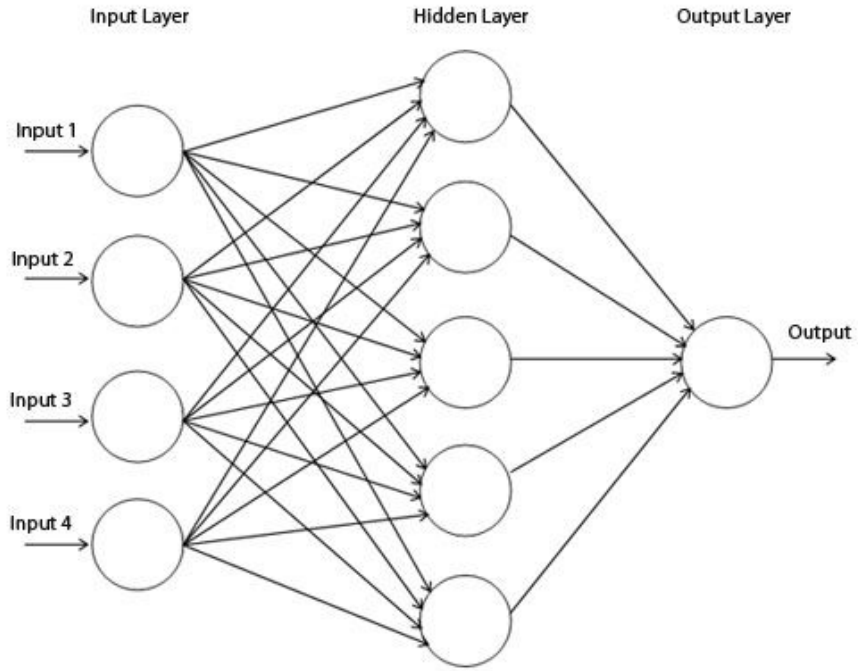
CS21si: AI for Social Good

Lecture 4: Convolutional Neural Networks

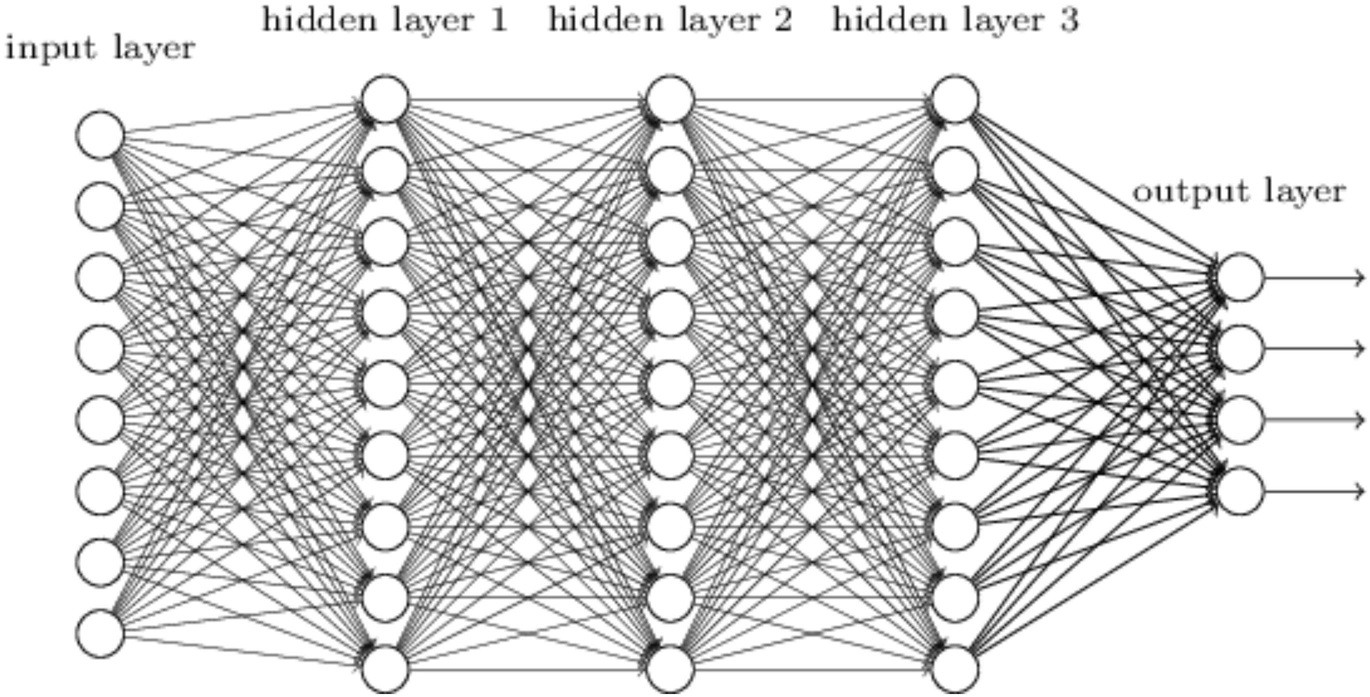
Plan for Today

- Review of deep neural networks
- Convolutional neural networks
- Implementing CNNs
- Adversarial attacks

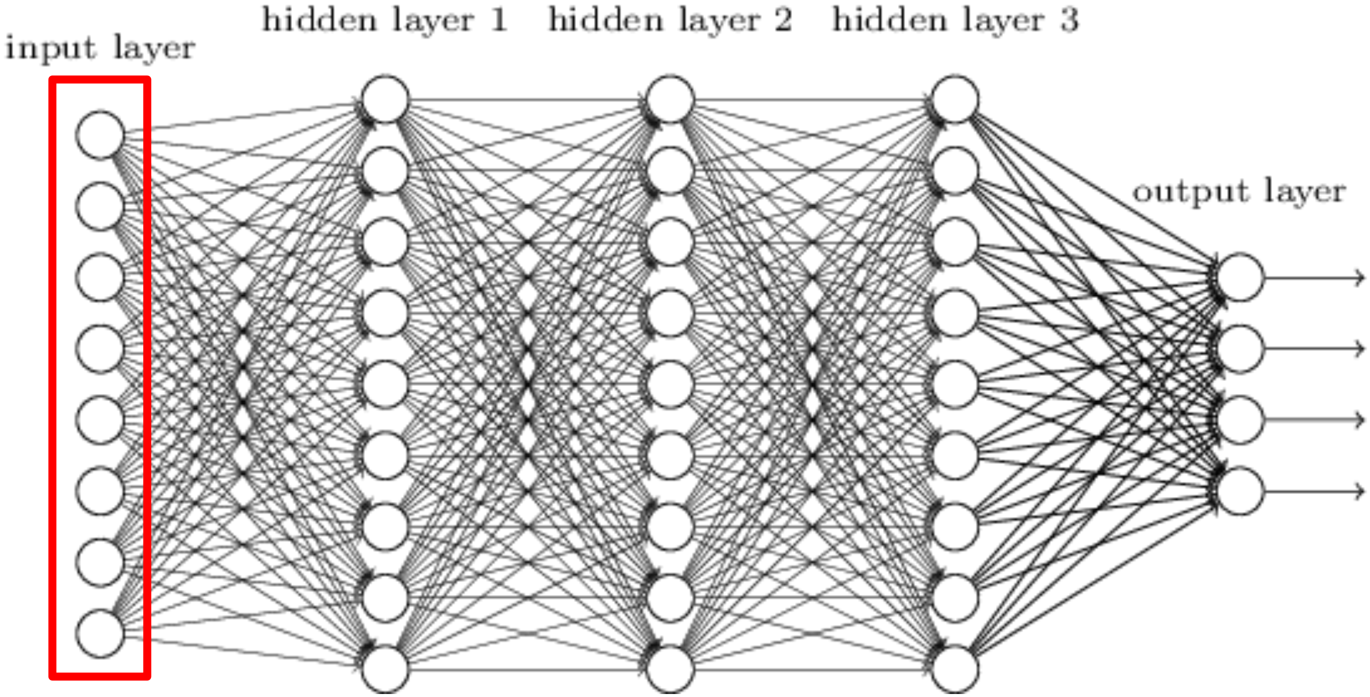
Deep Neural Networks



Deep Neural Networks



Deep Neural Networks

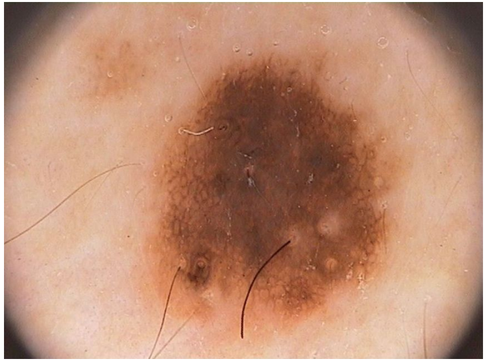


What sort of input can I give an neural net?

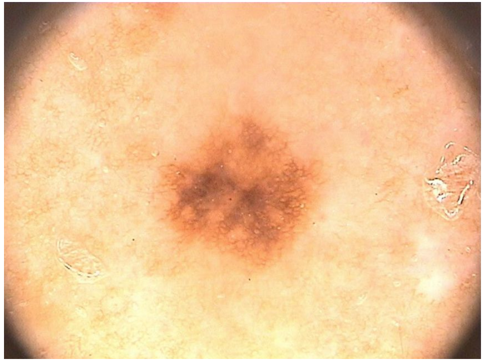
How do I handle image data?

ADDI - Automated Diagnosis for Dermoscopic Images

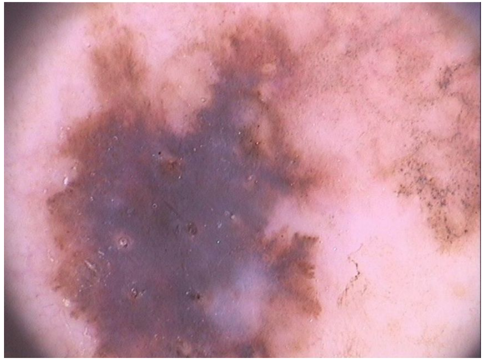
ADDI - Automated Diagnosis for Dermoscopic Images



Common Nevus

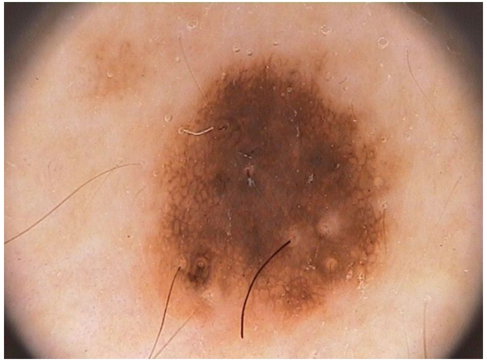


Atypical Nevus

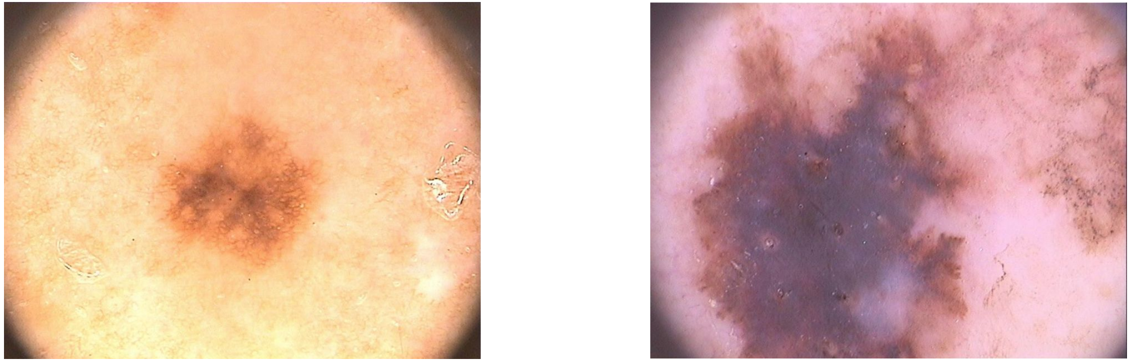


Melanoma

ADDI - Automated Diagnosis for Dermoscopic Images



Normal



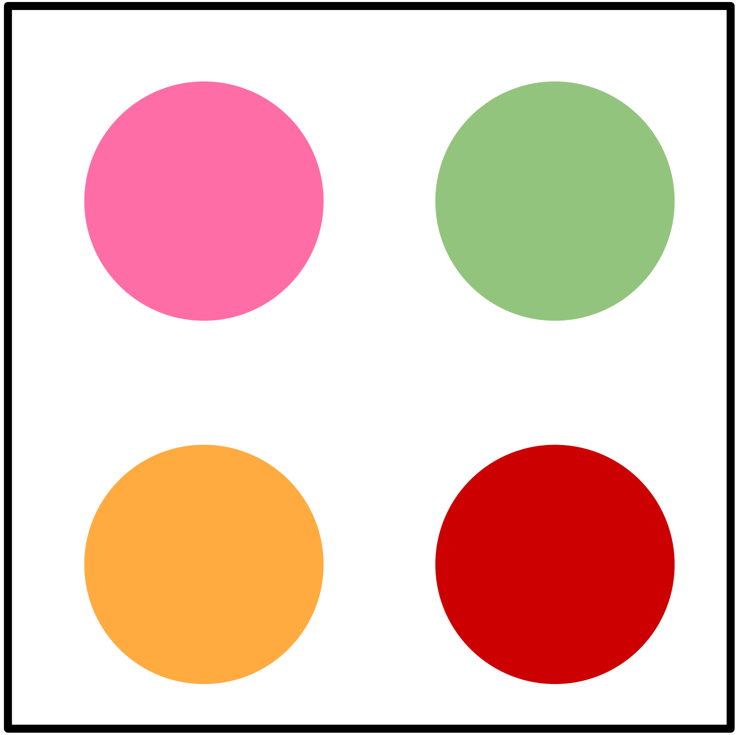
Abnormal

Jupyter Exercises 1: Visualize the Data

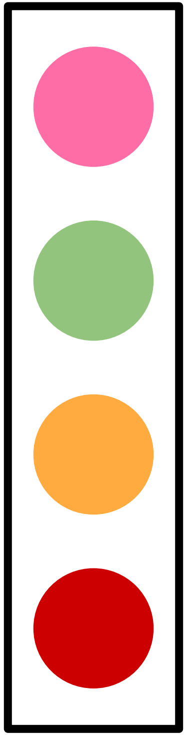
How do I handle image data?

One Idea...

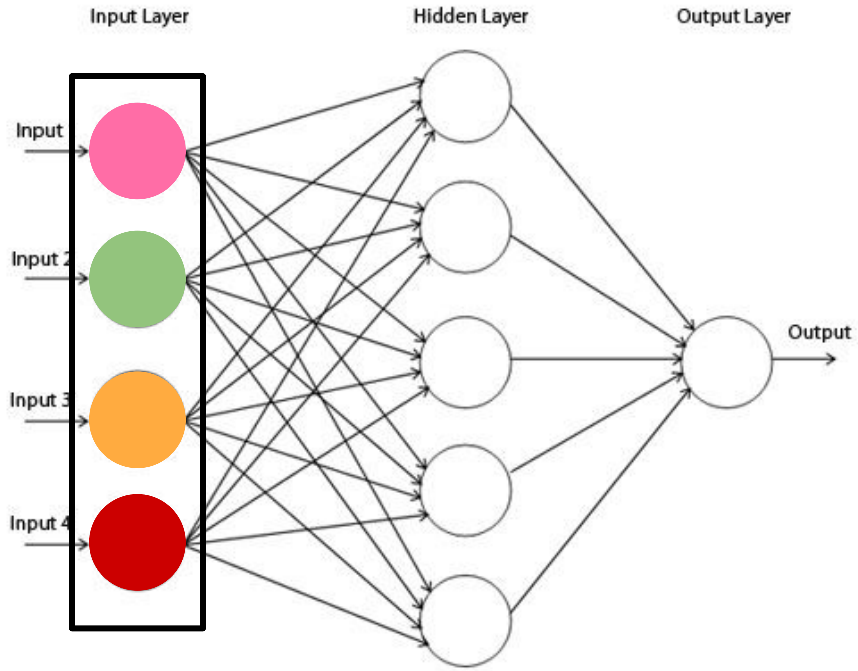
One Idea...



One Idea…

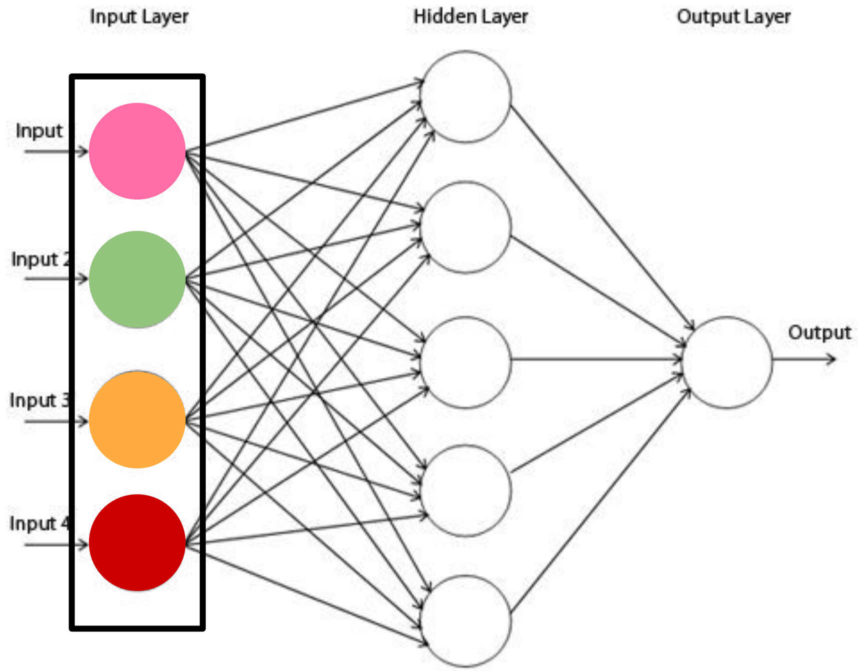


One Idea...


Input Layer
Hidden Layer
Output Layer
Input
Input 2
Input 3
Input 4
Output

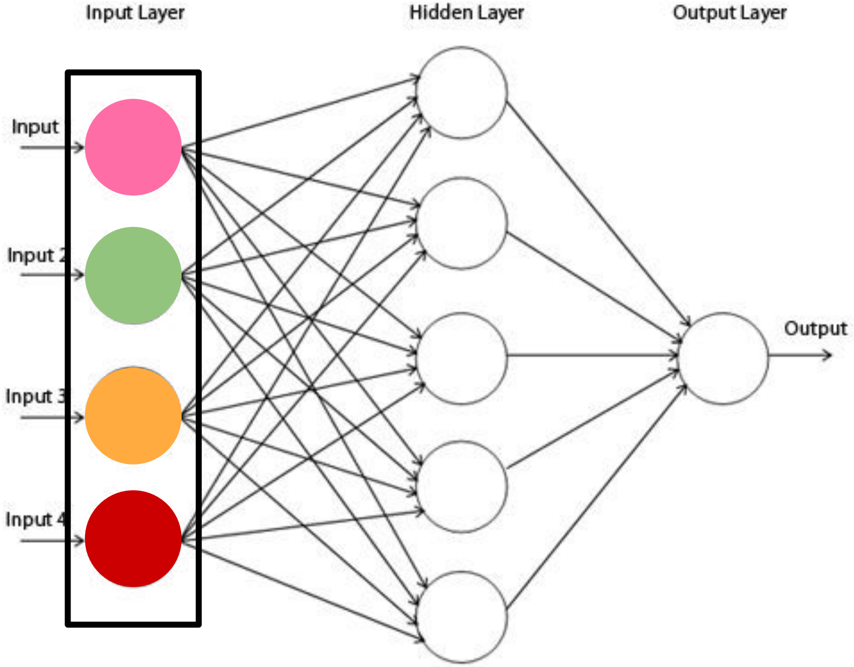
Questions?

Issue with Our Idea



Issue with Our Idea

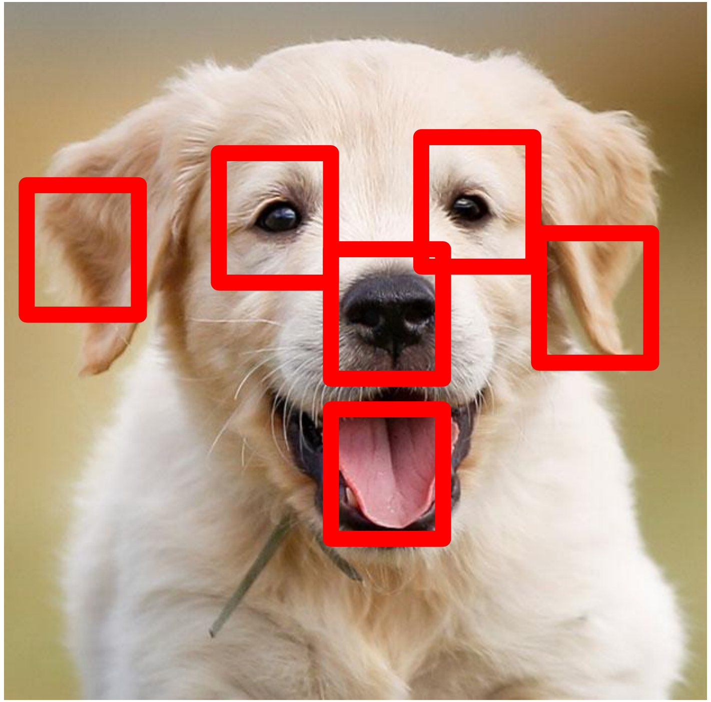
Assumes independence of features (pixels)!



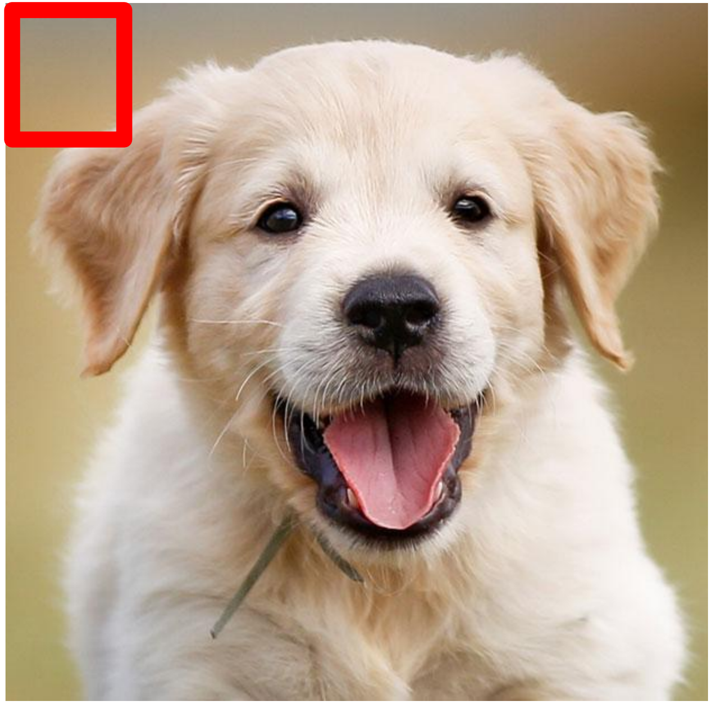
Key Insight: we want to maintain spatial dependence

Windowing

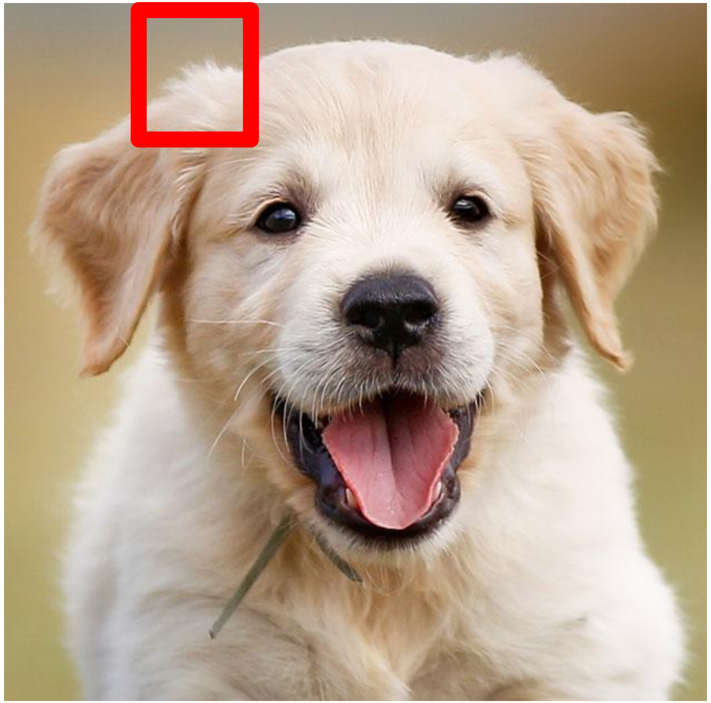
Windowing



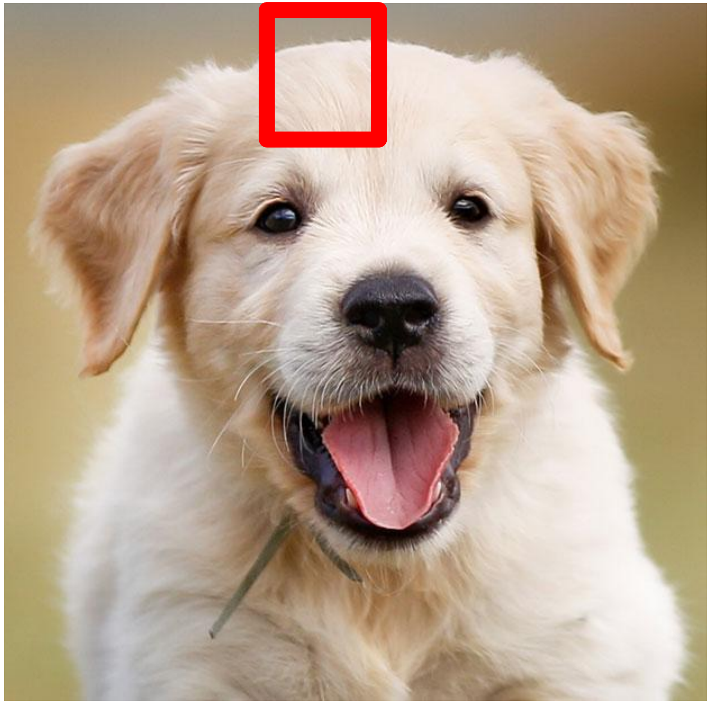
Windowing



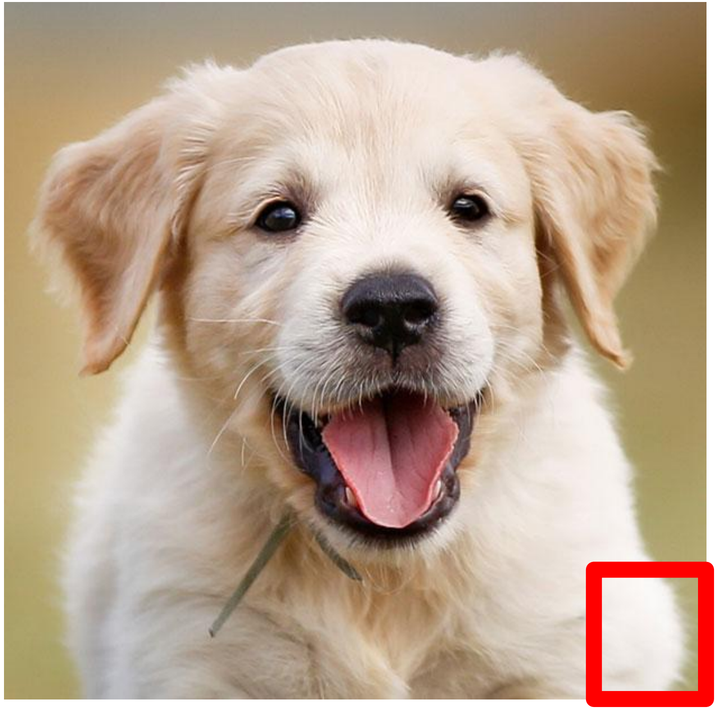
Windowing



Windowing

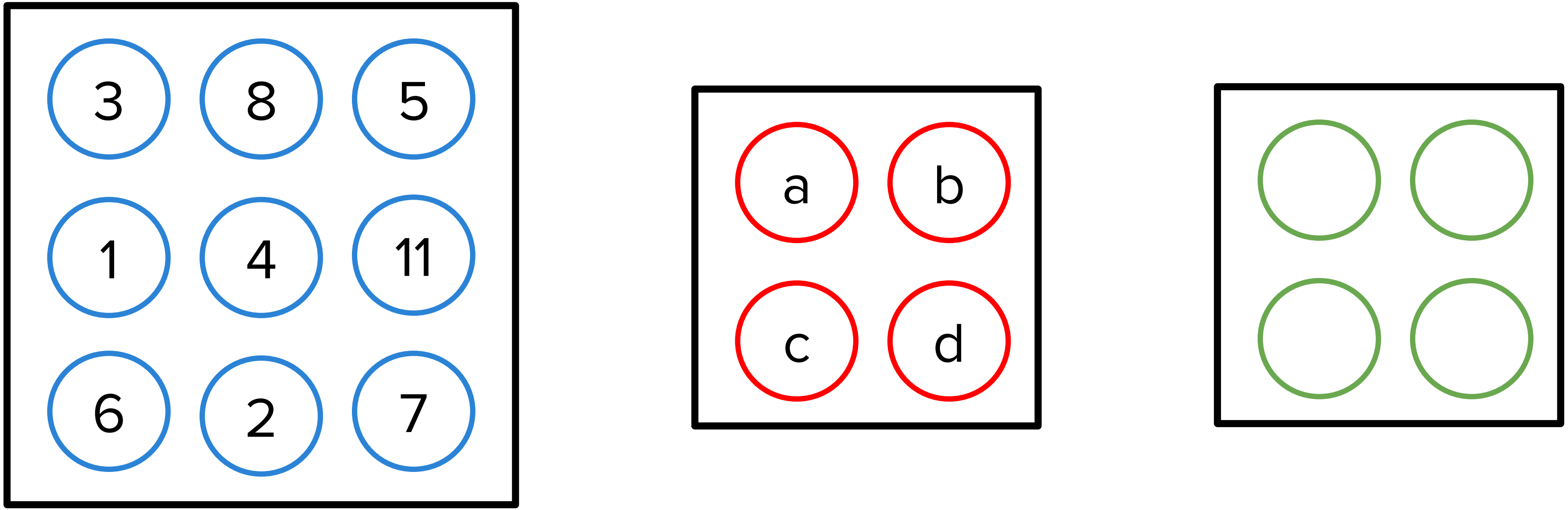


Windowing



Questions?

Convolutional Layer



3	8	5
1	4	11
6	2	7

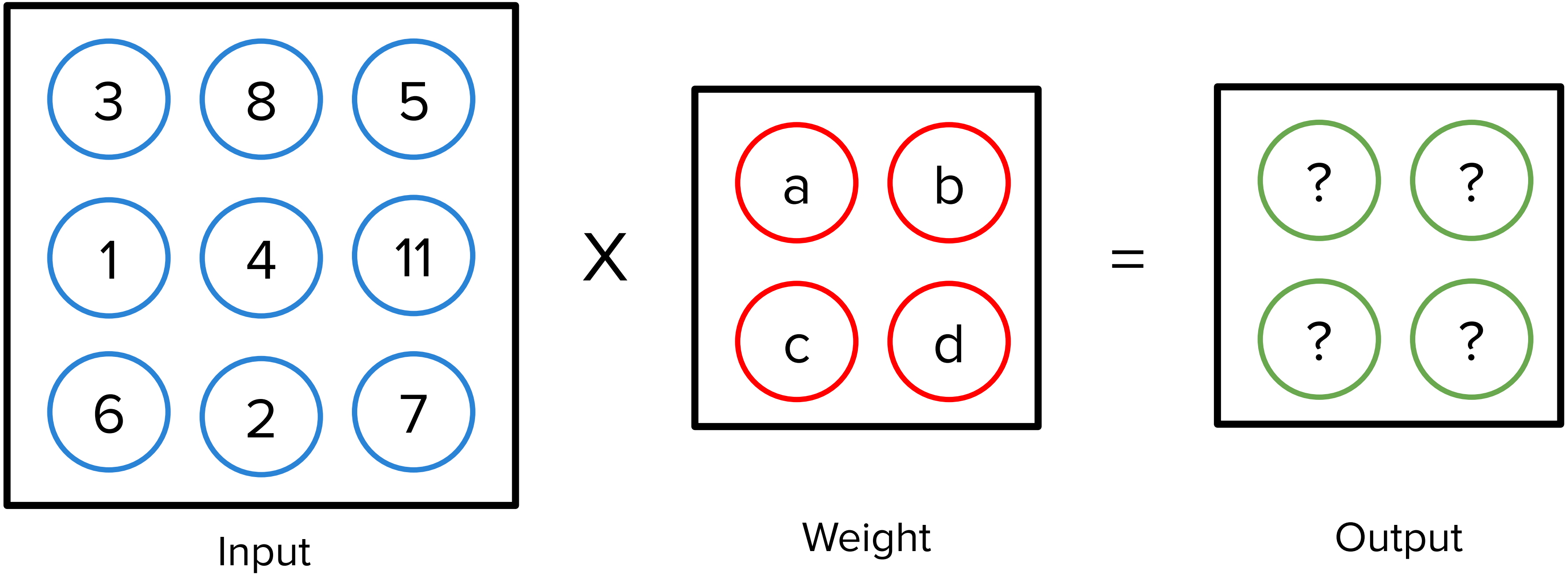
Input

a	b
c	d

Weight

Output

Convolutional Layer



3	8	5
1	4	11
6	2	7

Input

X

a	b
c	d

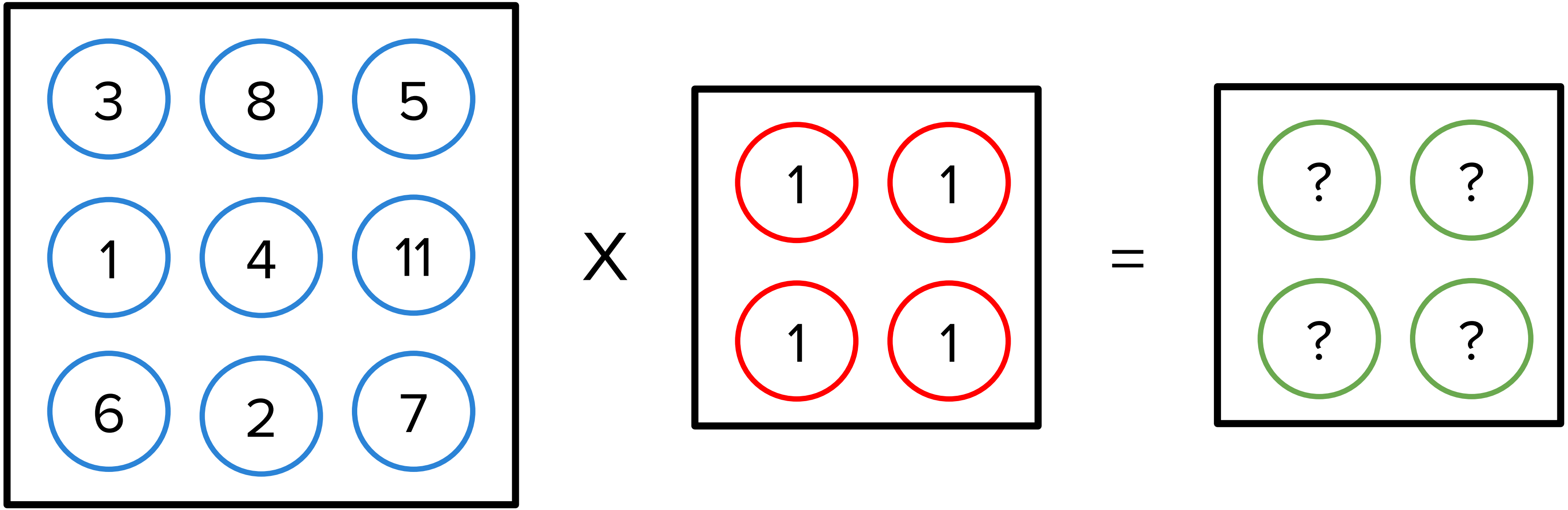
Weight

=

?	?
?	?

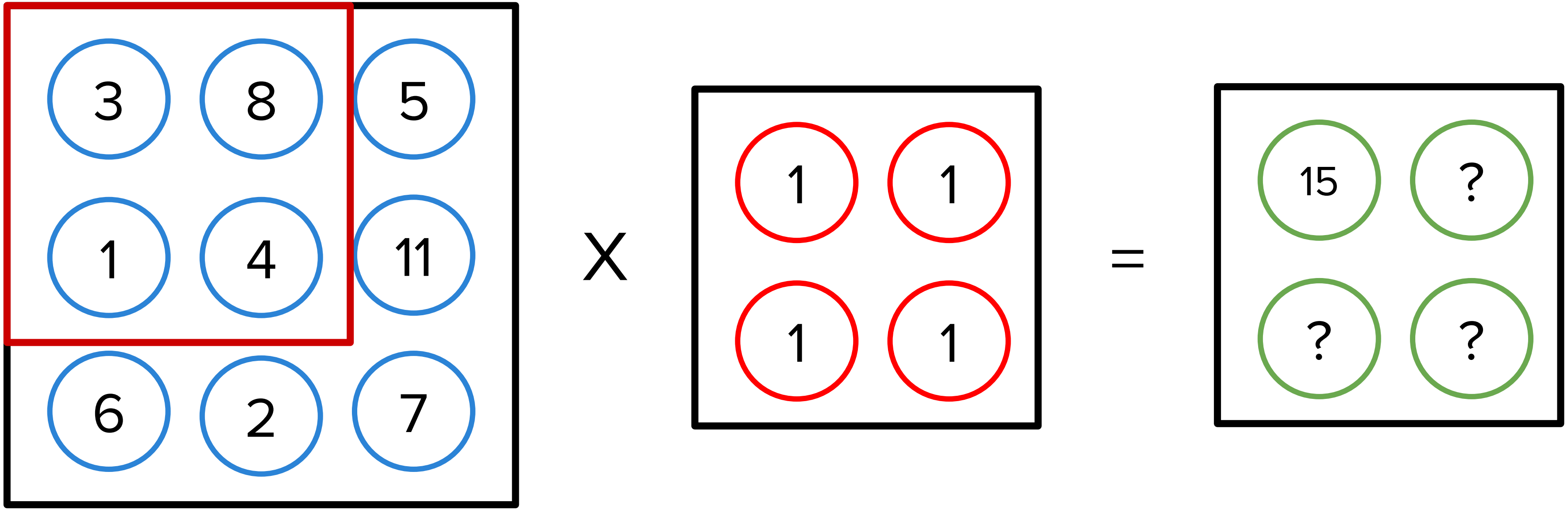
Output

Convolutional Layer

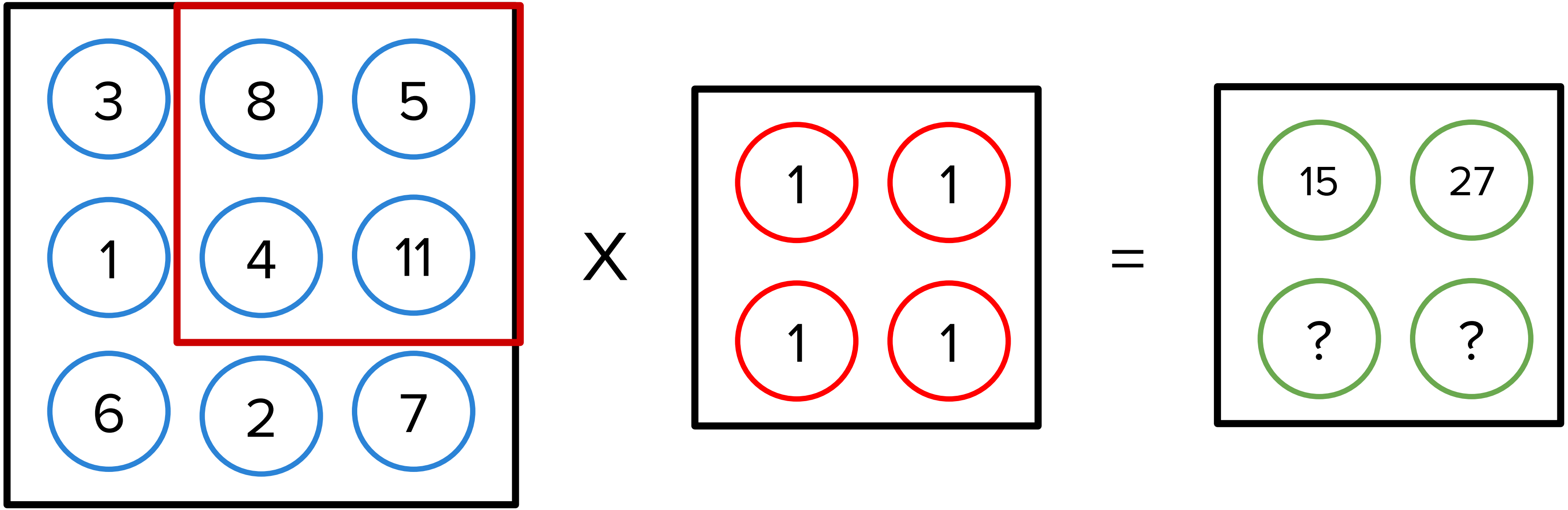


$$\begin{bmatrix}3 & 8 & 5\\ 1 & 4 & 11\\ 6 & 2 & 7\end{bmatrix} \text{ X } \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}? & ?\\ ? & ?\end{bmatrix}$$

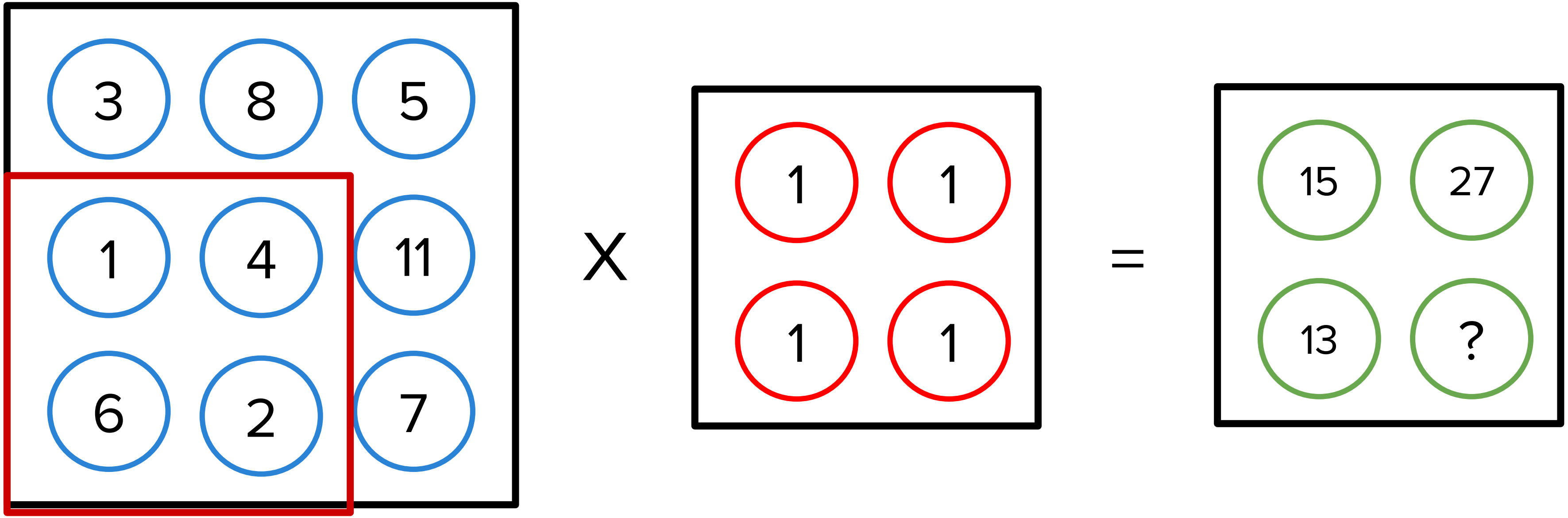
Convolutional Layer



Convolutional Layer

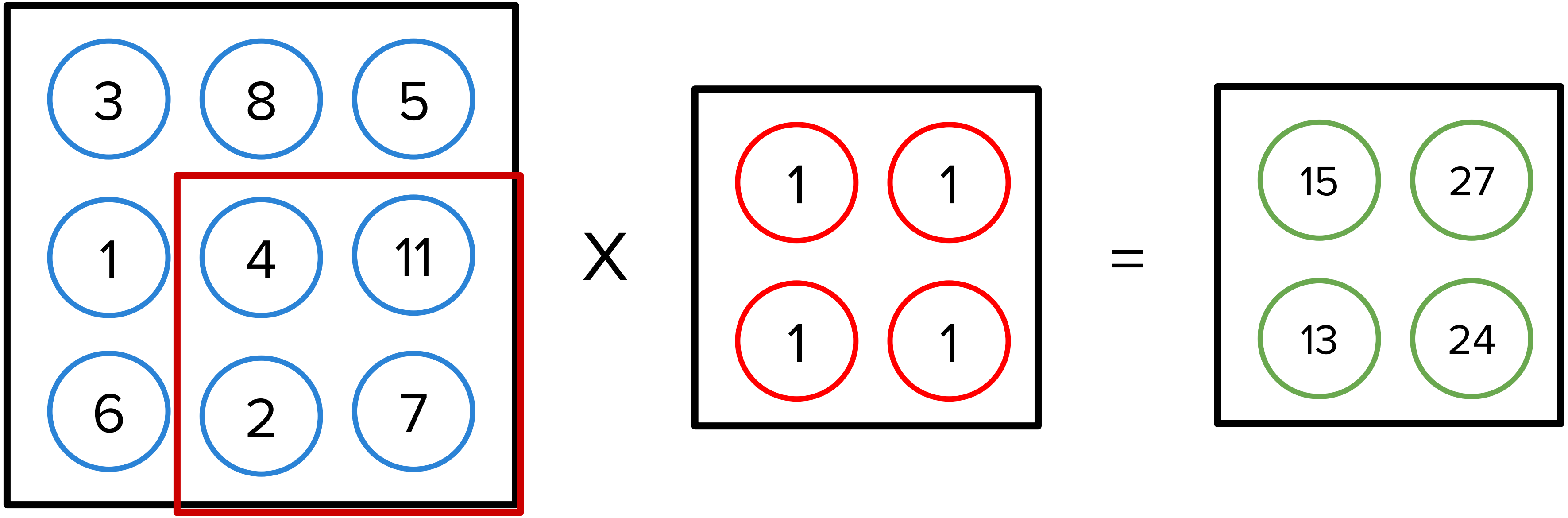


Convolutional Layer



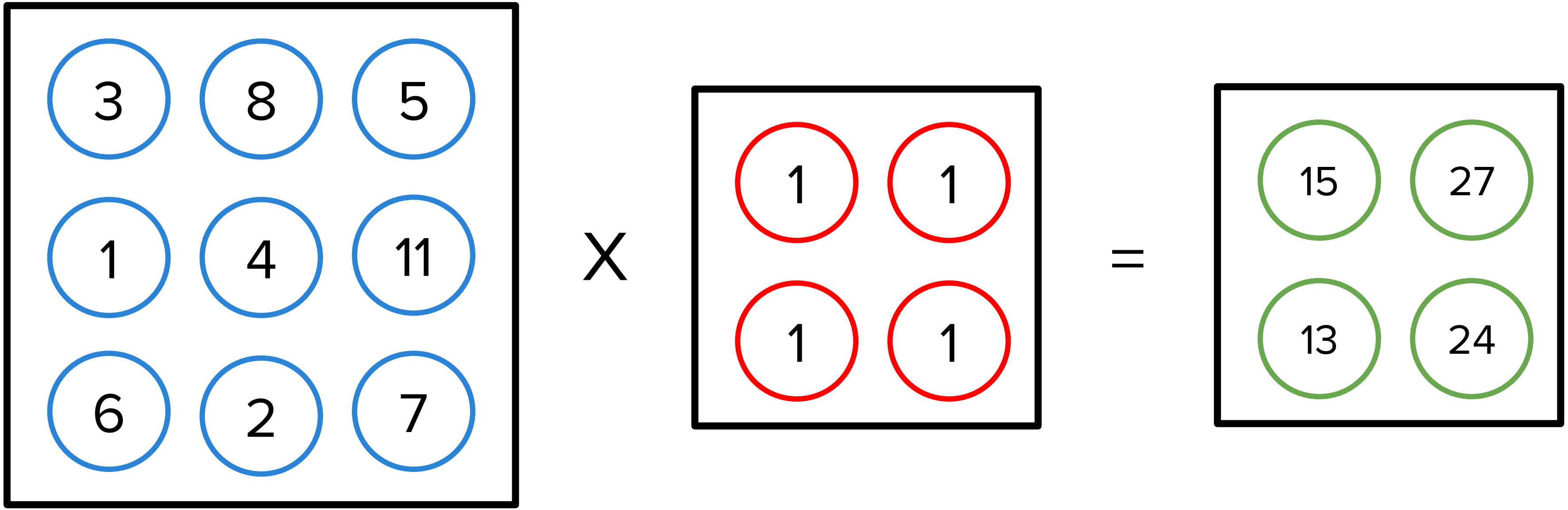
$$\begin{matrix}3 & 8 & 5\\ 1 & 4 & 11\\ 6 & 2 & 7\end{matrix} \quad X \quad \begin{matrix}1 & 1\\ 1 & 1\end{matrix} \quad = \quad \begin{matrix}15 & 27\\ 13 & ?\end{matrix}$$

Convolutional Layer



$$\begin{bmatrix}3 & 8 & 5\\ 1 & 4 & 11\\ 6 & 2 & 7\end{bmatrix} \text{X} \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}15 & 27\\ 13 & 24\end{bmatrix}$$

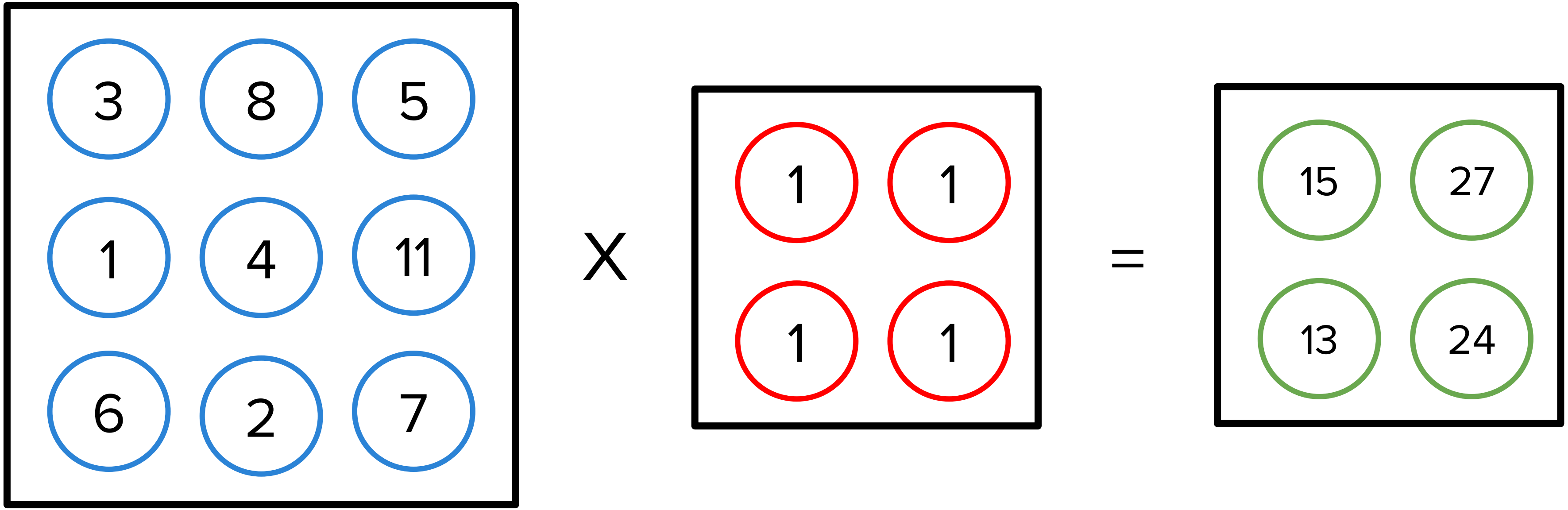
Convolutional Layer



$$\begin{bmatrix}3 & 8 & 5\\ 1 & 4 & 11\\ 6 & 2 & 7\end{bmatrix} \text{X} \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}15 & 27\\ 13 & 24\end{bmatrix}$$

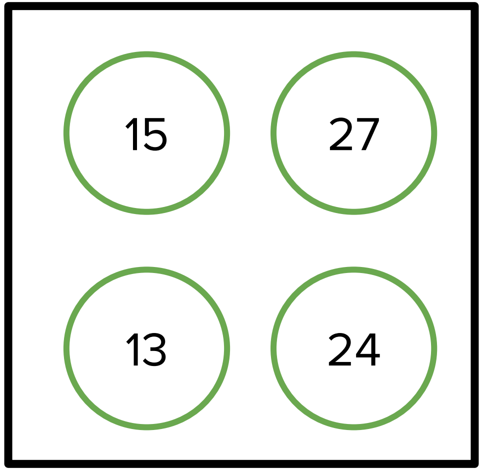
Questions?

Convolutional Output



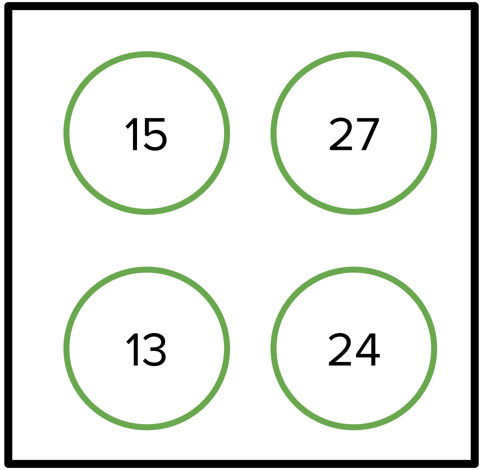
$$\begin{bmatrix}3 & 8 & 5\\ 1 & 4 & 11\\ 6 & 2 & 7\end{bmatrix} \text{X} \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}15 & 27\\ 13 & 24\end{bmatrix}$$

Convolutional Output



15	27
13	24

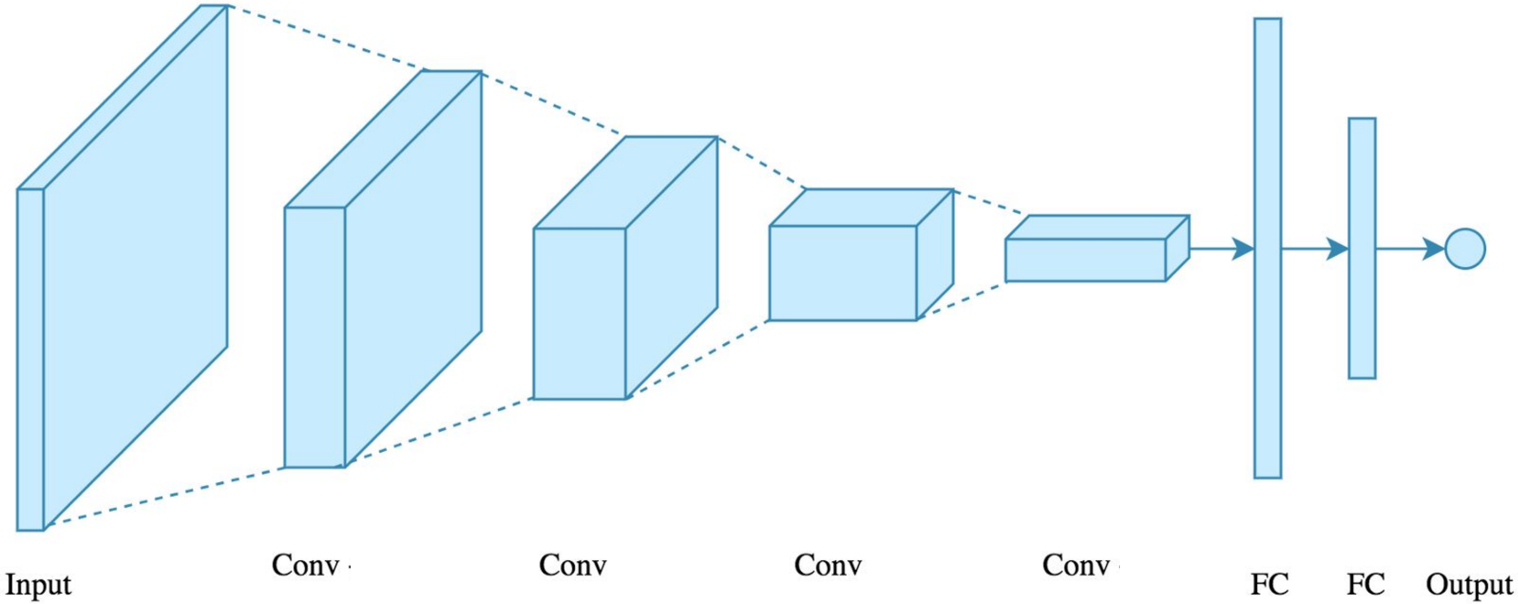
Convolutional Output



15	27
13	24

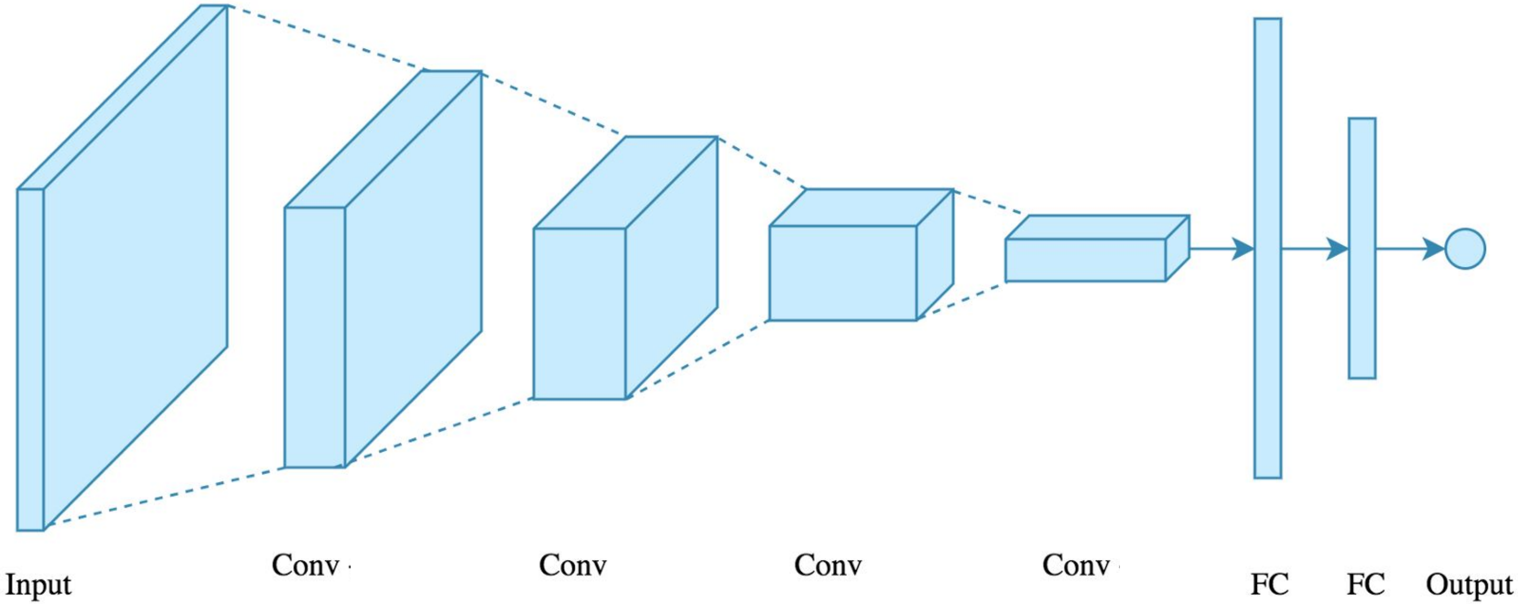
Can be fed as input to another convolutional layer!

Convolutional Neural Network



Questions?

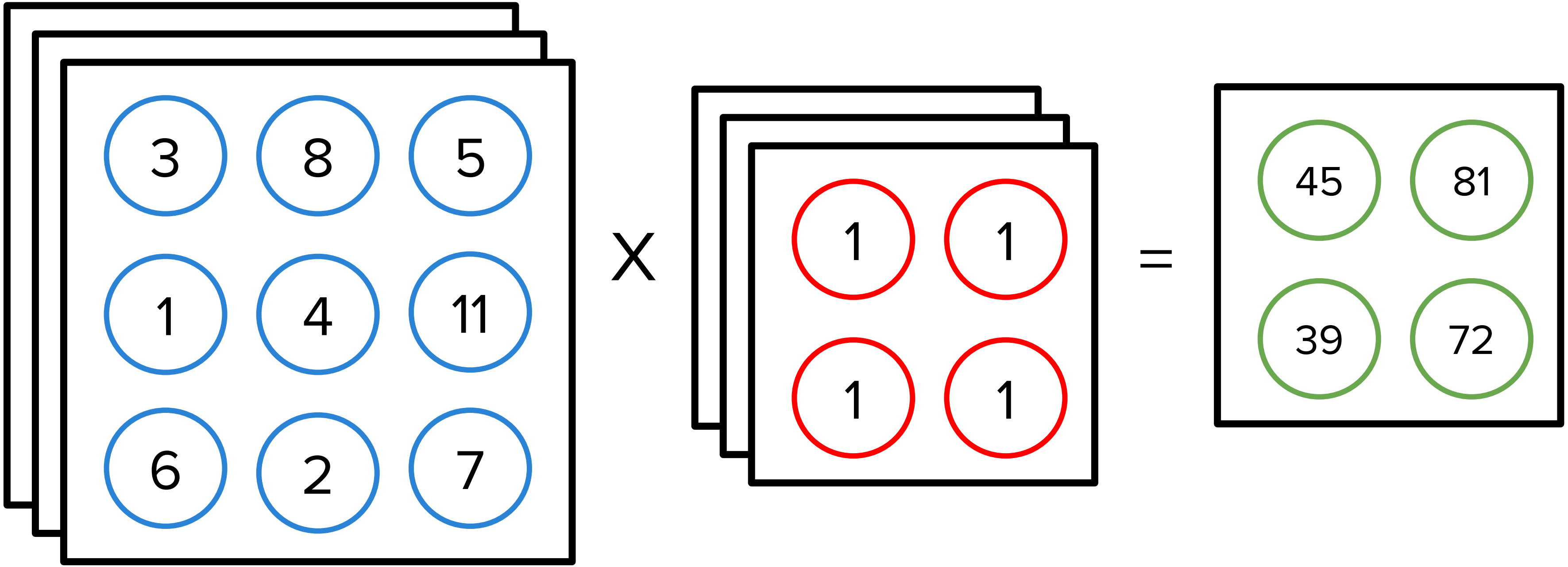
Convolutional Neural Network



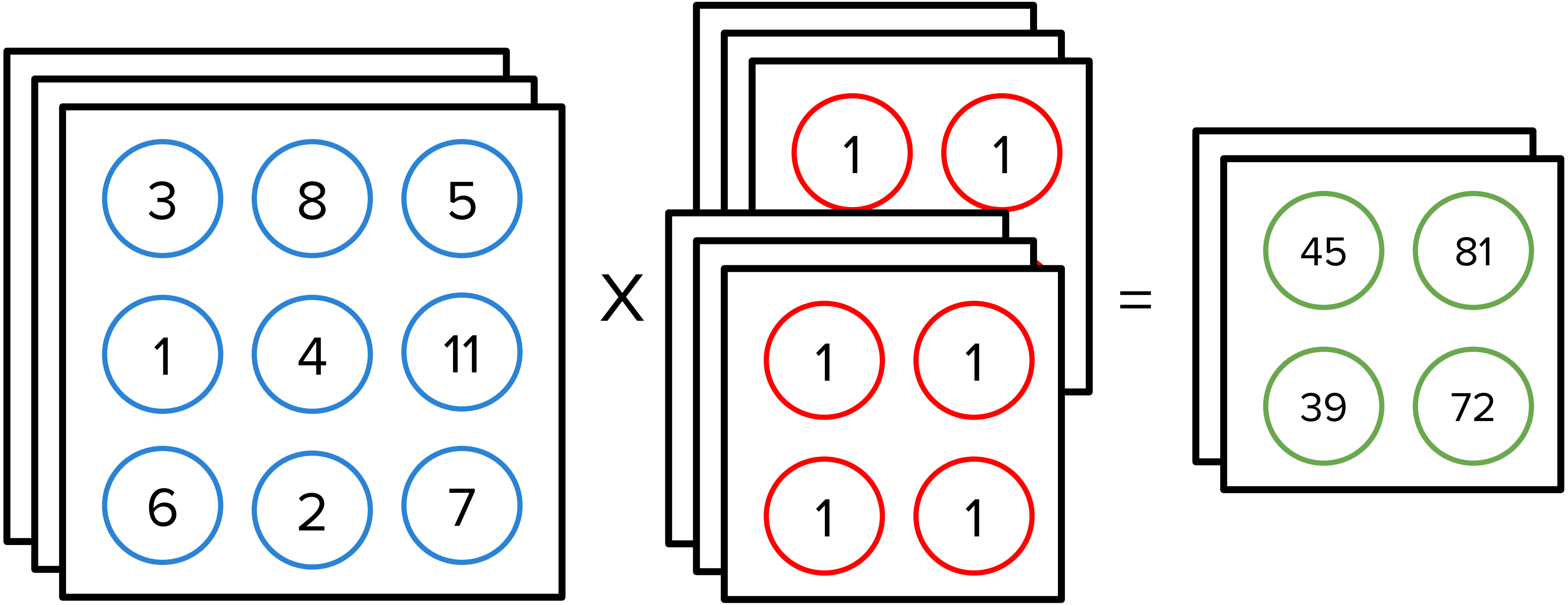
Input/Output Channel Size

$$\begin{bmatrix}3 & 8 & 5\\ 1 & 4 & 11\\ 6 & 2 & 7\end{bmatrix} \text{X} \begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} = \begin{bmatrix}15 & 27\\ 13 & 24\end{bmatrix}$$

Input/Output Channel Size



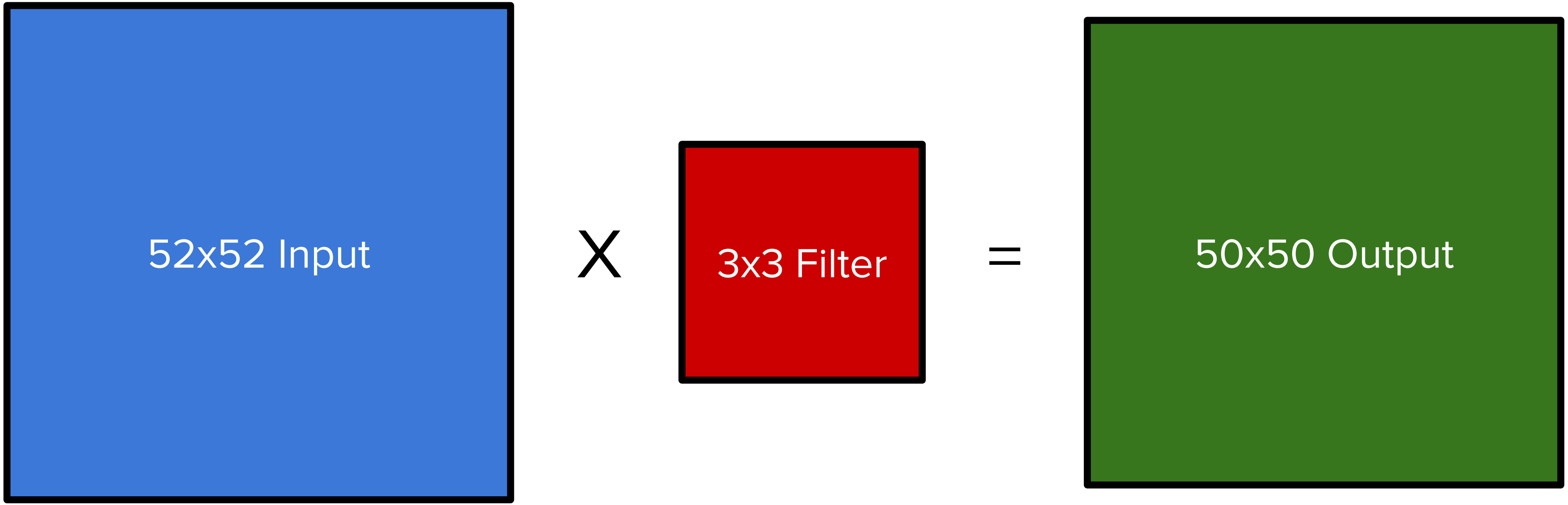
Input/Output Channel Size



Questions?

Jupyter Exercises 2: Simple CNN

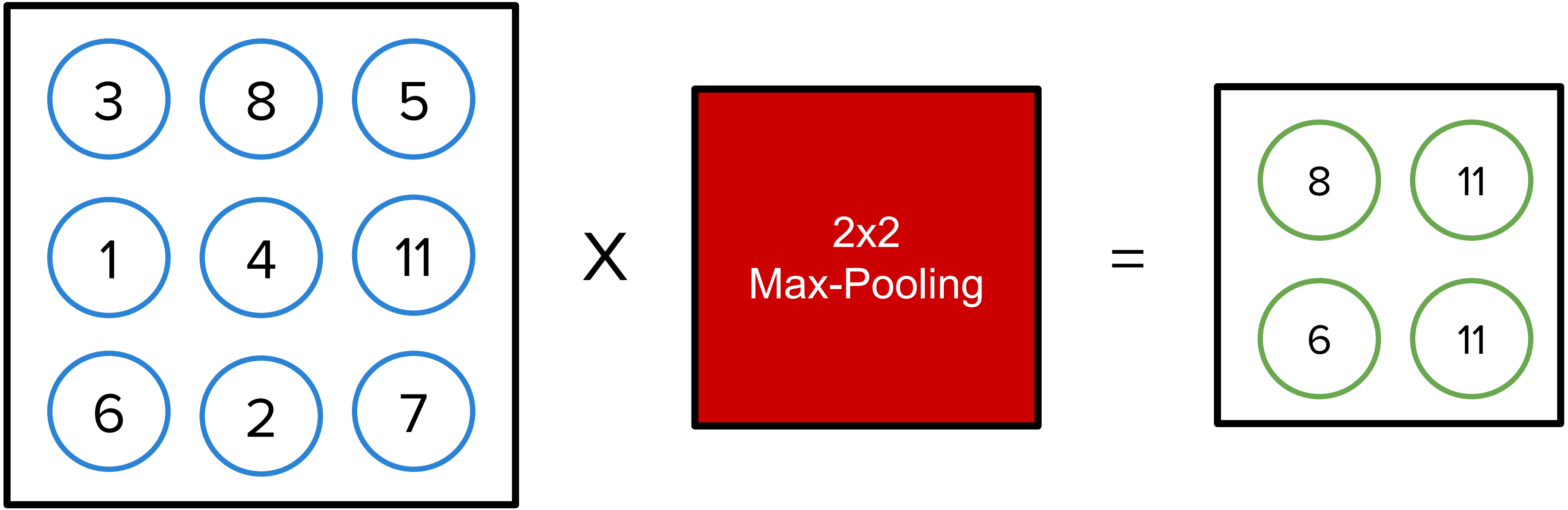
Problem: Output Size



52x52 Input X 3x3 Filter = 50x50 Output

We want to make the output smaller without losing info

Solution: Max-Pooling



3	8	5
1	4	11
6	2	7

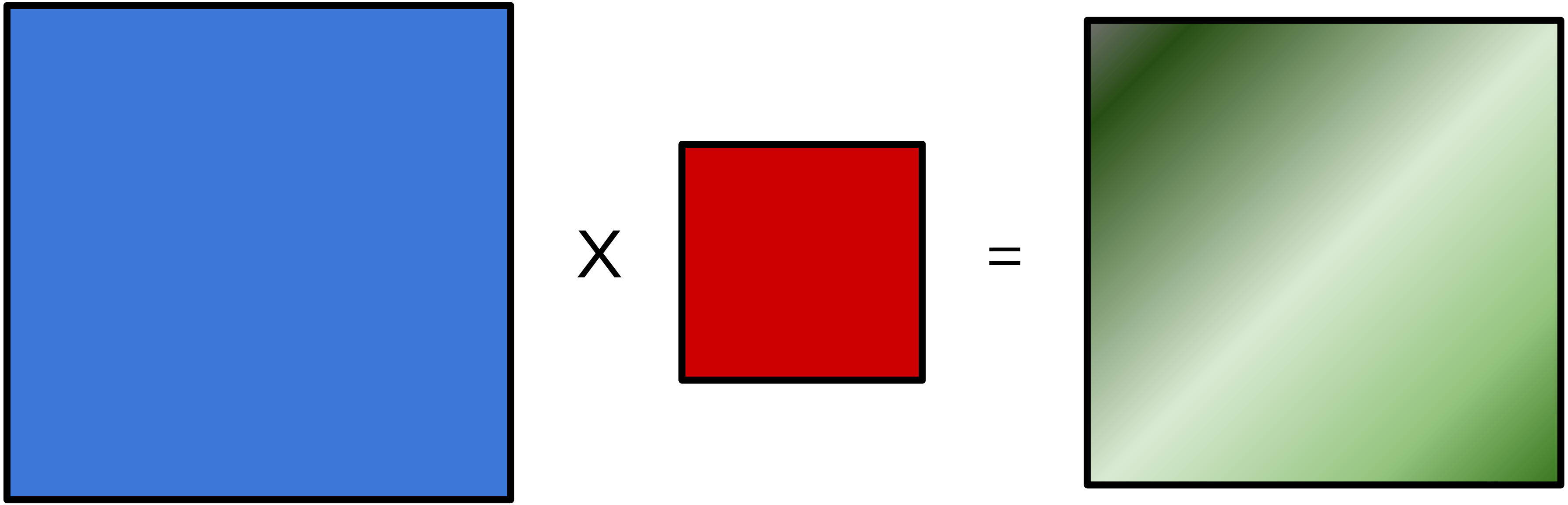
X

2x2
Max-Pooling

=

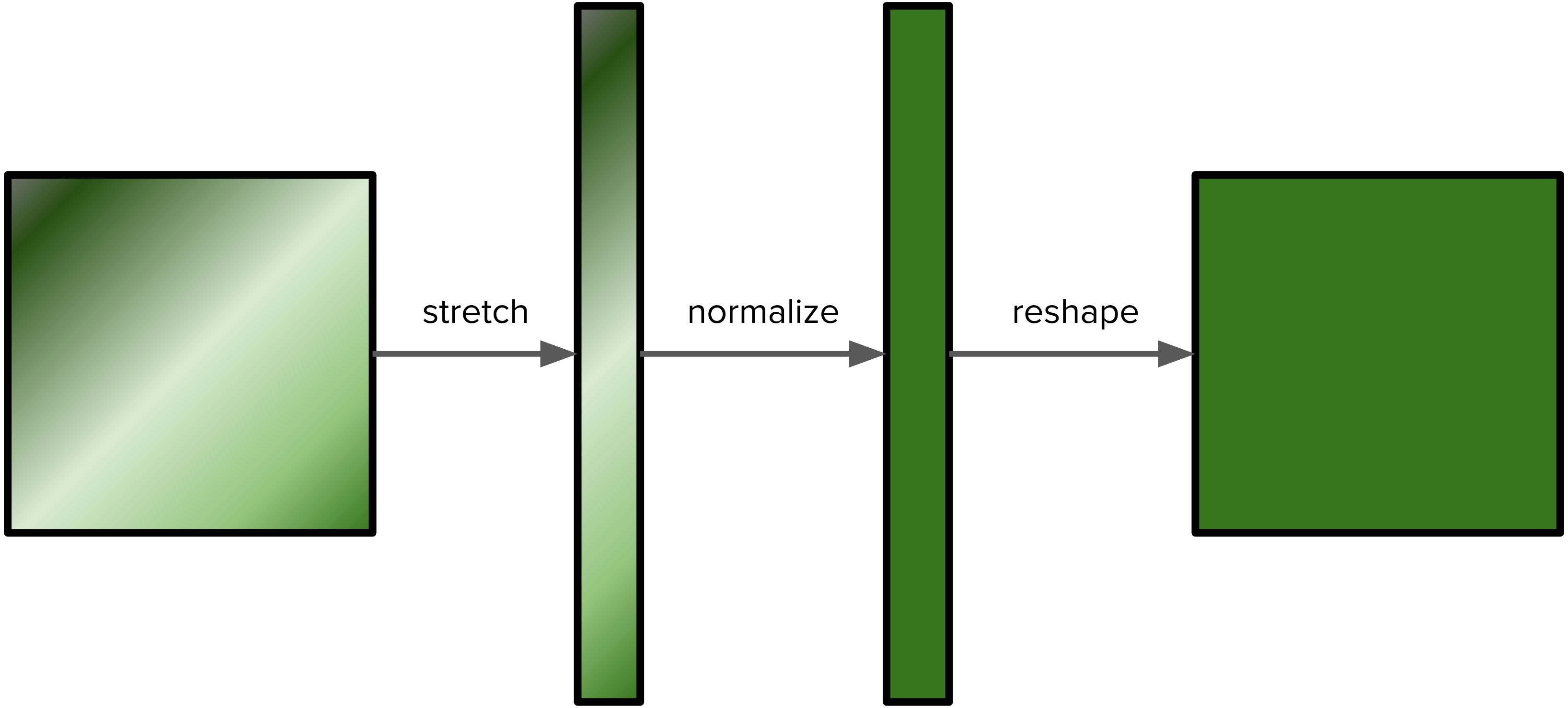
8	11
6	11

Problem: Covariate Shift



We want to normalize our convolutional output

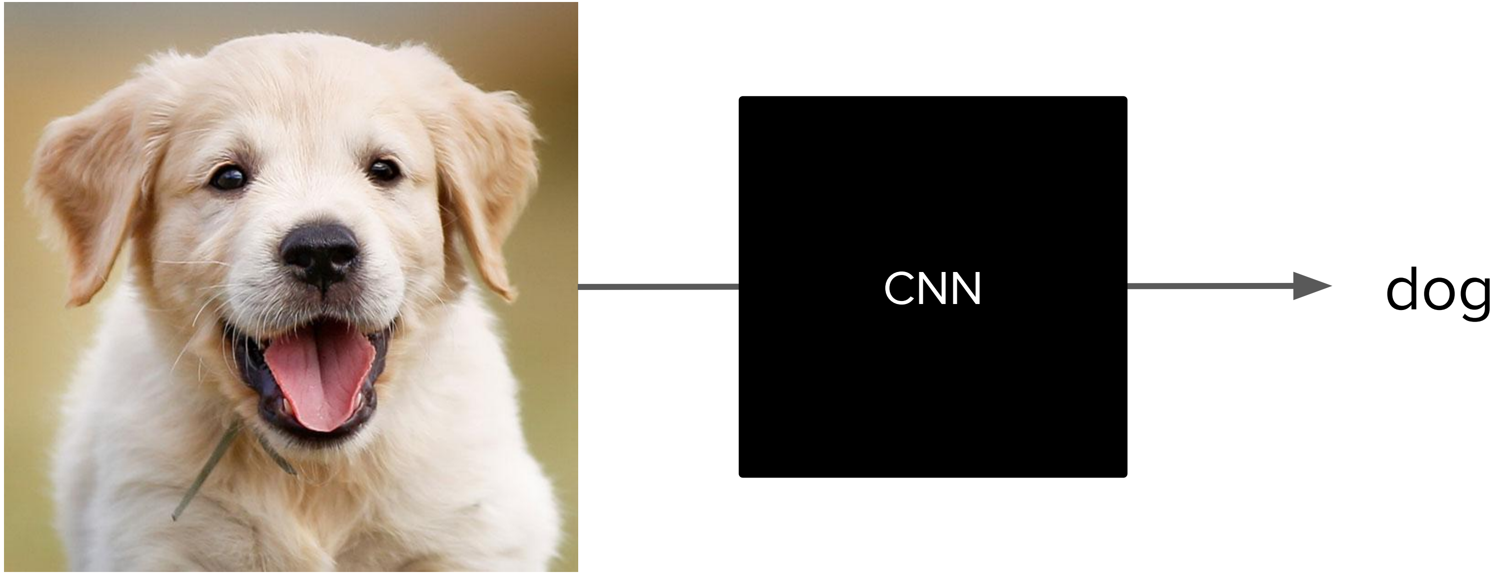
Solution: Spatial Batch Normalization



Questions?

Jupyter Exercises 3: Advanced CNN

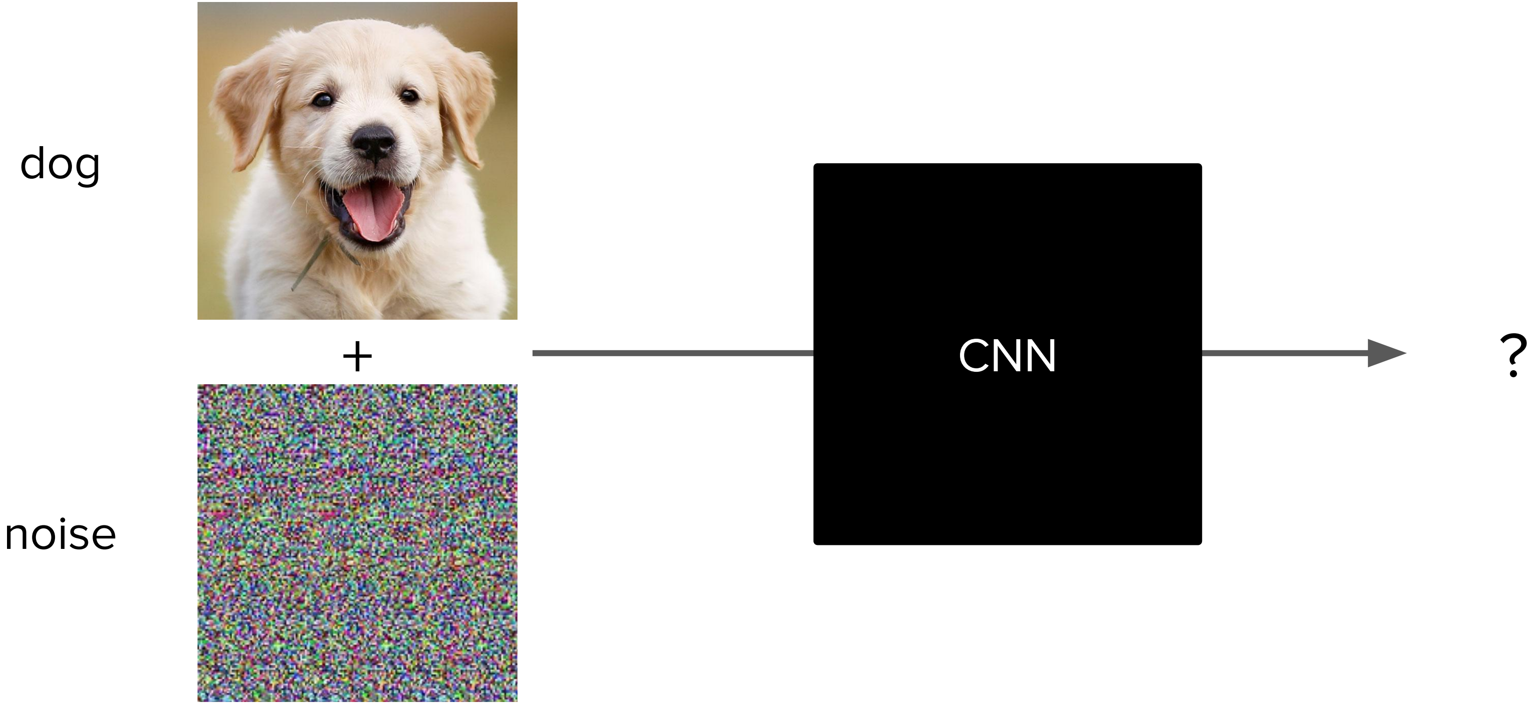
Adversarial Attacks



CNN

dog

Adversarial Attacks



dog

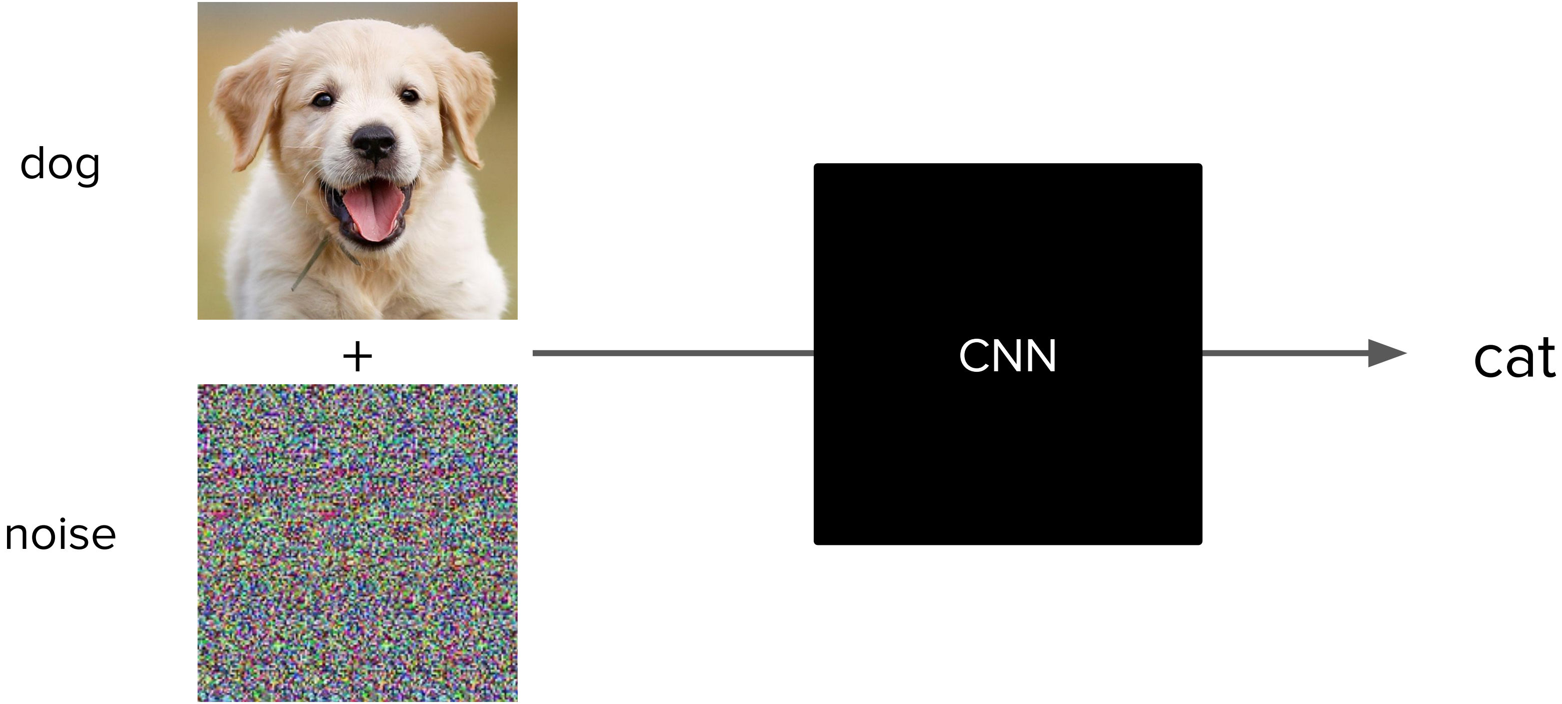
+

noise

CNN

?

Adversarial Attacks



dog

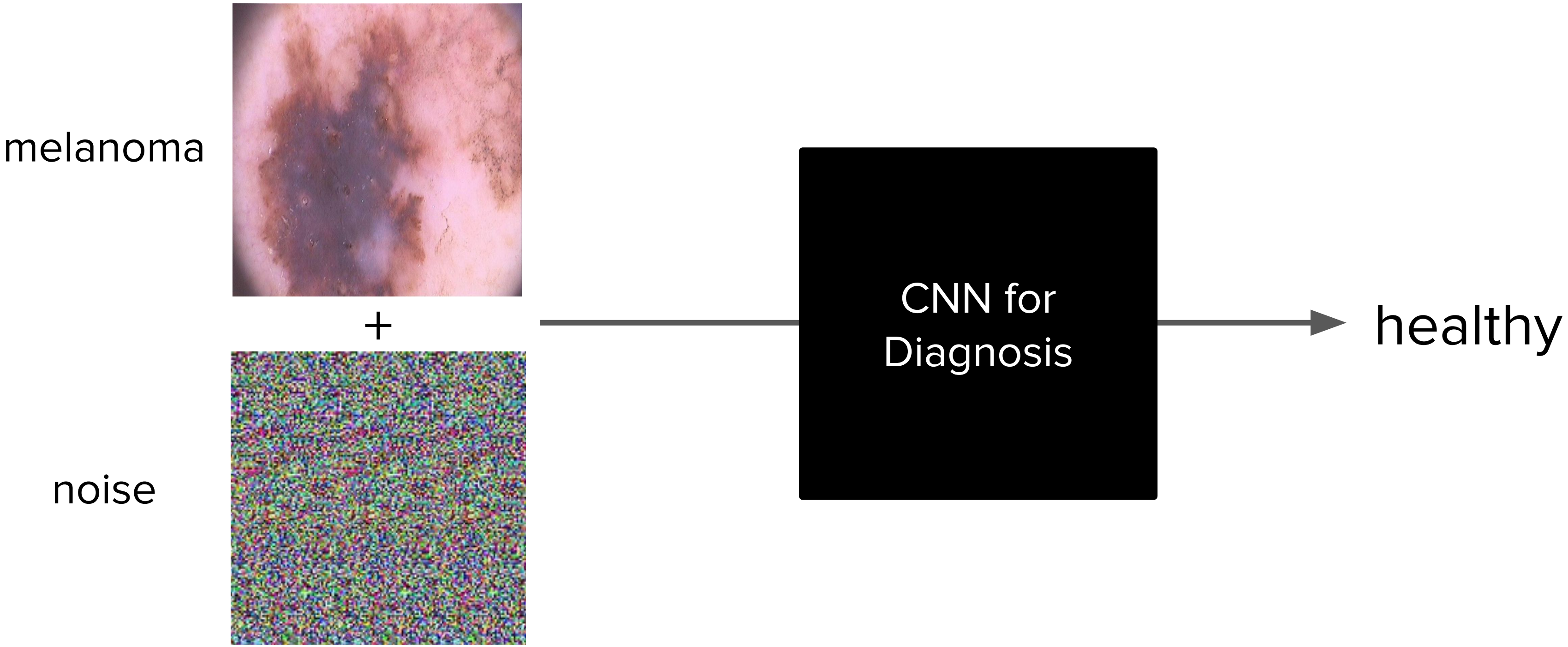
noise

+

CNN

cat

Adversarial Attacks



melanoma

+

noise

CNN for Diagnosis

healthy

Homework:
Adversarial Attacks

Summary of Today

- Reviewed of deep neural networks
- Learned about convolutional neural networks
- Implemented CNNs using max-pool and batch-norm
- Learned about adversarial attacks

Questions?