CNN theory

梁鼎

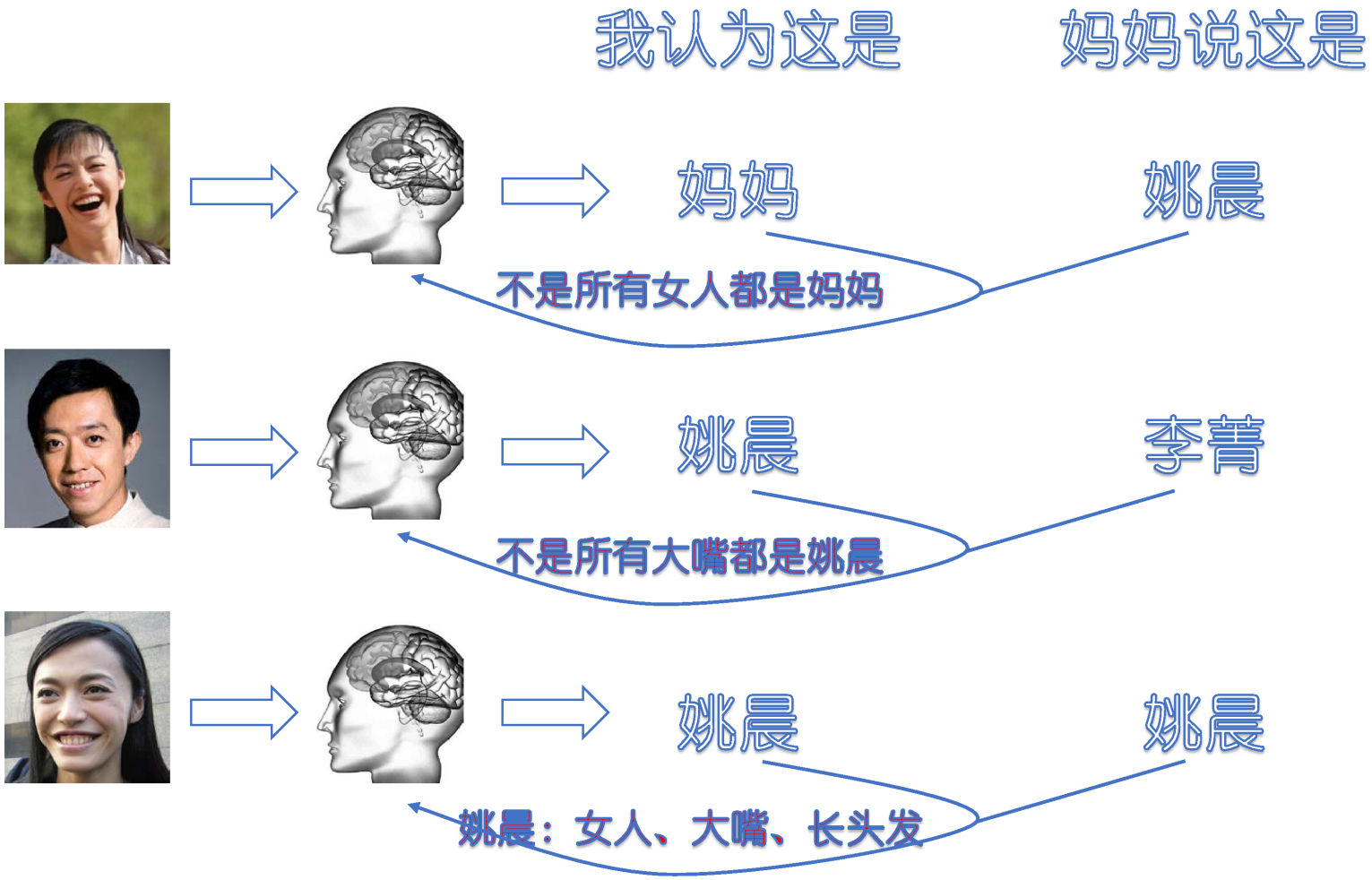
July 13, 2015

Contents I

Attention

- Only basic theory
- No formulas
- Mainly images

How do human learn from data



What we want

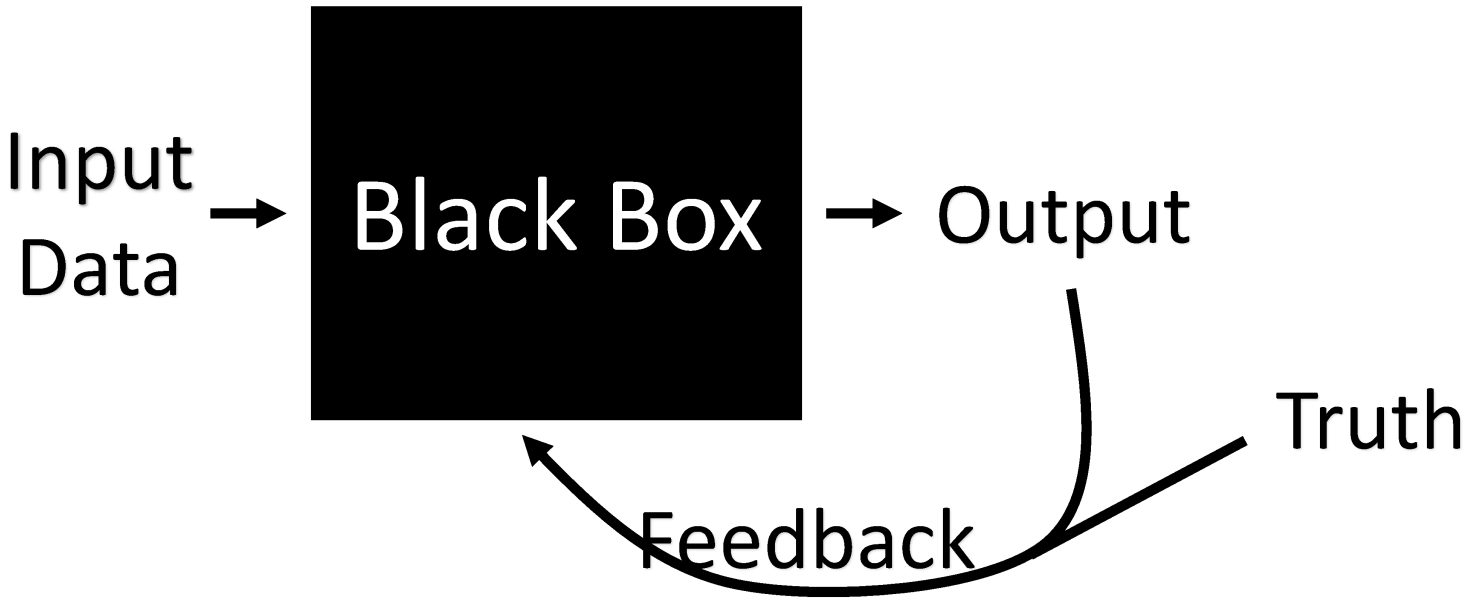


Figure: Make machine learn as people

Pipe

Figure: Control the pipes to let water flow into correct target

Minist Demo

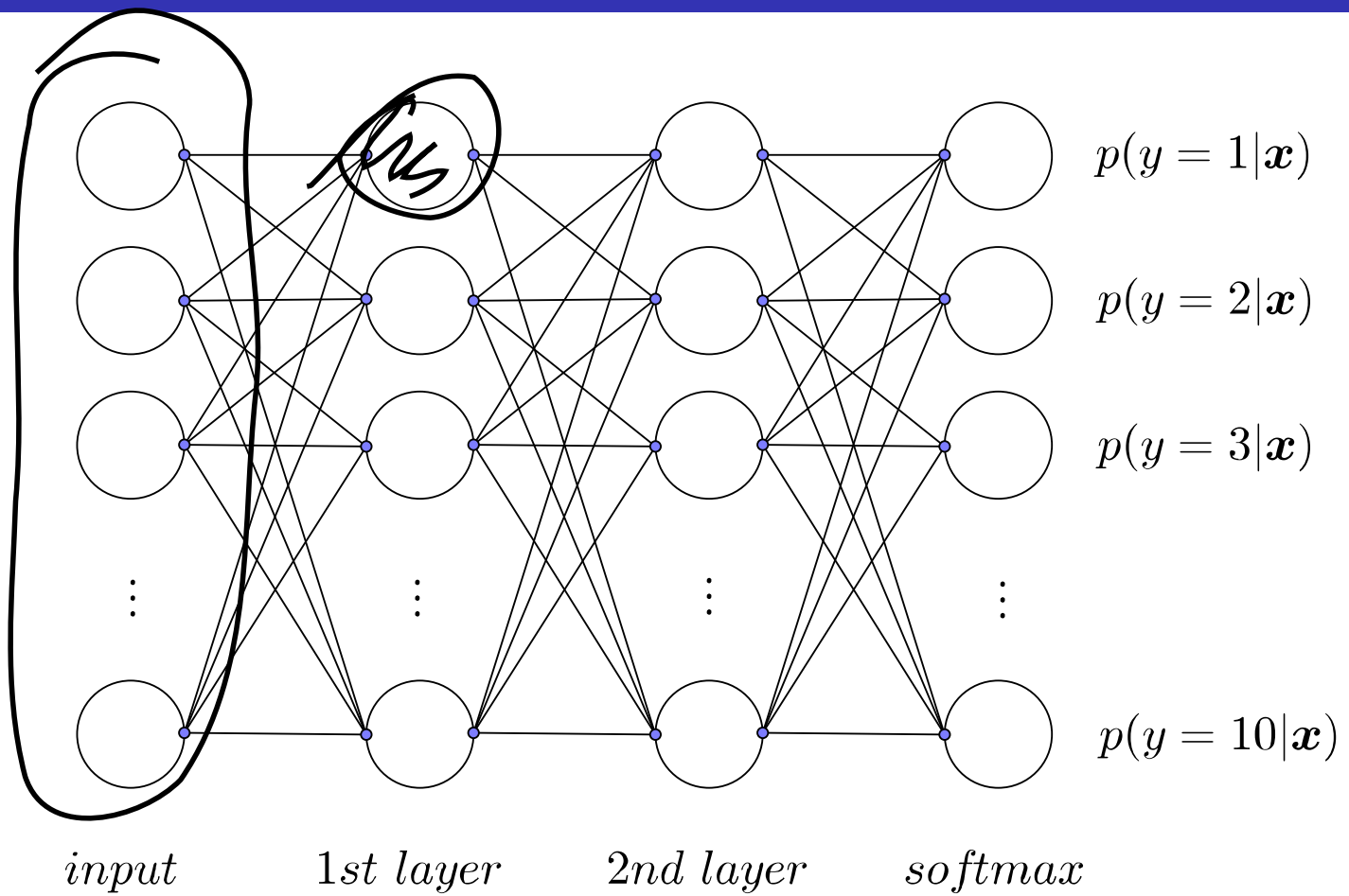
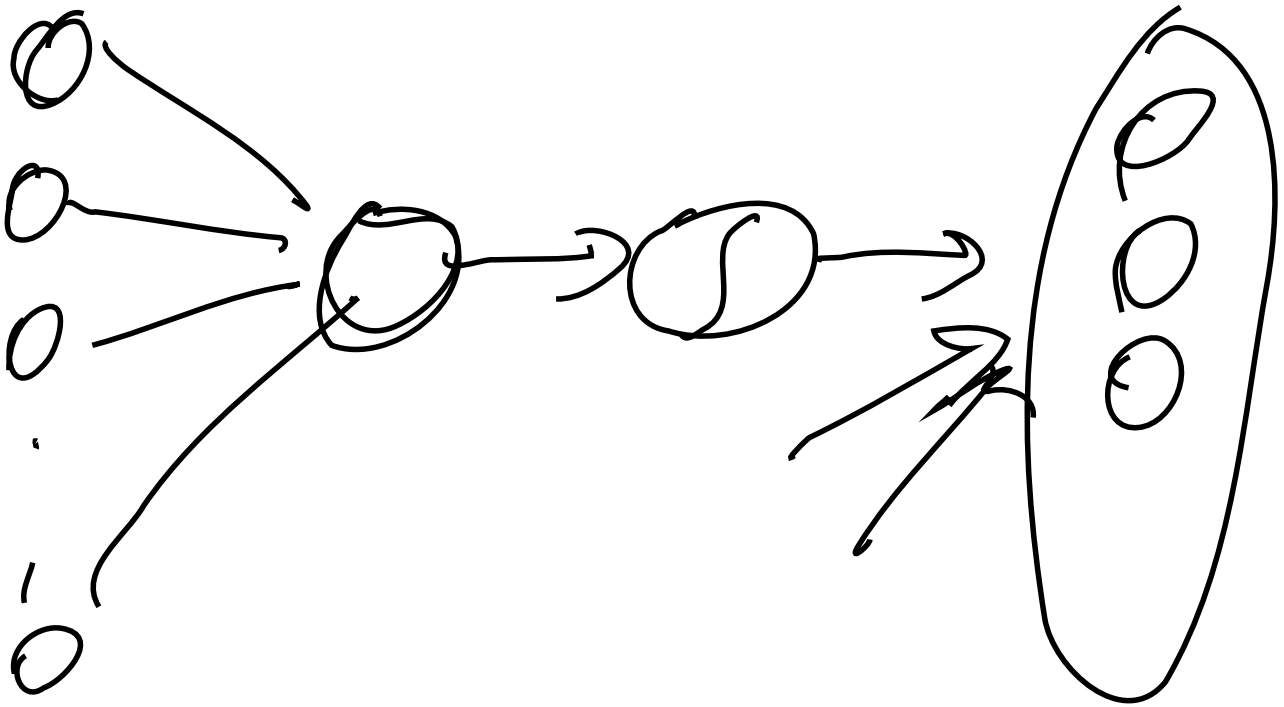


Figure: A demo neural network for digits recognition. Paramters are those pipes, controling the input data (water) flowing into every exit.

Neurons

1. Linear combination of all the neuron outputs of last layer
2. Pass through an active unit, bringing nonlinear components



Training steps

1. Forward
2. Backpropagation
3. Update parameters

CNN structure

CNN is an extended Nueral Network mainly designed for 2D input. More wonderful layers and extend 1D features to 2D feature maps.

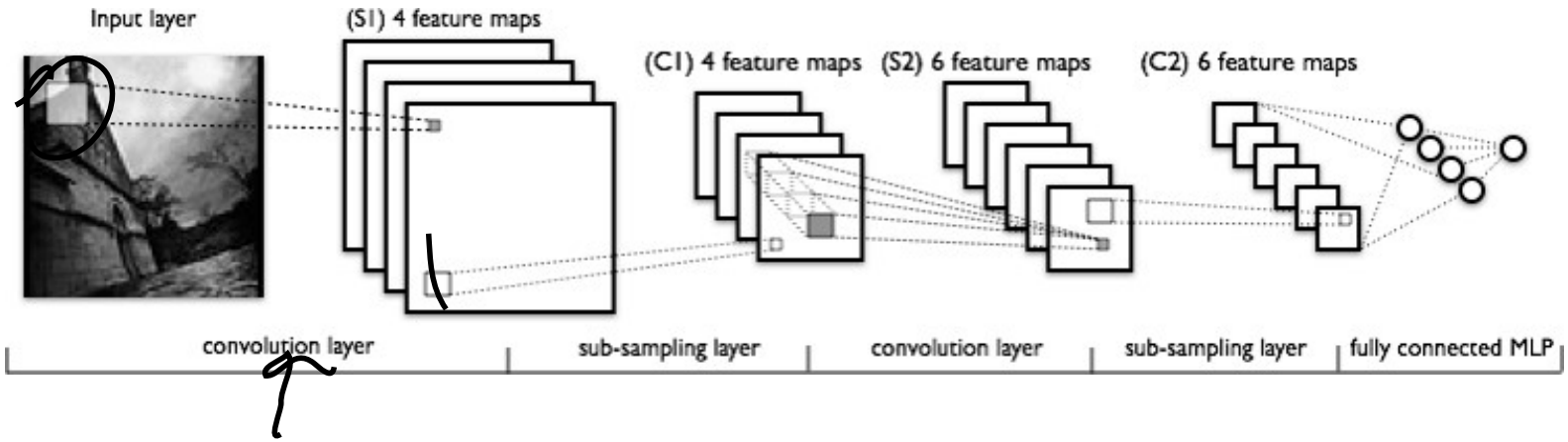
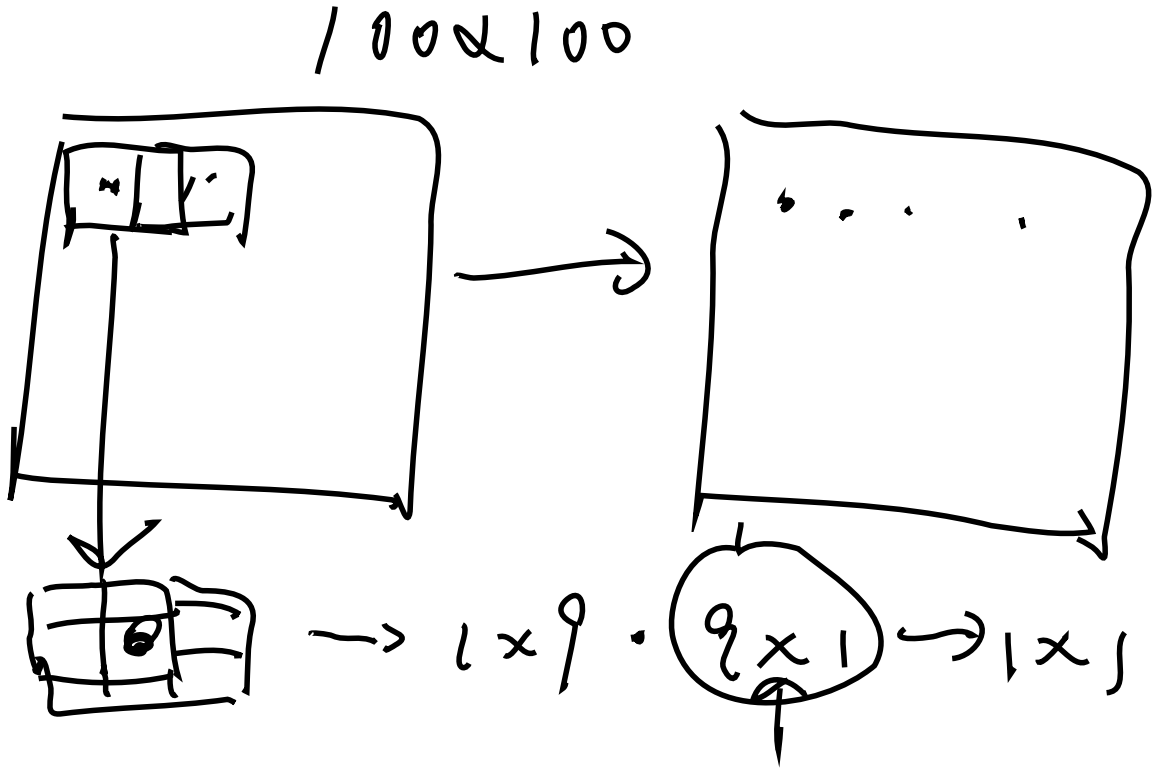


Figure: A simple CNN.

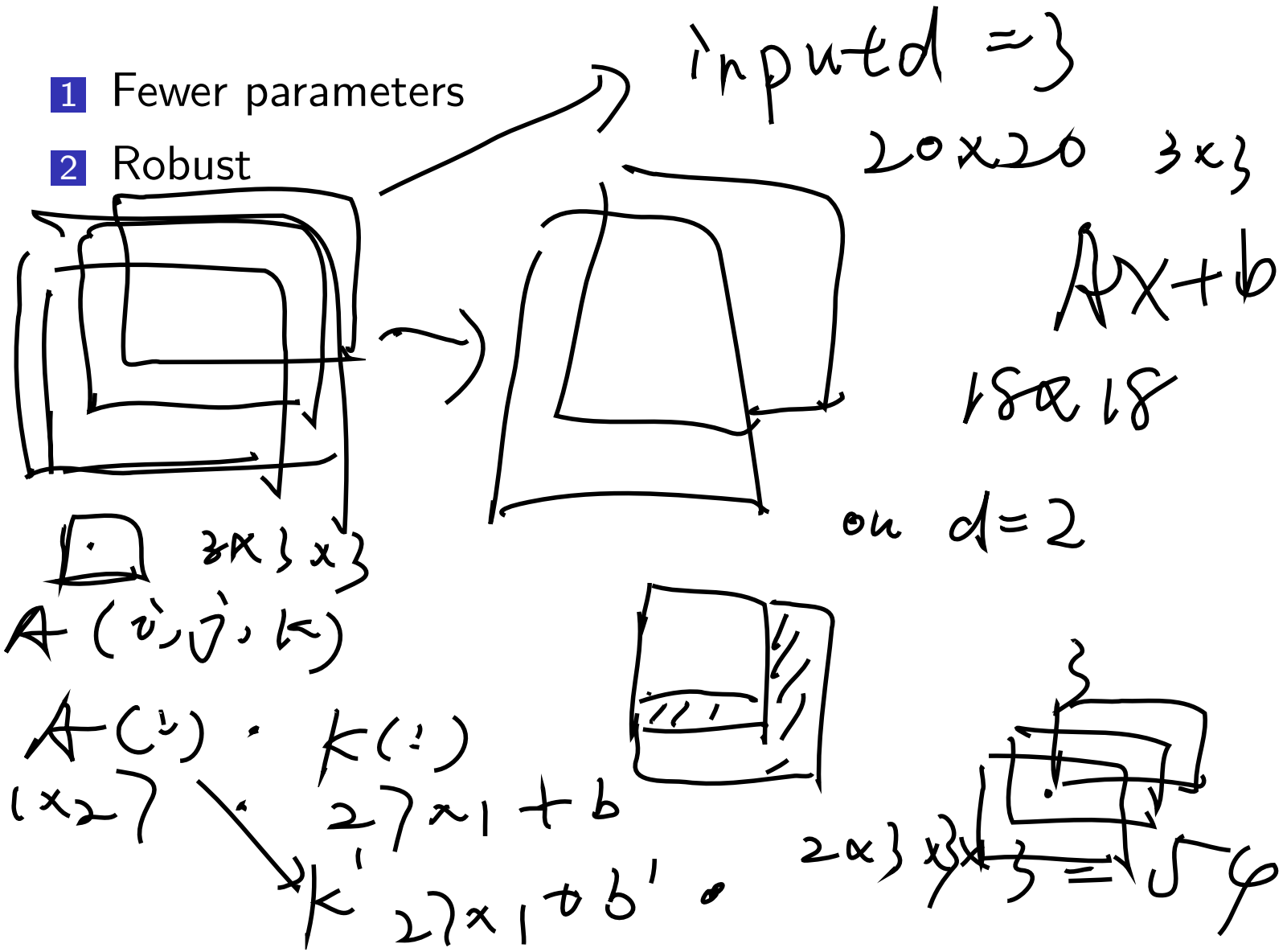
Convolution

1. Actually is still inner product
2. Combined with location infomation
3. Done locally



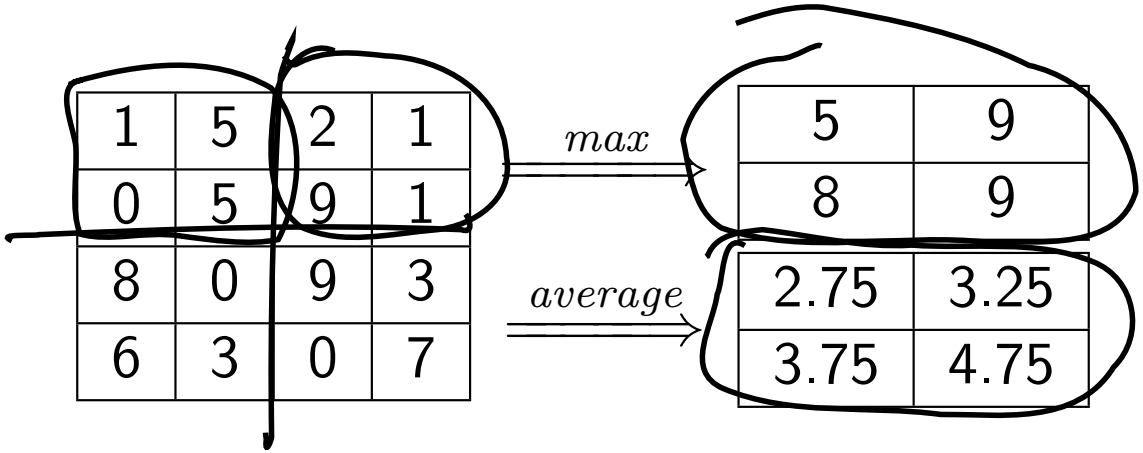
Convolution

1. Fewer parameters
2. Robust

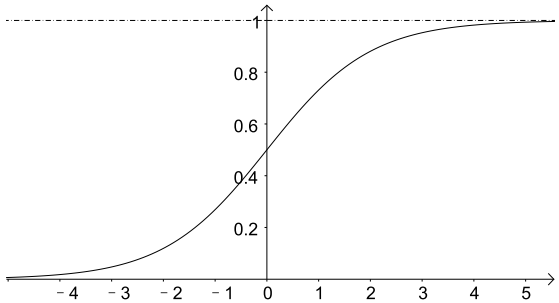
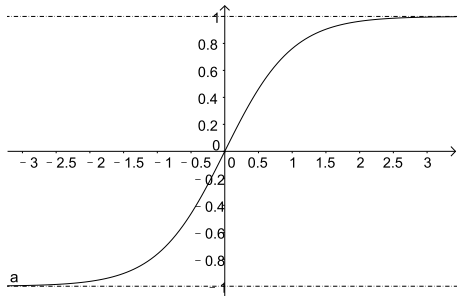
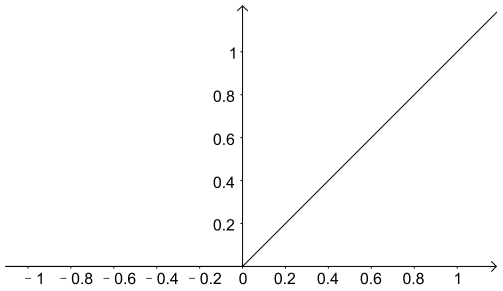


Pooling (Subsampling)

1. Preventing from overfitting
2. Translation invariant
3. Max or average and so on
4. Simple



Nonlinear

Name	Function	Range	Plot
Sigmoid	$\frac{1}{1+e^{-x}}$	$(0,1)$	
TanH	$\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$	$(-1,1)$	
ReLU	$\max(0,x)$	$(0,+\infty)$	

Innerproduct

1. Traditional Neural Networks
2. Large Parameter Space

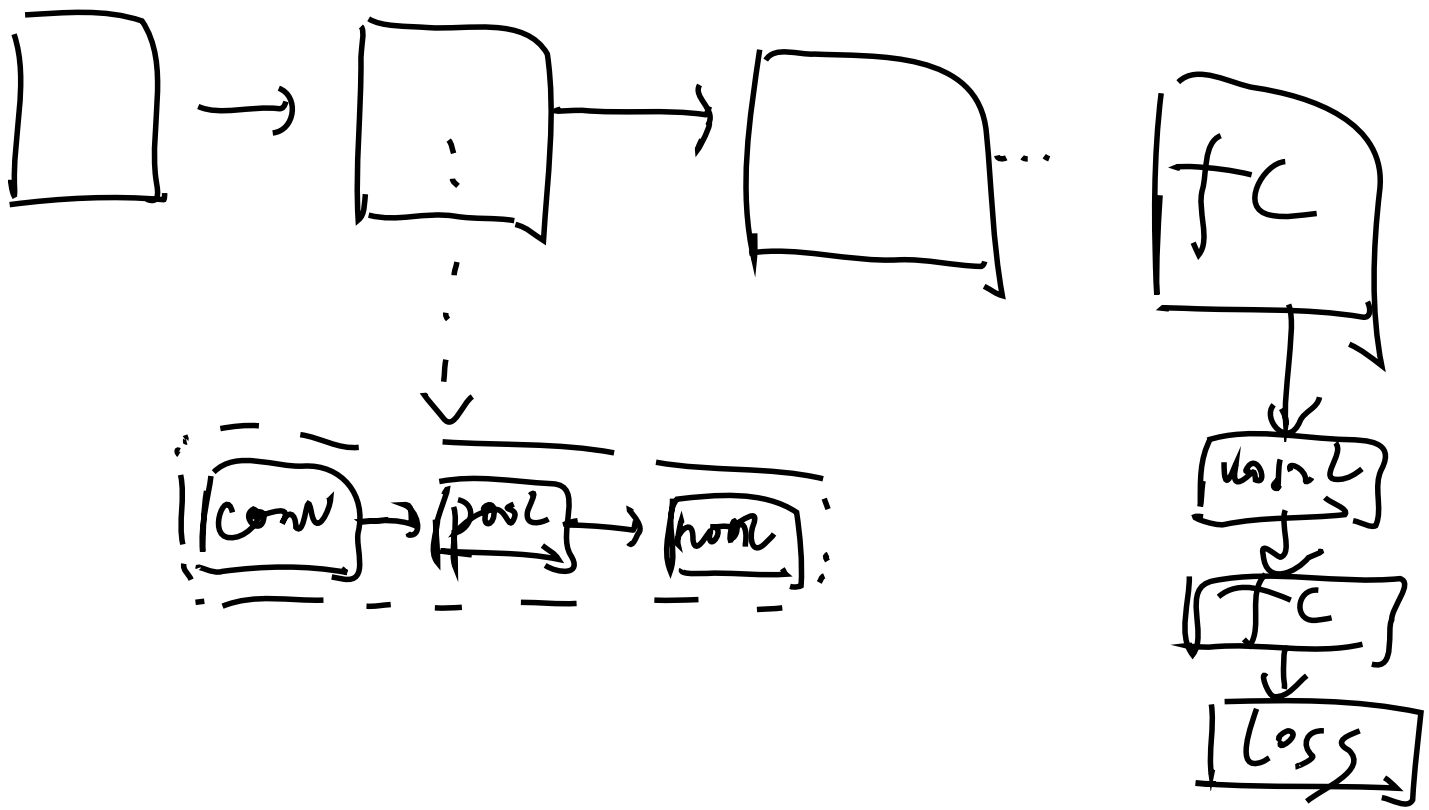
Loss Layer

1. Softmax Loss – Classification
2. Euclidean (L2) Loss – Regression

$$\frac{e^{-\theta x}}{e^{-\theta_1 x} + \cdots e^{-\theta_k x}}$$

$$\|y - y'\|_2$$

Build your own CNN



Difference

1. Dimension
2. Fewer parameters
3. More types of layers
4. Avoiding overfitting

Thanks!