Computer Vision

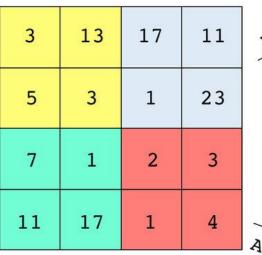
CNN-Convolution

I(0,0)	I(1,0)	I(2,0)	I(3,0)	I(4,0)	I(5,0)	I(6,0)								
I(0,1)	I(1,1)	I(2,1)	I(3,1)	I(4,1)	I(5,1)	I(6,1)						O(0,0)		
I(0,2)	I(1,2)	I(2,2)	I(3,2)	I(4,2)	I(5,2)	I(6,2)		H(0,0)	H(1,0)	H(2,0)				
I(0,3)	I(1,3)	I(2,3)	I(3,3)	I(4,3)	I(5,3)	I(6,3)	×	H(0,1)	H(1,1)	H(2,1)	=			
I(0,4)	I(1,4)	I(2,4)	I(3,4)	I(4,4)	I(5,4)	I(6,4)		H(0,2)	H(1,2)	H(2,2)				
I(0,5)	I(1,5)	I(2,5)	I(3,5)	I(4,5)	I(5,5)	I(6,5)		F	$\mathrm{ilt}\epsilon$	er				
I(0,6)	I(1,6)	I(2,6)	I(3,6)	I(4,6)	I(5,6)	I(6,6)								

Input image

Output image

CNN- Pooling

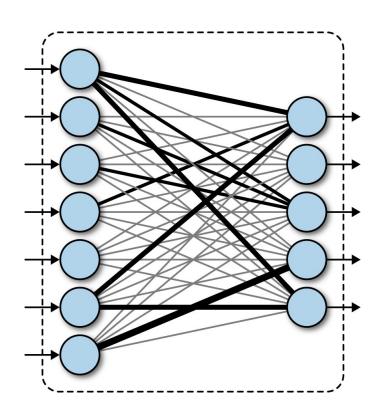


13 23
Max Pooling 7 17 4

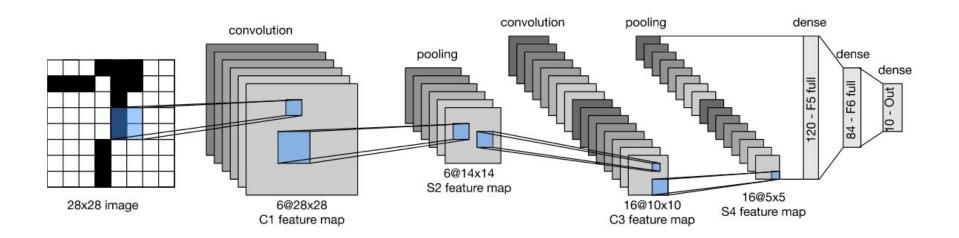
Avergae Pooling

6 139 2.5

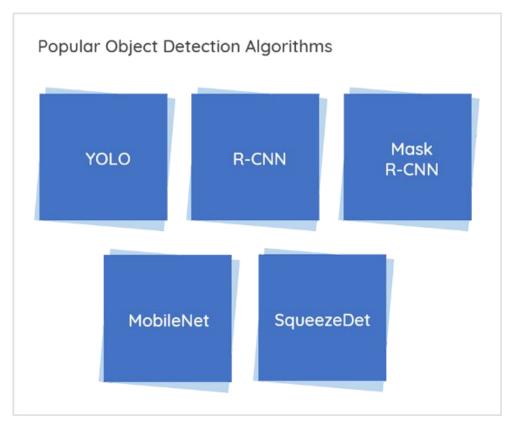
CNN- Fully Connected Layer



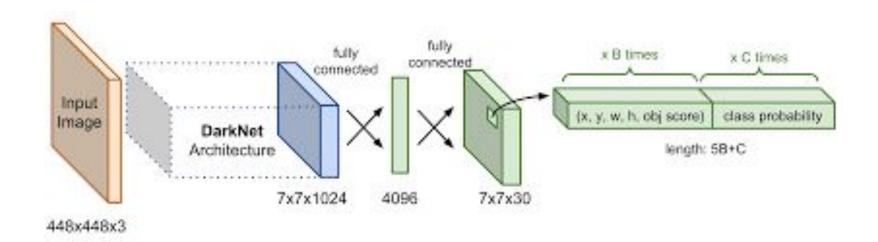
CNN- Lenet5



Vision- Object Detection



YOLO Object detection



Coding

- 1. for Lenet5 with MNIST
- 2. YOLO Object detection