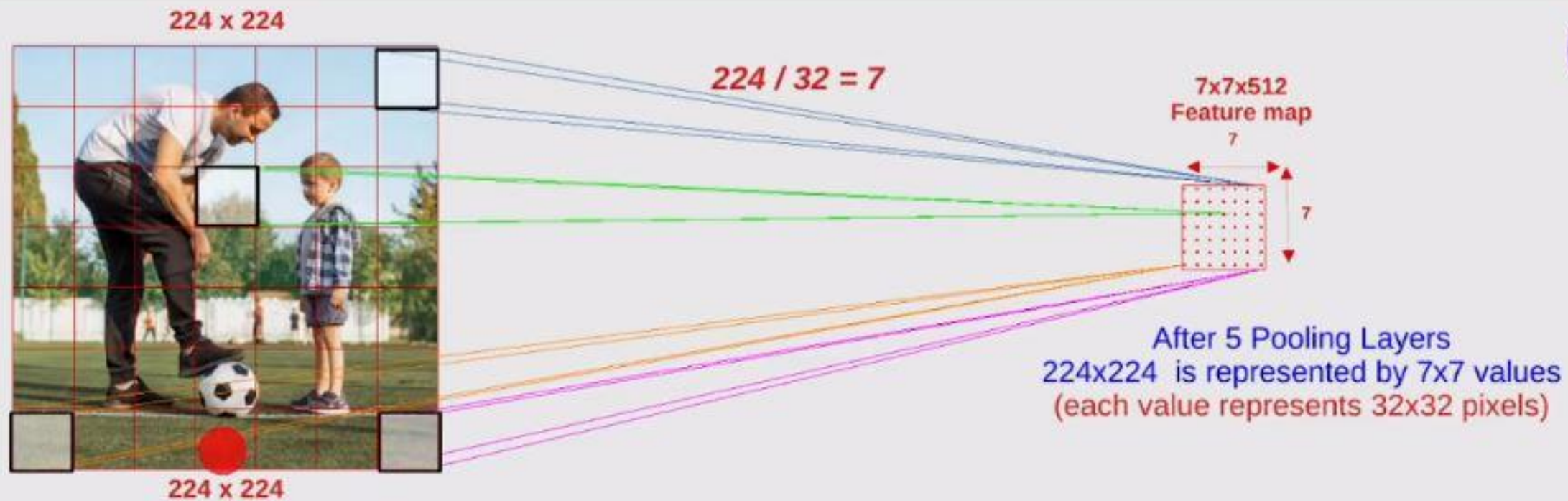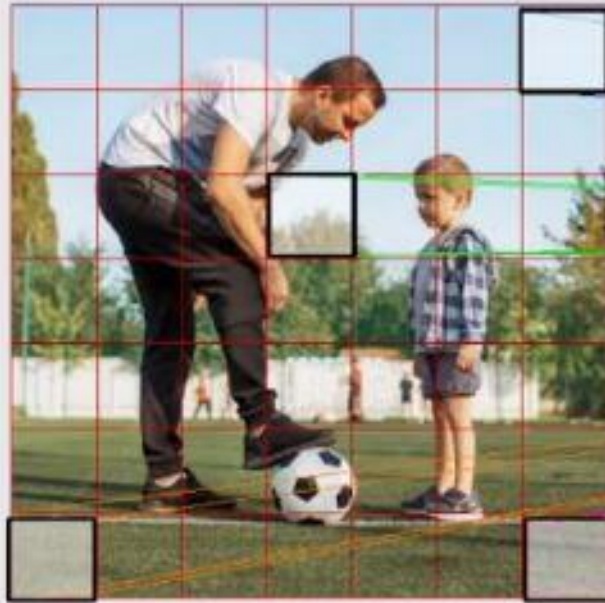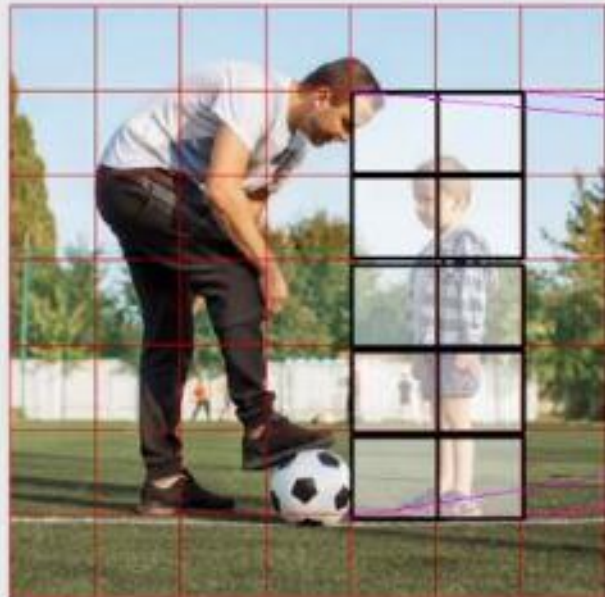# ROI Pooling

and

ROI Align Pooling

224 x 224

7x7x512
Feature map
7
7

After 5 Pooling Layers
224x224 is represented by 7x7 values
(each value represents 32x32 pixels)

224 x 224

7x7x512
Feature map
7
7

10 values in Feature Map are
representing the Boy

95

Effect of Shift and Crop of Feature Map

224 x 224

7x7x512
Feature map

7

7
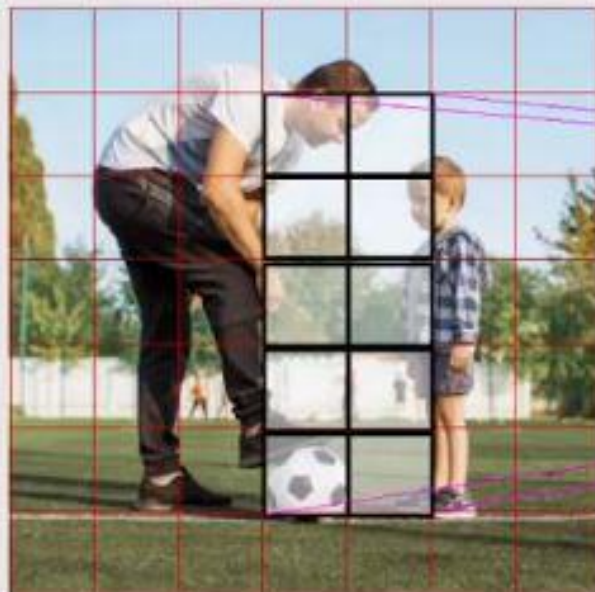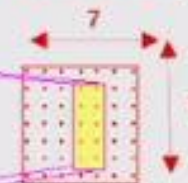
Effect of shift of values in feature map.
Missing Portion of the Boy
Adding portion of the father
[Error due to *Quantization of Origin coordinates*]

7x7x512
Feature map

7

7
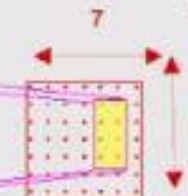
Effect of cropping some values in feature map.
Missing Portion of the Boy

[Error due to *Quantization of Width and Height*]

Effect of Shift and Crop of Feature Map

224 x 224
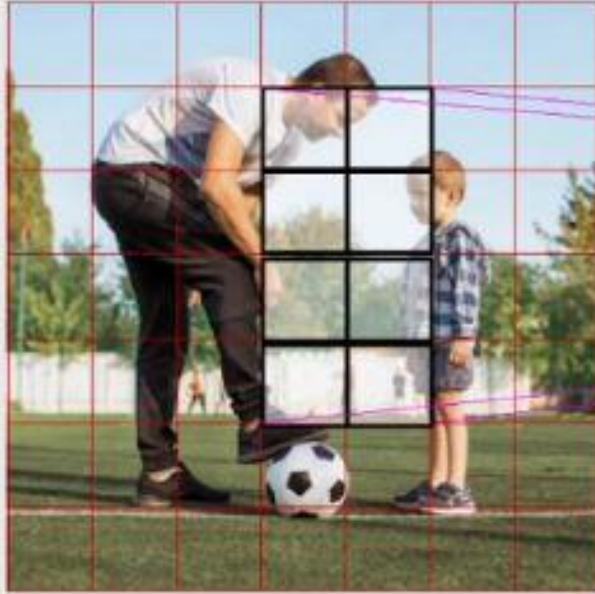
7x7x Chs
Feature map
7

7
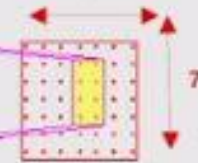
Double Quantization Errors

[Error due to Quantization of Origin coordinates]
[Error due to Quantization of Width and Height]

Quantization is a source of ERROR
But

Why do we need to Quantize ?

Region Of Interest (ROI) Pooling to 2 x 2 size

98

7x7xChs Feature map

7

7
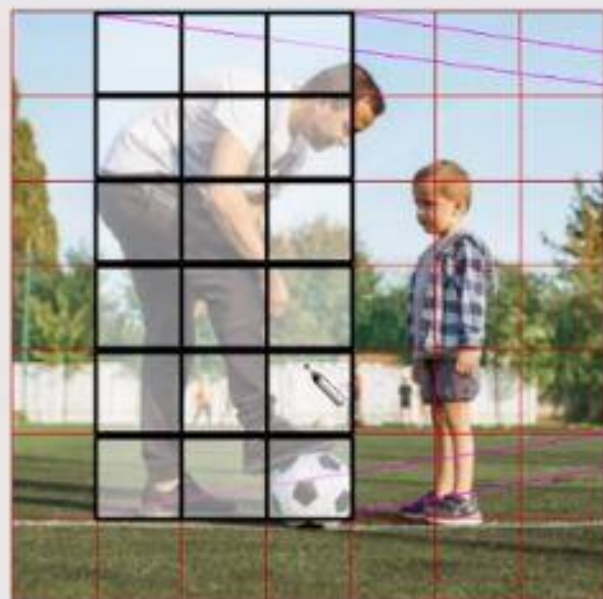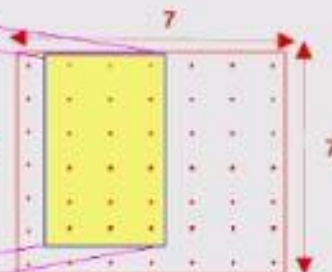
ROI Pooling 2 x 2

Object Feature Map
W=3, H=6
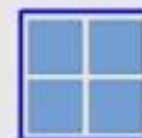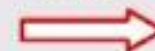
Size after Pooling
W=2, H=3

7x7xChs
Feature map

7

7
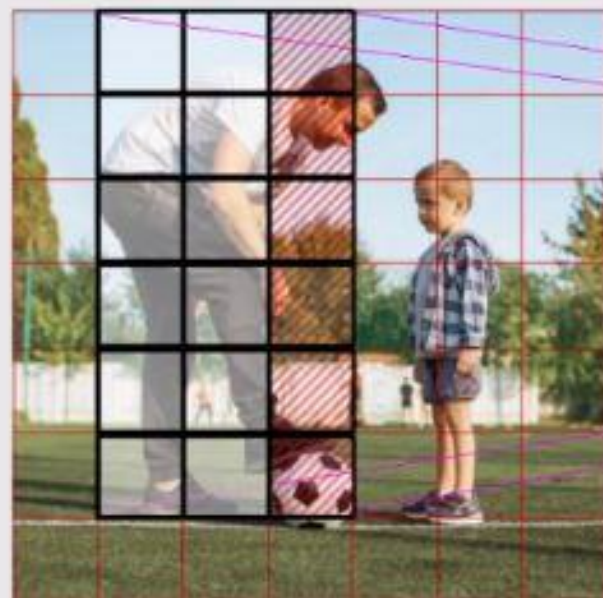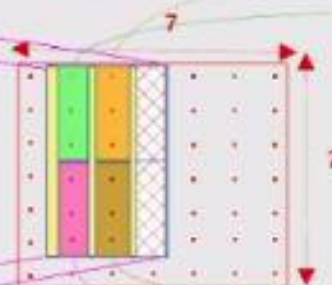
ROI Pooling
2 x 2

Object Feature Map
W=3, H=6
3/2=1 , 6/2=3
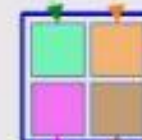
(Width Quantization error)

Region Of Interest (ROI) Pooling to 2 x 2 size

100

7x7xChs
Feature map

7

7
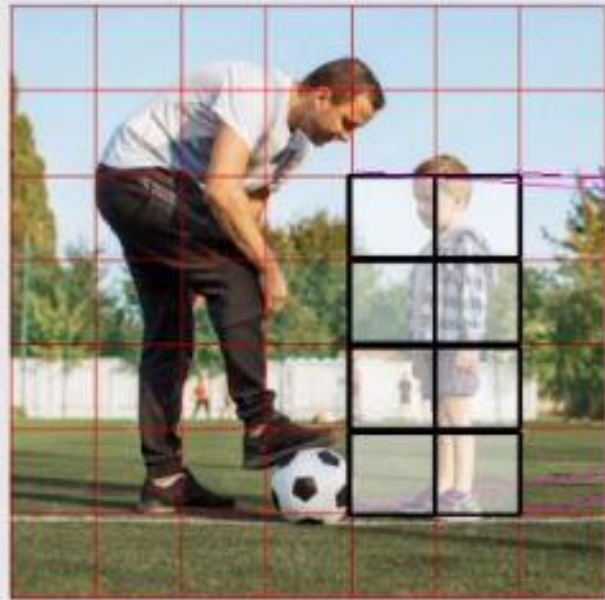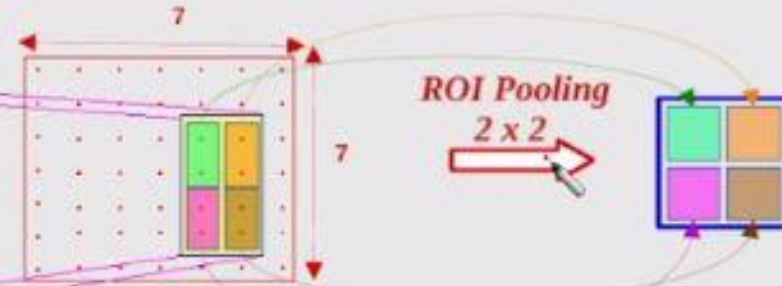
ROI Pooling
2 x 2

Object Feature Map
W=2, H=4
2/2=1 , 4/2=3
(NO Quantization error BUT !!!)

Size after Pooling
W=2, H=3
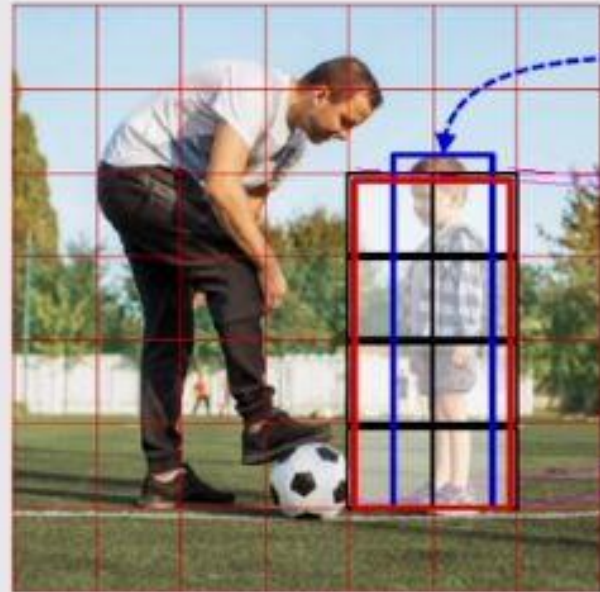
Ground Truth
Box

7x7xChs
Feature map
7

7
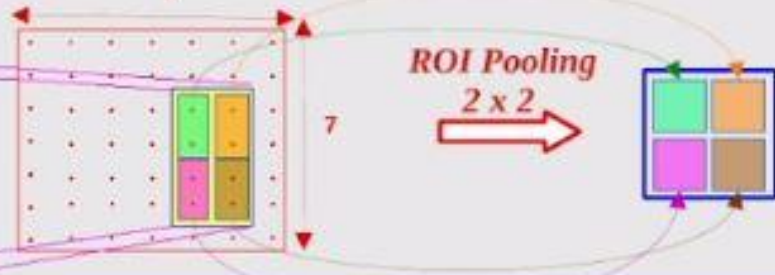
ROI Pooling
2 x 2

Box is NOT Properly aligned
Ground Truth Box (Blue)
Box Corresponding to selected values in Feature Map (Red)
Remember: Each value in Feature Map corresponds to 16 x 16 cell

ROI Pooling: Problem Definition

800 x 576 Image

25 x 18 Feature Map

A Pre-Trained CNN Model (with 5 Pooling Layers)

Region of Interest (ROI)
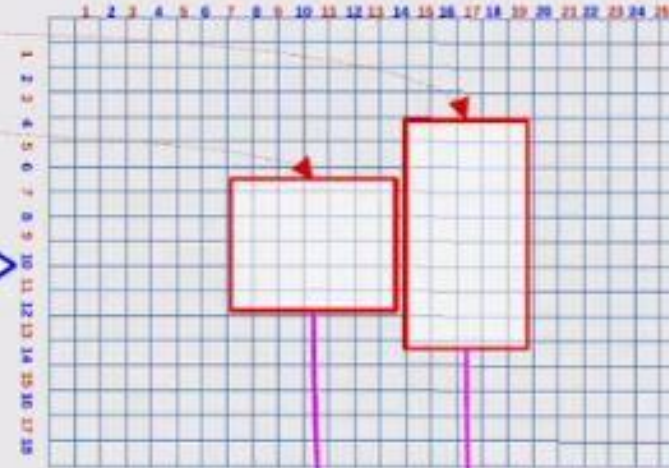
ROI Pooling
Output: 3x3
(irrespective of input size)
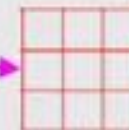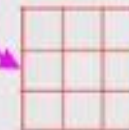
Fully Connected NN Classifier

Feature Map (800/32 , 576/32) = 25 x 18

0,0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.2	0.1	0.8	0.3	0.5	0.5	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6	0.5	0.1
2	0.5	0.3	0.6	0.8	0.1	0.2	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.2
3	0.4	0.5	0.7	0.2	0.4	0.9	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9	0.3	0.5
4	0.5	0.1	0.2	0.9	0.8	0.7	0.2	0.1	0.8	0.3	0.5	0.5	0.9	0.7	0.7	0.2	0.1	0.3	0.6	0.5	0.9	0.7	0.7	0.2	0.1
5	0.3	0.2	0.5	0.6	0.2	0.4	0.5	0.3	0.6	0.8	0.1	0.2	0.1	0.2	0.1	0.9	1.0	1.0	0.7	0.3	0.1	0.2	0.1	0.9	1.0
6	0.4	0.2	0.3	0.4	0.6	0.2	0.4	0.5	0.7	0.2	0.4	0.9	0.3	1.0	0.2	0.2	0.4	0.1	0.9	0.3	0.3	1.0	0.2	0.2	0.4
7	0.2	0.1	0.1	0.7	0.3	0.6	0.5	0.1	0.2	0.9	0.8	0.7	0.4	0.5	0.5	0.3	0.1	0.7	0.7	0.2	0.2	0.1	0.8	0.3	0.5
8	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.2	0.5	0.6	0.2	0.4	0.4	0.2	1.0	0.9	0.8	0.2	0.1	0.9	0.5	0.3	0.6	0.8	0.1
9	0.5	0.4	0.1	0.3	0.1	0.9	0.3	0.5	0.7	0.4	0.3	0.2	0.5	0.8	0.8	0.1	0.4	1.0	0.2	0.2	0.4	0.5	0.7	0.2	0.4
10	0.9	0.7	0.7	0.2	0.1	0.2	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3					0.3	0.5	0.1	0.2	0.9	0.8
11	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.6	0.8	0.3	0.2	0.6	0.6	0.2
12	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9	0.6	0.7	0.2	0.3	0.5	0.7	0.4	0.3
13	0.4	0.5	0.5	0.3	0.1	0.9	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.1	0.2	0.9	0.8	0.8	0.5	0.4	0.4
14	0.4	0.2	1.0	0.9	0.8	0.6	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	0.5	0.6	0.9	0.1	0.9	0.7	0.3
15	0.5	0.8	0.8	0.1	0.4	0.4	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.7	0.4	0.2	0.3	0.4	0.6	0.2
16	0.1	0.1	0.7	0.3	0.6	0.4	0.3	0.6	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6
17	1.0	0.2	0.6	1.0	0.7	0.7	1.0	0.7	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7
18	0.4	0.1	0.3	0.1	0.9	0.6	0.1	0.9	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.8	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9

Feature Map of the whole image   Width = 800, Height=576 , CNN with 5 Pooling Layers

Feature Map of selected object (Region of Interest [ROI]))

0,0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.2	0.1	0.8	0.3	0.5	0.5	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6	0.5	0.1
2	0.5	0.3	0.6	0.8	0.1	0.2	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.2
3	0.4	0.5	0.7	0.2	0.4	0.9	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9	0.3	0.5
4	0.5	0.1	0.2	0.9	0.8	0.7					0.5	0.5	0.9	0.7	0.7	0.2	0.1	0.3	0.6	0.5	0.9	0.7	0.7	0.2	0.1
5	0.3	0.2	0.5	0.6	0.2	0.4					0.1	0.2	0.1	0.2	0.1	0.9	1.0	1.0	0.7	0.3	0.1	0.2	0.1	0.9	1.0
6	0.4	0.2	0.3	0.4	0.6	0.2	0.4	0.5			0.9	0.3	1.0	0.2	0.2	0.4	0.1	0.9	0.3	0.3	1.0	0.2	0.2	0.4	
7	0.2	0.1	0.1	0.7	0.3	0.6	0.5	0.1	0.2	0.9	0.8	0.7	0.4	0.5	0.5	0.3	0.1	0.7	0.7	0.2	0.2	0.1	0.8	0.3	0.5
8	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.2	0.5	0.6	0.2	0.4	0.4	0.2	1.0	0.9	0.8	0.2	0.1	0.9	0.5	0.3	0.6	0.8	0.1
9	0.5	0.4	0.1	0.3	0.1	0.9	0.3	0.5	0.7	0.4	0.3	0.2	0.5	0.8	0.8	0.1	0.4	1.0	0.2	0.2	0.4	0.5	0.7	0.2	0.4
10	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6	0.1	0.8	0.3	0.5	0.1	0.2	0.9	0.8
11	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.6	0.8	0.3	0.2	0.5	0.6	0.2
12	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9	0.5	0.7	0.2	0.3	0.5	0.7	0.4	0.3
13	0.4	0.5	0.5	0.3	0.1	0.9	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.1	0.1	0.2	0.9	0.8	0.8	0.5	0.4	0.4
14	0.4	0.2	1.0	0.9	0.8	0.6	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.2	0.5	0.6	0.9	0.1	0.9	0.7	0.3
15	0.5	0.8	0.8	0.1	0.4	0.4	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6					3	0.4	0.6	0.2	
16	0.1	0.1	0.7	0.3	0.6	0.4	0.3	0.6	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8			1	0.7	0.3	0.6			
17	1.0	0.2	0.6	1.0	0.7	0.7	1.0	0.7	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7
18	0.4	0.1	0.3	0.1	0.9	0.6	0.1	0.9	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9

300/32=9.375
210/32=6.563

(9.375,6.563)

(16.563,12.188)

7.188

(300+230)/32=16.563
(210+180)/32=12.188

Feature Map Corresponding to Object at x,y= (300,210) w,h=(230,180)

ROI Pooling Output is 3x3

0,0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.2	0.1	0.8	0.3	0.5	0.5	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6	0.5	0.1
2	0.5	0.3	0.6	0.8	0.1	0.2	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.2
3	0.4	0.5	0.7	0.2	0.4	0.9	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9	0.3	0.5
4	0.5	0.1	0.2	0.9	0.8	0.7	0.2	0.1	0.8	0.3	0.5	0.5	0.9	0.7	0.7	0.2	0.1	0.3	0.6	0.5	0.9	0.7	0.7	0.2	0.1
5	0.3	0.2	0.5	0.6	0.2	0.4	0.5	0.3	0.6	0.8	0.1	0.2	0.1	0.2	0.1	0.9	1.0	1.0	0.7	0.3	0.1	0.2	0.1	0.9	1.0
6	0.4	0.2	0.3	0.4	0.6	0.2	0.4	0.5	(9.375,6.563)			0.9	0.3	1.0	0.2	0.2	0.4	0.1	0.9	0.3	0.3	1.0	0.2	0.2	0.4
7	0.2	0.1	0.1	0.7	0.3	0.6	0.5	0.1	0.2	0.9	0.8	0.7	0.4	0.5	0.5	0.3	0.1	0.7	0.7	0.2	0.2	0.1	0.8	0.3	0.5
8	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.2	0.5	0.6	0.2	0.4	0.4	0.2	1.0	0.9	0.8	0.2	0.1	0.9	0.5	0.3	0.6	0.8	0.1
9	0.5	0.4	0.1	0.3	0.1	0.9	0.3	0.5	0.7	0.4	0.3	0.2	0.5	0.8	0.8	0.1	0.4	1.0	0.2	0.2	0.4	0.5	0.7	0.2	0.4
10	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6	0.1	0.8	0.3	0.5	0.1	0.2	0.9	0.8
11	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.6	0.8	0.3	0.2	0.5	0.6	0.2
12	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9	0.5	0.7	0.2	0.3	0.5	0.7	0.4	0.3
13	0.4	0.5	0.5	0.3	0.1	0.9	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	(16.563,12.188)		0.9	0.8	0.8	0.5	0.4	0.4	
14	0.4	0.2	1.0	0.9	0.8	0.6	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.2	0.5	0.6	0.9	0.1	0.9	0.7	0.3
15	0.5	0.8	0.8	0.1	0.4	0.4	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.7	0.4	0.2	0.3	0.4	0.6	0.2
16	0.1	0.1	0.7	0.3	0.6	0.4	0.3	0.6	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6
17	1.0	0.2	0.6	1.0	0.7	0.7	1.0	0.7	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7
18	0.4	0.1	0.3	0.1	0.9	0.6	0.1	0.9	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9

Feature Map of ROI (object) will be pooled to 3x3 (Align Pooling)

ROI Pooling Output is 3x3 (Irrespective of Object Size / Aspect Ratio)

106

Object 1

Object 2

(9.375,6.563)

(16.563,12.188)
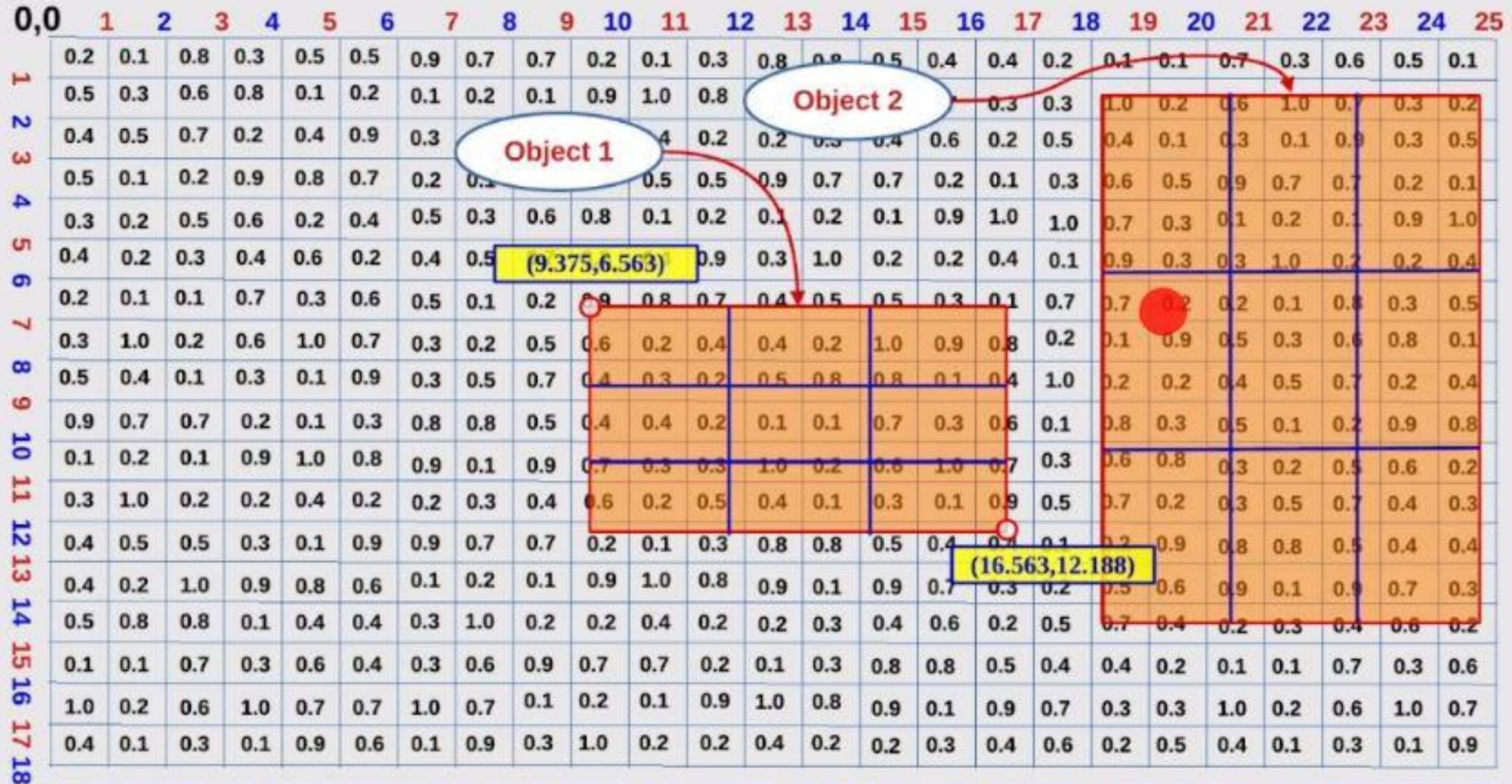
Feature Map Corresponding to two Objects with Different Sizes (all will be Pooled to 3x3)

Sampling Points "S_x" represent values inside each cell *(4 Points)*

0,0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.2	0.1	0.8	0.3	0.5	0.5	0.9	0.7	0.7	0.2	0.1	0.3	0.8		0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6	0.5	0.1
2	0.5	0.3	0.6	0.8	0.1	0.2	0.1	0.2	0.1	0.9	1.0	0.8				0.3	0.3								
3	0.4	0.5	0.7	0.2	0.4	0.9	0.3				0.2	0.2		0.4	0.6	0.2	0.5								
4	0.5	0.1	0.2	0.9	0.8	0.7	0.2			0.5	0.5	0.9	0.7	0.7	0.2	0.1	0.3								
5	0.3	0.2	0.5	0.6	0.2	0.4	0.5	0.3	0.6	0.8	0.1	0.2	0.1	0.2	0.1	0.9	1.0	1.0							
6	0.4	0.2	0.3	0.4	0.6	0.2	0.4	0.5	(9.375,6.563)			0.9	0.3	1.0	0.2	0.2	0.4	0.1							
7	0.2	0.1	0.1	0.7	0.3	0.6	0.5	0.1	0.2		0.8	0.7	0.4	0.5	0.5	0.3	0.1	0.7							
8	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.2	0.5	0.6	0.2	0.4	0.4	0.2	1.0	0.9	0.8	0.2							
9	0.5	0.4	0.1	0.3	0.1	0.9	0.3	0.5	0.7	0.4	0.3	0.2	0.5	0.8	0.8	0.1	0.4	1.0							
10	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6	0.1							
11	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7	0.3							
12	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.6	0.4	0.3	0.1	0.9	0.5							
13	0.4	0.5	0.5	0.3	0.1	0.9	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	(16.563,12.188)								
14	0.4	0.2	1.0	0.9	0.8	0.6	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.2							
15	0.5	0.8	0.8	0.1	0.4	0.4	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5							
16	0.1	0.1	0.7	0.3	0.6	0.4	0.3	0.6	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6
17	1.0	0.2	0.6	1.0	0.7	0.7	1.0	0.7	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7
18	0.4	0.1	0.3	0.1	0.9	0.6	0.1	0.9	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9

Object 1
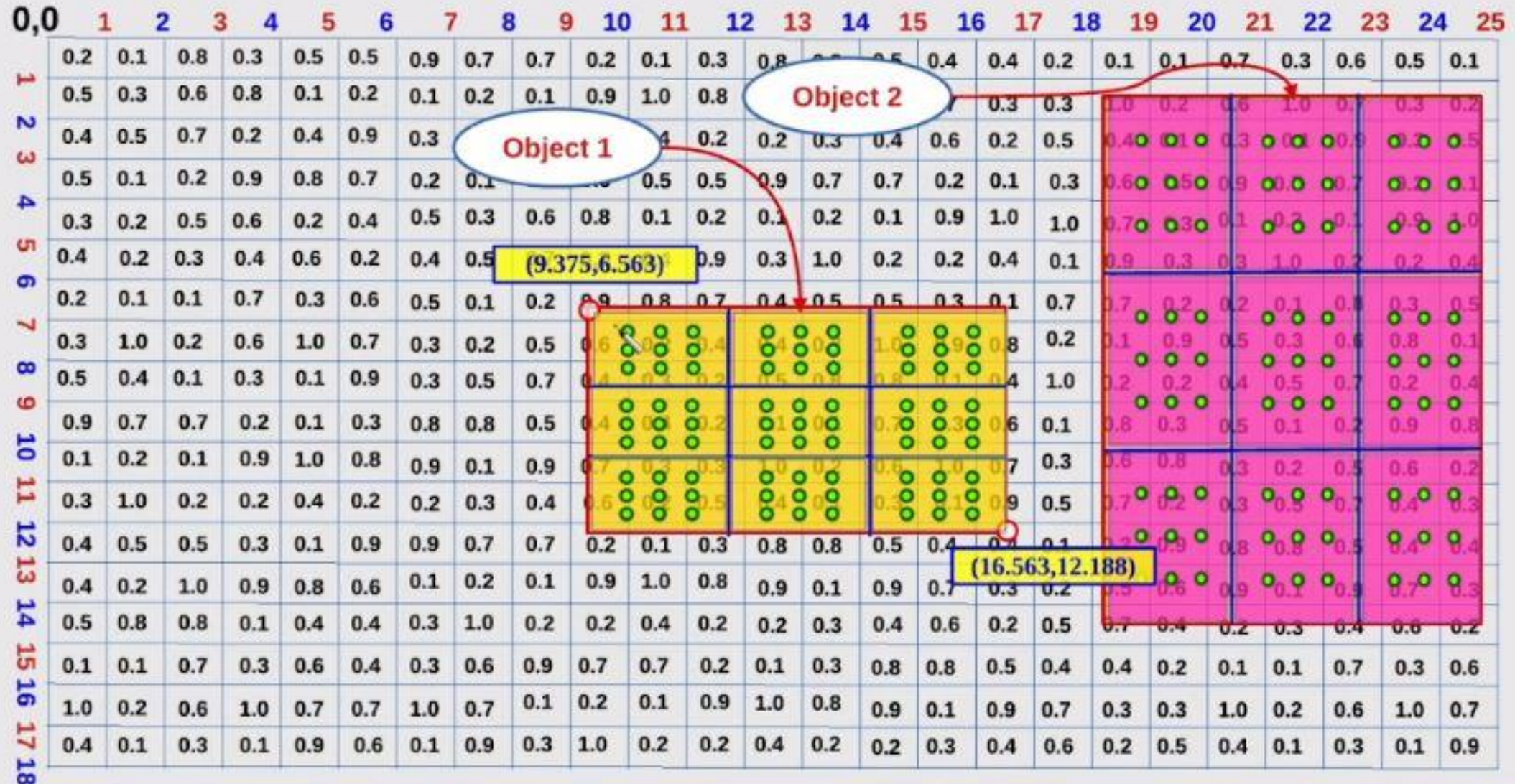
Object 2

Select <u>2x2</u> Sampling Points *(Calculate Corresponding Values then apply Max Pooling)*

Sampling Points "S_x" represent values inside each cell *(9 Points)*
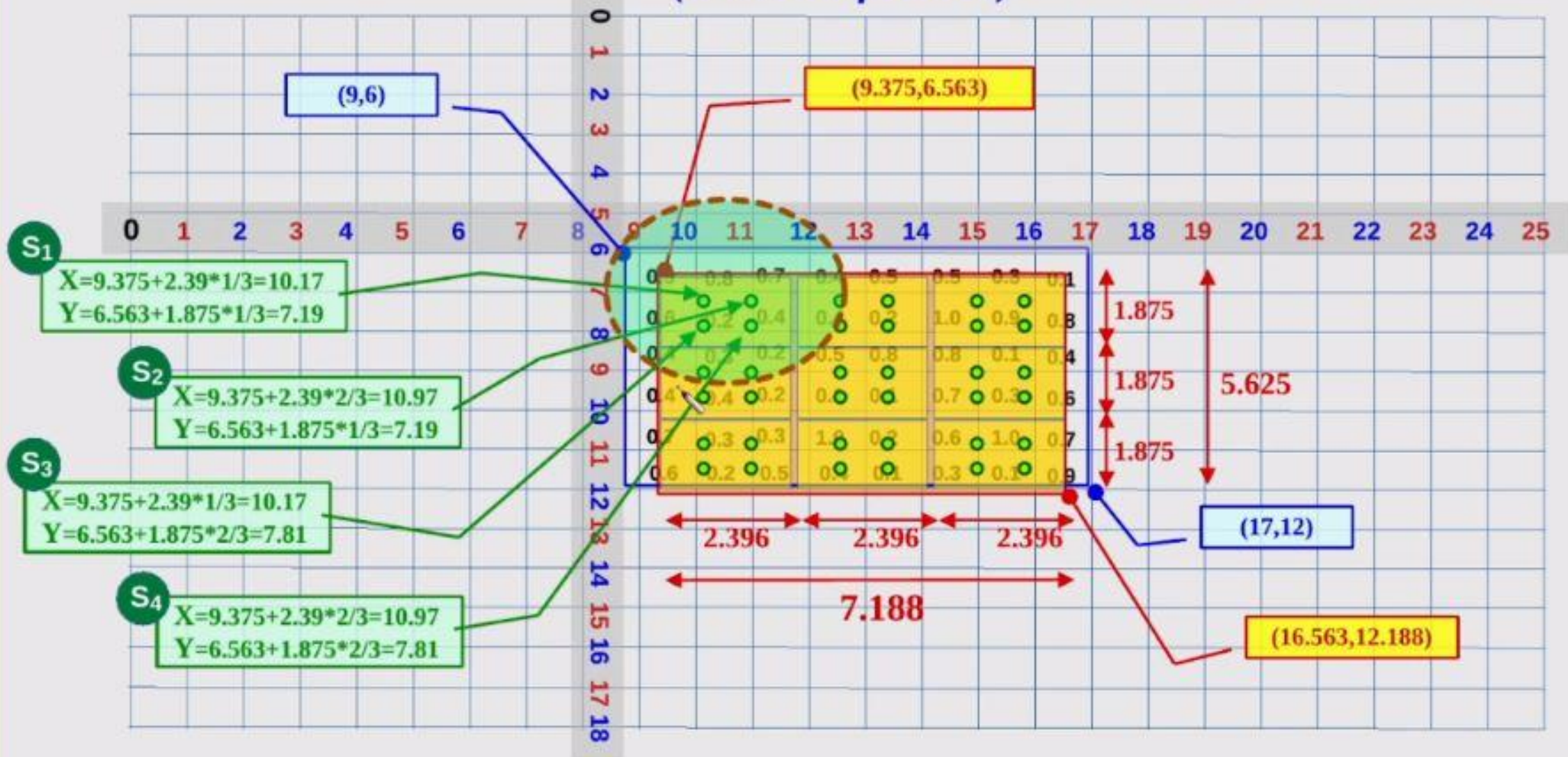
0,0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.2	0.1	0.8	0.3	0.5	0.5	0.9	0.7	0.7	0.2	0.1	0.3	0.8		0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6	0.5	0.1
2	0.5	0.3	0.6	0.8	0.1	0.2	0.1	0.2	0.1	0.9	1.0	0.8				0.3	0.3								
3	0.4	0.5	0.7	0.2	0.4	0.9	0.3				0.2	0.2	0.3	0.4	0.6	0.2	0.5								
4	0.5	0.1	0.2	0.9	0.8	0.7	0.2	0.1		0.5	0.5	0.9	0.7	0.7	0.2	0.1	0.3								
5	0.3	0.2	0.5	0.6	0.2	0.4	0.5	0.3	0.6	0.8	0.1	0.2	0.1	0.2	0.1	0.9	1.0	1.0							
6	0.4	0.2	0.3	0.4	0.6	0.2	0.4	0.5	(9.375,6.563)		0.9	0.3	1.0	0.2	0.2	0.4	0.1								
7	0.2	0.1	0.1	0.7	0.3	0.6	0.5	0.1	0.2	0.9	0.8	0.7	0.4	0.5	0.5	0.3	0.1	0.7							
8	0.3	1.0	0.2	0.6	1.0	0.7	0.3	0.2	0.5								0.8	0.2							
9	0.5	0.4	0.1	0.3	0.1	0.9	0.3	0.5	0.7									1.0							
10	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5									0.1							
11	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9									0.3							
12	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4									0.5							
13	0.4	0.5	0.5	0.3	0.1	0.9	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	(16.563,12.188)								
14	0.4	0.2	1.0	0.9	0.8	0.6	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7									
15	0.5	0.8	0.8	0.1	0.4	0.4	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5							
16	0.1	0.1	0.7	0.3	0.6	0.4	0.3	0.6	0.9	0.7	0.7	0.2	0.1	0.3	0.8	0.8	0.5	0.4	0.4	0.2	0.1	0.1	0.7	0.3	0.6
17	1.0	0.2	0.6	1.0	0.7	0.7	1.0	0.7	0.1	0.2	0.1	0.9	1.0	0.8	0.9	0.1	0.9	0.7	0.3	0.3	1.0	0.2	0.6	1.0	0.7
18	0.4	0.1	0.3	0.1	0.9	0.6	0.1	0.9	0.3	1.0	0.2	0.2	0.4	0.2	0.2	0.3	0.4	0.6	0.2	0.5	0.4	0.1	0.3	0.1	0.9

Object 1

Object 2

Select 3x3 Sampling Points *(Calculate Corresponding Values then apply Max Pooling)*
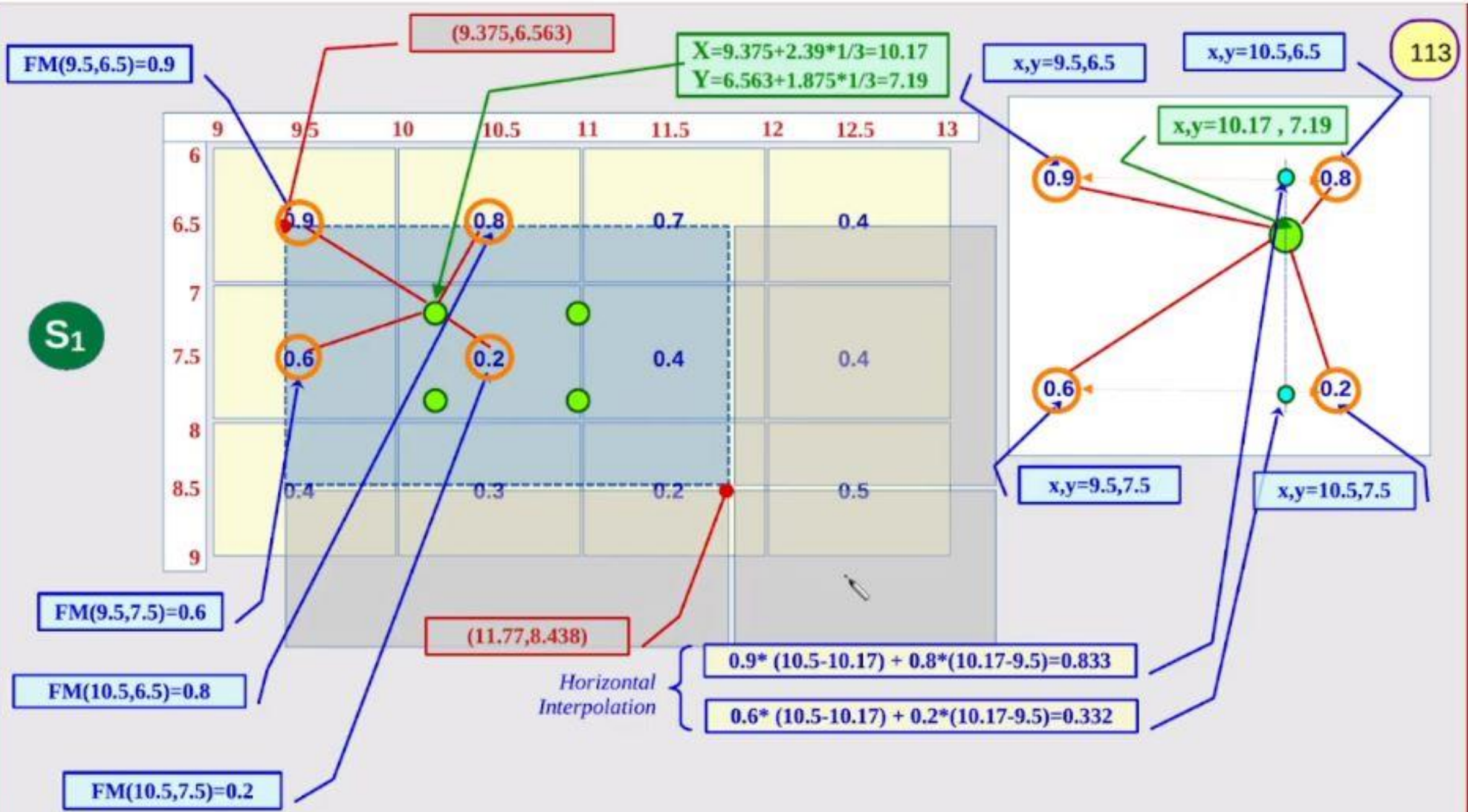
Sampling Points Coordinates

109

(300/32 , 210/32)

(9.375,6.563)

(9,6)

S₁
X=9.375+2.39*1/3=10.17
Y=6.563+1.875*1/3=7.19

S₂
X=9.375+2.39*2/3=10.97
Y=6.563+1.875*1/3=7.19

S₃
X=9.375+2.39*1/3=10.17
Y=6.563+1.875*2/3=7.81

S₄
X=9.375+2.39*2/3=10.97
Y=6.563+1.875*2/3=7.81

1.875

1.875

5.623 (180/32)

1.875

(17,12)

2.396   2.396   2.396

7.188

(230/32)

(16.563,12.188)

Feature Map Corresponding to Object at   x,y= (300,210)   w,h=(230,180) , 5 Times Pooling

Sampling Points Coordinates
(Focus on first cell)

(9,6)

(9.375,6.563)

S_1

X=9.375+2.39*1/3=10.17
Y=6.563+1.875*1/3=7.19

S_2

X=9.375+2.39*2/3=10.97
Y=6.563+1.875*1/3=7.19

S_3

X=9.375+2.39*1/3=10.17
Y=6.563+1.875*2/3=7.81

S_4

X=9.375+2.39*2/3=10.97
Y=6.563+1.875*2/3=7.81

1.875
1.875
1.875
5.625

2.396   2.396   2.396

7.188

(17,12)

(16.563,12.188)

110

FM(9.5,6.5)=0.9

(9.375,6.563)

$X=9.375+2.39*1/3=10.17$
$Y=6.563+1.875*1/3=7.19$

x,y=9.5,6.5

x,y=10.5,6.5

x,y=10.175,7.19

S_1

FM(9.5,7.5)=0.6

x,y=9.5,7.5

x,y=10.5,7.5

FM(10.5,6.5)=0.8

(11.77,8.438)

Horizontal Interpolation

$Prev\_Val*(X_{next}-X)+Next\_Val*(X-X_{Prev})$

$Prev\_Val*(X_{next}-X)+Next\_Val*(X-X_{Prev})$

FM(10.5,7.5)=0.2

Prev_Val   Next_Val

X_Prev   X   X_Next

FM(9.5,6.5)=0.9

(9.375,6.563)

X=9.375+2.39*1/3=10.17
Y=6.563+1.875*1/3=7.19

x,y=9.5,6.5

x,y=10.5,6.5

113

x,y=10.17 , 7.19

S₁

FM(9.5,7.5)=0.6

(11.77,8.438)

FM(10.5,6.5)=0.8

x,y=9.5,7.5

x,y=10.5,7.5

Horizontal Interpolation

0.9* (10.5-10.17) + 0.8*(10.17-9.5)=0.833

0.6* (10.5-10.17) + 0.2*(10.17-9.5)=0.332

FM(10.5,7.5)=0.2

FM(9.5,7.5)=0.6

(9.375,6.563)

X=9.375+2.39*1/3=10.17
Y=6.563+1.875*2/3=7.81

x,y=9.5,7.5

x,y=10.5,7.5

116

x,y=10.17 , 7.81

S₃

FM(9.5,8.5)=0.4

(11.77,8.438)

FM(10.5,7.5)=0.2

x,y=9.5,8.5

x,y=10.5,8.5

Horizontal Interpolation

0.6* (10.5-10.17) + 0.2*(10.17-9.5)=0.332

0.4* (10.5-10.17) + 0.3*(10.17-9.5)=0.333

Vertical Interpolation

0.332* (8.5-7.81) + 0.333*(7.81-7.5)=0.332

FM(10.5,8.5)=0.3

FM(10.5,7.5)=0.2

(9.375,6.563)

X=9.375+2.39*2/3=10.97
Y=6.563+1.875*2/3=7.81

x,y=10.5,7.5

x,y=11.5,7.5

117

x,y=10.97 , 7.81

S₄

FM(10.5,8.5)=0.3

(11.77,8.438)

x,y=10.5,8.5

x,y=11.5,8.5

FM(11.5,7.5)=0.4

Horizontal Interpolation
0.2* (11.5-10.97) + 0.4*(10.97-10.5)=0.294
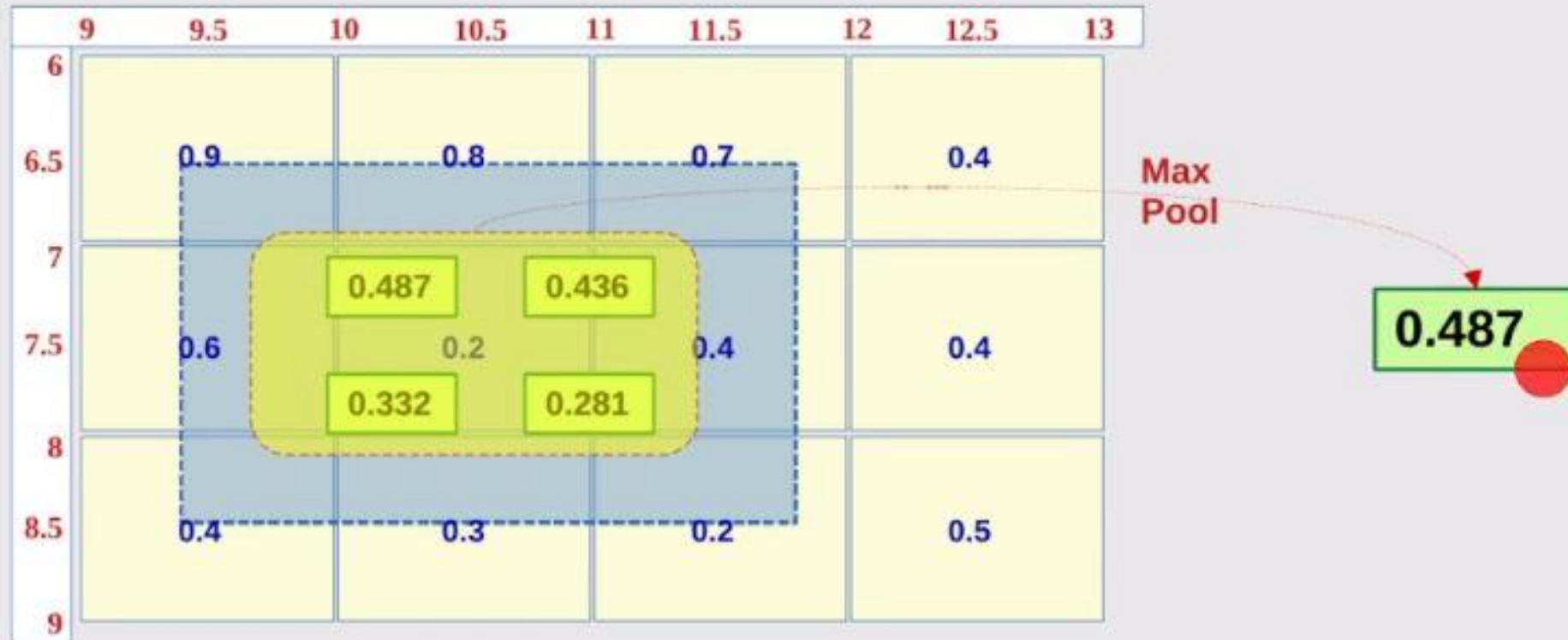0.3* (11.5-10.97) + 0.2*(10.97-10.5)=0.253

FM(11.5,8.5)=0.2

Vertical Interpolation
0.294* (8.5-7.81) + 0.253*(7.81-7.5)=0.281

Apply Max-Pooling on the Calculated Sampling Points

*Repeat for All other Cells Within the Region Of Interest (ROI)
to obtain 3x3 values*
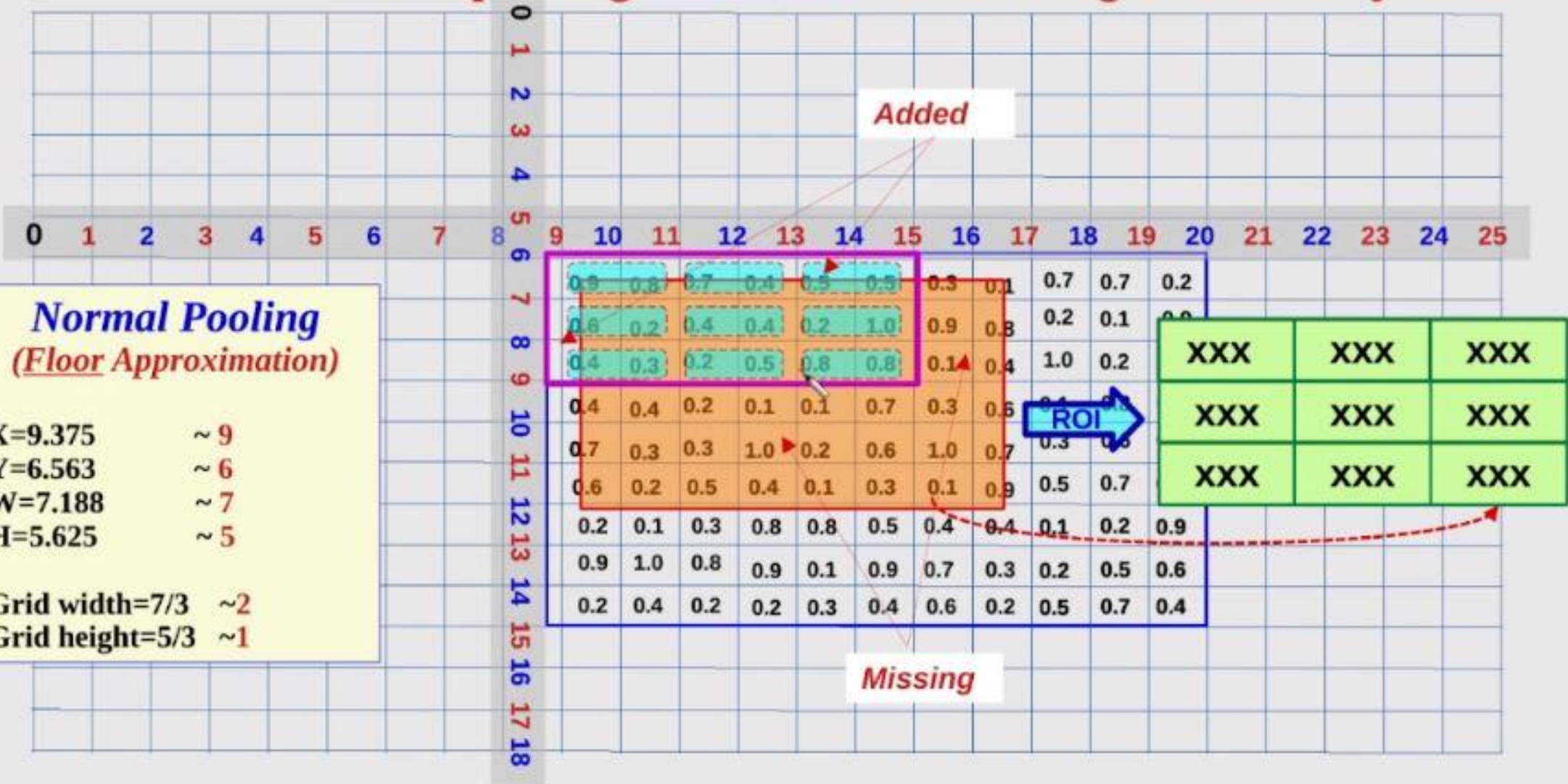
3x3 ROI Align for selected Object
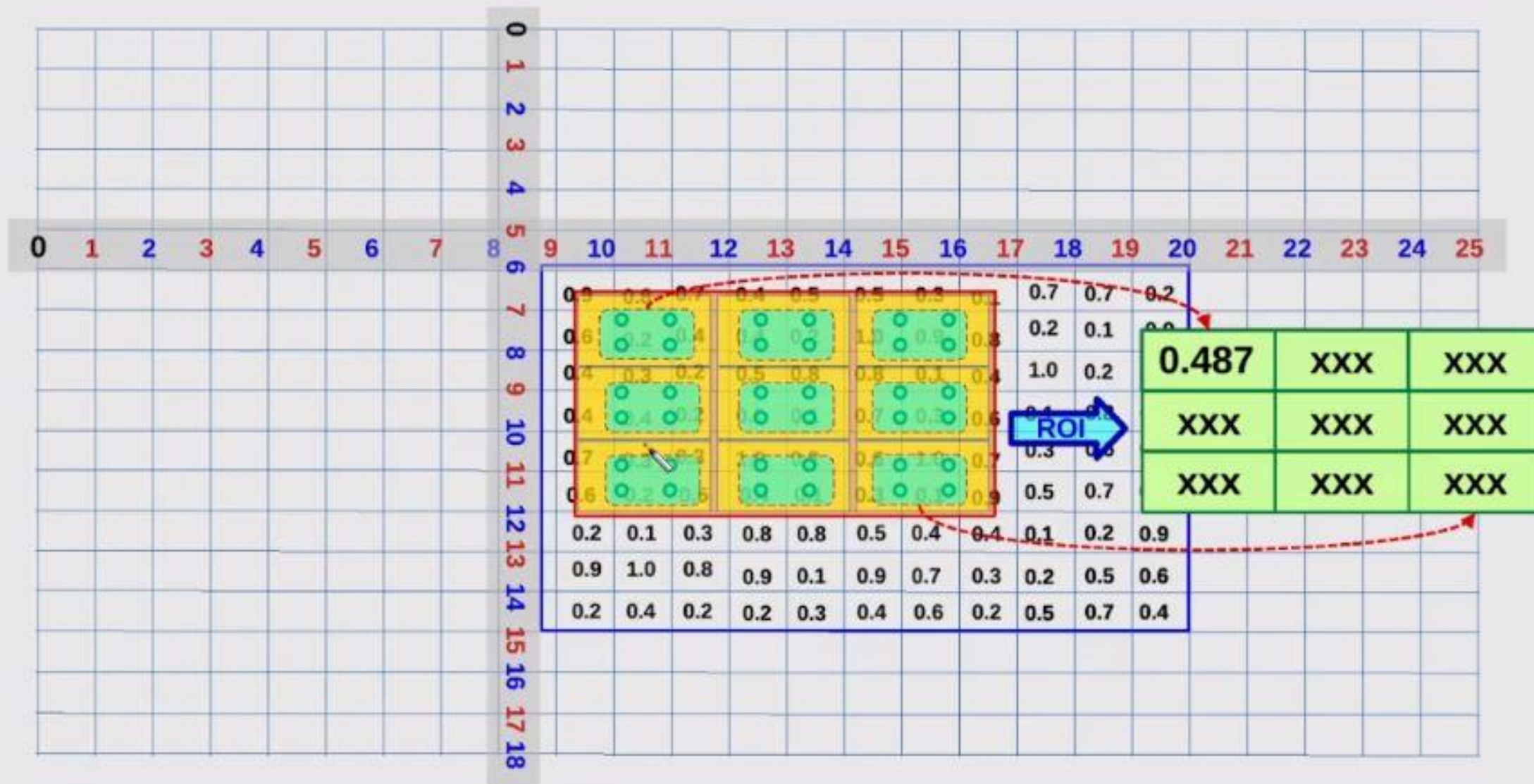
Remember:

Corresponding _Traditional_ ROI Pooling for same Object

Missing

Normal Pooling
(Round Approximation)

X=9.375   ~ 9
Y=6.563   ~ 7
W=7.188   ~ 7
H=5.625   ~ 6

Grid width=7/3   ~2
Grid height=6/3   ~2

ROI

Added
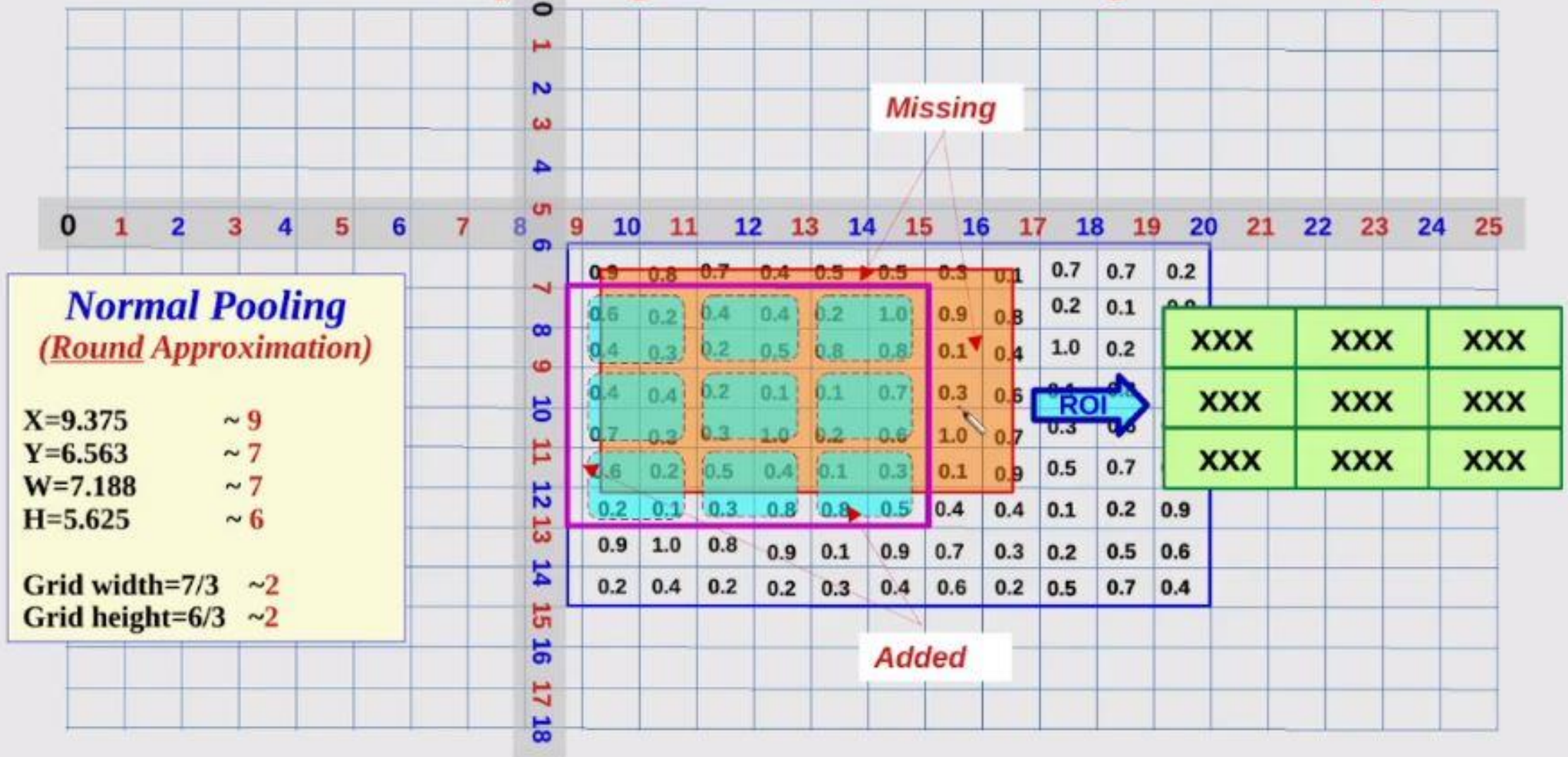
Max – Pooling is applied on Channel Basis
(repeat pooling for Each Channel)

A Pre-Trained CNN Model
(with 5 Pooling Layers)

N Channels

Pooling

Flatten

Same Vector Size
For different proposals Sizes

N Channels

N Channels

Classifier
(Detector)

122
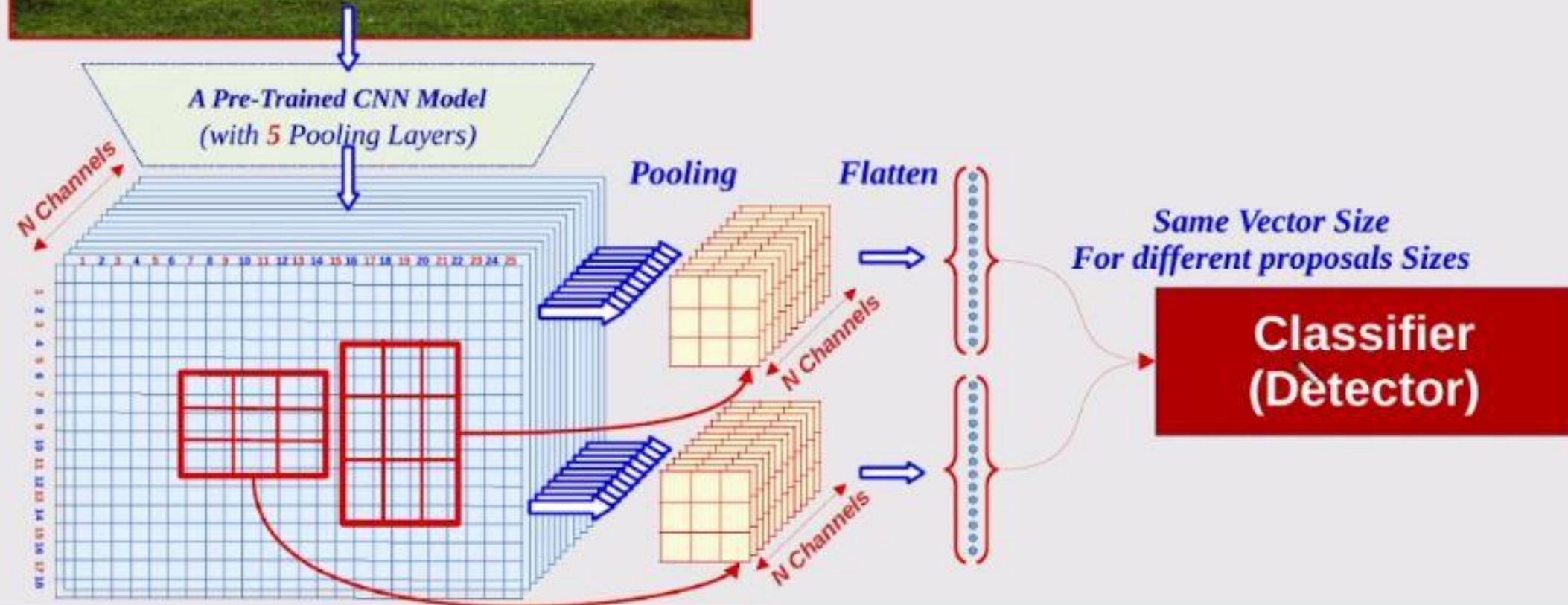
Architecture of Detector Network *(for Fast RCNN, Faster RCNN)* (123)

One *proposed region*
(After ROI Pooling)

3x3x512

*3x3x512
=4608*

**Fully
connected**

Fully Connected

Fully Connected

**Fully
connected**

**Fully
connected**

Softmax Layer

*1000 class x 4
values/class = 4000*

Regressor
*Output
[d$_x$, d$_y$, d$_w$, d$_h$]
For each class*

Classification
*output
(Probabilities)*

*1000 class +
background = 1001*

Example:
Imagenet (1000 Class)
Fixed ROI 3x3 x 512 Ch
3 Hidden Layers