deep learning book IAN J GOODFELLOW

THE CONVOLUTIONAL OPERATION

$S(i, j) = (I * K)(i, j) = \square m \square n I(i + m, j + n)K(m, n)$

- for less computation. (the diagram in the book)
- for feature extraction.
- **sparse interactions**
- **parameter sharing**
- **equivariant representation**
- receptive fields of s3, reduced computation
- the representation of the features with each conv op.
- final diagram.
- example.

POOLING OPERATION

- to reduce the computational cost for next layers.
- gaussian distribution and variance and strong priors.
- diagram.
- example

variants of conv op

- single kernel extracts one feature. we need single layer to extract many features at many locations.
- 4-D kernel tensor
  - $Z\ i,j,k = \square l,m,n V\ l,j+m-1,k+n-1\ K\ i,l,m,n$
- i,j,k,l positions.
- K i,j,k,l giving the connection strength between a unit in channel i of the output and a unit in channel j of the input, with an offset of k rows and l columns between the output unit and the input unit.

STRIDE

- we want to skip some computations for better computation so we introduce **stride**

- $Z\ i,j,k = c(K,V,s)\ i,j,k = \square l,m,n \square\ V\ l,(j-1)\times s+m,(k-1)\times s+n\ K\ i,l,m,n \square$

POOLING

- we require balance in reduced computations and good accuracy.
- the diagram.
- the example.

STRUCTURED OUTPUTS

- S i,j,k where i is the class j and k are probability of the pixel.
- to say this in simple we find the features in individual pixels then neighboring pixels then final class prediction.

DATA TYPES

1-D = single channel (1-D audio form). multi channel (skeleton animation data).

2-D = single channel (2-D audio that is Fourier transformed). multi channel (color image data).

3-D = single channel (volumetric data-medical imaging). multi channel (color video data).

SUPERVISED AND UNSUPERVISED TECHNIQUES IN CNN's KERNEL(conv op)

- unsupervised learning is used to reduce cost function.
  - in terms of kernels:
  - initialize them randomly
  - design them by hand.
  - use unsupervised learning = apply knn to each patch then combine all the centroid for convolutional op.
- random filters work really good with cnn.
- greedy layer pretraining , extracts the features in isolation then the output is given as input to the next layer