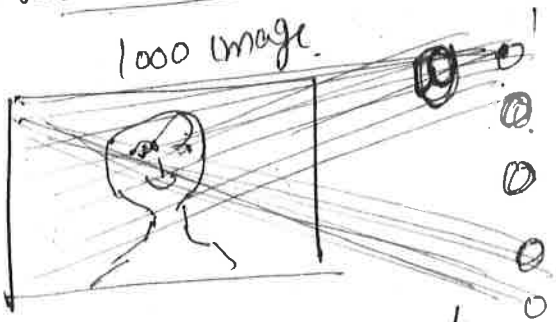
Deep Learning
(Machine Learning)

COL 774

April 29, 2017

(1) Convolutional Neural Networks:



Too many parameters.

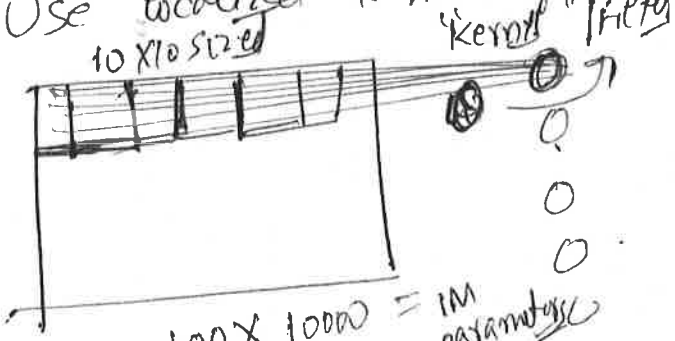
If we have 1000 hidden units

1 billion parameters
Too many!

(2)

Idea:-
Image has a spatial arrangement
⇒ Exploit compositionality
→ Face is made of multiple parts.

Use localized kernels:-



100 X 10000 = 1M parameters
(for a single kernel).

(3)

Further:- the parameters are tied:-

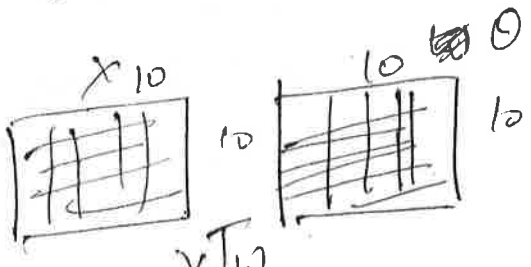
10x10 = 100 parameters + 1 for bias

∴ There can be multiple kernels

[depth d] such kernels ⇒ (100+1)d parameters

"Each kernel learns a different kind of pattern"

Given a kernel W:-



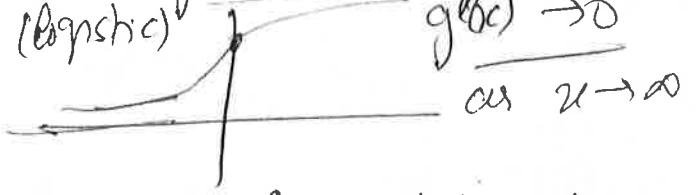
(4)

$h_\theta(x) = g(\theta^T x)$

⇐

g function:- $g(x) = \frac{1}{1+e^{-x}}$

(a) sigmoid (logistic)



$g(x) \to 0$ as $x \to -\infty$

(b) tanh:- $g(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

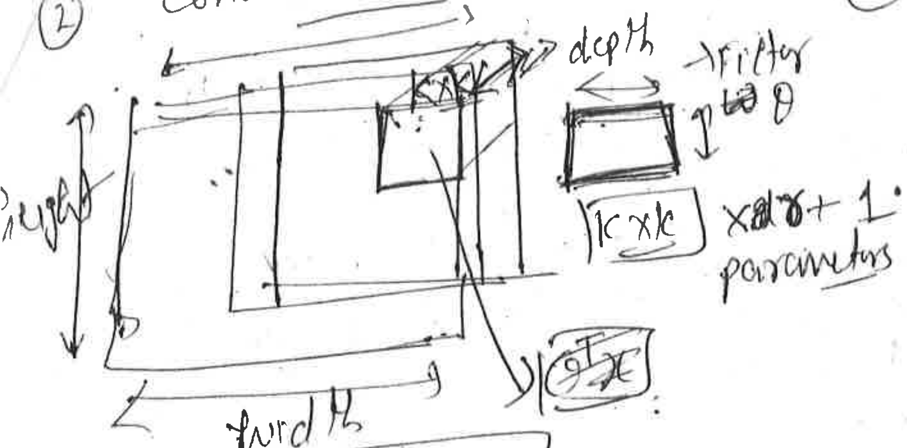
$g(x) \to c$ as $x \to \infty$

$g'(x) = \frac{1}{(e^x+e^{-x})^2}$

(c) $g(x) = \ln(1+e^x)$

$g'(x) = \frac{e^x}{1+e^x}$

Convolution:-



stride = S

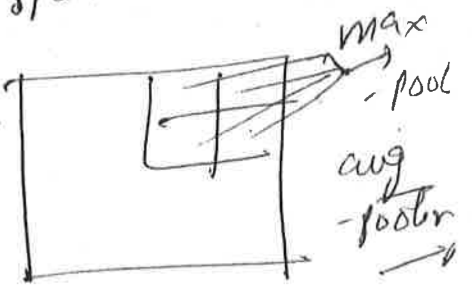
t - kernels

Output:- H × W

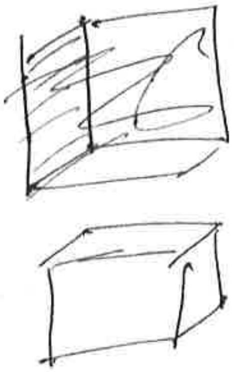
$$\left[\frac{(h-k)}{S}+1\right] \times \left[\frac{(w-k)}{S}+1\right] \times t$$

Pooling:-

Simply reduce the spatial size / down sample

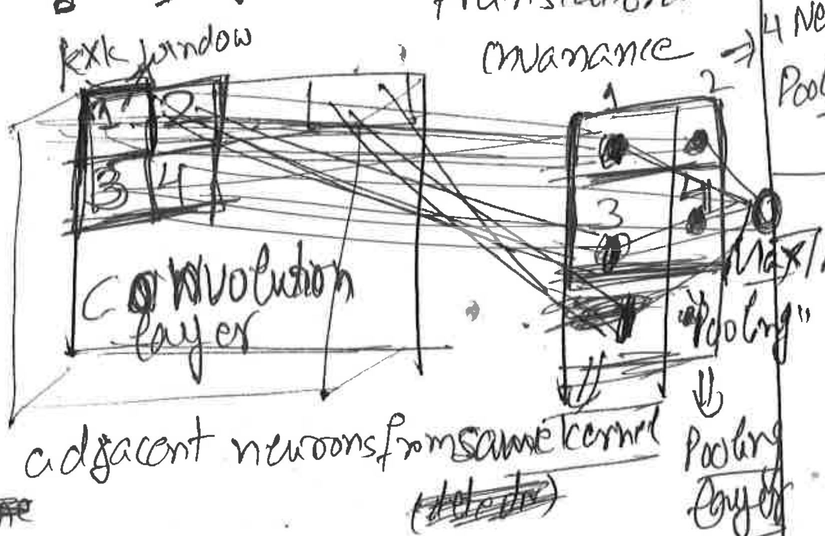


Fully connected layers:-



(Max) Pooling:-

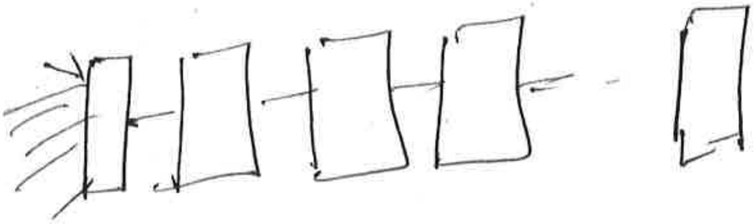
Takes care of translational invariance



adjacent neurons from same kernel

Multiple such layers:-

"conv + Pool" or "conv"



Imagenet Classification (Alexn

ILSVRC (Imagenet (2012)

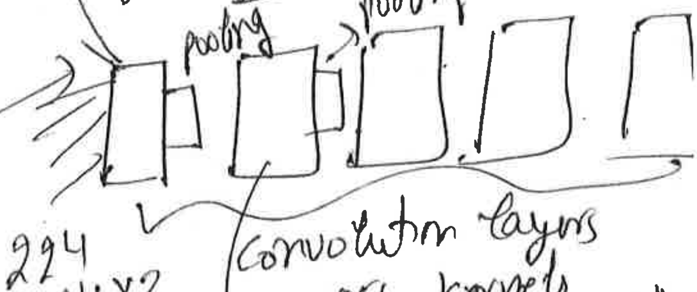
large scale visual recognition challenge)

label 1000

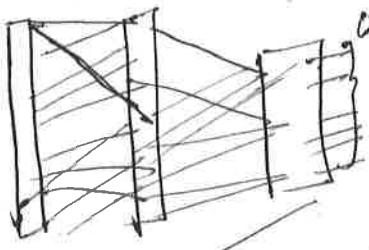
↳ Top 5 error rate of 15.4%

Train – 1.2M
Test – 150,000

Second Best :- 26.2%



Fully Connected Layer



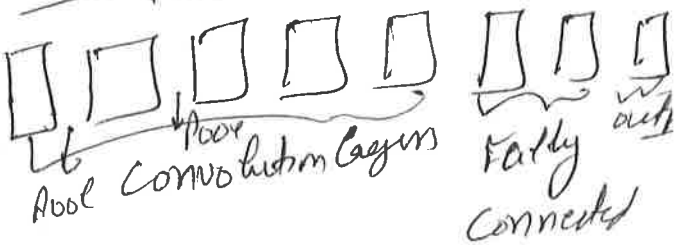
↓ implements softmax

$O^{(L)} =$ [illegible]

L: Number of labels

↓ generalization of Logistic

Alexnet [2012]



Dropout — (hidden layer)

Randomly omit each to compute with $P = 0.5$

↓ Avoids overfitting