ALEXNET

31 October 2023 07:08

Layers	No. of Neurons	Filter size	Stride	padding	size of feature map	Activation function
					227x227x3 (3 → RGB)	
Conv 1	96	11x11	4	—	55x55x96	Relu.
Max pool	—	3x3	2	—	27x27x96.	
Conv 2	256	5x5	1	2	27x27x256	Relu.
Maxpool	—	3x3	2	—	13x13x256	—
Conv 3	384	3x3	1	1	13x13x384	Relu
Conv 4	384	3x3	1	1	13x13x384	Relu.
Conv 5	256	3x3	1	1	13x13x256	Relu.
Maxpool	—	3x3	1	1	6x6x256.	
Flatten					9216	

—x0x—

Dense 1 4096. (0.5) }
Dense 2 4096

Output layer 1000 Neurons. Soft max.

Advantages.

1. Milestone of Deep Learning CNN image classification.
2. Convolution + pooling dropout, gpu, Relu.
3. Alexnet: Input image can be feed directly.
4. visual patterns were captured

Disadvantages.

1. Later on vgg 16, 19, google net Resnet.
2. conv. filter 5x5, 11x11, 3x3.
3. Training time ⑦ 2weeks.

Vgg 16.

face recognition
Computer vision
Image classification.

visual geometry group. 16 → No. of layers.

conv filter 3x3 }
stride 1x1 } uniform
padding same }

Max pool 2x2 }
Stride 2x2 } uniform.

Layers	filter no.	filter size	stride	padding	feature map	Activation
Input	224x224x3 → RGB.					
Conv1	64	3x3	1	1	224x224x64	Relu.
Conv2	64	3x3	1	1	224x224x64	Relu.
Maxp.	—	2x2	—	—	112x112x64	—
conv3	128	3x3	1	1	112x112x128	Relu.
conv4	128	3x3	1	1	112x112x128	Relu.
Maxp.	—	2x2	—	—	56x56x128	—
conv 5	256	3x3	1	1	56x56x256	Relu.
Conv 6	256	3x3	1	1	56x56x256	Relu
Conv 7	256	3x3	1	1	56x56x256	Relu.
Maxp.	—	2x2	-	-	28x28x256.	
Conv 8	512	3x3	1	1	28x28x512	Relu.
Conv 9	512	3x3	1	1	28x28x512	Relu
Conv 10	512	3x3	1	1	28x28x512	Relu.
Maxp.	—	2x2	—	—	14x14x512	
Conv 11	512	3x3	1	1	14x14x512	Relu
Conv 12	512	3x3	1	1	14x14x512	Relu
Conv 13	512	3x3	1	1	14x14x512	Relu.
Maxp.		2x2			7x7x512	
Flatten					25088.	
FC 1					4096	
FC 2					4096	
output					1000 Softmax.	

2x1 (Conv1–Conv2)
2x1 (Conv3–Conv4)
3x1 (Conv5–Conv7)
3x1 (Conv8–Conv10)
3x1 (Conv11–Conv13)
3 (FC)

224x224x3
vgg 16

2 (→ C)	2	3	3	3	3
64 (→ M)	128	256	512	512	4096 4096 1000.

f = 3x3   MP = 2x2
s = 1   s = 2
p = 1

$$\frac{n+2p-f}{s}+1$$

n → image, p → padding, f → filter, s → stride