

AlexNet

컴퓨터 비전 분야의 올림픽이라 할 수 있는 ILSVRC(ImageNet Large Scale Visual Recognition Challenge)의 2012년 대회에서 1위를 차지했다.

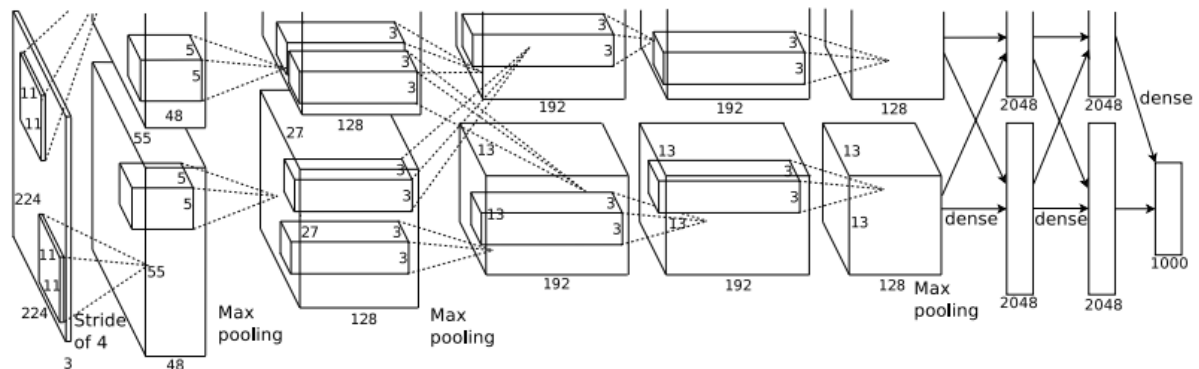
이 AlexNet 덕분에 딥러닝, 특히 CNN이 세간의 주목을 받게 되었으며, CNN 구조의 GPU 구현과 dropout적용이 보편화 되었다.

논문:

ImageNet Classification with Deep Convolutional
Neural Networks

AlexNet Architecture

AlexNet은 일부가 max-pooling layer가 적용된 5개의 convolution layer와 3개의 fully-connected layer로 이루어져 있다.



AlexNet은 [Input layer-Conv1-MaxPool1-Norm1-Conv2-Maxpool2-Norm2-Conv3-conv4-Conv5-Maxpool3-FC1-FC2-output layer]으로 구성되어 있다.

input layer

224 x 224 x3 크기의 이미지이다.

Conv1

96kernels of size 11x11, stride=4, padding=0

input:224x224x3

output:(224-11)/4+1=55>>55x55x96

MaxPool1

3x3kernels,stride2

input:55x55x96

output:27x27x96

Norm1

LRN을 사용한 normalization layer이다. LRN? 밑에서!!

normalization을 수행하는 layer이다.

input:27x27x96

output:27x27x96

Conv2

256kernels of size 5x5, stride=1, padding=2

input:27x27x96

output:27x27x256

MaxPool2

3x3 kernels,stride=2

input:27x27x256

output:13x13x256

Conv3

384kernels of size 3x3,stride=1,padding=1

input:13x13x256

output:13x13x384

Conv4

384kernels of size 3x3,stride=1,padding=1

input:13x13x384

output:13x13x384

Conv5

256 kernels of size 3x3, stride=1, padding=1

input: 13x13x384

output: 13x13x256

MaxPool3

3x3 kernels, stride=2

input: 13x13x256

output: 6x6x256

FC1

Fully connected layer with 4096 neurons

input: 6x6x256

output: 4096

FC2

Fully connected layer with 4096 neurons

input: 4096

output: 4096

Output layer

fully connected layer with 1000-way softmax

input: 4096

output: 1000

AlexNet의 구조에 적용된 특징

1. first use of ReLU

activation function으로 ReLU함수를 적용하였다.

tanh나 sigmoid보다 학습속도가 몇배는 빠르다고한다.

2. Training on Multiple GPUs

network를 2개의 GPU로 나누어서 학습시켰다.(병렬구조)

이를 GPU parallelization이라고 한다.

3. Local Response Normalization(LRN)

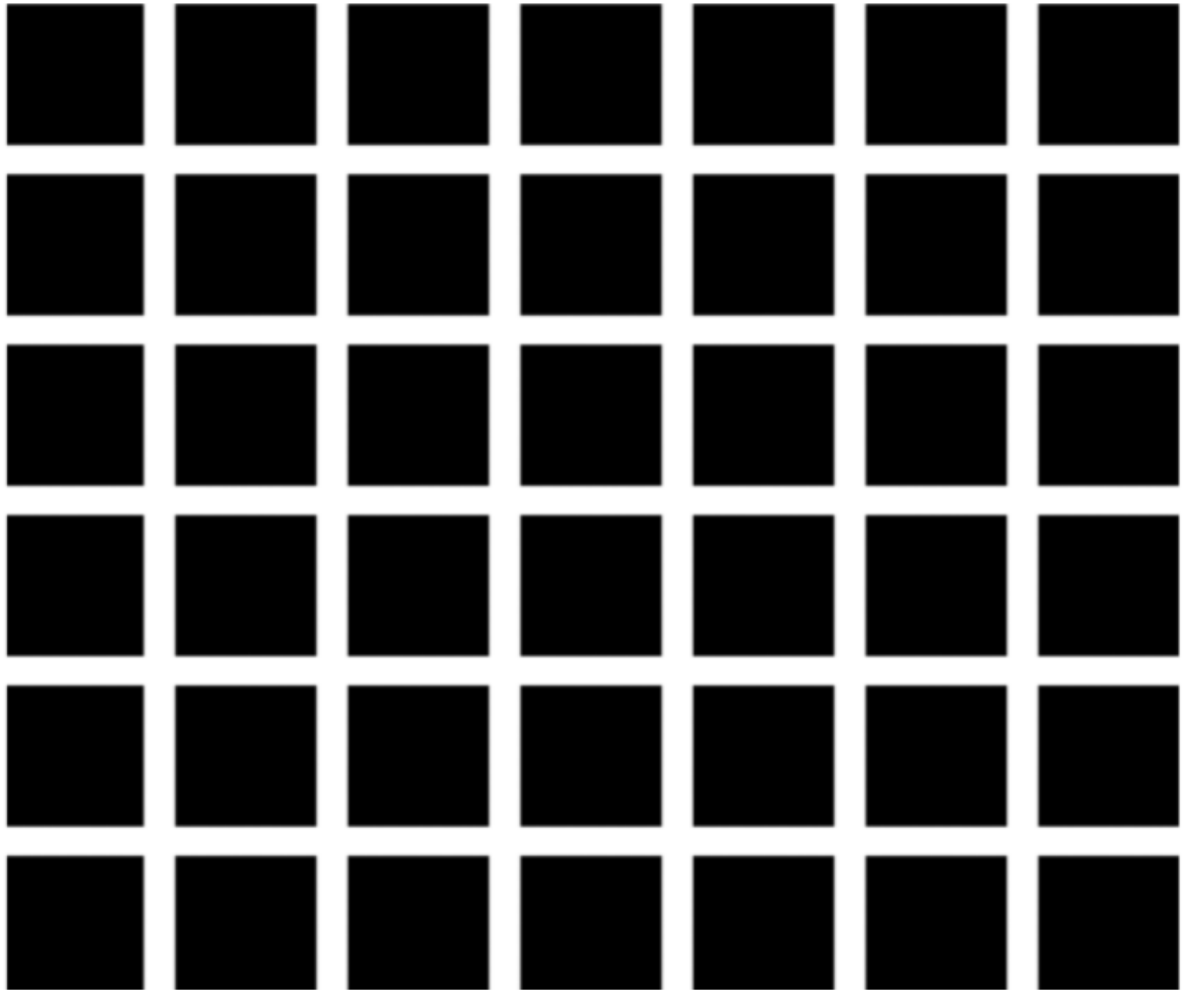
LRN은 generalization을 목적으로 한다. sigmoid 나 tanh 함수는 입력 data의 속성이 서로 편차가 심하면 saturating되는 현상이 심해져 gradient vanishing를 유발할 수 있게 된다.

반면에 ReLU는 non-saturating nonlinearity함수이기 때문에 saturating을 예방하기 위한 입력 normalization이 필요로 하지 않는 성질을 갖고 있다.

ReLU는 양수값을 받으면 그 값을 그대로 neuron에 전달하기 때문에 너무 큰 값이 전달되어 주변의 낮은 값이 neuron에 전달되는 것을 막을 수 있다.이것을 예방하기 위한 normalization이 LRN 이다.

논문에서는 LRN을 측면 억제(lateral inhibition)의 형태로 구현된다고 나와 있다.

측면억제는 강한 자극이 주변의 약한 자극을 전달하는 것을 막는 효과를 말한다.



출처: <https://ko.wikipedia.org/wiki/착시>

위 그림은 측면 억제(側面抑制)의 유명한 그림인 헤르만 격자다. 검은 사각형안에 흰색의 선이 지나가고 있습니다. 신기한 것은 흰색의 선에 집중하지 않을 때 회색의 점이 보이는데 이러한 현상이 측면 억제에 의해 발생하는 것이다. 이는 흰색으로 둘러싸인 측면에서 억제를 발생시키기 때문에 흰색이 더 반감되어 보인다.

Reducing Overfitting

1. Data Augmentation

하나의 이미지를 가지고 여러 장의 비슷한 이미지를 만들어 냄

2. Dropout