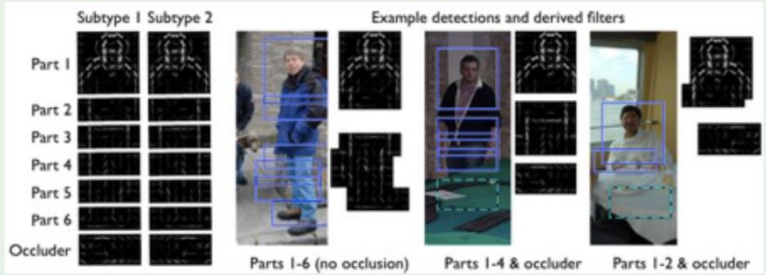
[FMR'08]
AP 0.27



[FGMR'10]
AP 0.36



[GFM voc-release5]
AP 0.45



[GFM'11]
AP 0.49

Slide by Ross Girshik

Autonomous Driving

Vehicle and pedestrian detection (HOG+SVM)
https://www.youtube.com/watch?v=itmV7druy9Y

https://www.youtube.com/watch?v=NpjixVTNmyw

https://www.youtube.com/watch?v=XRYFqd8ygXI