xkcd.com/1897/

Convolutional Neural Networks

Slides adapted from Bryan Pardo, Mark Dredze, and Hung-Yi Lee

Logistics

- Quiz 10 due this Friday (Apr. 25)
- Last assignment – assignment 5 due next Monday (Apr. 28)

- The last quiz will be released in a moment, it's going to be free response questions on today's lecture. I will give you some time to finish it here.

- Last class "online asynchronous"

Logistics

- Course evaluation: fill out your course evaluation, **upload a screenshot showing you have completed it for CSC532 C**; then you earn 1 bonus point; if more than 90% of students finish this, then everyone gets 1 additional point. So in total, 2 bonus points available.

- https://uiscoursesurvey.apps.uis.edu

- Course evaluation bonus assignments has been available on Canvas, and it's due **23:45, May 3**.
- https://uispringfield.instructure.com/courses/15935/assignments/275198

Logistics

- Last course survey (1 bonus point) will be released online around May. 5, and due May 9
- Participation scores (attendance and Ed Discussion) will be released around May 9, right after the end of all the participation activities.

- Final Project will be due on 23:59, May 9. I need time to review your projects and submit final grades on time. Thus, please submit your project on time.
- Late final projects will get **a penalty of 15% per day up to three days after the due date.** After that, you can no longer submit the project.

Image Classification

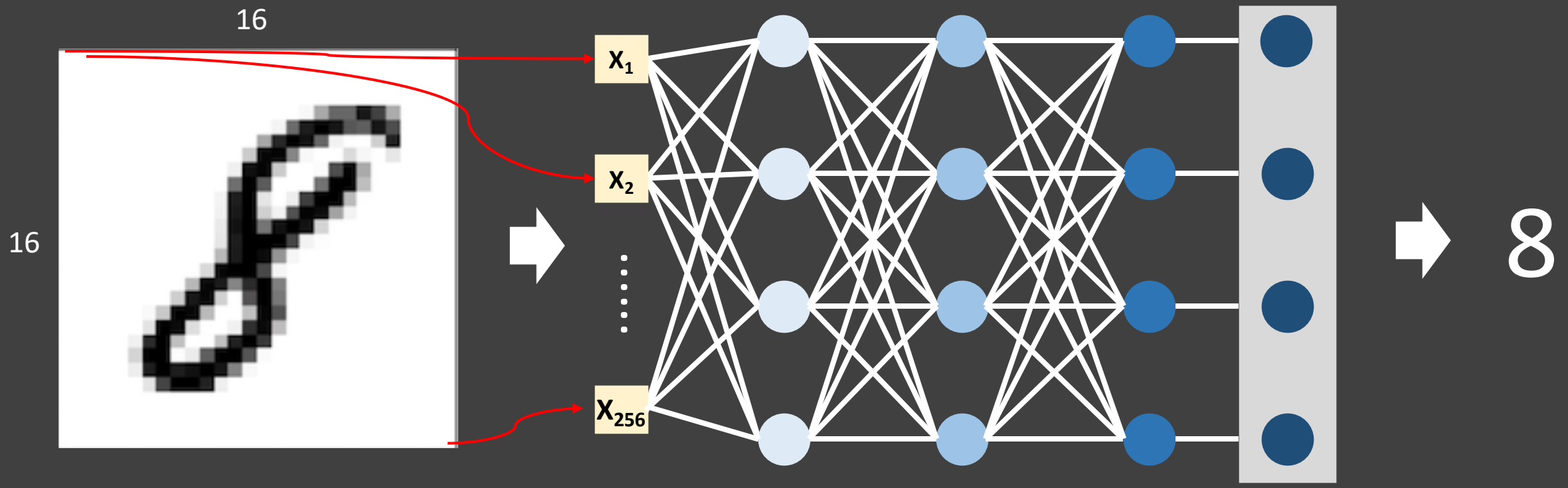
Recognize which digit it is!

$Y = f($  $)$

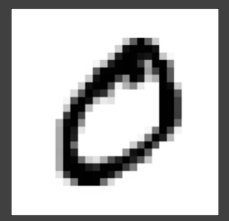
$P(Y = 0 \mid$  $)$ ... $P(Y = 9 \mid$  $)$

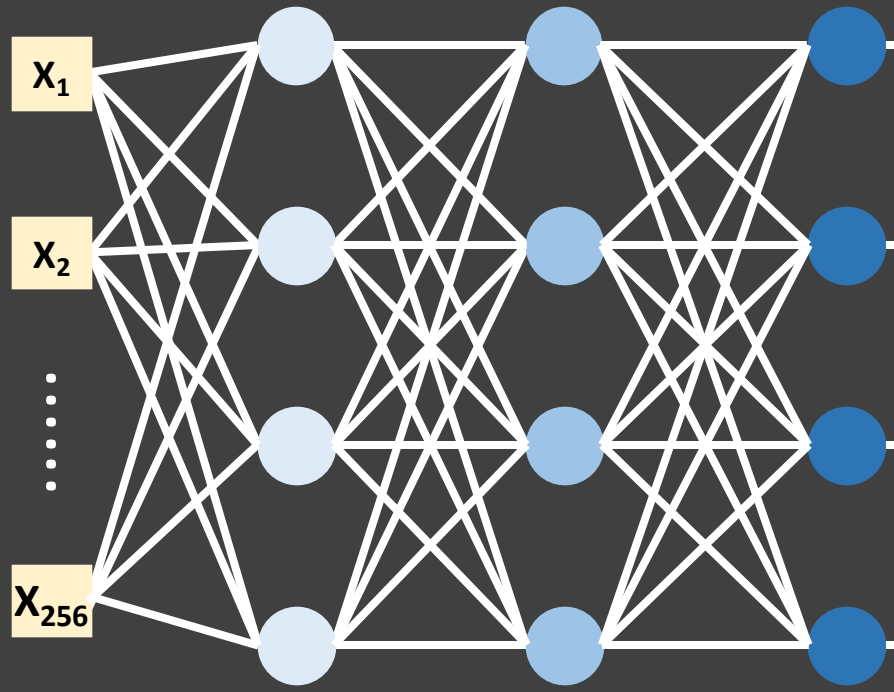
Please note that the slides for CNNs are adapted from
https://www.youtube.com/watch?v=FrKWiRv254g

Image Classification



Neural Network -- Inference





Probability of Input is 0

Probability of Input is 1

Find the Difference between Prediction and Truth

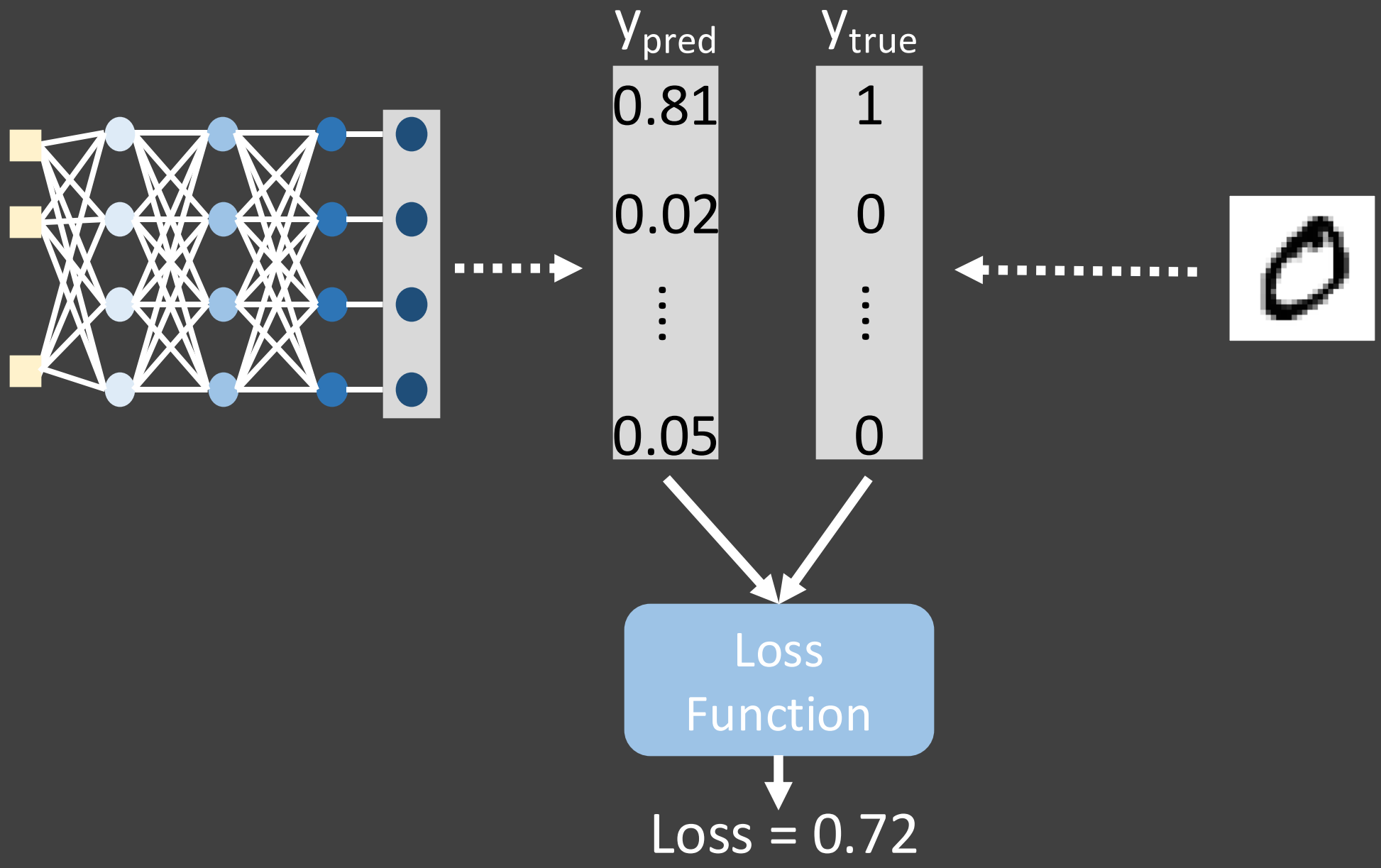
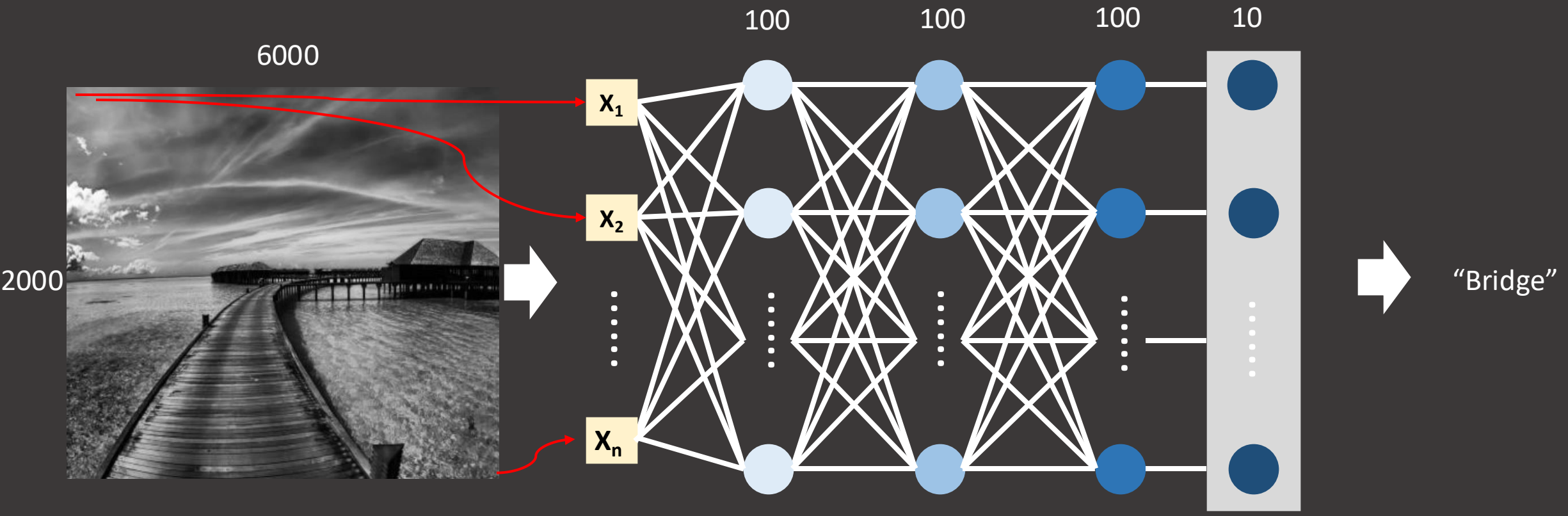


Image Classification



Input nodes = 6000 * 2000
Hidden nodes per layer = 100
Too many parameters!!!

CNN saves the world

- Some patterns are much smaller than the whole image

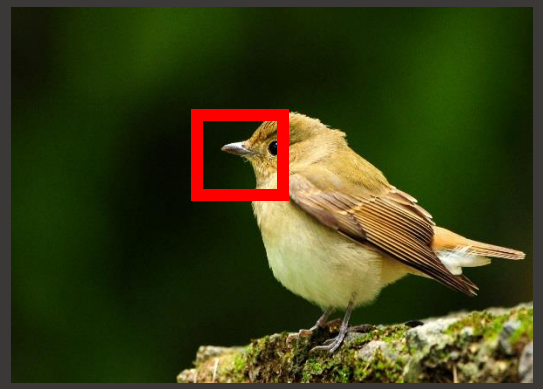
A neuron does not have to see the whole image to discover the pattern.

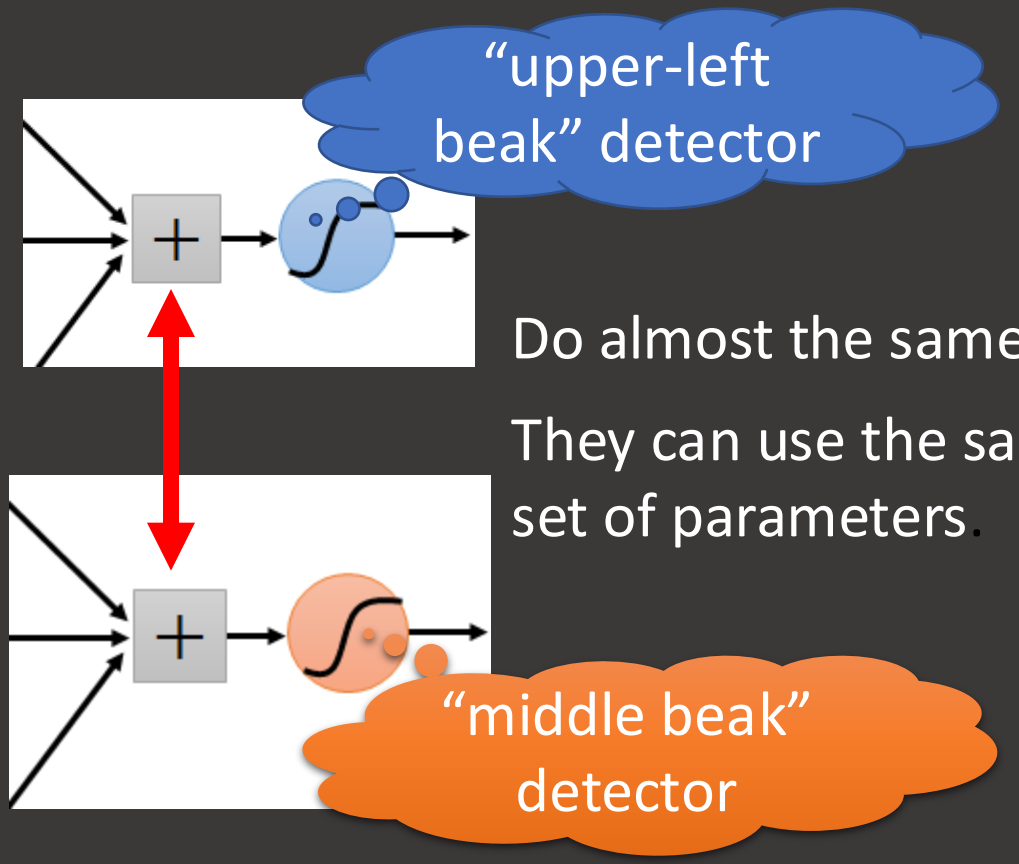
Connecting to small region with less parameters

Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

CNN saves the world

- The same patterns appear in different regions.





Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

CNN saves the world

- Subsampling the pixels will not change the object

bird

subsampling

bird

We can subsample the pixels to make image smaller

→ Less parameters for the network to process the image

Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

The whole CNN

Convolution

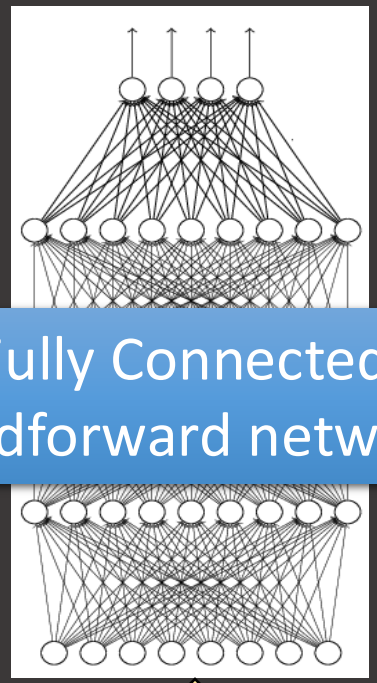
Max Pooling

Convolution

Max Pooling

Can repeat many times

Flatten



Fully Connected Feedforward network

cat dog ……

Slide adopted from https://www.youtube.com/watch?v=FrKWiRv254g

The whole CNN

Property 1

- Some patterns are much smaller than the whole image

Property 2

- The same patterns appear in different regions.

Property 3

- Subsampling the pixels will not change the object

Convolution

Max Pooling

Convolution

Max Pooling

Can repeat many times

Flatten

Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

The whole CNN

Convolution

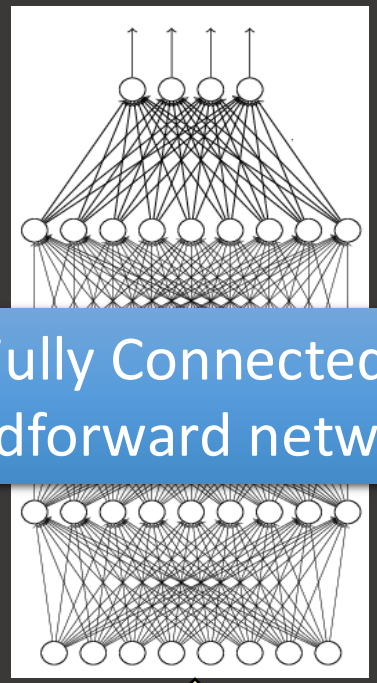
Max Pooling

Convolution

Max Pooling

Can repeat many times

Flatten



Fully Connected Feedforward network

cat dog ……

Slide adopted from https://www.youtube.com/watch?v=FrKWiRv254g

CNN – Convolution

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

Those are the network parameters to be learned.

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1 Matrix

-1	1	-1
-1	1	-1
-1	1	-1

Filter 2 Matrix

⋮

Property 1

Each filter detects a small pattern (3 x 3).

Slide adopted from https://www.youtube.com/watch?v=FrKWiRv254g

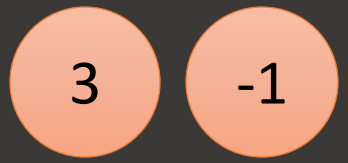
CNN – Convolution

1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

stride=1

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0



6 x 6 image

Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

CNN – Convolution

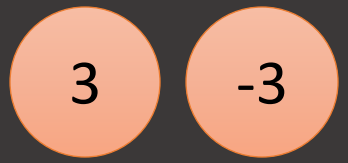
1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

If stride=2

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

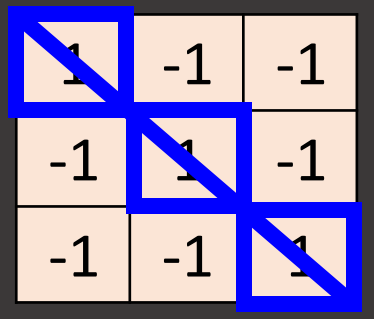
6 x 6 image



We set stride=1 below

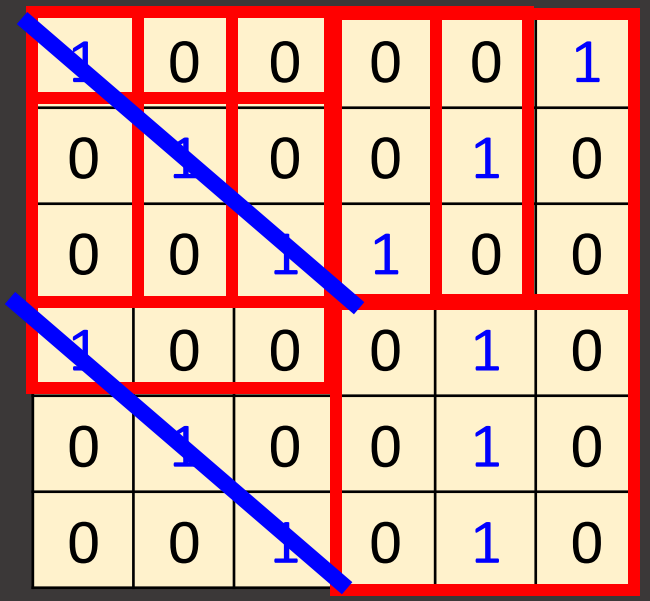
Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

CNN – Convolution

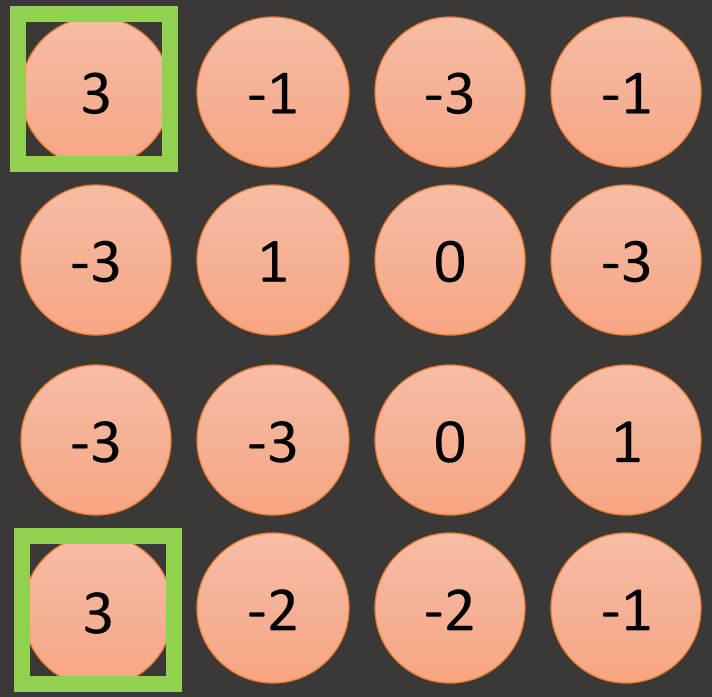


Filter 1

stride=1



6 x 6 image



Property 2

Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

CNN – Convolution

-1	1	-1
-1	1	-1
-1	1	-1

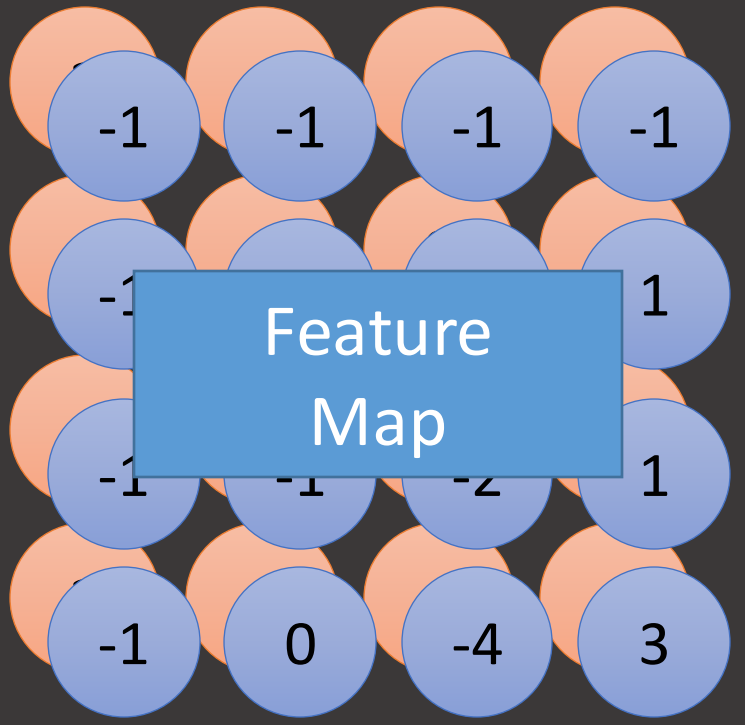
Filter 2

stride=1

1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

6 x 6 image

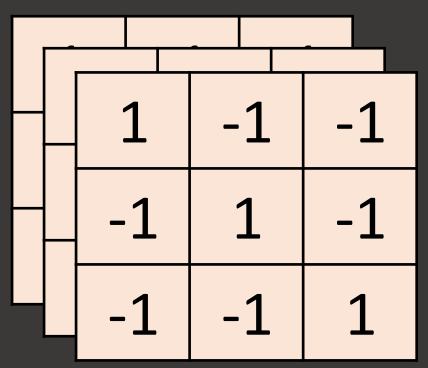
Do the same process for every filter



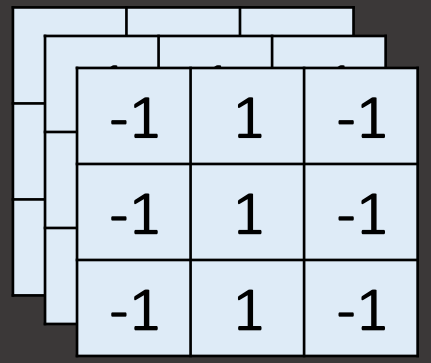
4 x 4 image

Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

CNN – Colorful image

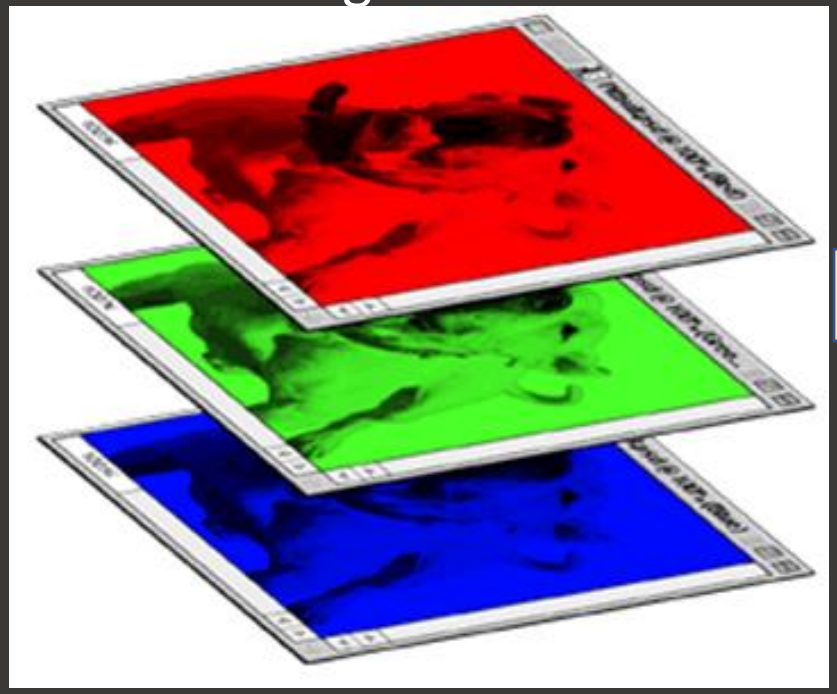


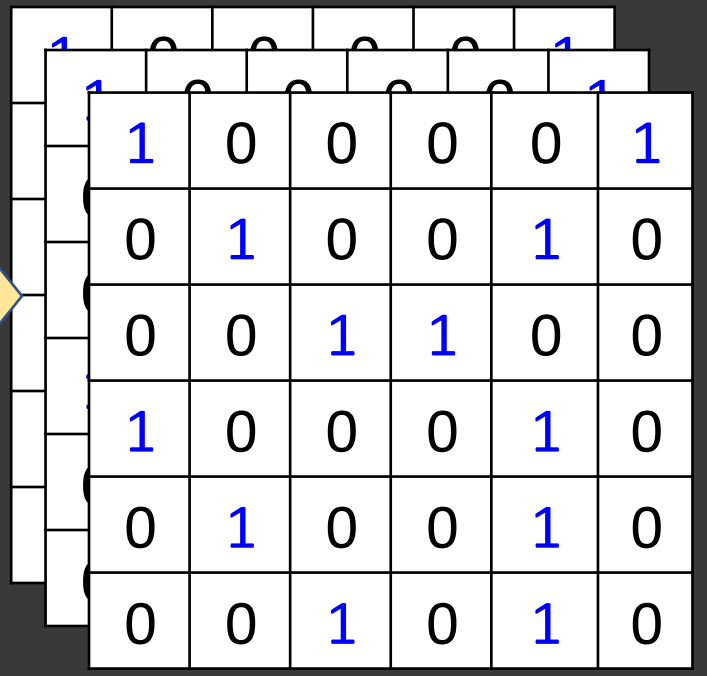
Filter 1



Filter 2

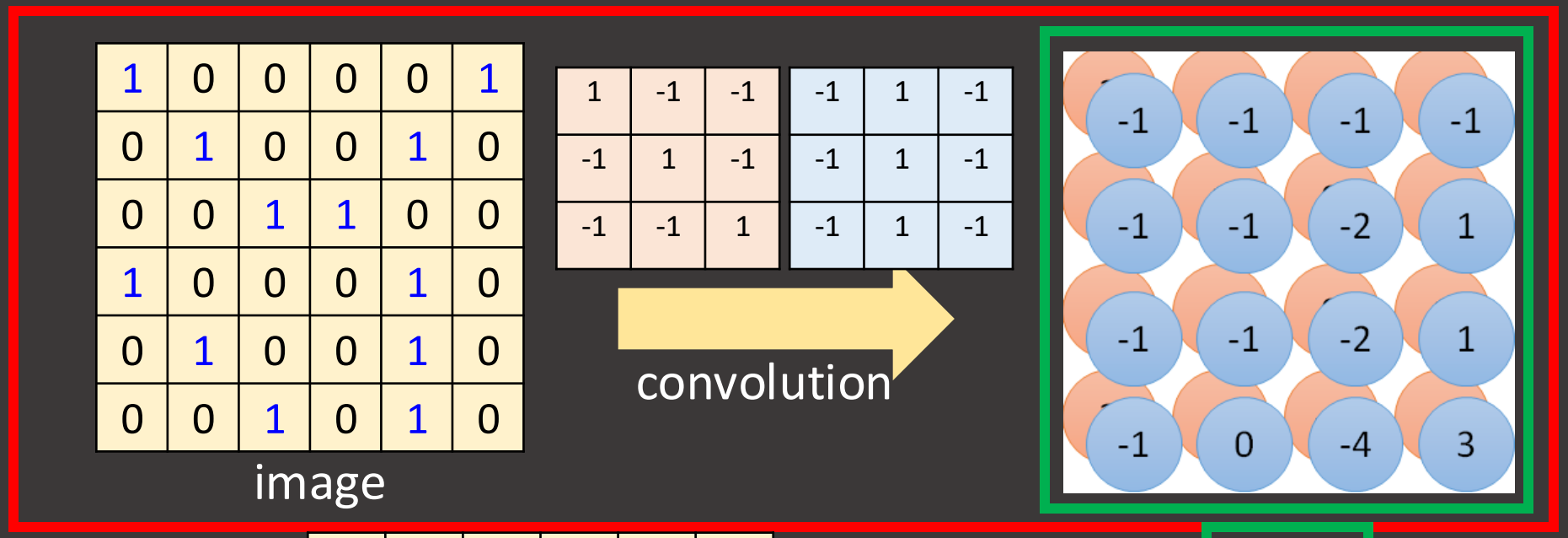
Colorful image





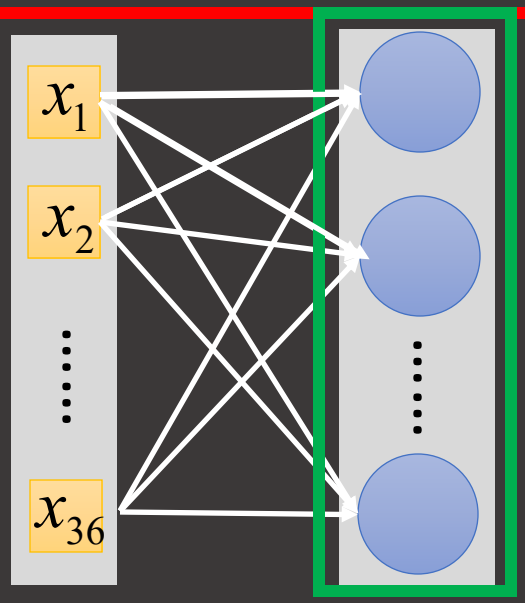
Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

Convolution v.s. Fully Connected

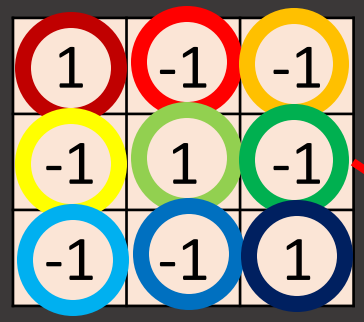


Fully-connected

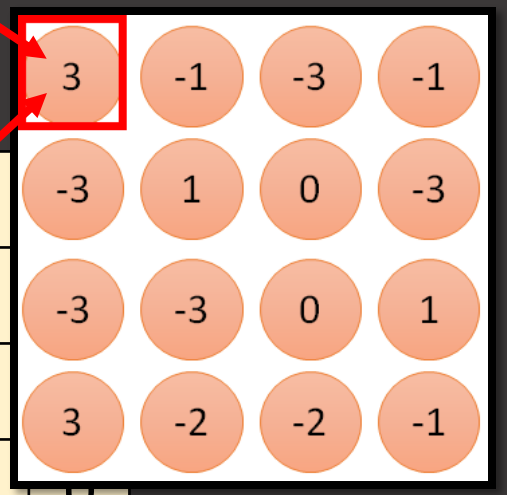
1	0	0	0	0	1
0	1	0	0	1	0
0	0	1	1	0	0
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0



Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g



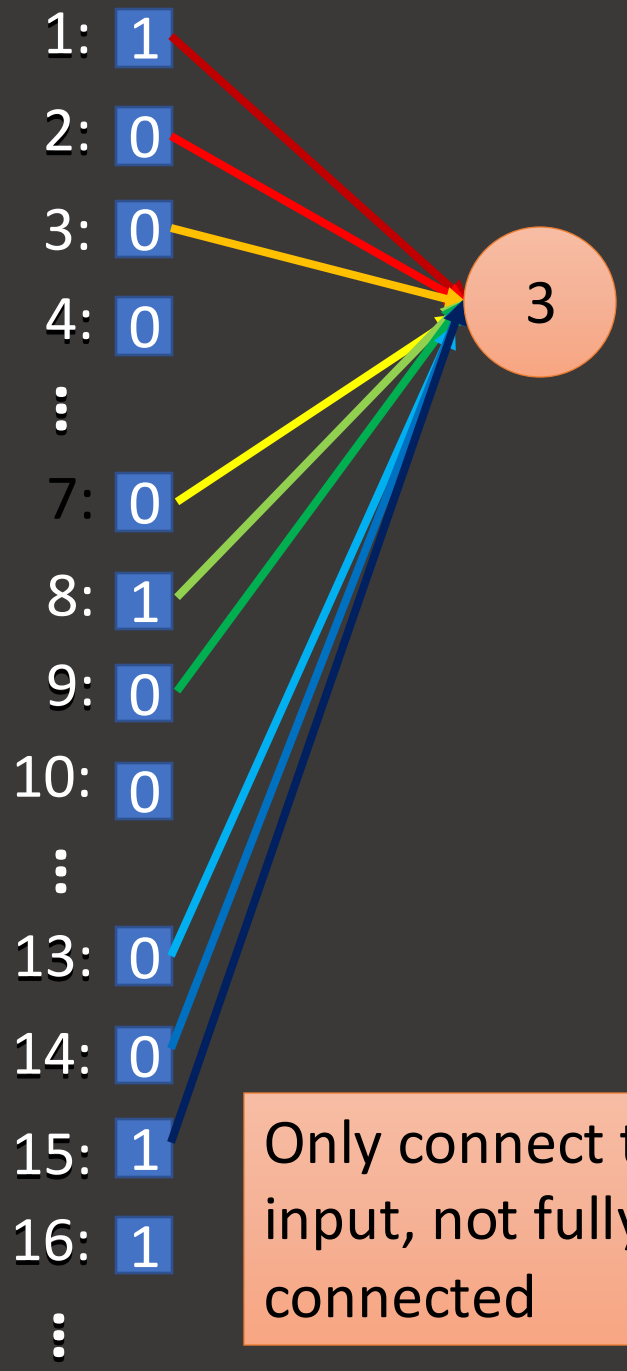
Filter 1



1	0	0	0	0	
0	1	0	0	1	
0	0	1	1	0	
1	0	0	0	1	0
0	1	0	0	1	0
0	0	1	0	1	0

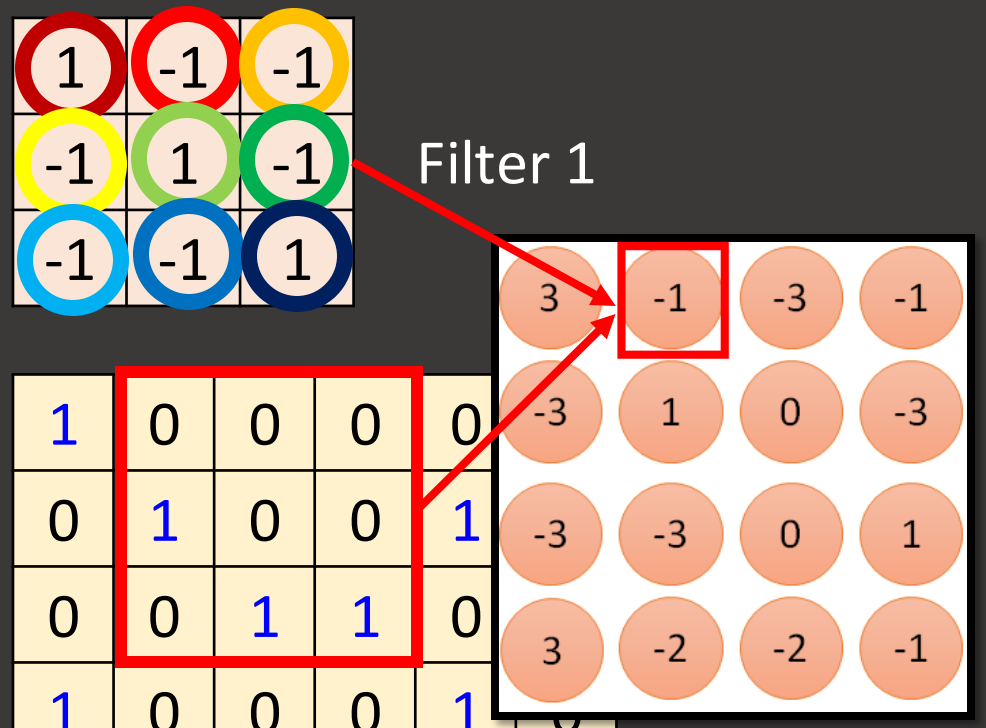
6 x 6 image

Less parameters!



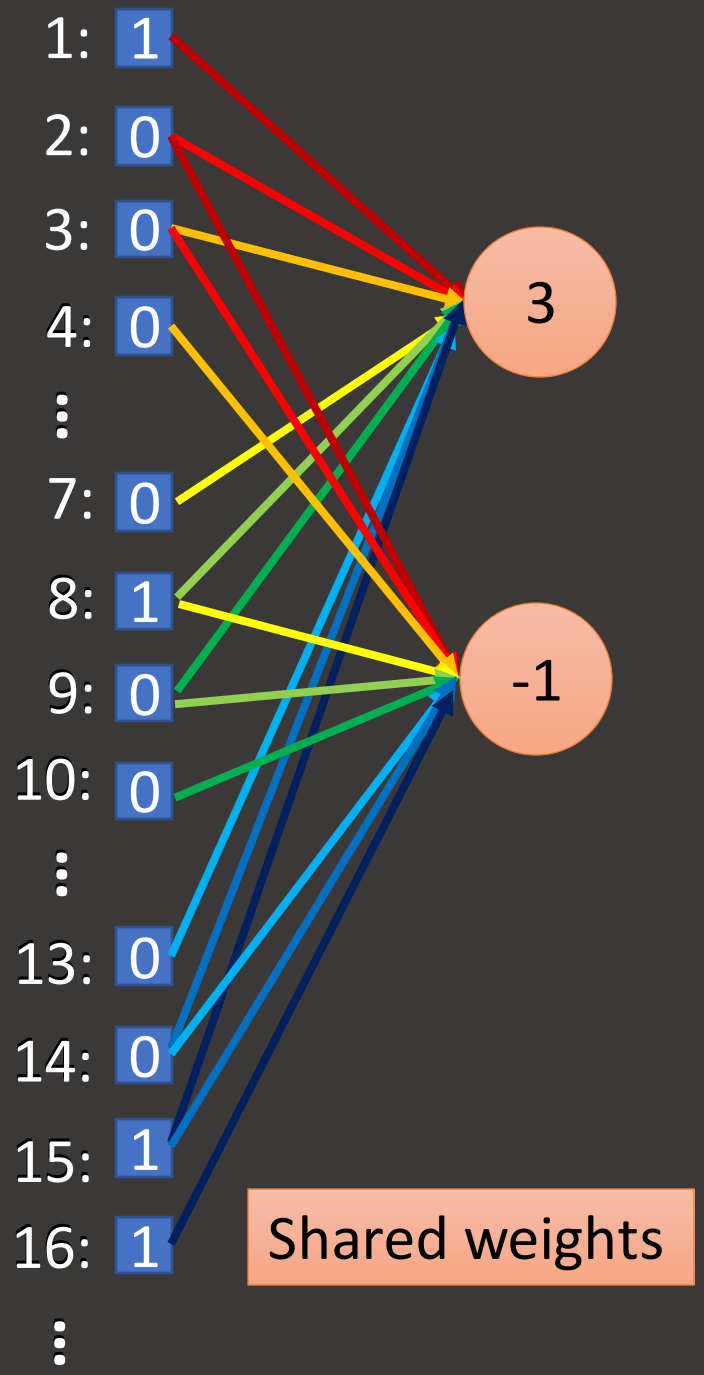
Only connect to 9 input, not fully connected

Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g



Less parameters!

Even less parameters!



Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

The whole CNN

Convolution

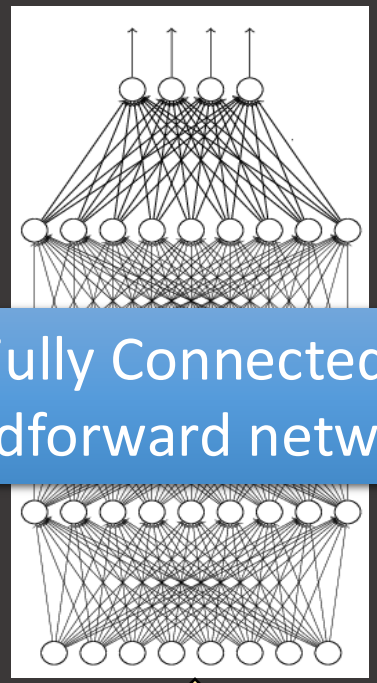
Max Pooling

Convolution

Max Pooling

Can repeat many times

Flatten



Fully Connected Feedforward network

cat dog ……

Slide adopted from https://www.youtube.com/watch?v=FrKWiRv254g

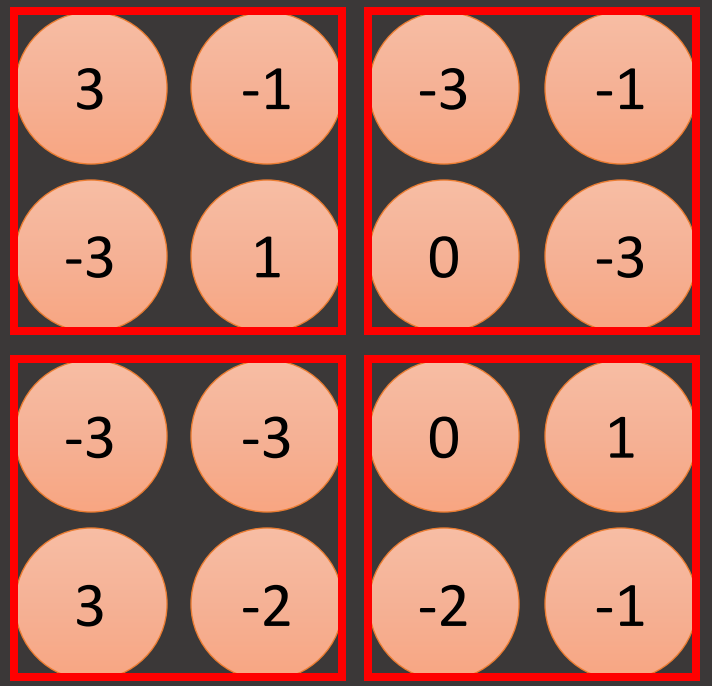
CNN – Max Pooling

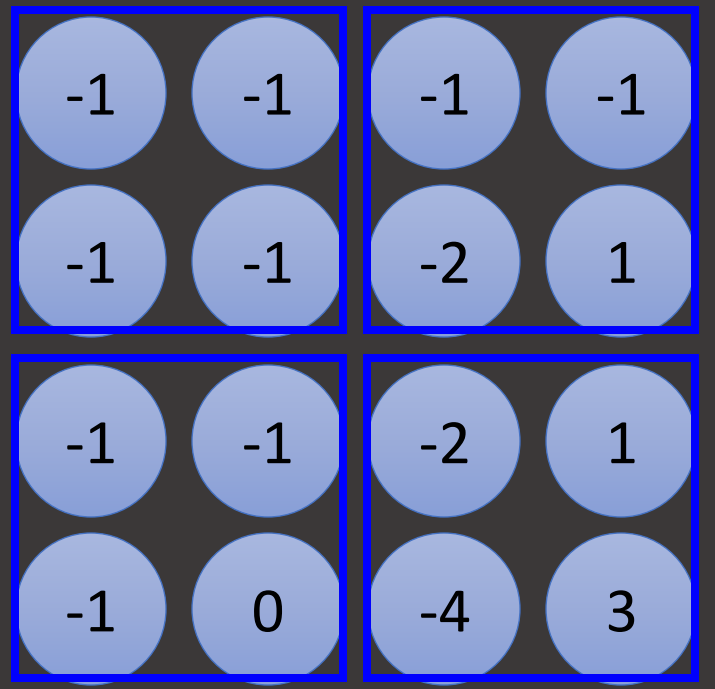
1	-1	-1
-1	1	-1
-1	-1	1

Filter 1

-1	1	-1
-1	1	-1
-1	1	-1

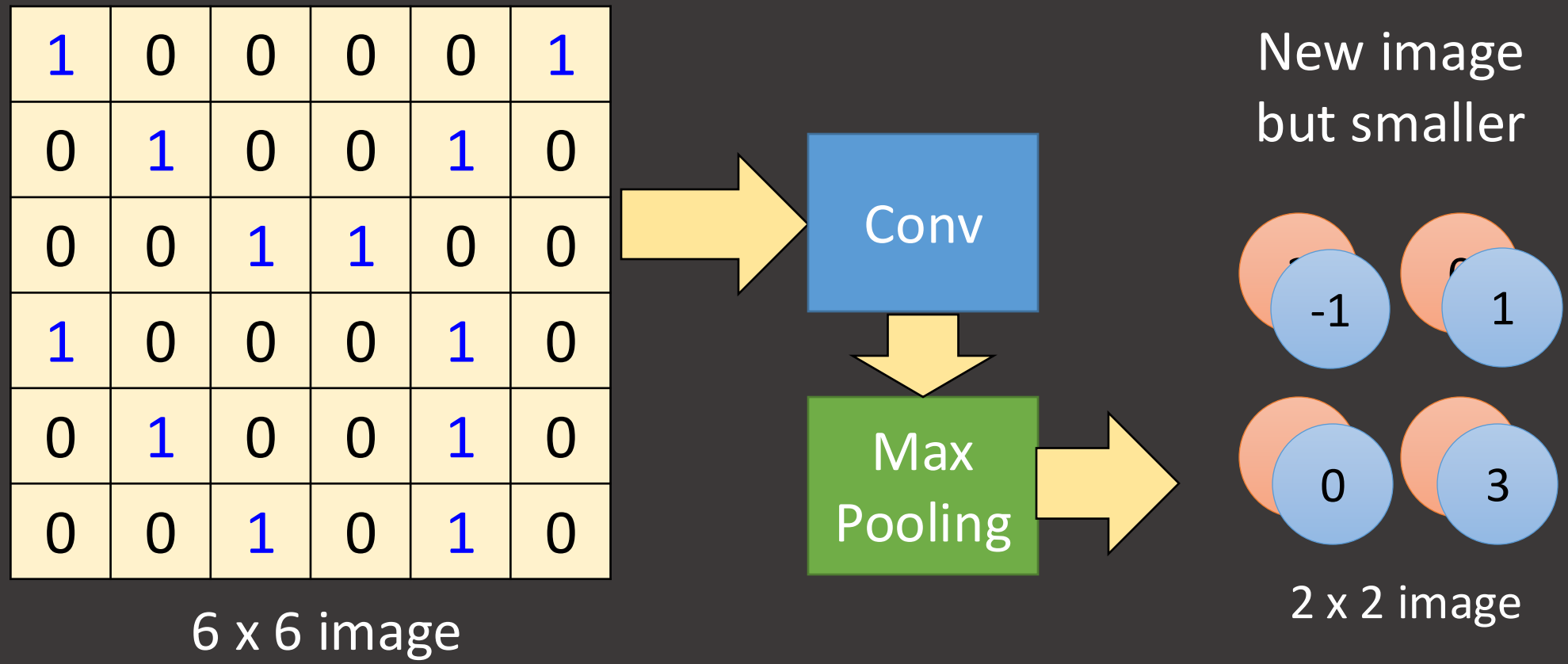
Filter 2





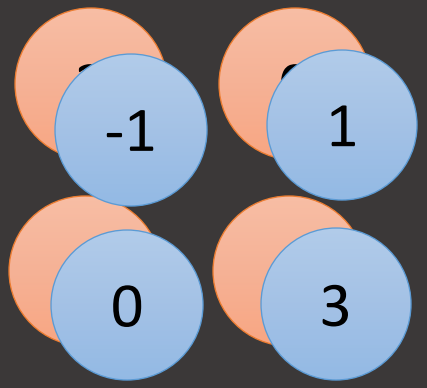
Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

CNN – Max Pooling



Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

The whole CNN



A new image

Convolution

Max Pooling

Convolution

Max Pooling

Can repeat many times

Smaller than the original image

The number of the channel is the number of filters

Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

The whole CNN

Convolution

Max Pooling

A new image

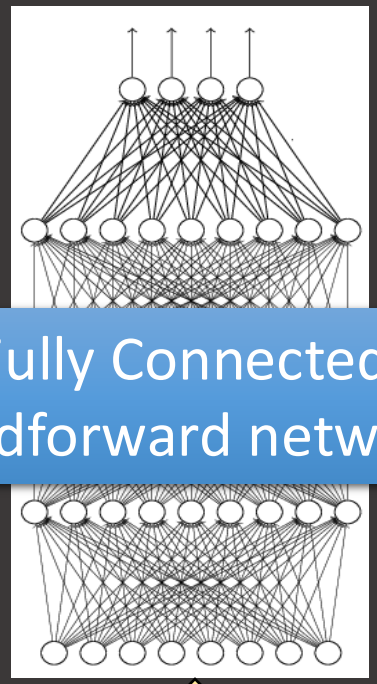
Convolution

Max Pooling

A new image

Flatten

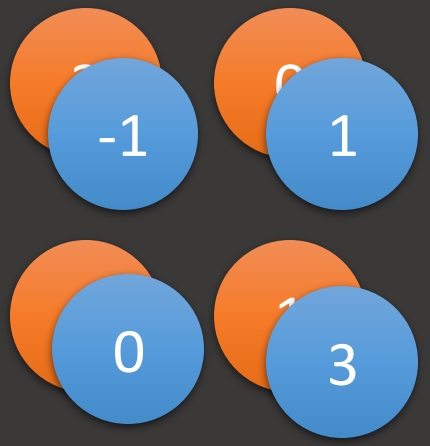
cat dog ……



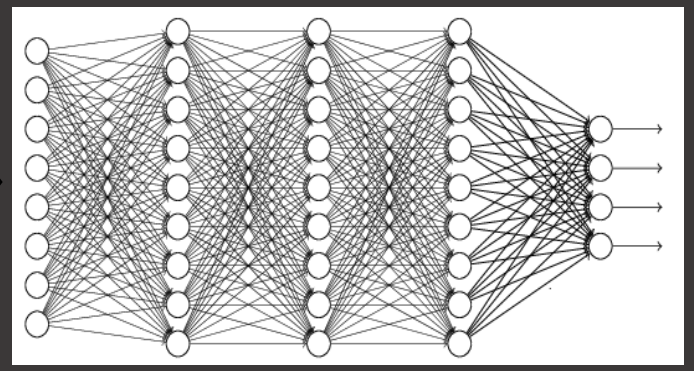
Fully Connected Feedforward network

Slide adopted from https://www.youtube.com/watch?v=FrKWiRv254g

Flatten



Flatten



Fully Connected Feedforward network

Slide adopted from
https://www.youtube.com/watch?v=FrKWiRv254g

How many layers?

- In general, more the conv layers, greater the extraction of features.
- Example – VGG 16

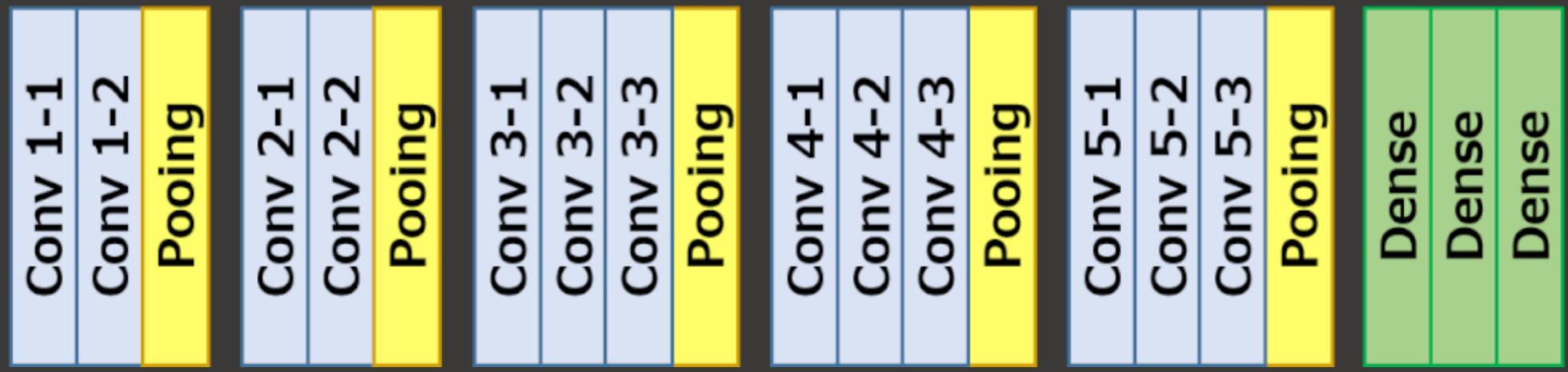


Image courtesy neurohive.io

How many layers?

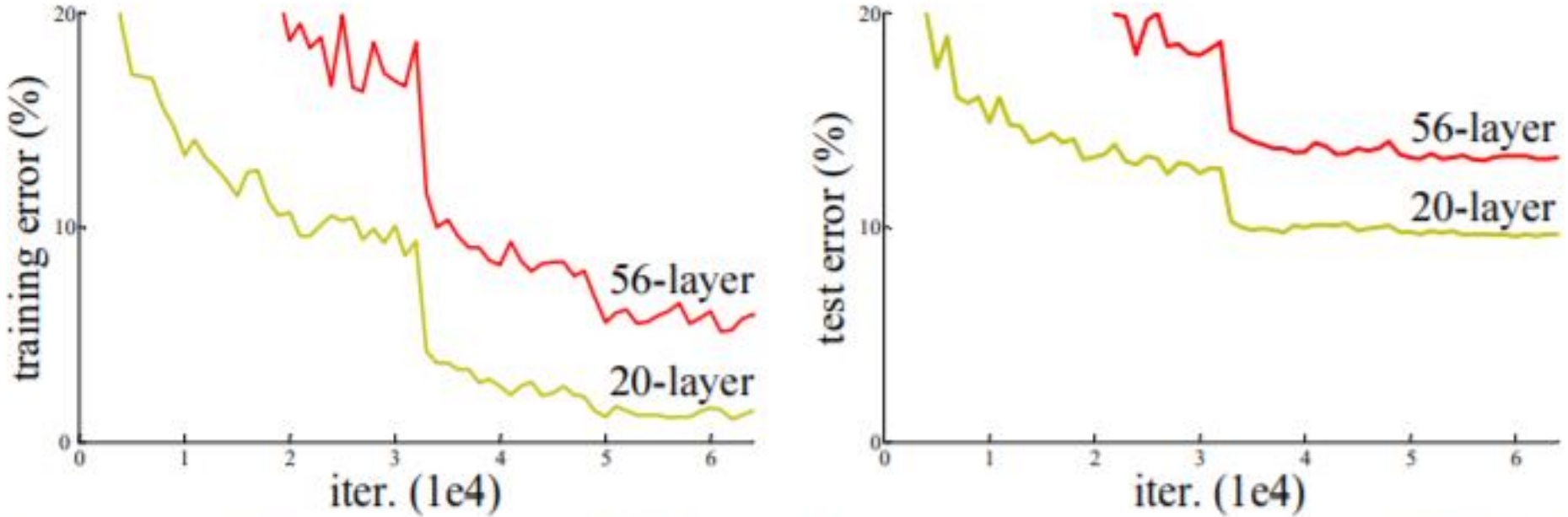
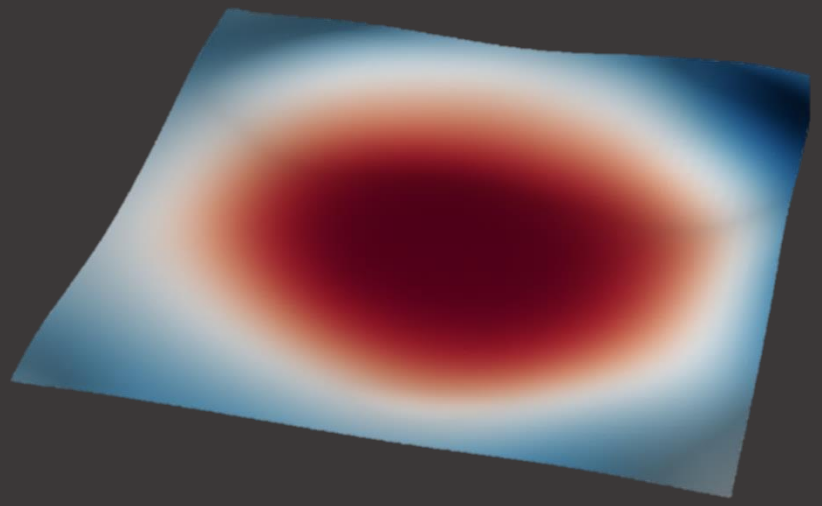


Figure 1. Training error (left) and test error (right) on CIFAR-10 with 20-layer and 56-layer "plain" networks. The deeper network has higher training error, and thus test error. Similar phenomena

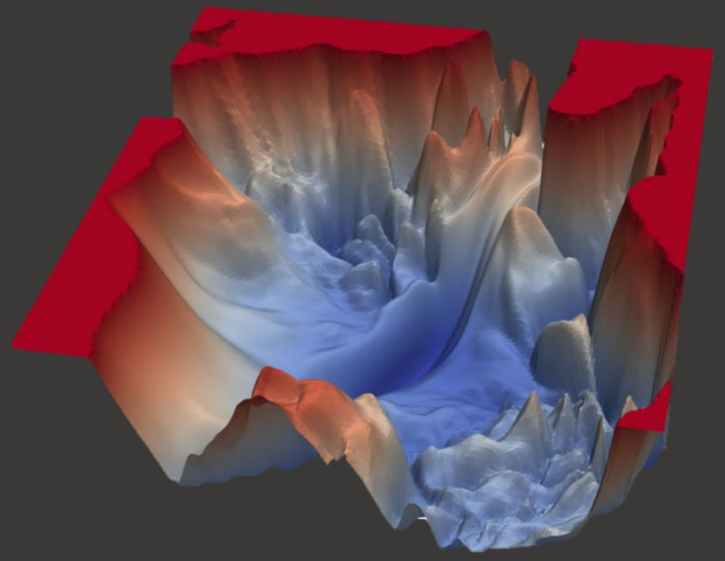
Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. "Deep Residual Learning for Image Recognition"

Why does this happen?

- Increasing the number of layers increases the non-linearity
- Therefore, the problem space becomes increasingly non-convex and thus more difficult to optimize



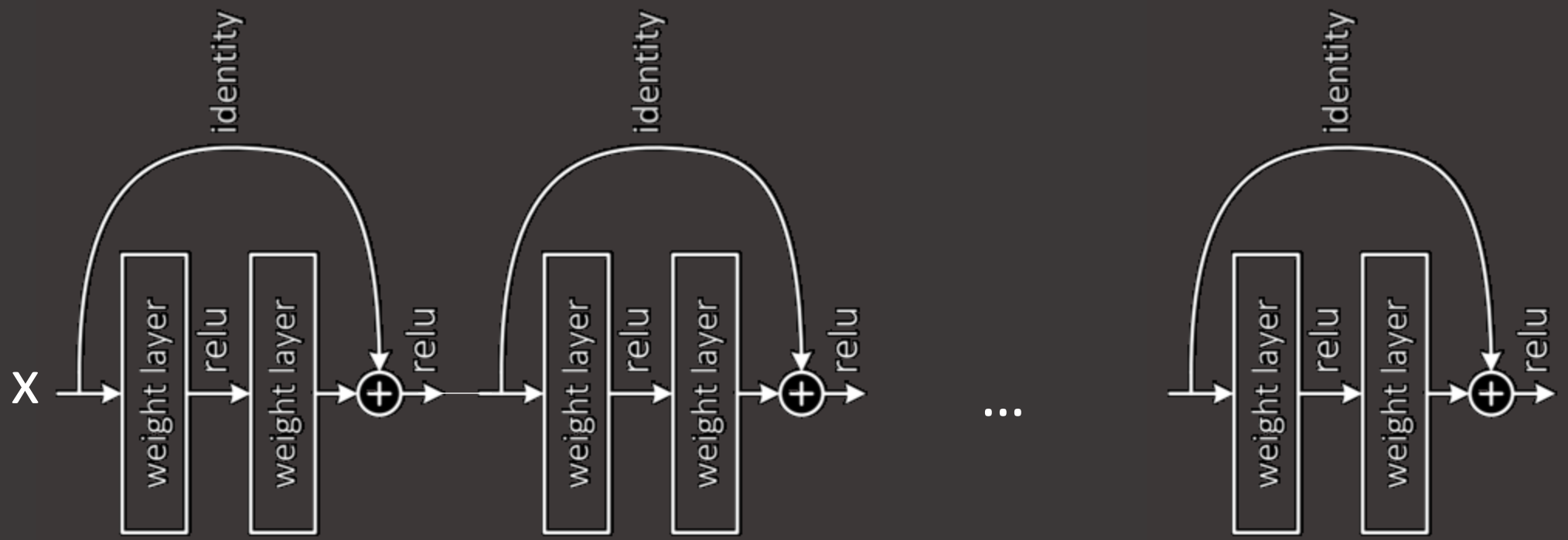
VGG-16 loss function



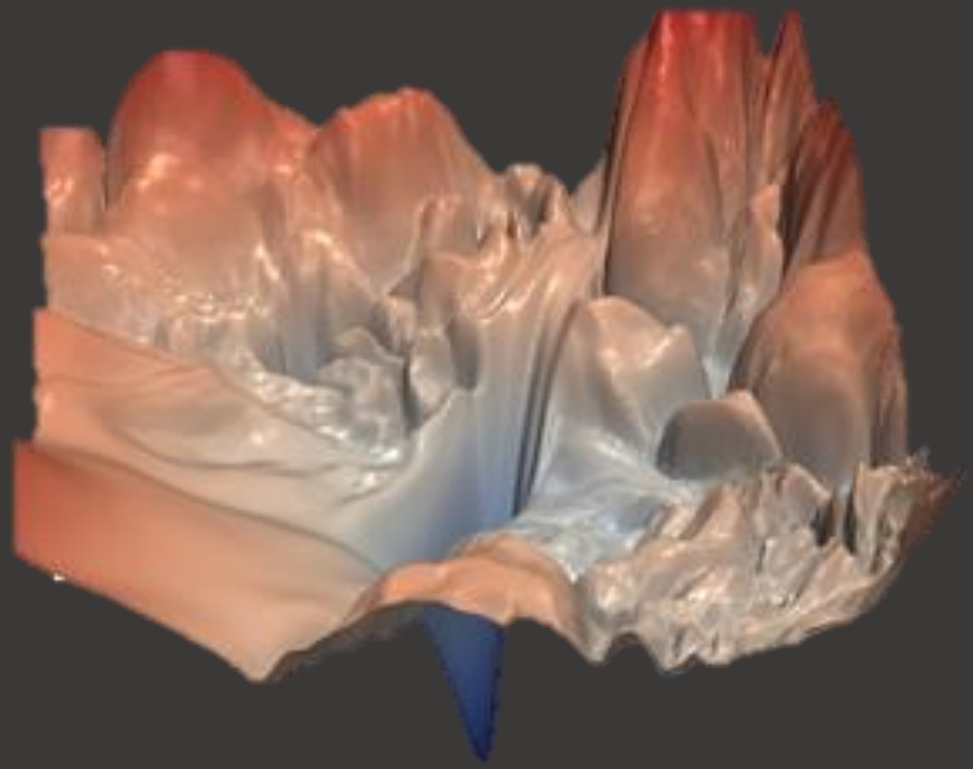
VGG-56 loss function

(Drawn using Loss Landscape Visualizer)

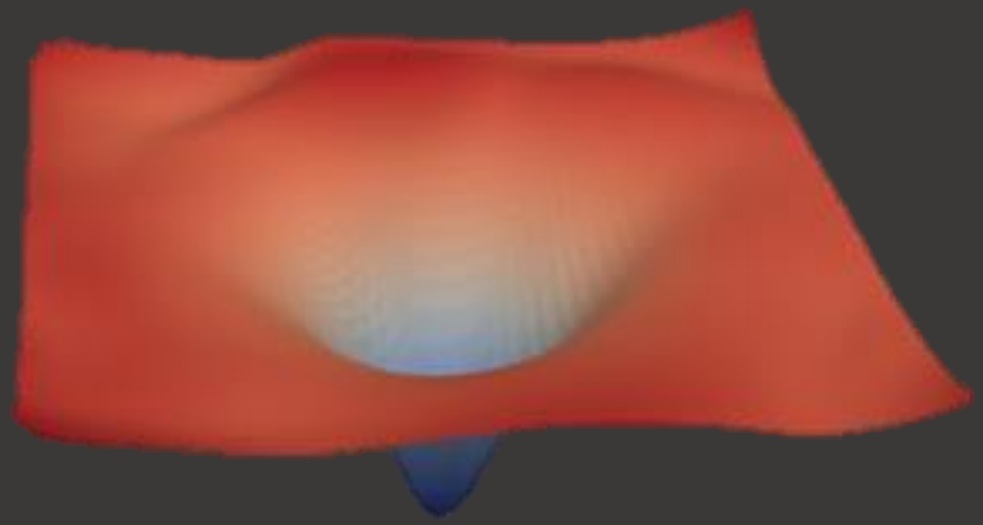
ResNet


x
identity
weight layer
relu
weight layer
relu
identity
weight layer
relu
weight layer
relu
...
identity
weight layer
relu
weight layer
relu

ResNet



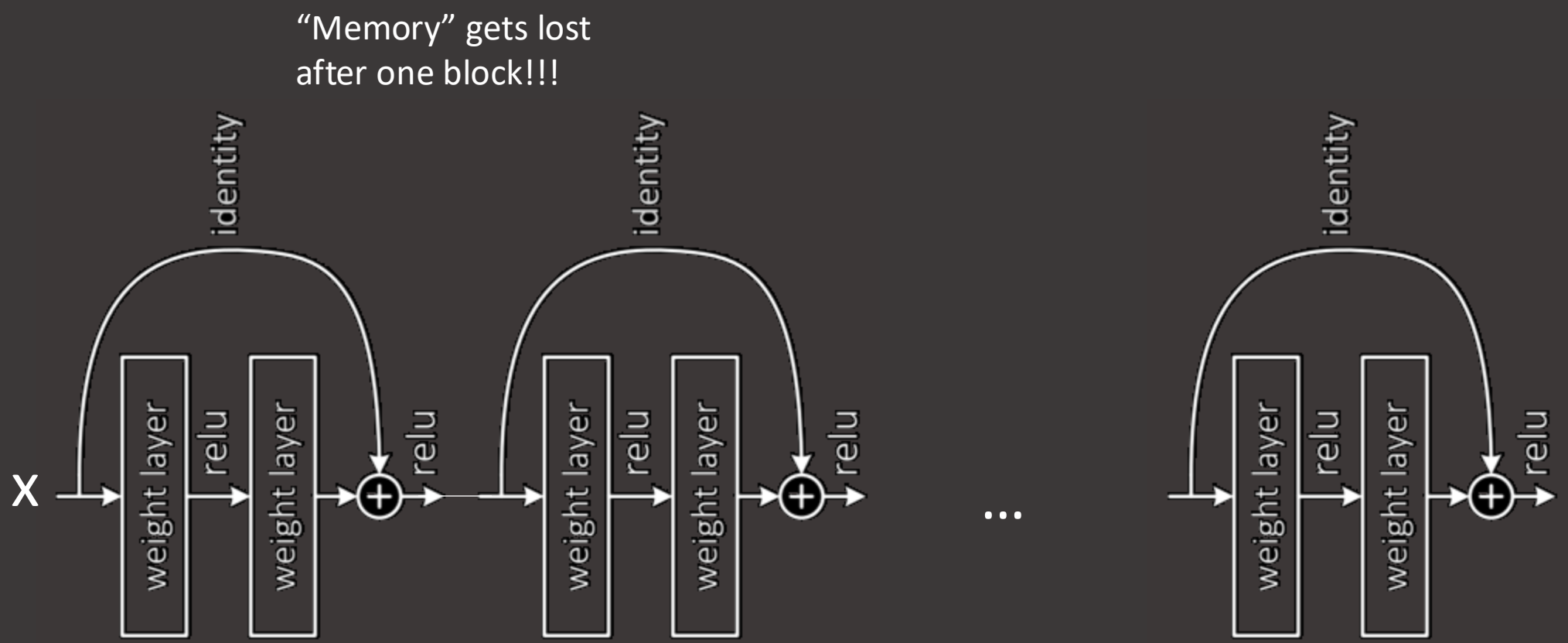
Without skip connections



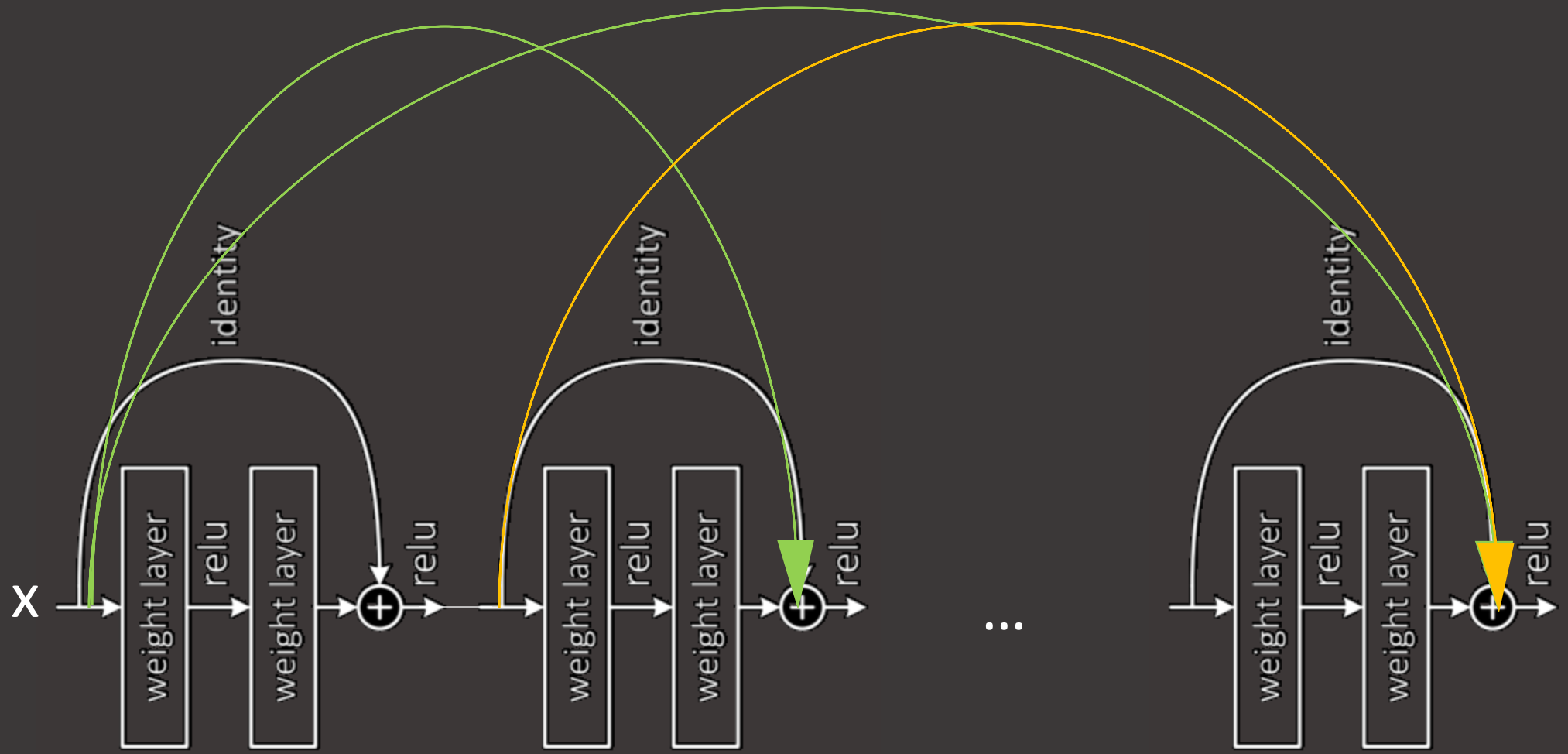
With skip connections

Hao Li, Zheng Xu, Gavin Taylor, Christoph Studer, Tom Goldstein. "Visualizing the Loss Landscape of Neural Nets"

ResNet

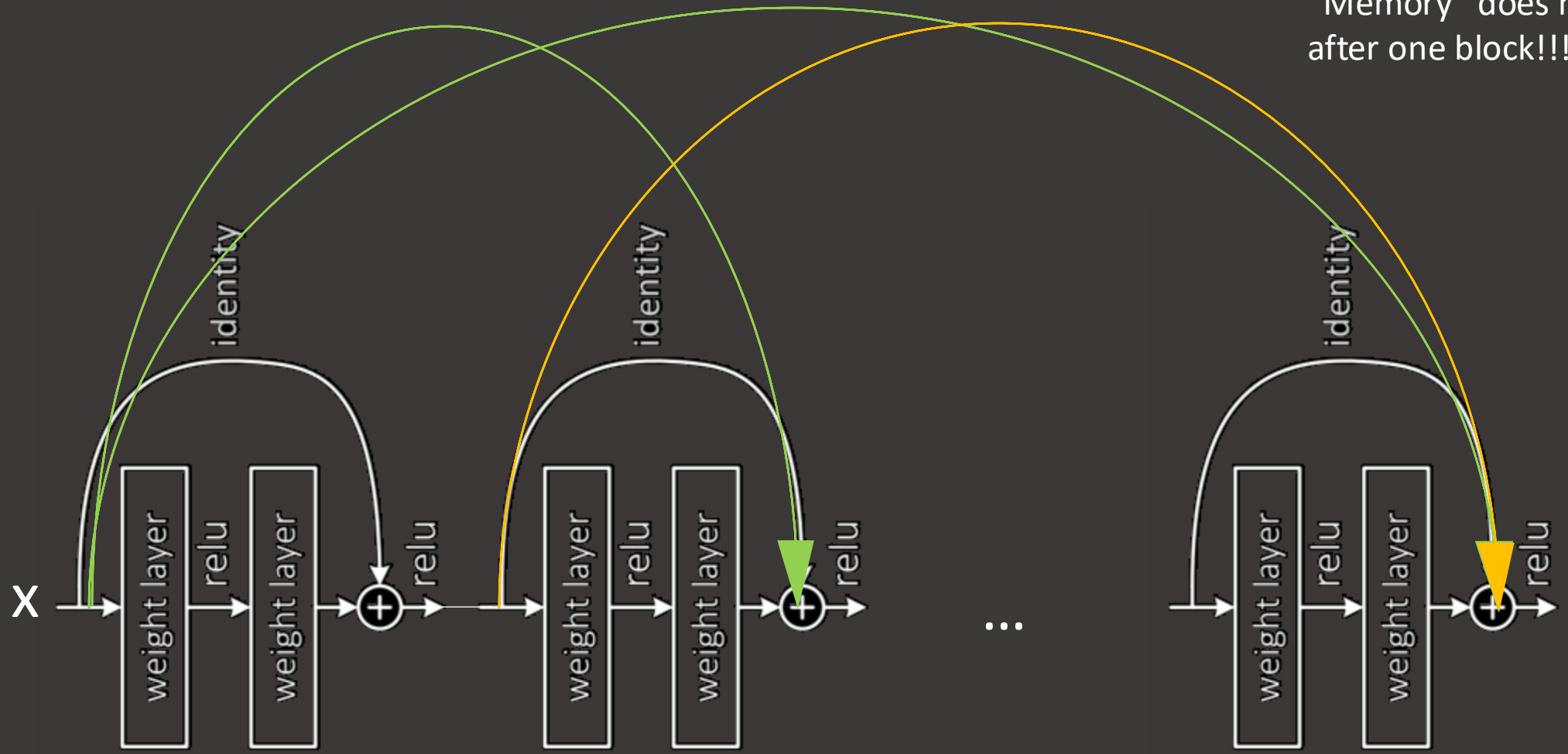

“Memory” gets lost
after one block!!!
identity
identity
identity
x
weight layer
relu
weight layer
relu
weight layer
relu
weight layer
relu
...
weight layer
relu
weight layer
relu

DenseNet



Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. InProceedings of the IEEE conference on computer vision and pattern recognition 2017 (pp. 4700-4708).

DenseNet

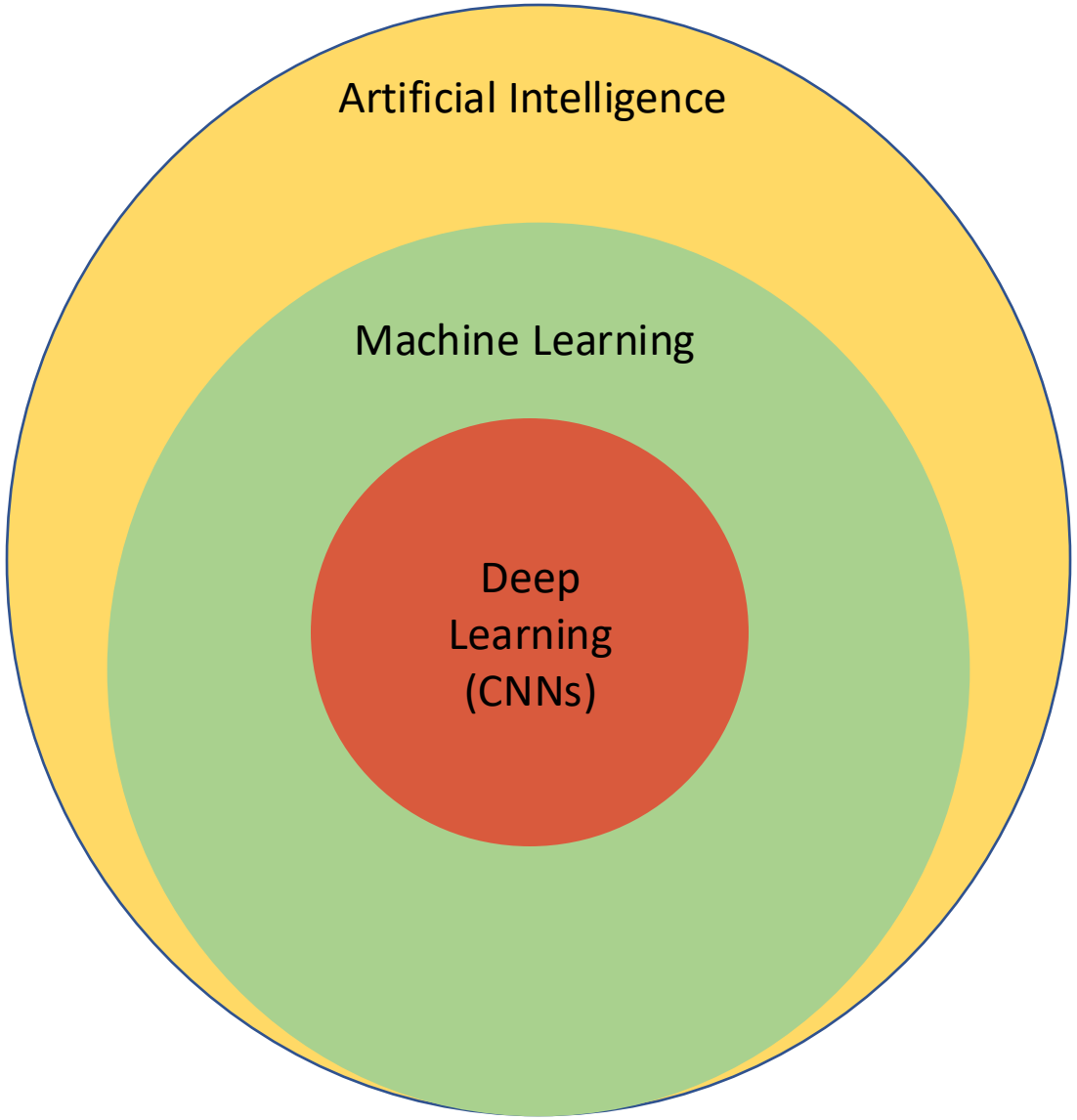


"Memory" does not get lost after one block!!!

Deep Learning

- ResNet and DenseNet could be hundreds or thousands of layers
- But that’s just the beginning of deep learning

What we have learned so far



Course recommendation: Intro to Reinforcement Learning

Online

Python-based

Dedicated 100-min Q&A Session (Office Hour) per week. Thus, not quite different from in-person course.

Quiz, 3-4 assignments, One group project (2~3 people)

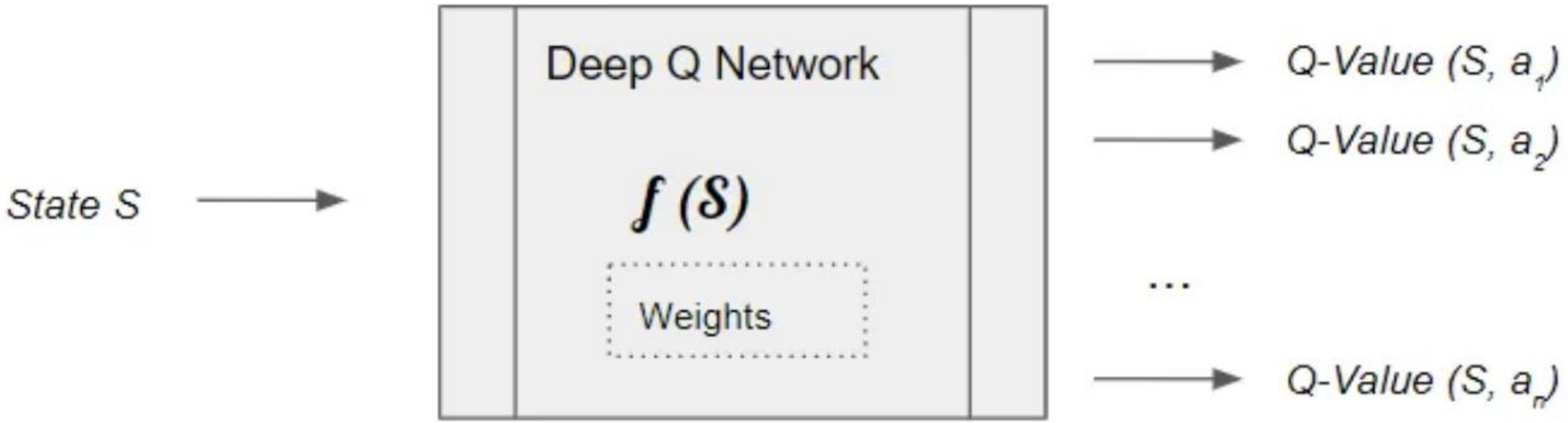
Similar logistics like my other courses.

Just need ML pre-requisite, we just remove the DL pre-requisite (need some time to take effect)

Deep RL: RoboCup

https://youtu.be/xkoXeF9oVH4

Deep RL: RoboCup



https://towardsdatascience.com/reinforcement-learning-explained-visually-part-5-deep-q-networks-step-by-step-5a5317197f4b