

Derin Öğrenme

Convolutional Neural Networks

Emir Öztürk

Görüntü tanıma - işleme

Wait, I know you

- Farklı alanlar
 - Sağlık
 - Otonom Uygulamalar
 - Güvenlik
- Yalnızca nesnelerin tanınması değil
 - Lokasyon belirleme
 - Hareket belirleme

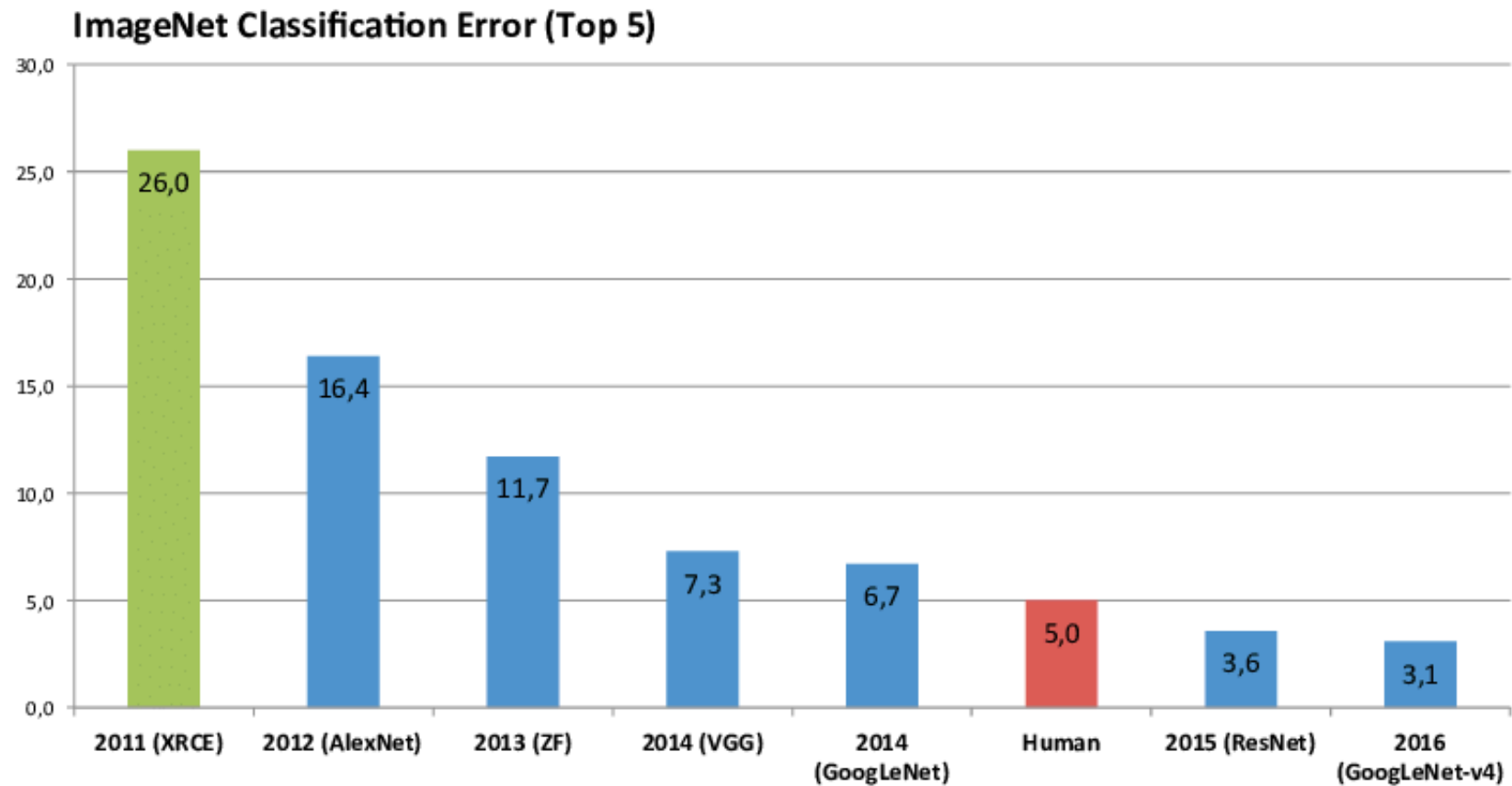
Görüntü Sınıflandırma

Getting Classy

- ImageNet Challenge
- Verinin çok ve erişilebilir olması
- Görüntü sınıflandırma problemi
- Eski yöntemlerde özellik çıkarımı ve etiketleme
 - Kural tabanlı

Eski yöntemler

Know your backstory

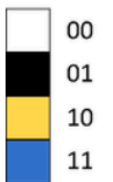
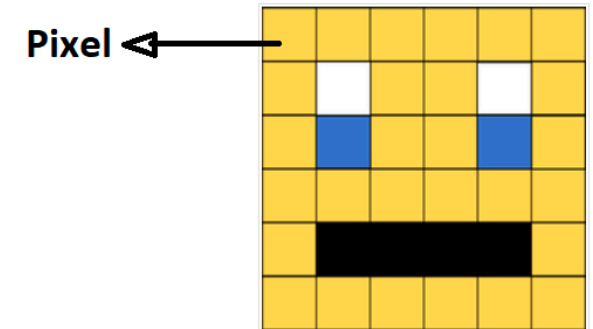


Resimlerin saklanması

There is spoon

- İki boyutlu
- Her renk için bir kanal
- Opacity için bir kanal

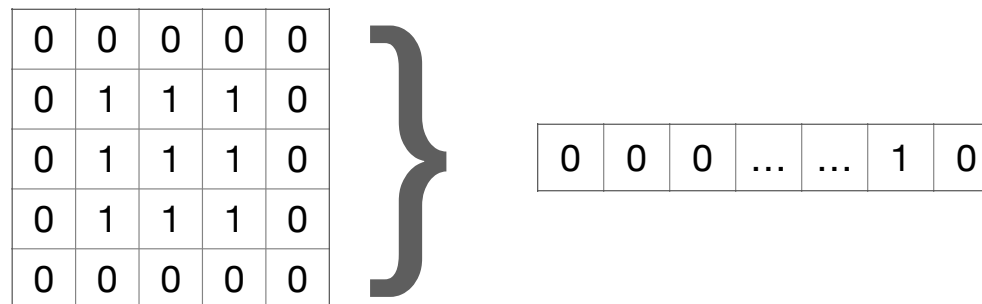
10	10	10	10	10	10	01010101010
10	00	10	10	00	10	00010100010
10	11	10	10	11	10	01110101110
10	10	10	10	10	10	01010101010
10	10	10	10	10	10	00101010110
10	01	01	01	01	10	01010101010
10	10	10	10	10	10	



Sinir ağına resimlerin verilmesi

You have to adapt

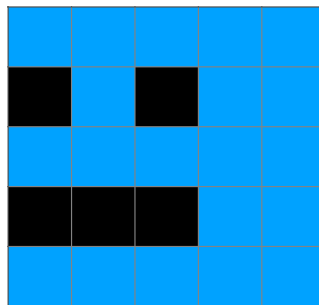
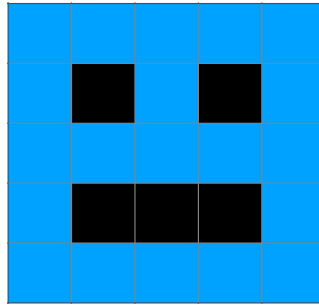
- Sinir ağına vektör verilmeli
- Resimler matris olarak saklanır
- Matristen vektöre dönüşüm yöntemleri



Matristen vektöre dönüşüm

The easiest way is not the best way

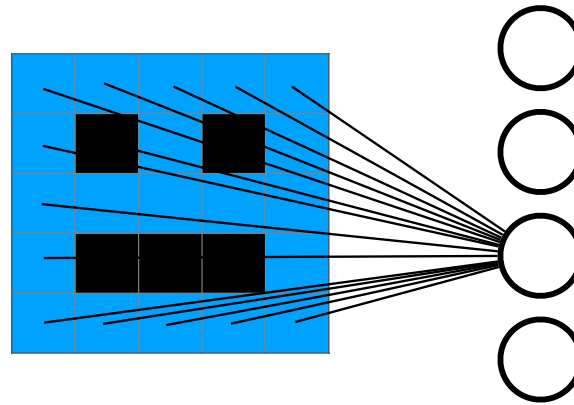
- Satır sıralı almak
- Problemler
 - Lokalite
 - Translation
 - Shifting



Tüm özelliklerin dense kullanılması

The hardest way is not also

- Lokalliğin kaybına sebep olur
- İlişkisiz verinin birbirine bağlanması
- Gereksiz karmaşıklık artışı



Lokalitenin sağlanması

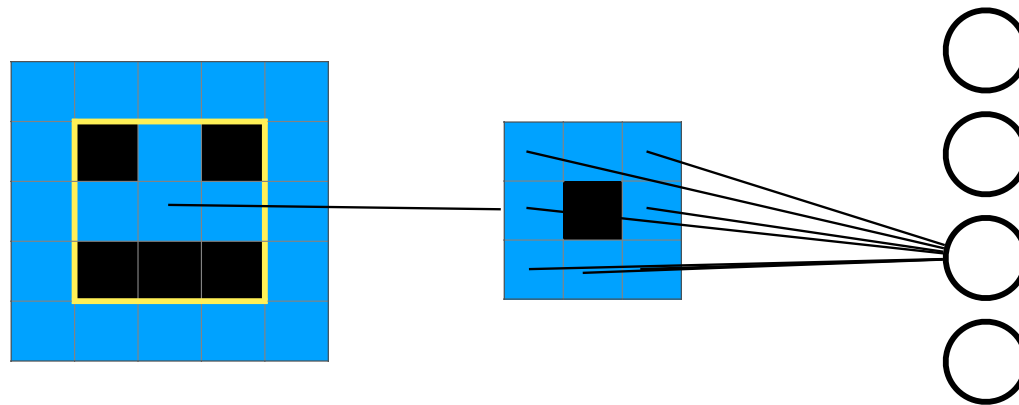
Know your neighbours

- Belirli tekrarlı grup özelliklerini tanımlamak
- Bu özelliklerin birleştirilmesi ile daha soyut özelliklere ulaşım
- Özelliklerin manual olarak elde edilmesine alternatif

Lokalitenin sağlanması

Kernel error

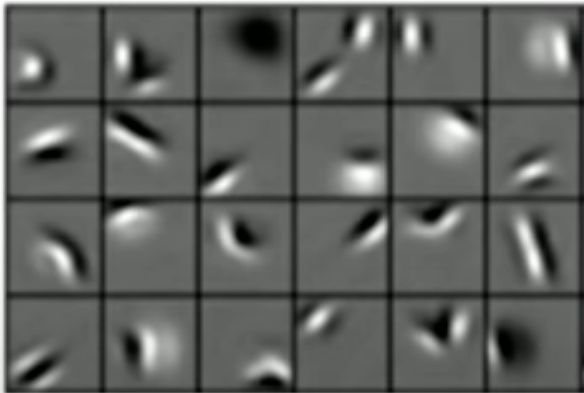
- Gruplar arası özellik çıkarımı
 - Farklı özelliklerin lokal olarak belirlenmesi
- Grupların boyutu
- Aranılan özelliklerin seçimi
- Kernel



Lokalitenin sağlanması

Cnn Ft. Recognition

Low level features



Edges, dark spots

Mid level features



Eyes, ears, nose

High level features



Facial structure

Kernel

We have to go deeper

- Farklı özellikler çıkartan farklı kernel'lar
- Her kernel'dan elde edilen farklı feature map
- Çıktı boyutu kernel sayısına göre değişir
- $(\text{En-KernelBoyutu} + 1) * (\text{boy-KernelBoyutu} + 1) * \text{kernel sayısı}$

1 _{x1}	1 _{x0}	1 _{x1}	0	0
0 _{x0}	1 _{x1}	1 _{x0}	1	0
0 _{x1}	0 _{x0}	1 _{x1}	1	1
0	0	1	1	0
0	1	1	0	0

Image

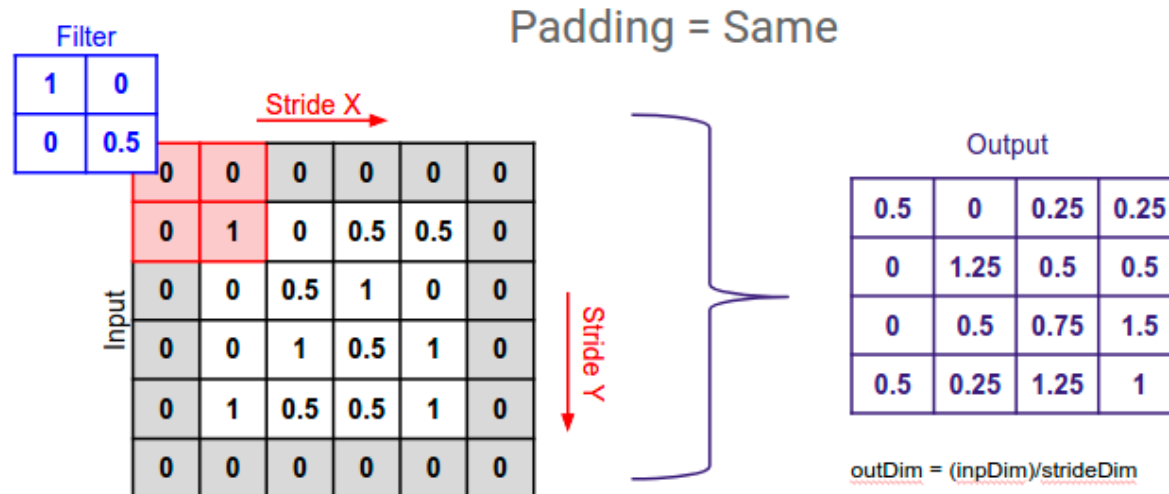
4		

Convolved
Feature

Kernel kullanımı

Not losing you again

- Eğer çıktı boyutu resim boyutu ile aynı olsun istenirse padding
- Değerlerin 0 verilmesi özellik yakalama açısından problemli olabilir



Kaynak kısıtı

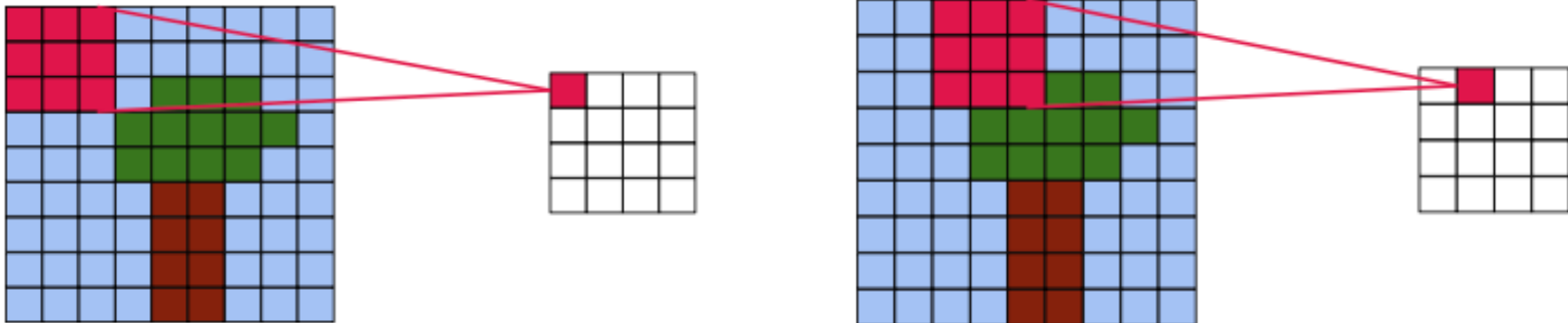
That is what engineering is

- Feature map küçültmek için
 - Kaydırma büyüklüğünü arttırmak
 - Kernel boyutunu arttırmak
 - Dilated convolution

Kaynak kısıtı - Kaydırma boyutunu arttırmak

Speeding up

- Gruplar içerisindeki özelliklerin tanınması aynı şekilde
- Atlama yapıldığı için bilgi kaybı



Kaynak kısıtı - Kernel boyutunu arttırmak

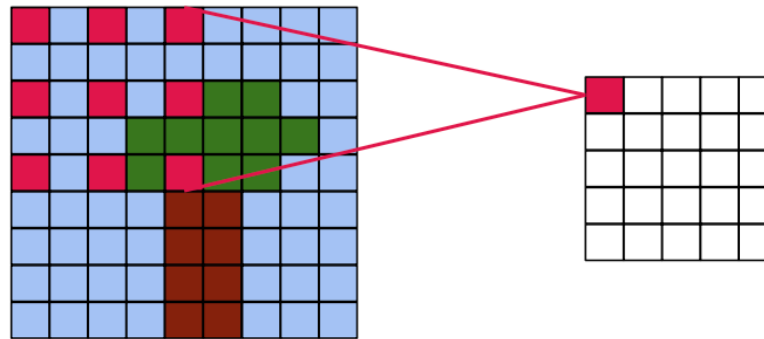
The big picture

- Daha büyük alanlar alındığı için alanlar arası ilişkisellik problemi
- Kernel'lar için işlem karmaşıklığı artar
- Daha küçük bir feature map

Kaynak kısıtı - Dilated convolution

Feeling complicated

- Kernel içerisinde yalnızca belirli noktaları alma
- Pencere kaydırma ile olan kaybı azaltmak
- Daha küçük bir feature map



İki boyutlu - üç boyutlu resim

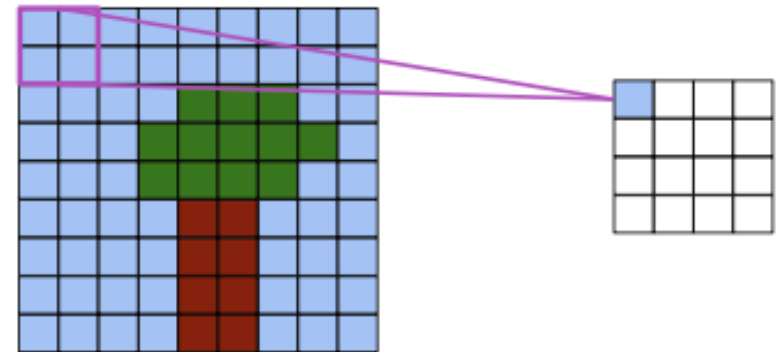
Why does it always have to be harder

- Önceki örnekler 2d resimler üzerine
- 3 boyut olduğunda her kanal farklı birer matris
- 3 boyuttan ikiye düşüş
 - Grayscale çalışmak
 - Her kanalın farklı ağırlıklar verilerek tek bir kernel ile kullanılması
- Depthwise convolution

Pooling

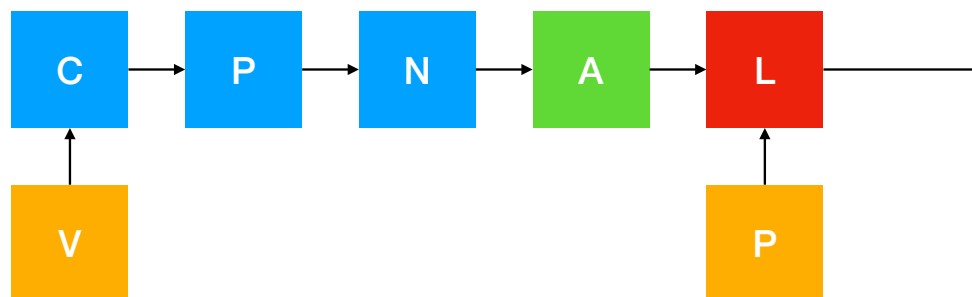
Not the pool you think of

- Veri boyutunun küçültülmesi
- Aranılan özelliklerde temsilci seçilmesi
 - Min
 - Avg
 - Max



En temel CNN

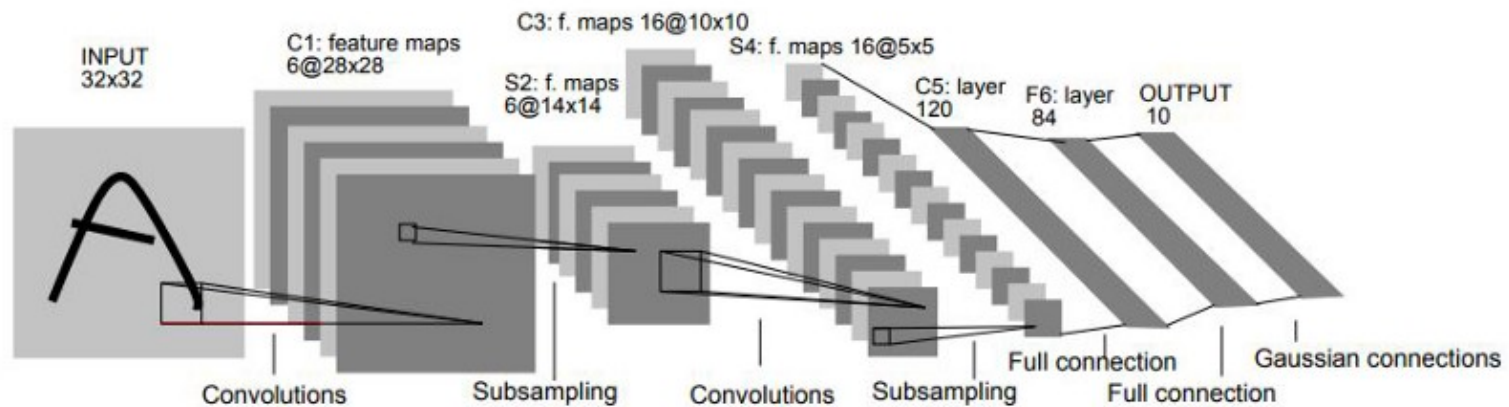
Do you remember the shape



CNN örnek modelleri

Example 1

- LeNet-5



Example 2

- AlexNet

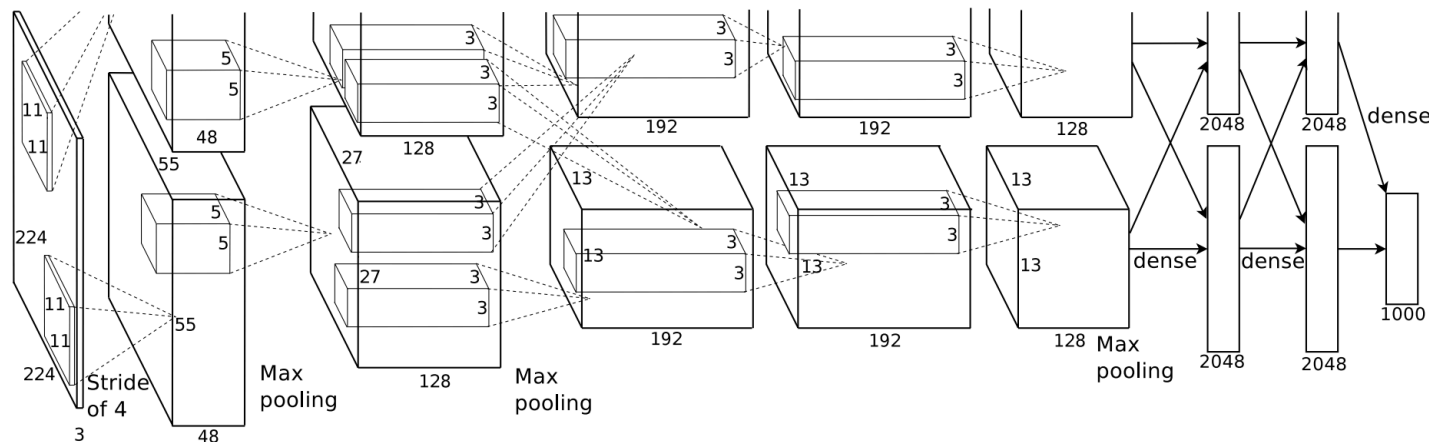
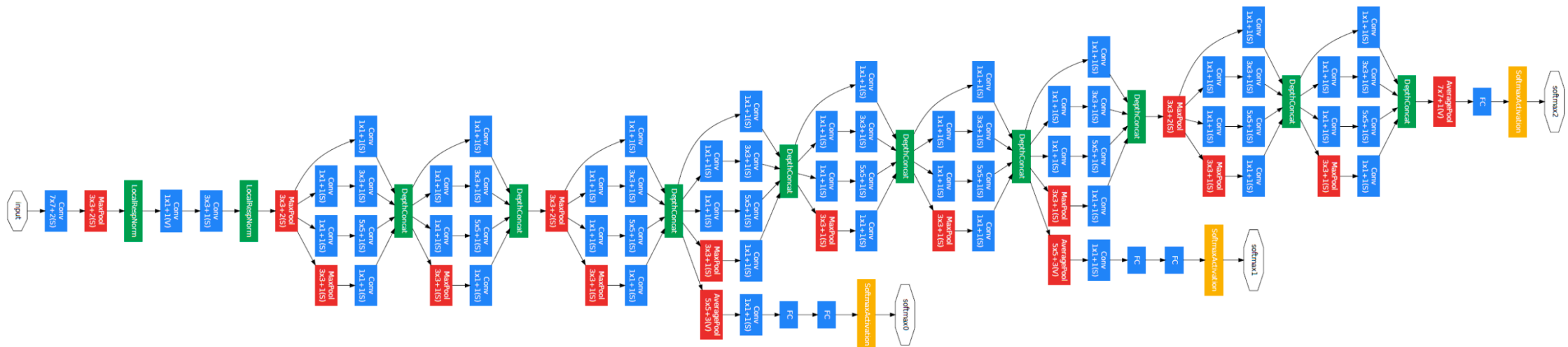


Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network’s input is 150,528-dimensional, and the number of neurons in the network’s remaining layers is given by 253,440–186,624–64,896–64,896–43,264–4096–4096–1000.

CNN örnek modelleri

Example 3

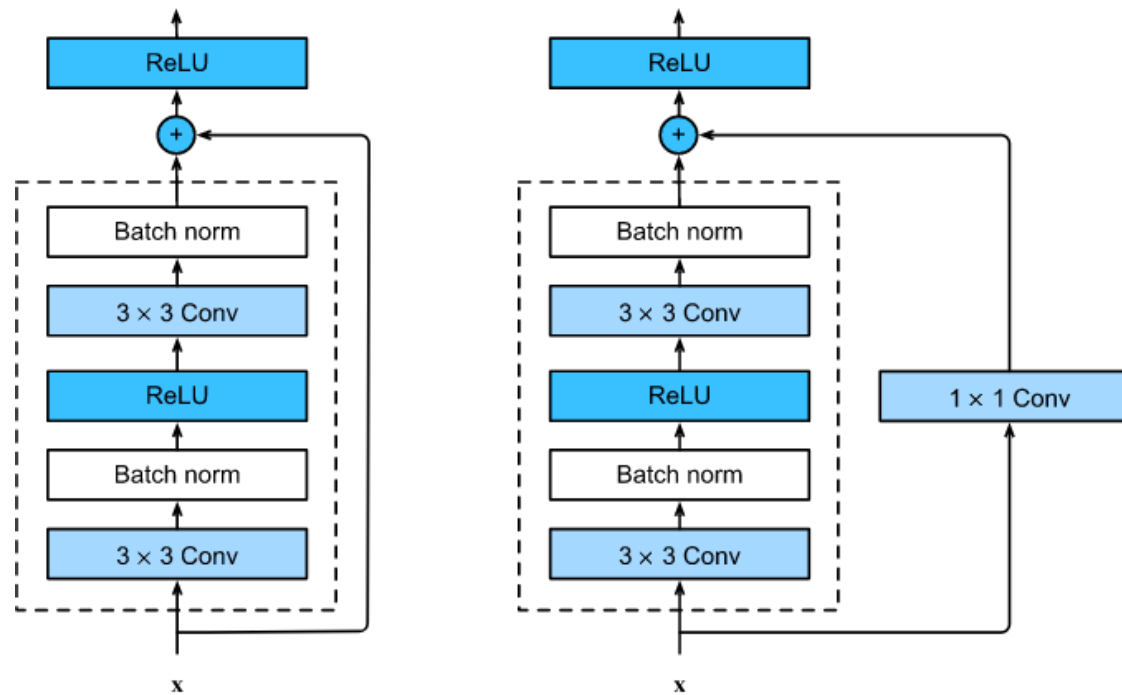
- GoogLeNet



CNN örnek modelleri

Rambo 4

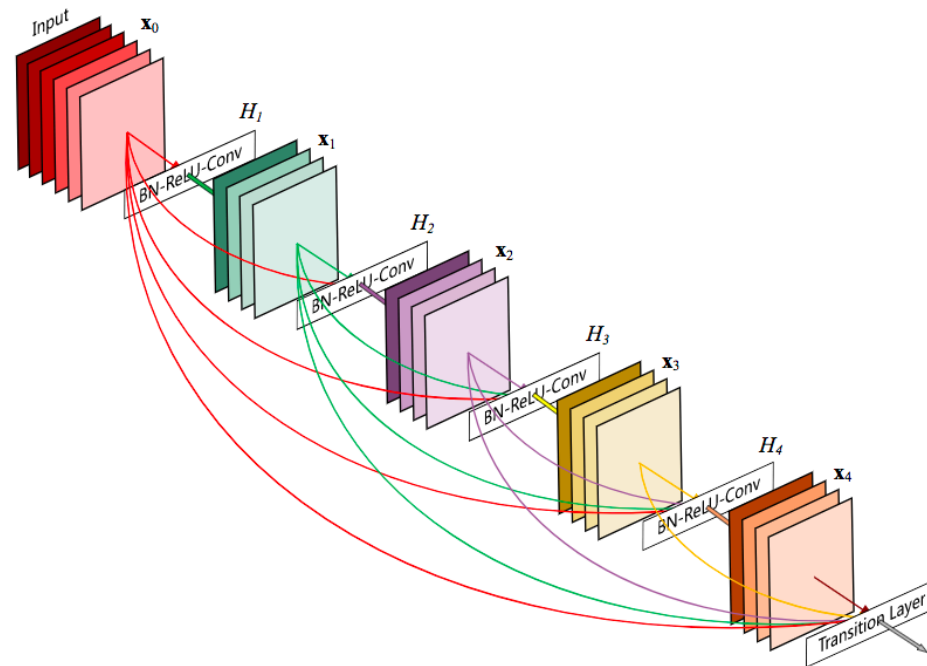
- Resnet



CNN örnek modelleri

There is no end to this

- DenseNet



CNN örnek modelleri

Are you serious?

- AmoebaNet

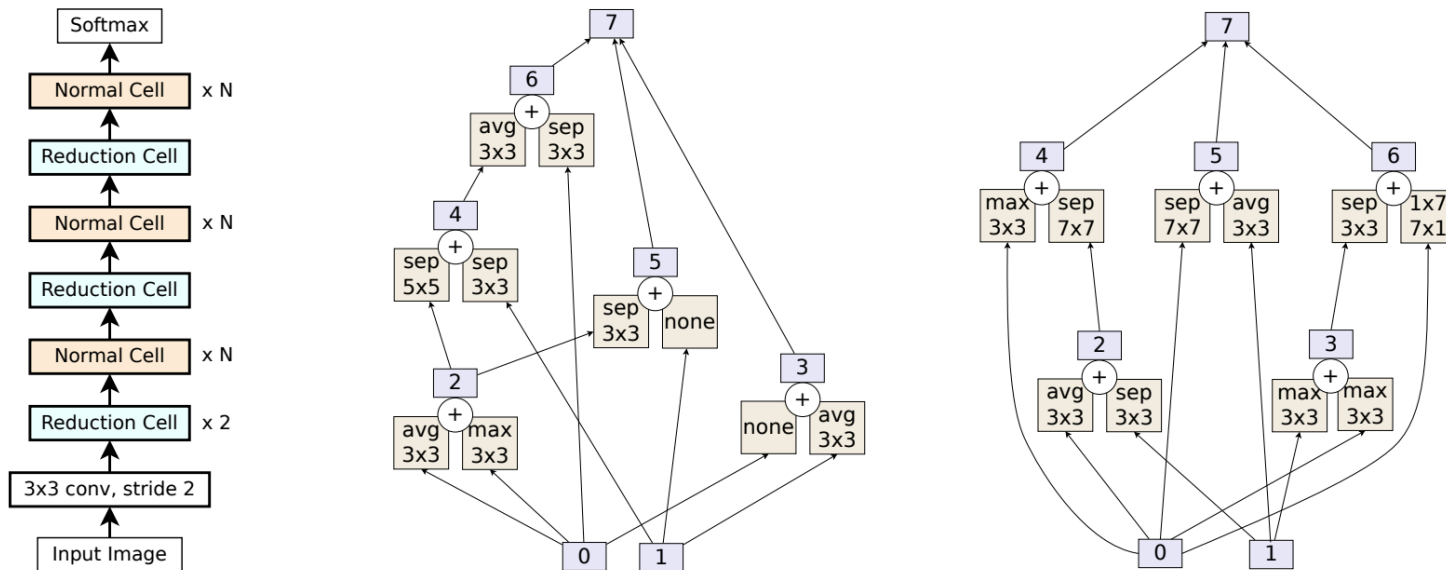
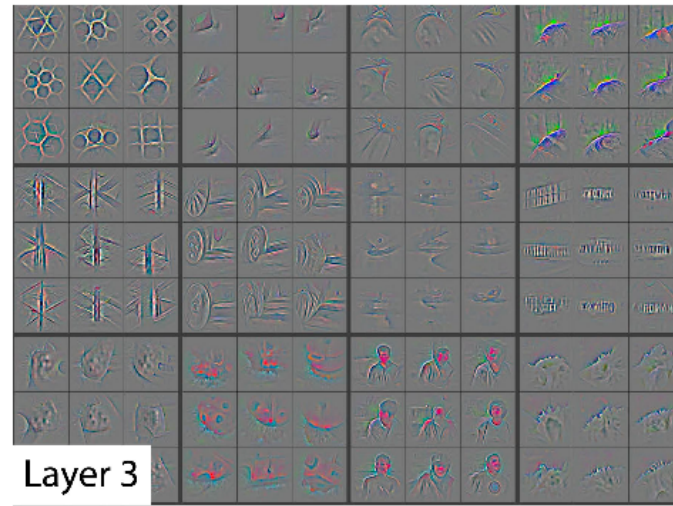
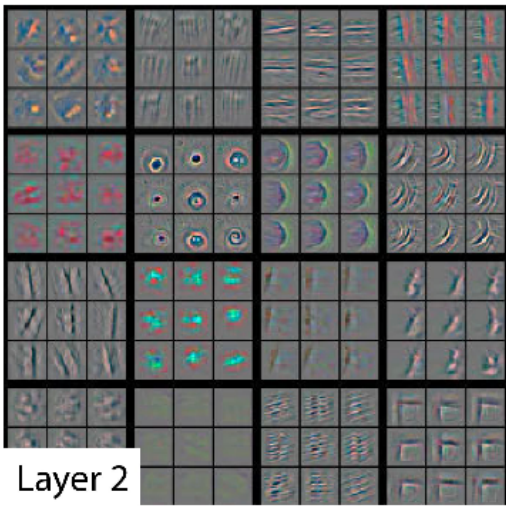


Figure 5: AmoebaNet-A architecture. The overall model [54] (LEFT) and the AmoebaNet-A normal cell (MIDDLE) and reduction cell (RIGHT).

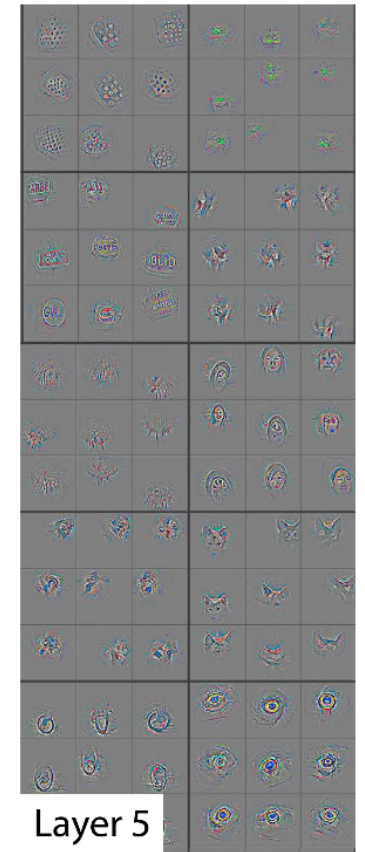
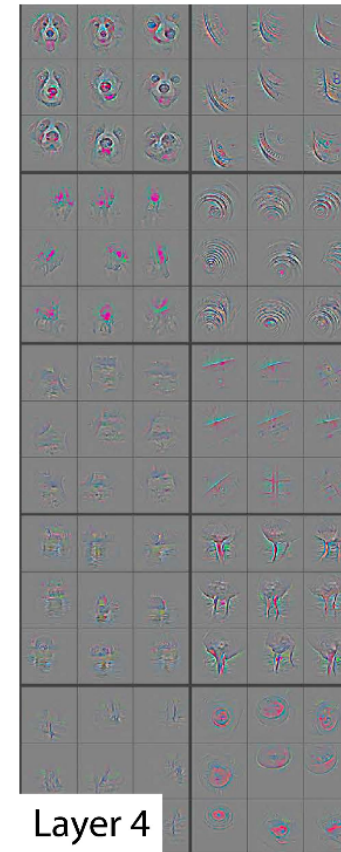
CNN'in öğrenme çıktıları

We have seen it before



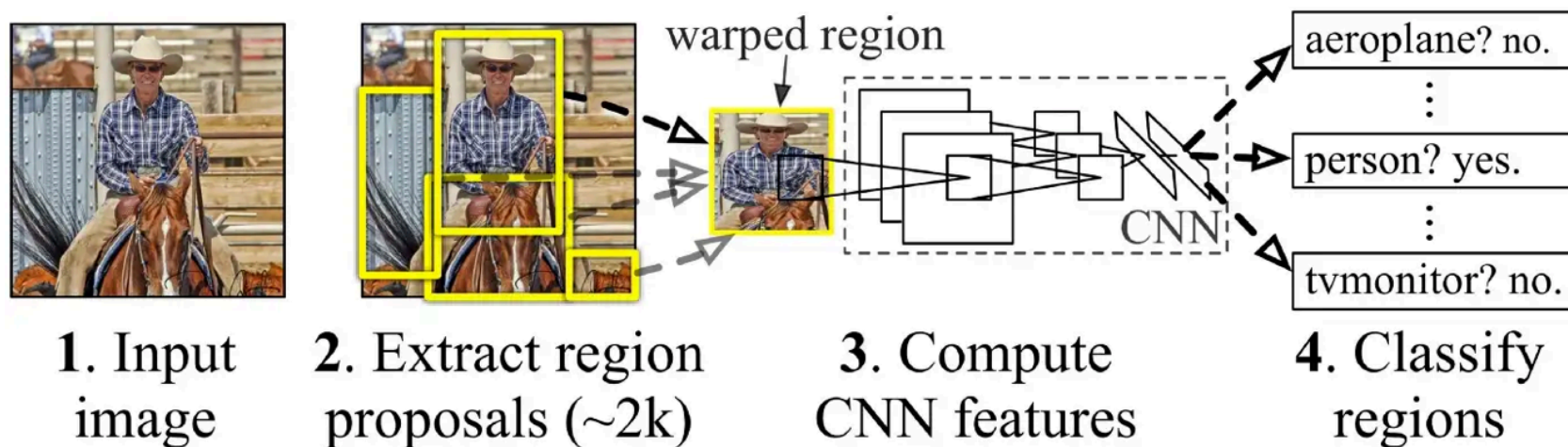
Figures from Zeiler et al. (2014)

<https://distill.pub/2017/feature-visualization/>



Object Detection - RCNN

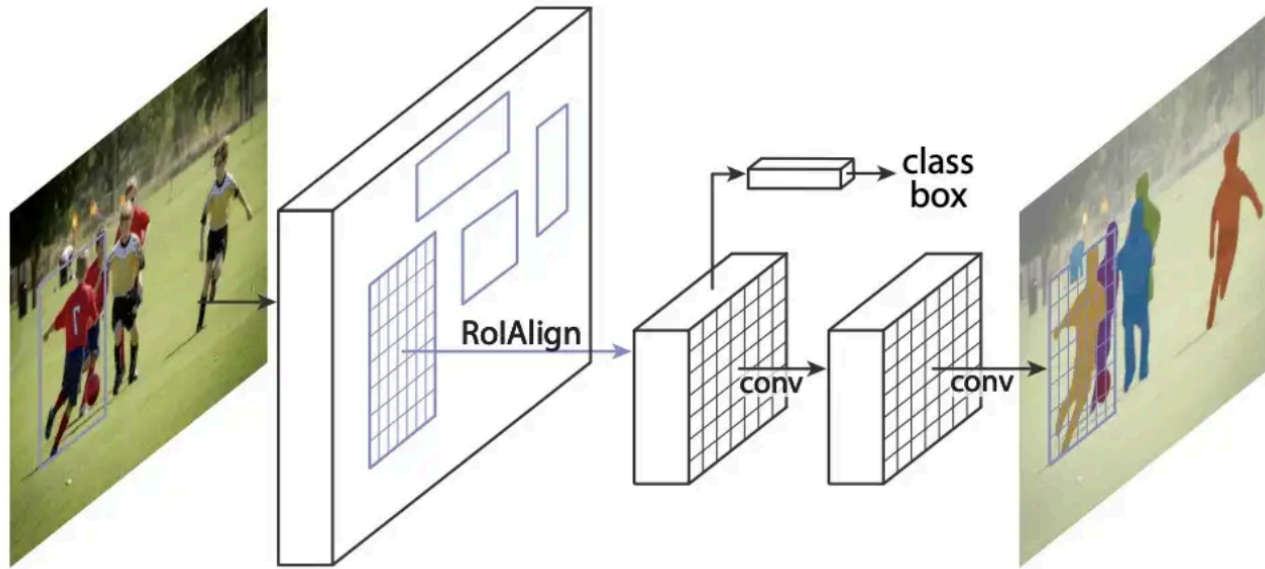
It's getting late



Object Detection - Mask R-CNN

But we are finishing

- Daha keskin konum tespiti
- Konum ve tanıma işini yapan iki NN



Tips and tricks

That time of the day

- Her layer lineer classifier olarak davranır.
- Model derinleştirildikçe iyi olma garantisi yoktur.
- Başlangıç değerleri uniform ya da çok küçük seçilmemelidir
 - Exploding - Vanishing gradients
- Normalizasyon
- Batch normalization
 - Her zaman iyi sonuç verme garantisi bulunmaz

Tips and tricks

And another

- Bir kernel'in birden fazla kanal özelliğini bir fonksiyonda öğrenmesi karmaşıklığı arttırabilir
- Azaltmak için depthwise convolution
 - Her kanala ayrı özellik çıkarımı
- Kanalları öncelikle arttırıp depthwise convolution kullanmak
 - SVM benzeri