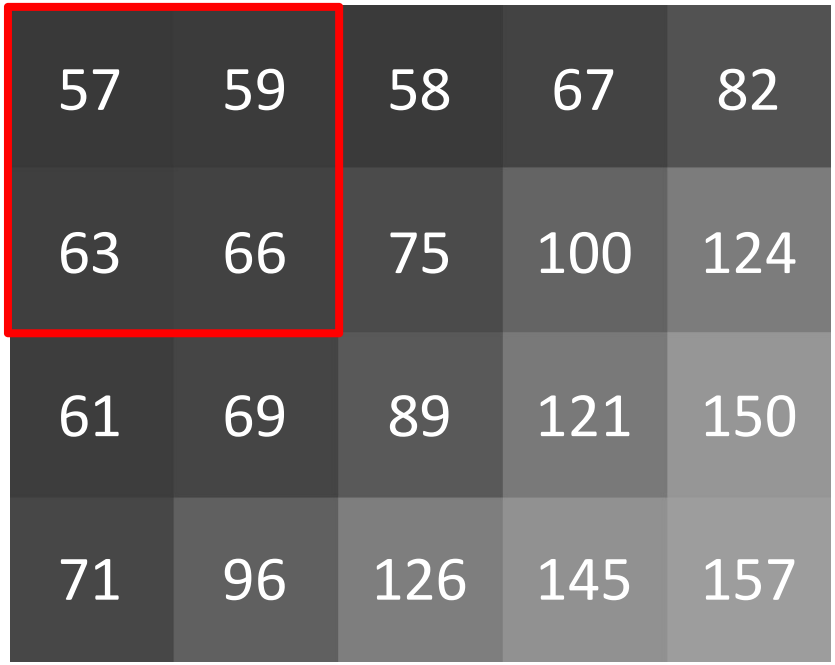
Computer Vision with Embedded Machine Learning

Pooling Layer

Average Pooling

Image



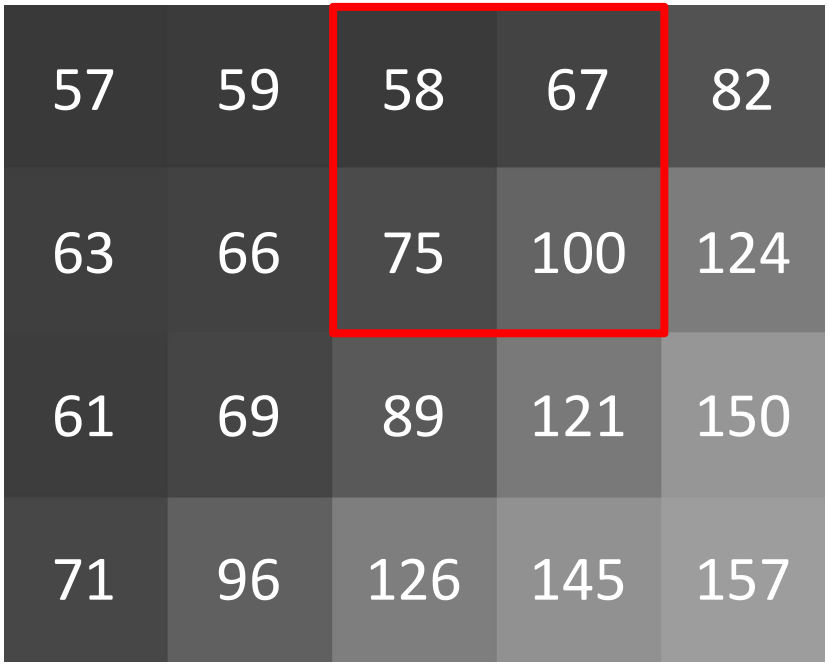
Window: 2x2
Stride: 2
Find average under window:
(57 + 59 + 63 + 66) / 4 = 61.25

Output

61	

Average Pooling

Image



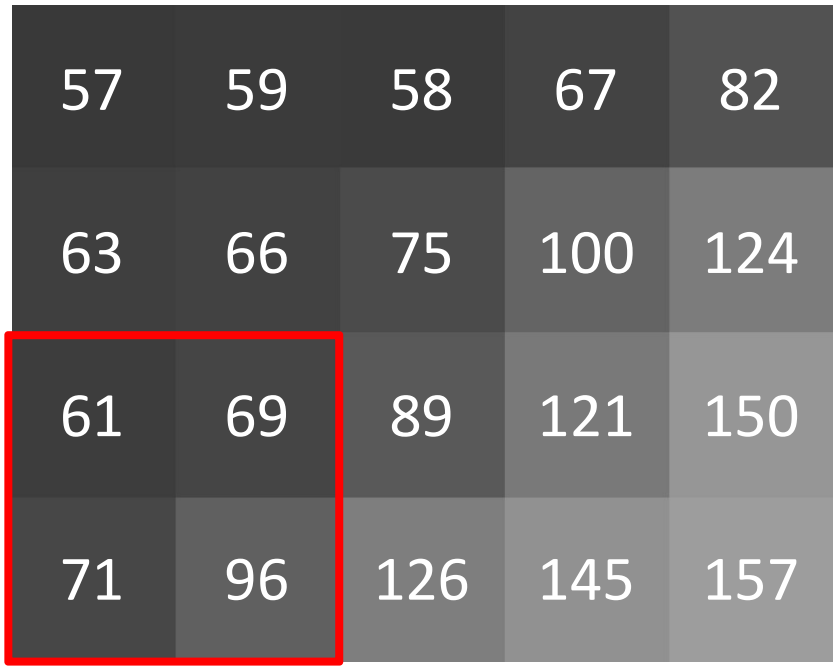
Window: 2x2
Stride: 2
Find average under window:
(58 + 67 + 75 + 100) / 4 = 75

Output

61	75

Average Pooling

Image



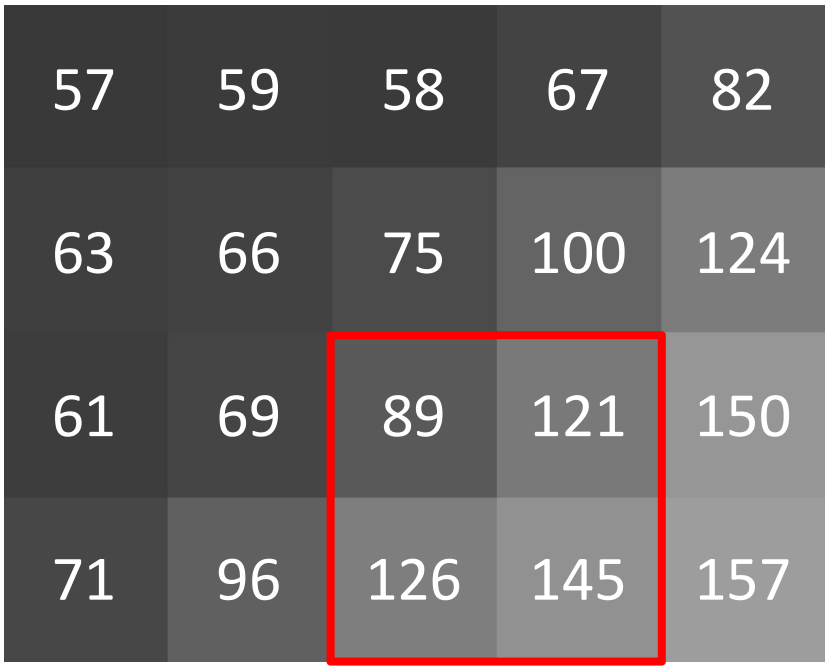
Window: 2x2
Stride: 2
Find average under window:
(61 + 69 + 71 + 96) / 4 = 74.25

Output

61	75
74	

Average Pooling

Image



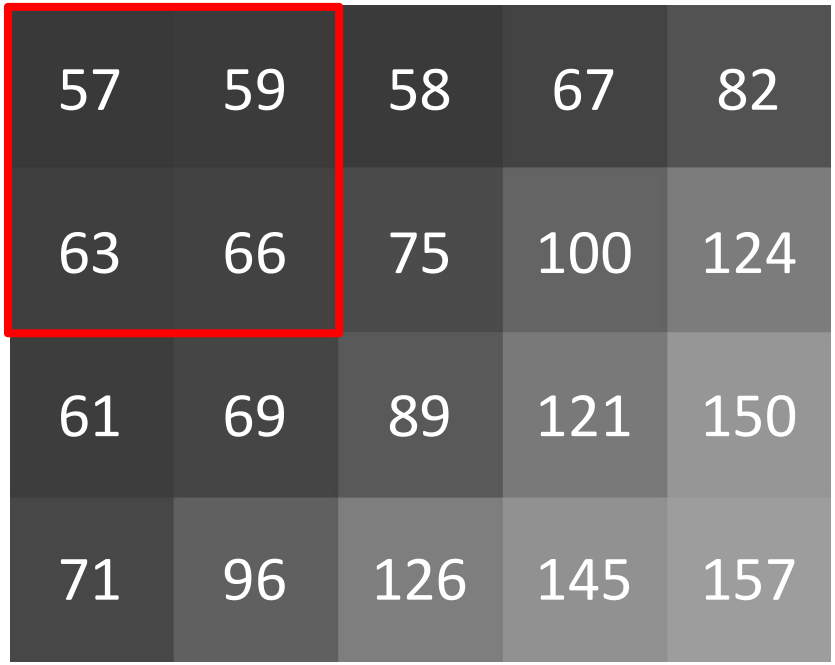
Window: 2x2
Stride: 2
Find average under window:
(89 + 121 + 126 + 145) / 4 = 120.25

Output

61	75
74	120

Max Pooling

Image



Window: 2x2
Stride: 2

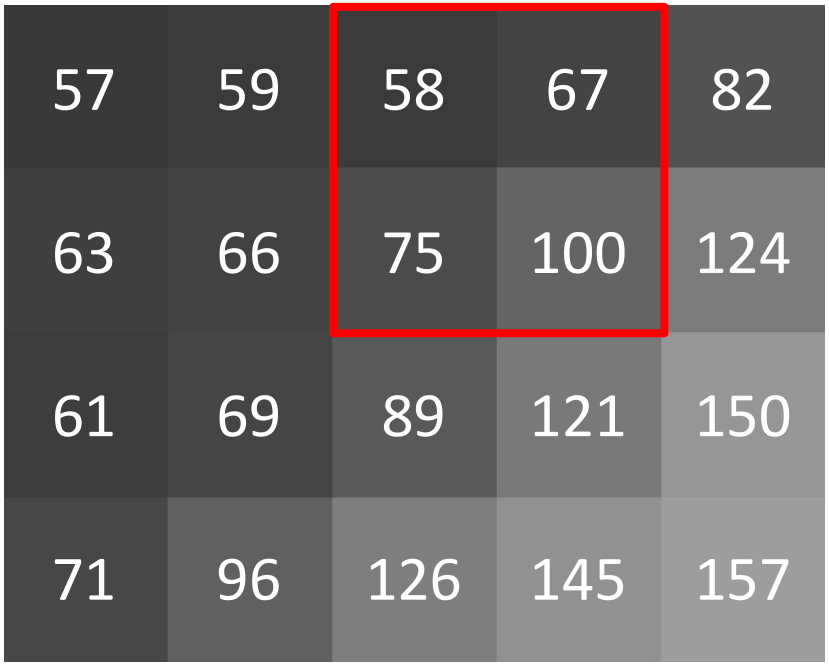
Find maximum value under window: 66

Output

66	

Max Pooling

Image



Window: 2x2
Stride: 2

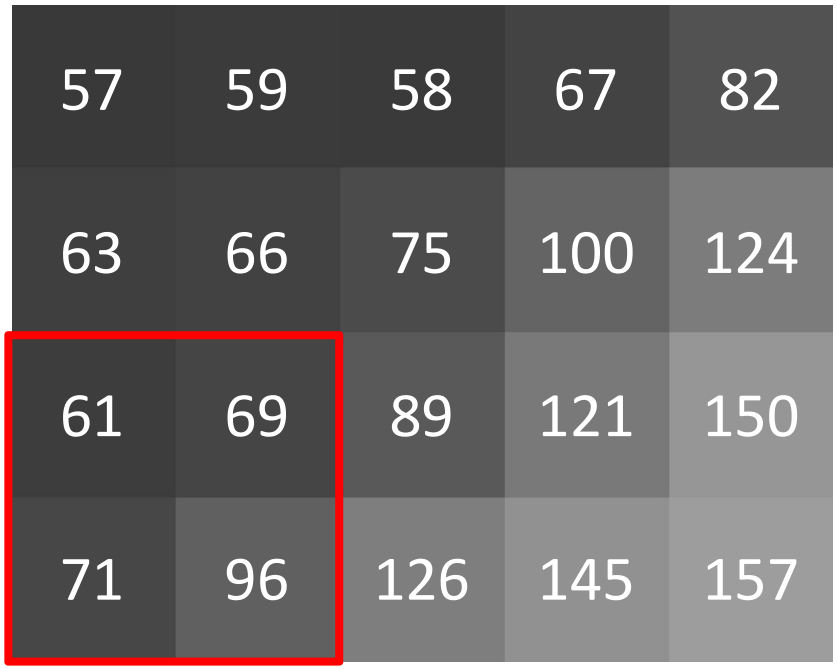
Find maximum value under window: 100

Output

66	100

Max Pooling

Image



Window: 2x2
Stride: 2

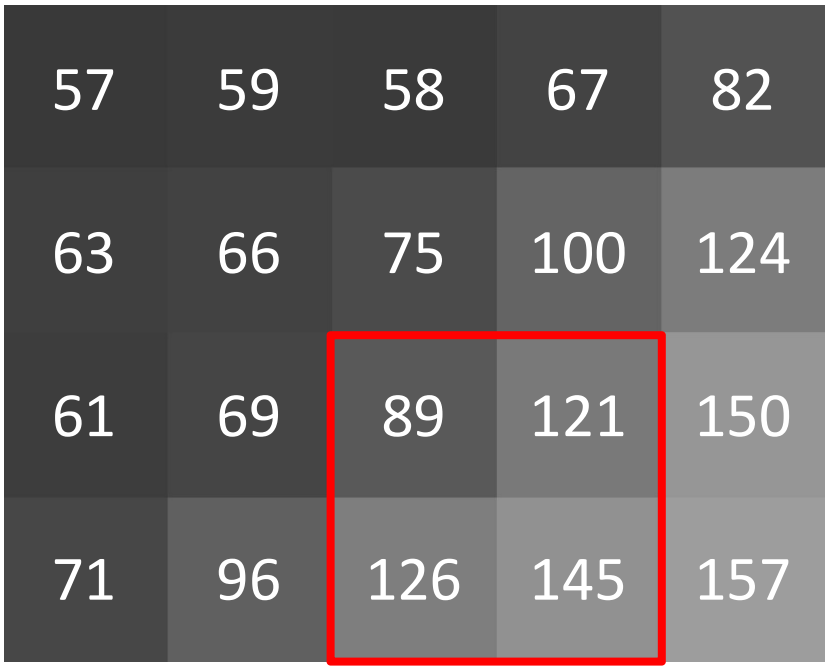
Find maximum value under window: 96

Output

66	100
96	

Max Pooling

Image



Window: 2x2
Stride: 2

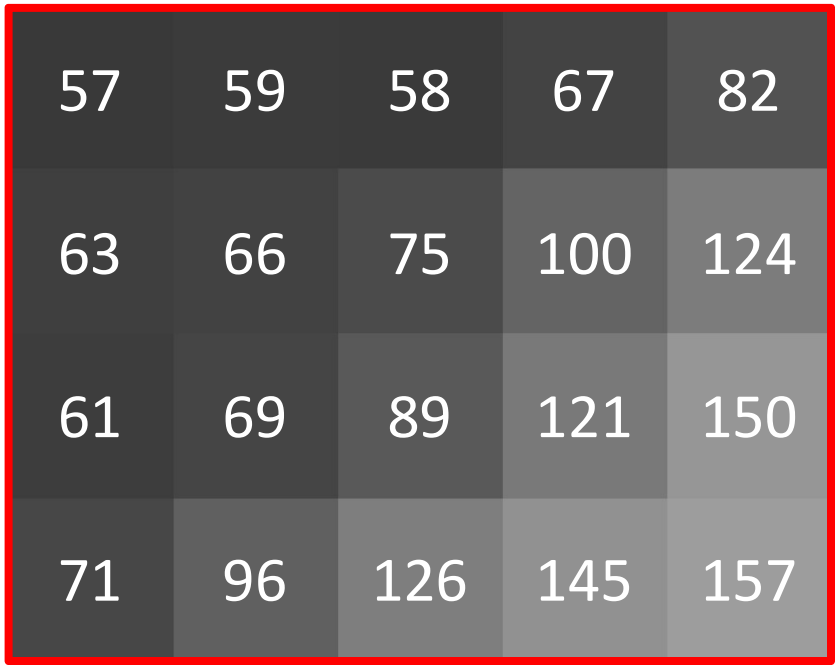
Find maximum value under window: 145

Output

66	100
96	145

Global Pooling

Image



Global average output: 91.8

92

Global max output: 157

157

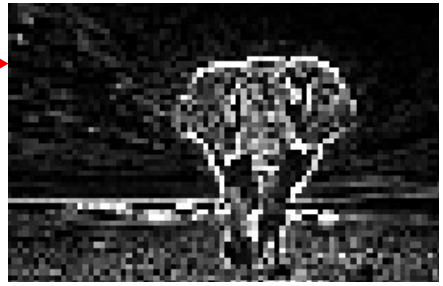
Filtered image
(198x128)

Average Pooling

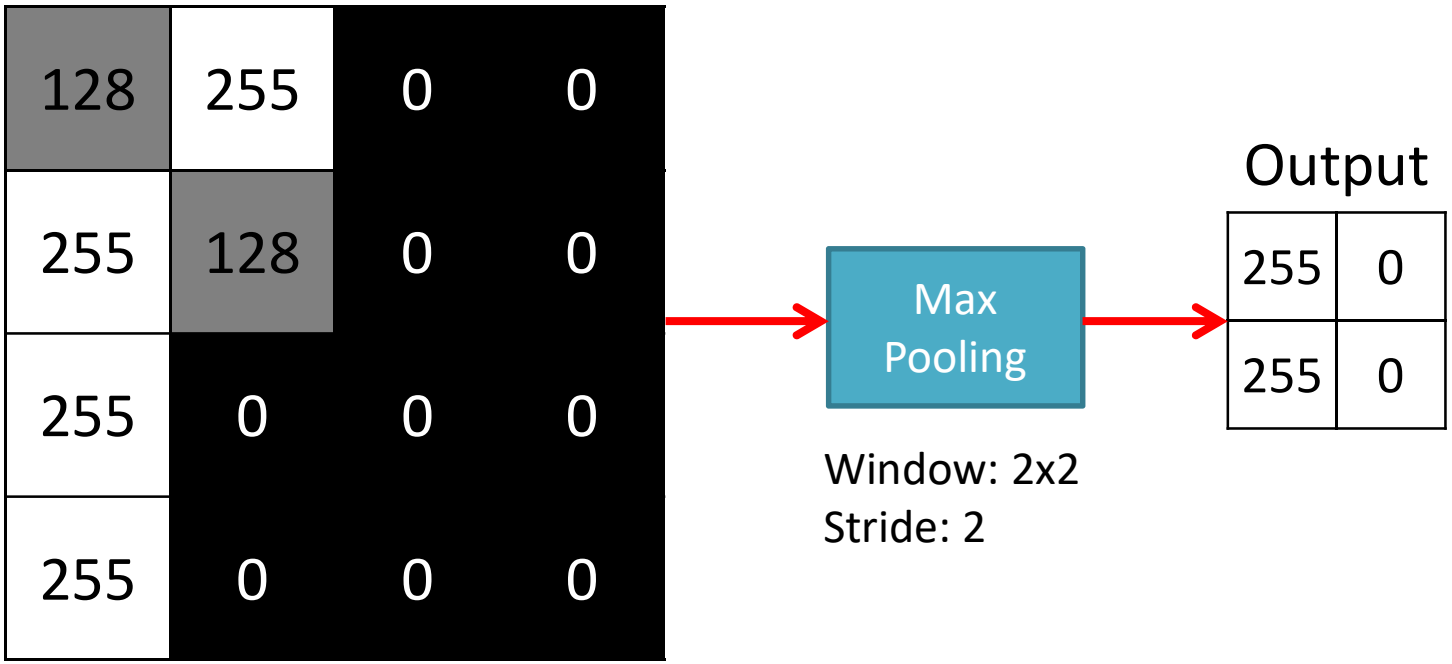
(99x64)

Max Pooling

(99x64)



Local Translation Invariance



Local Translation Invariance