Applied Deep Learning

Dr. Philippe Blaettchen
Bayes Business School (formerly Cass)

www.bayes.city.ac.uk

Learning objectives of today

Goals:

- Discuss the importance of computer vision for different business applications
- Understand how convolution works and how it can be implemented in TensorFlow
- Apply CNNs to medical diagnosis

How will we do this?

- We will start with a look at Zebra Medical Vision, which provides diagnostic services based on medical imaging
- We will see how convolutional layers drastically simplify the task of analyzing images
- We then turn to the impact of developments in computer vision more broadly

But first: the group assignment

The importance of transparency

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The importance of transparency

Trust

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The importance of transparency

Trust

Fairness

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The importance of transparency

Trust

Fairness

Regulation

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The importance of transparency

Trust

Fairness

Regulation

Informativeness

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

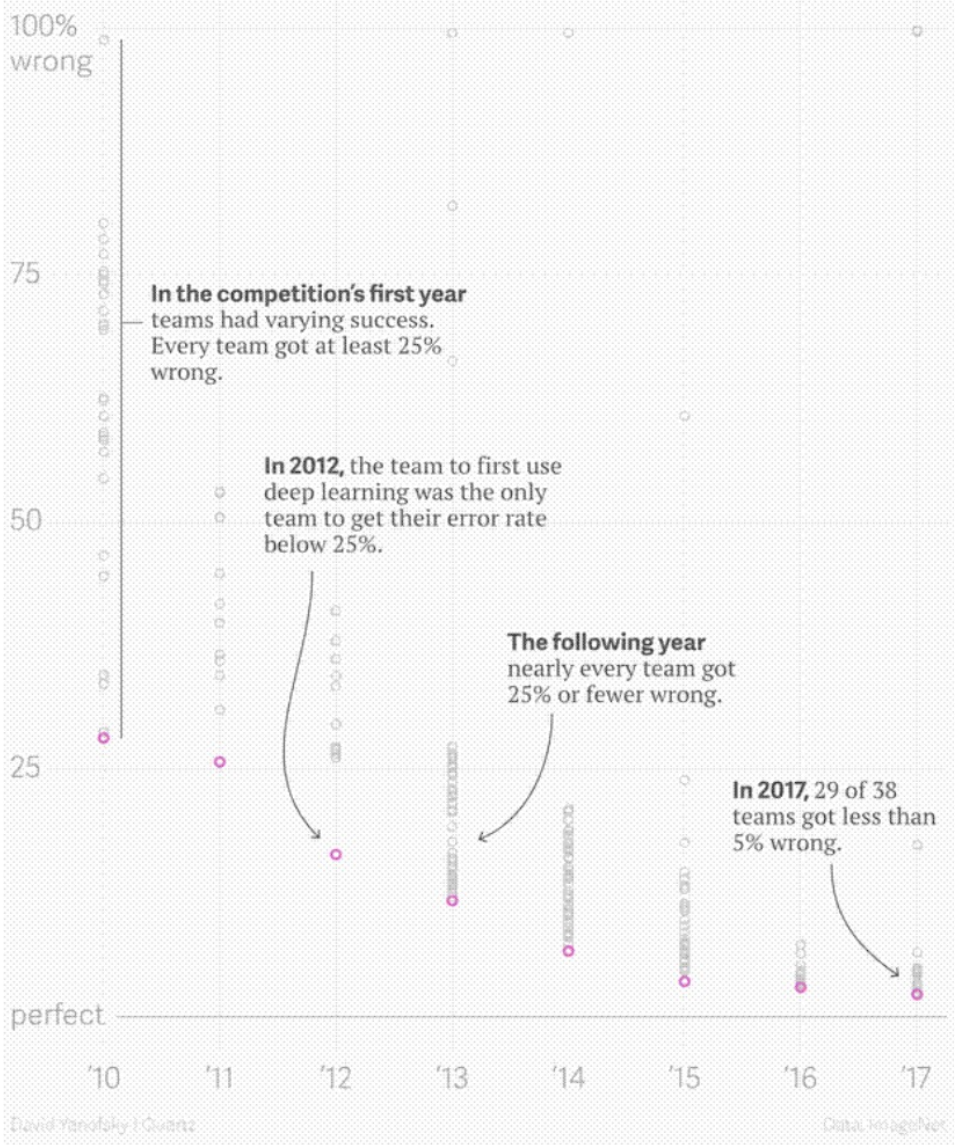
Zebra Medical Vision

The startup

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

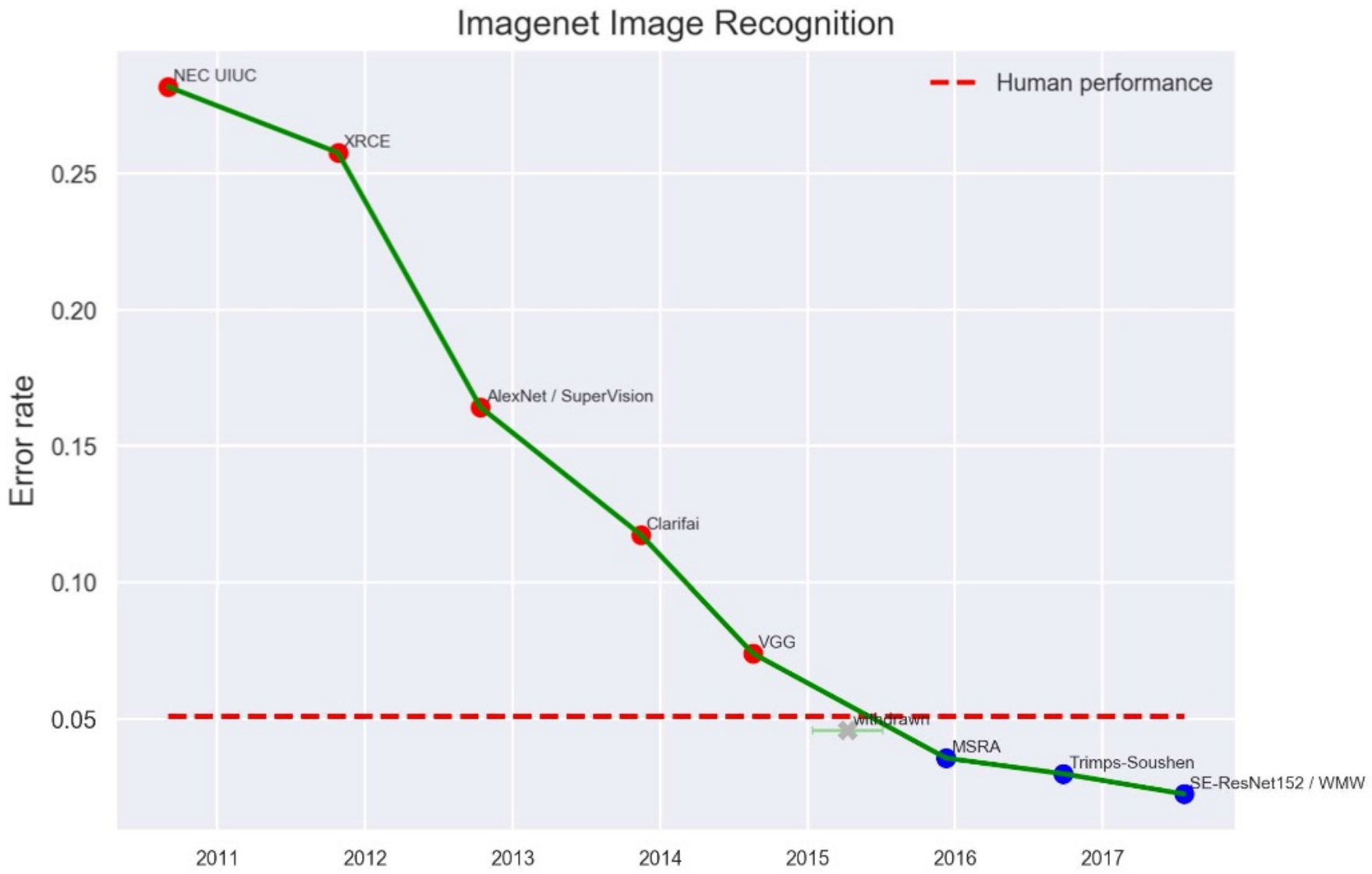
Why now? Why not a decade ago?

ImageNet Large Scale Visual Recognition Challenge Results



BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

ImageNet Large Scale Visual Recognition Challenge Results



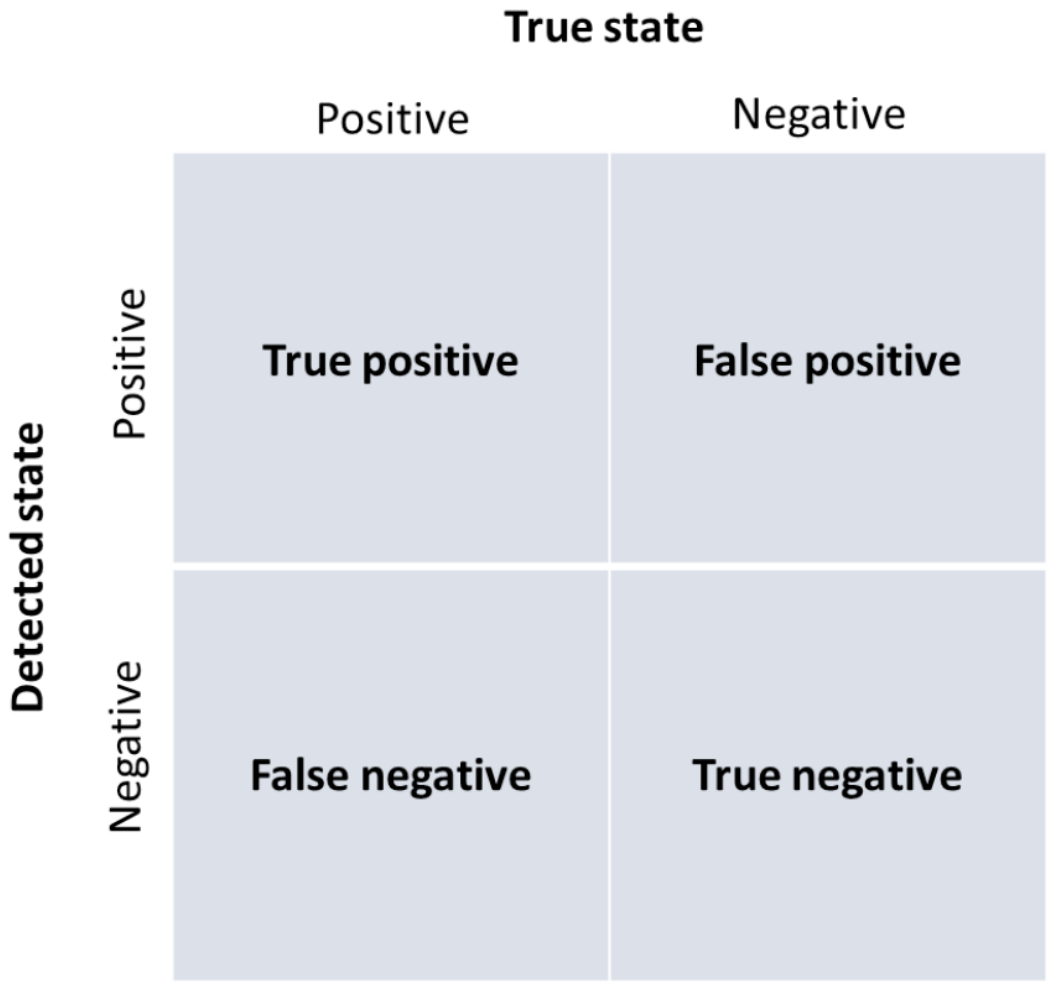
BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

Why now? Why not a decade ago?

What is Zebra fundamentally doing?

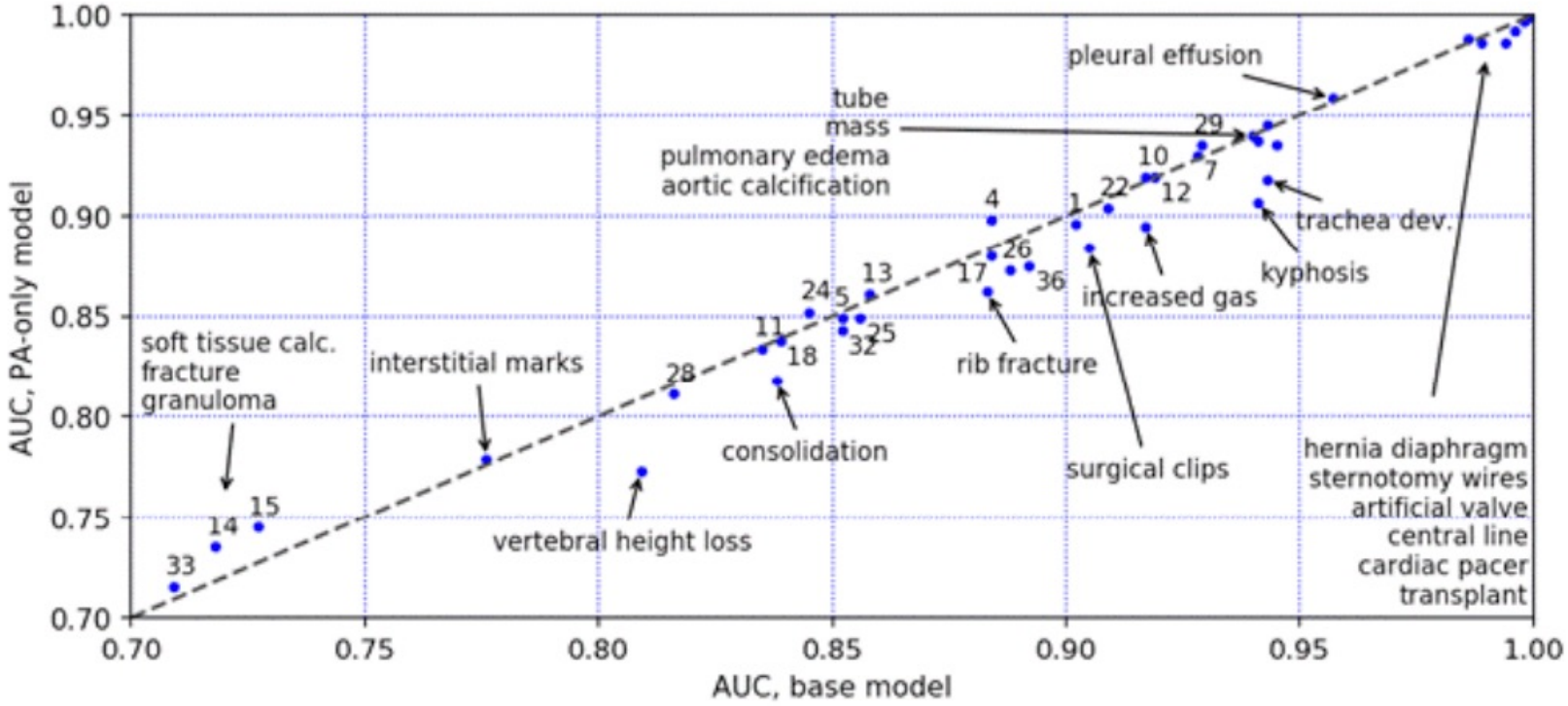
BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

What is Zebra fundamentally doing?



BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

Predictions based on lung X-rays



BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

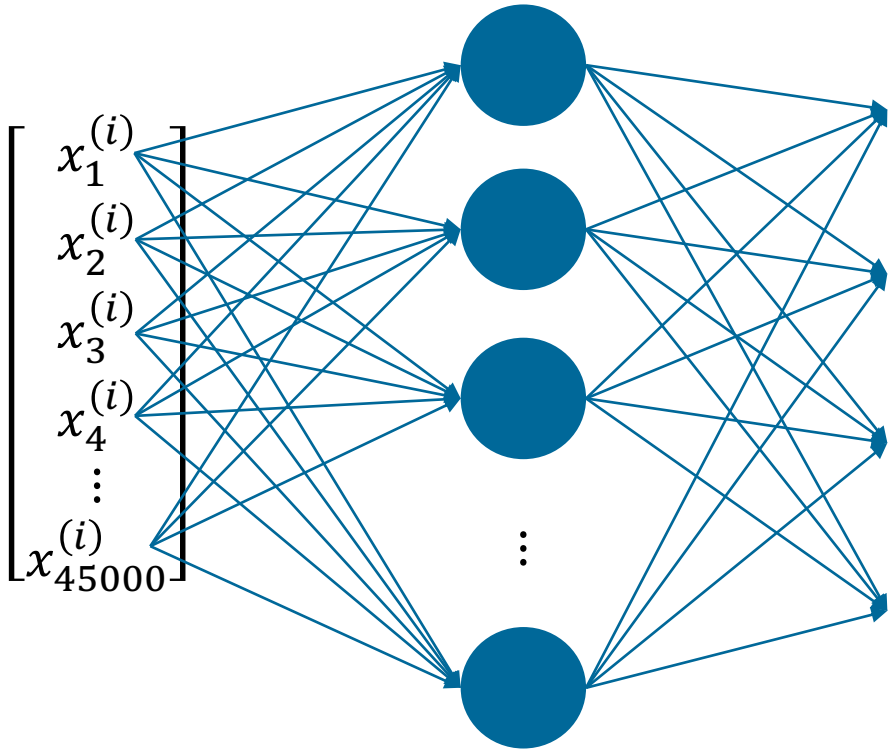
Let’s try classifying lung X-rays ourselves

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

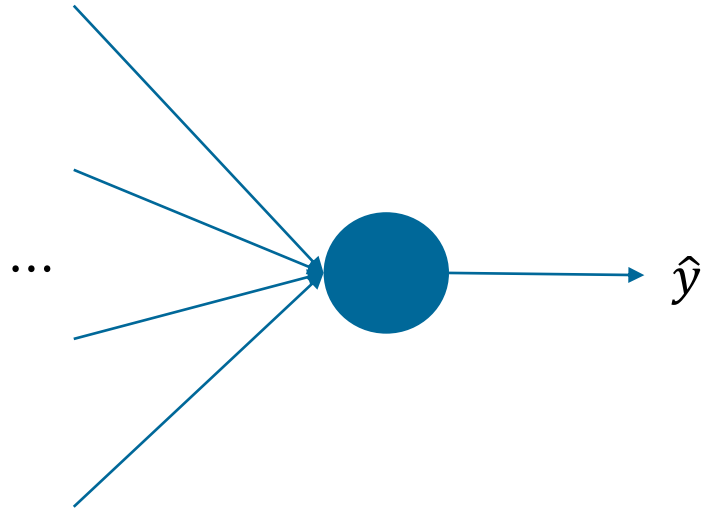
Basics of convolution

Challenges to deep learning using large images

150x100x3



$y = 1$, if it's a cat
$y = 0$, if it's a BA student

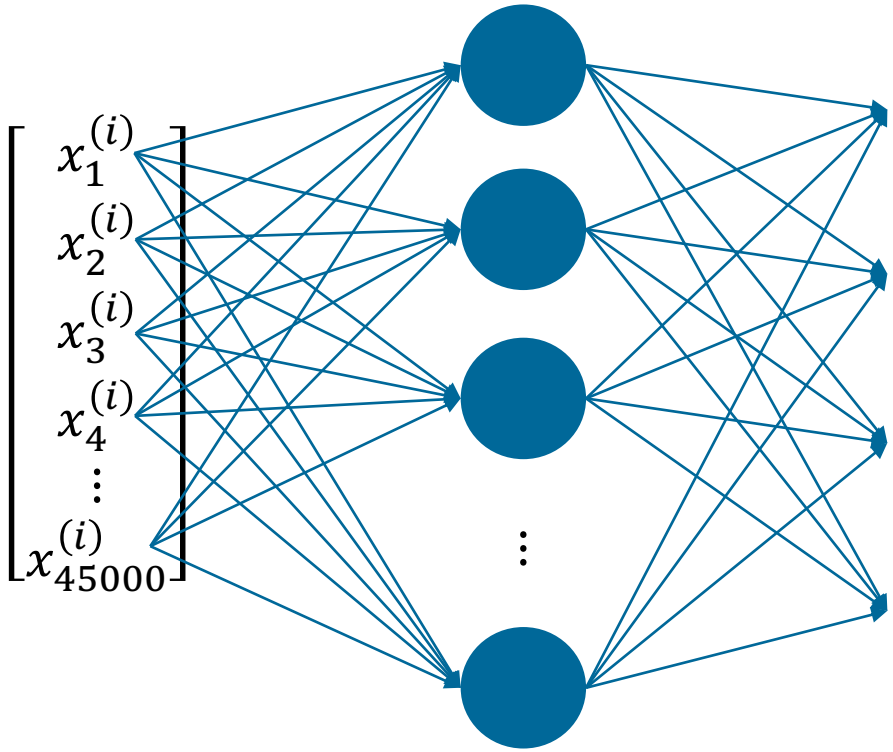


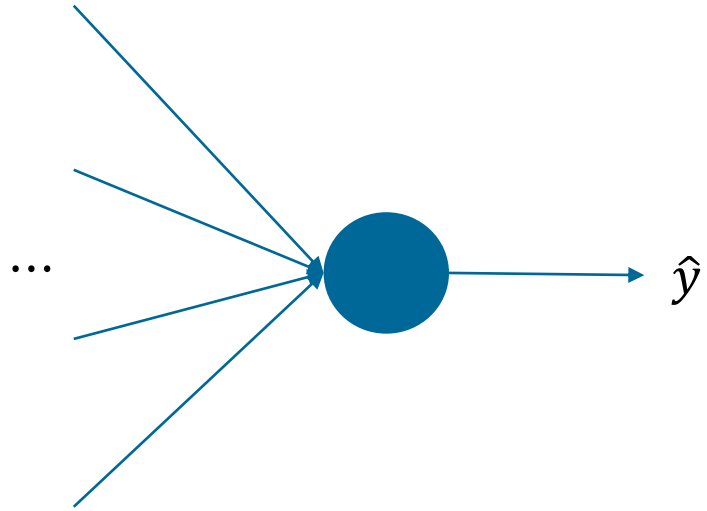
BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

Challenges to deep learning using large images

$y = 1$, if it's a cat
$y = 0$, if it's a BA student

150x100x3

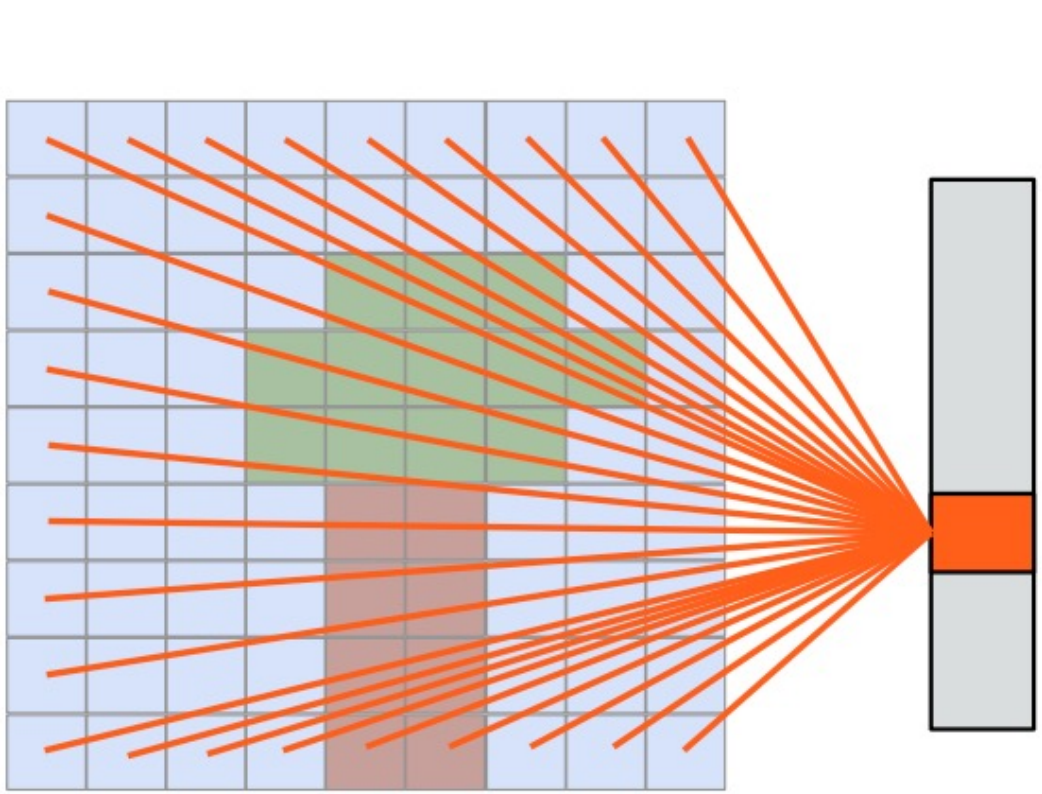


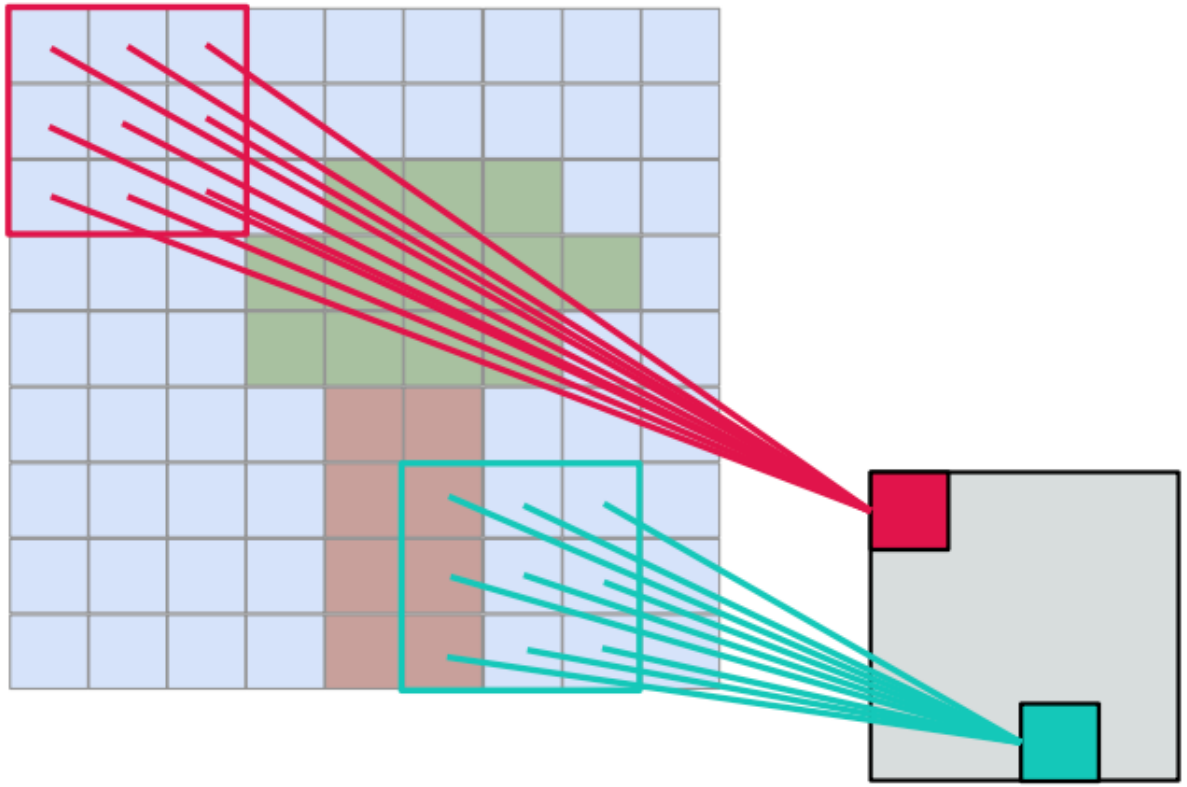


Say we have 100 neurons at the first layer.
Then we need 100*(45000+1)=4.5 mio parameters, just for the first layer!

BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

From fully connected to locally connected





Source: Dieleman

An image filter

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

*

1	0	-1
1	0	-1
1	0	-1

=

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

An image filter

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

*

1	0	-1
1	0	-1
1	0	-1

=

0			

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

An image filter

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

*

1	0	-1
1	0	-1
1	0	-1

=

0			

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

An image filter

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

*

1	0	-1
1	0	-1
1	0	-1

=

0	30		

An image filter

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

*

1	0	-1
1	0	-1
1	0	-1

=

0	30		

An image filter

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

*

1	0	-1
1	0	-1
1	0	-1

=

0	30	30	

An image filter

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

*

1	0	-1
1	0	-1
1	0	-1

=

0	30	30	

An image filter

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

*

1	0	-1
1	0	-1
1	0	-1

=

0	30	30	0

An image filter

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

*

1	0	-1
1	0	-1
1	0	-1

=

0	30	30	0
0	30	30	0
0	30	30	0

An image filter

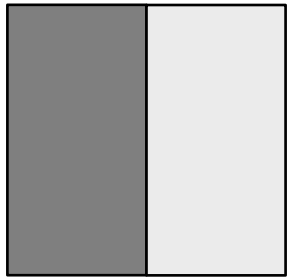
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

*

1	0	-1
1	0	-1
1	0	-1

=

0	30	30	0
0	30	30	0
0	30	30	0



BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

An image filter

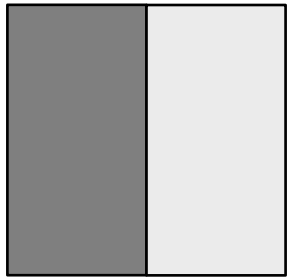
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

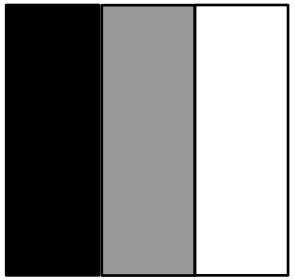
*

1	0	-1
1	0	-1
1	0	-1

=

0	30	30	0
0	30	30	0
0	30	30	0





BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

An image filter

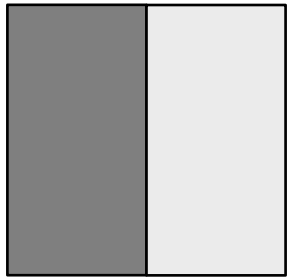
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

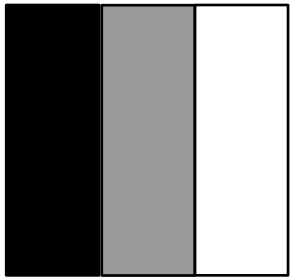
*

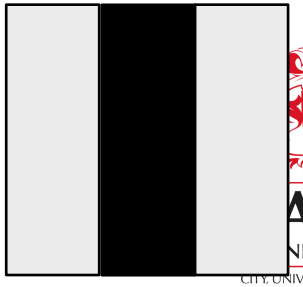
1	0	-1
1	0	-1
1	0	-1

=

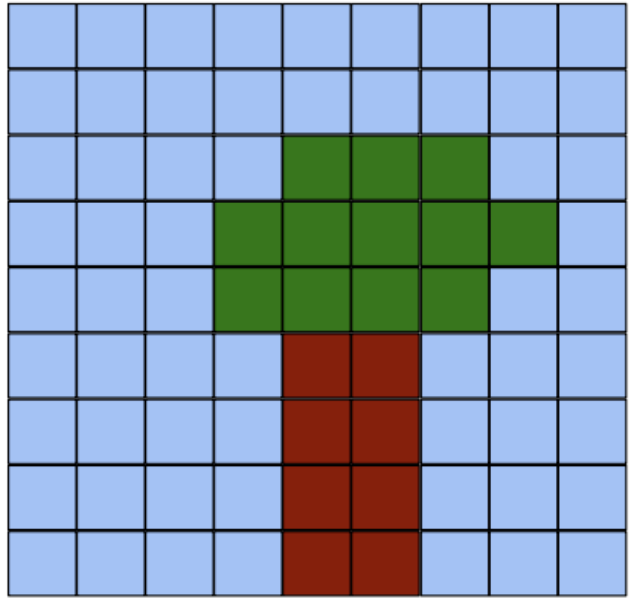
0	30	30	0
0	30	30	0
0	30	30	0







Convolutional layers: learning the filter

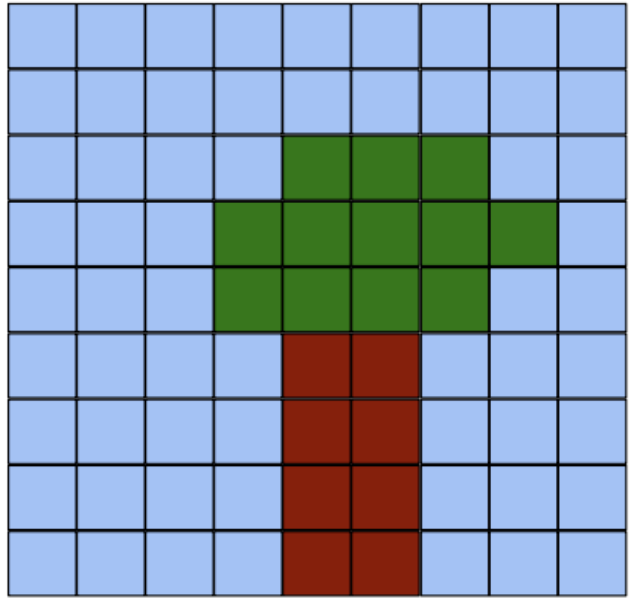


$*$

$w_{1,1}$	$w_{1,2}$	$w_{1,3}$
$w_{2,1}$	$w_{2,2}$	$w_{2,3}$
$w_{3,1}$	$w_{3,2}$	$w_{3,3}$

Source: Dieleman

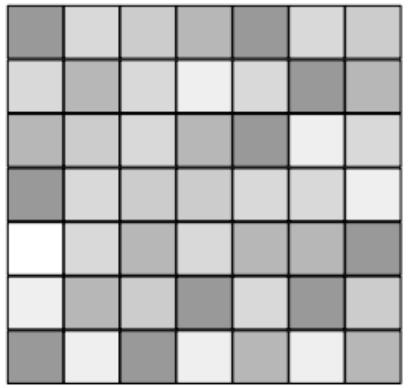
Convolutional layers: learning the filter



$*$

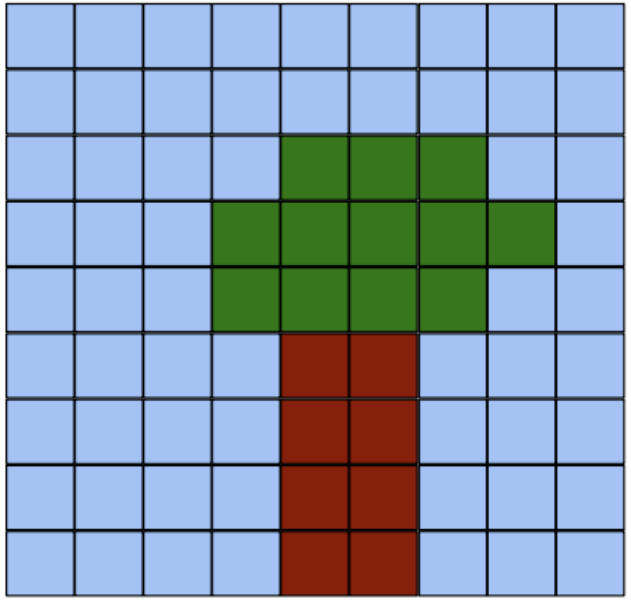
$w_{1,1}$	$w_{1,2}$	$w_{1,3}$
$w_{2,1}$	$w_{2,2}$	$w_{2,3}$
$w_{3,1}$	$w_{3,2}$	$w_{3,3}$

$+ \quad b \quad =$



Source: Dieleman

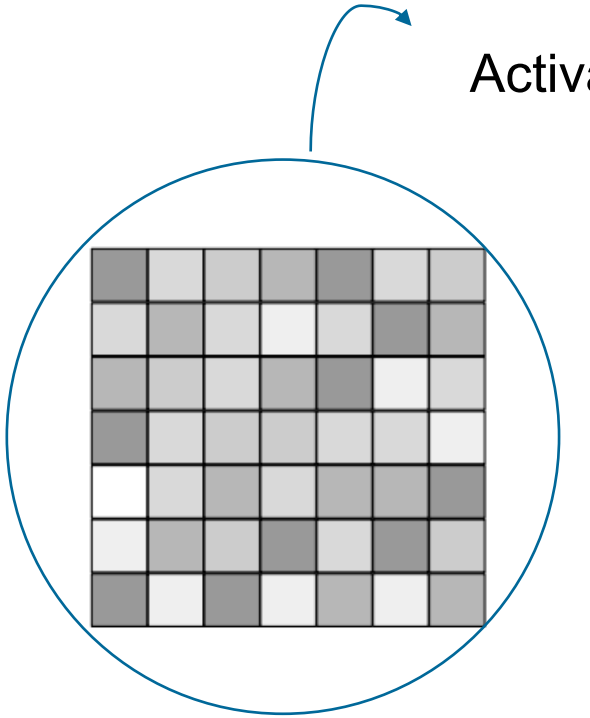
Convolutional layers: learning the filter



$*$

$w_{1,1}$	$w_{1,2}$	$w_{1,3}$
$w_{2,1}$	$w_{2,2}$	$w_{2,3}$
$w_{3,1}$	$w_{3,2}$	$w_{3,3}$

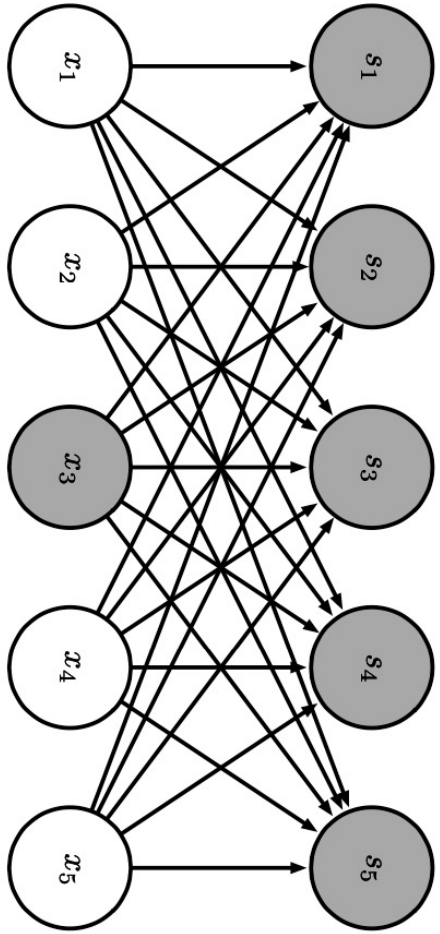
$+ \quad b \quad =$

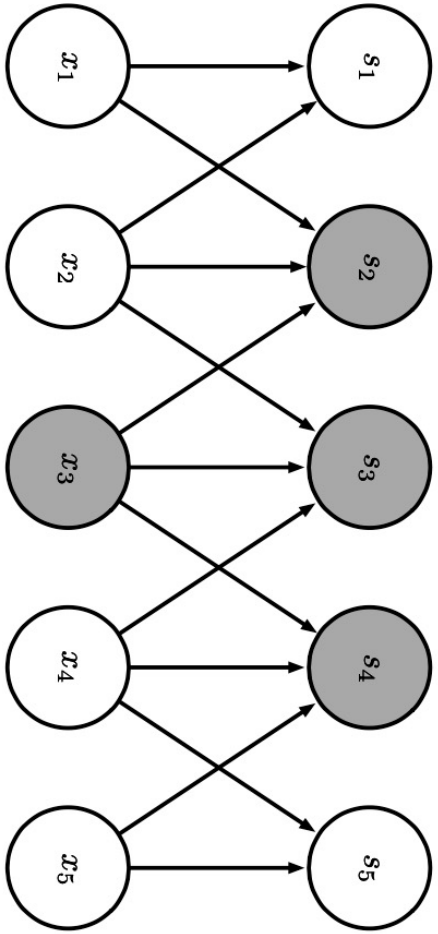


Activation function

Source: Dieleman

Less parameters required

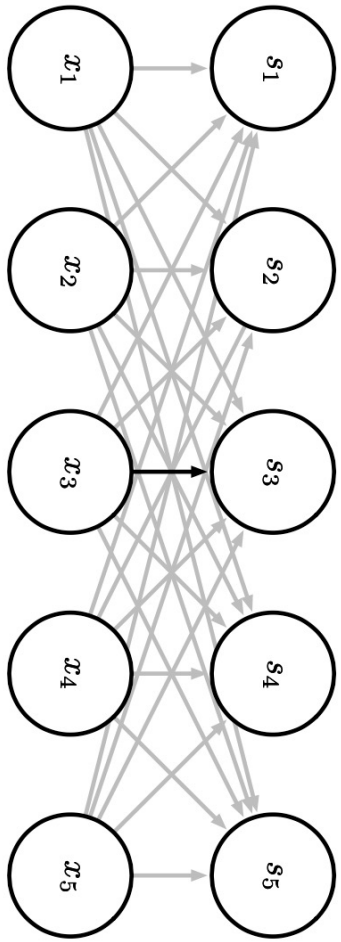


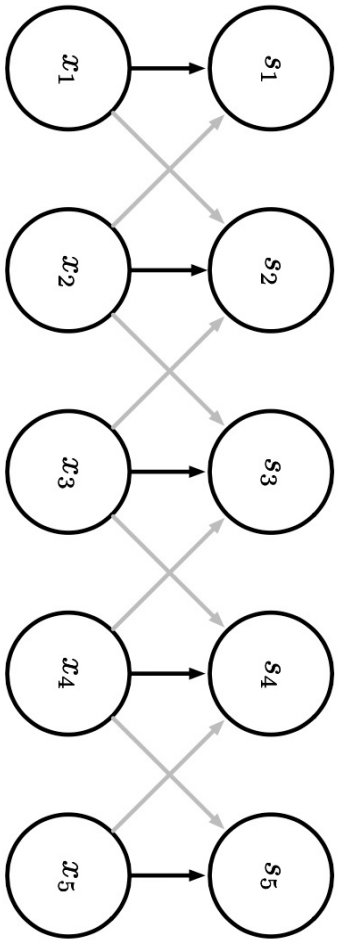


Source: Goodfellow

BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

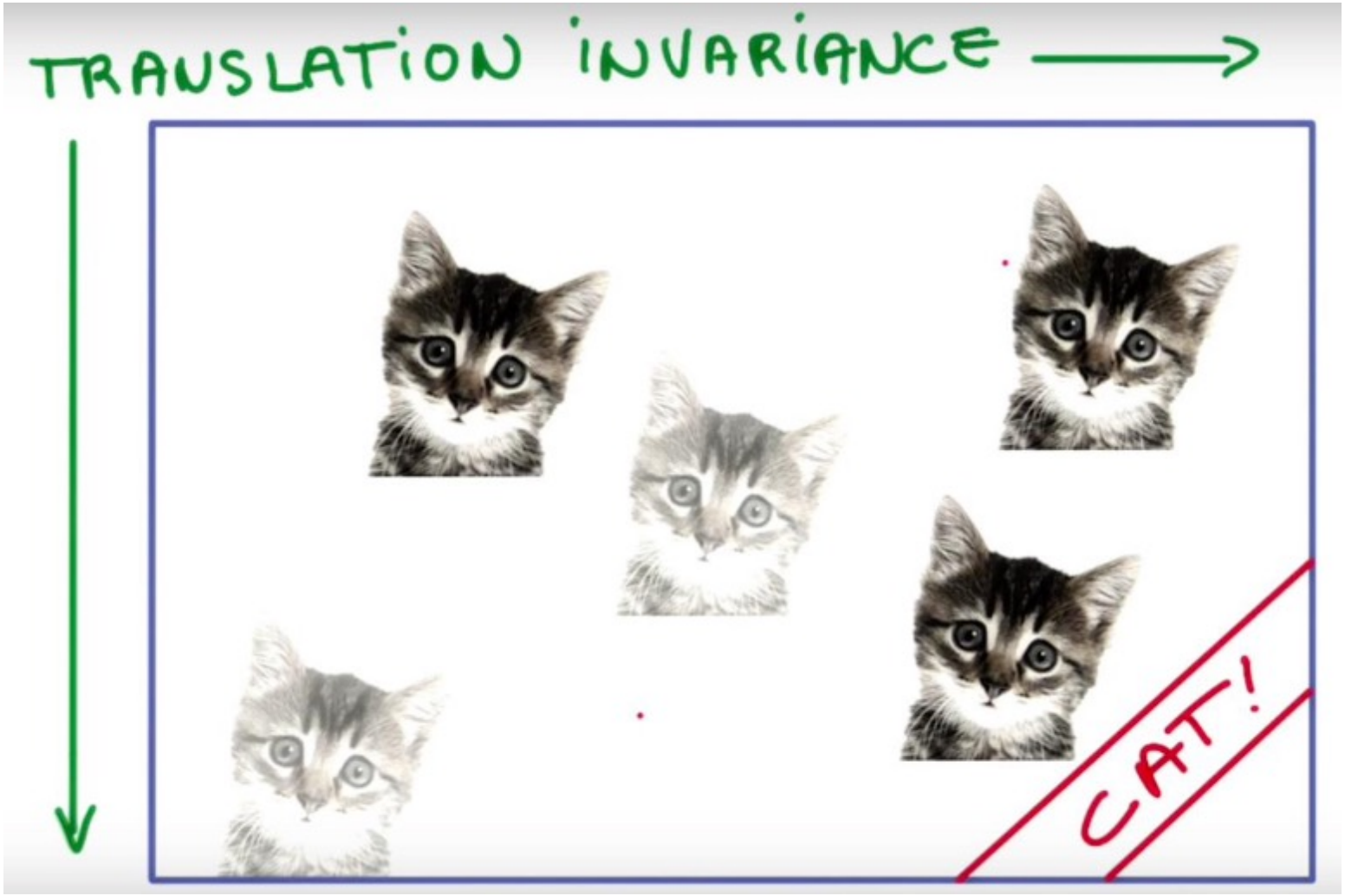
Parameters are shared





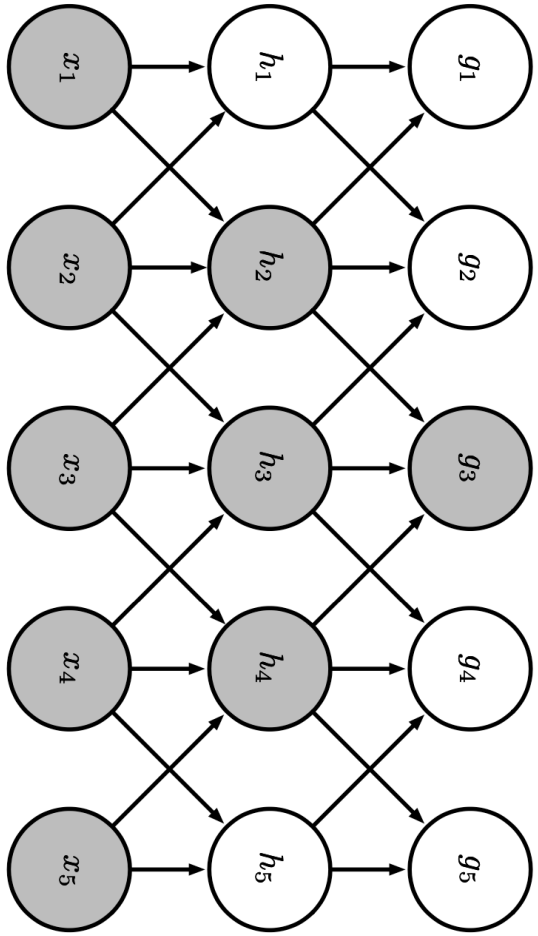
Source: Goodfellow

Shared parameters lead to translation invariance



Source: Bhaskhar

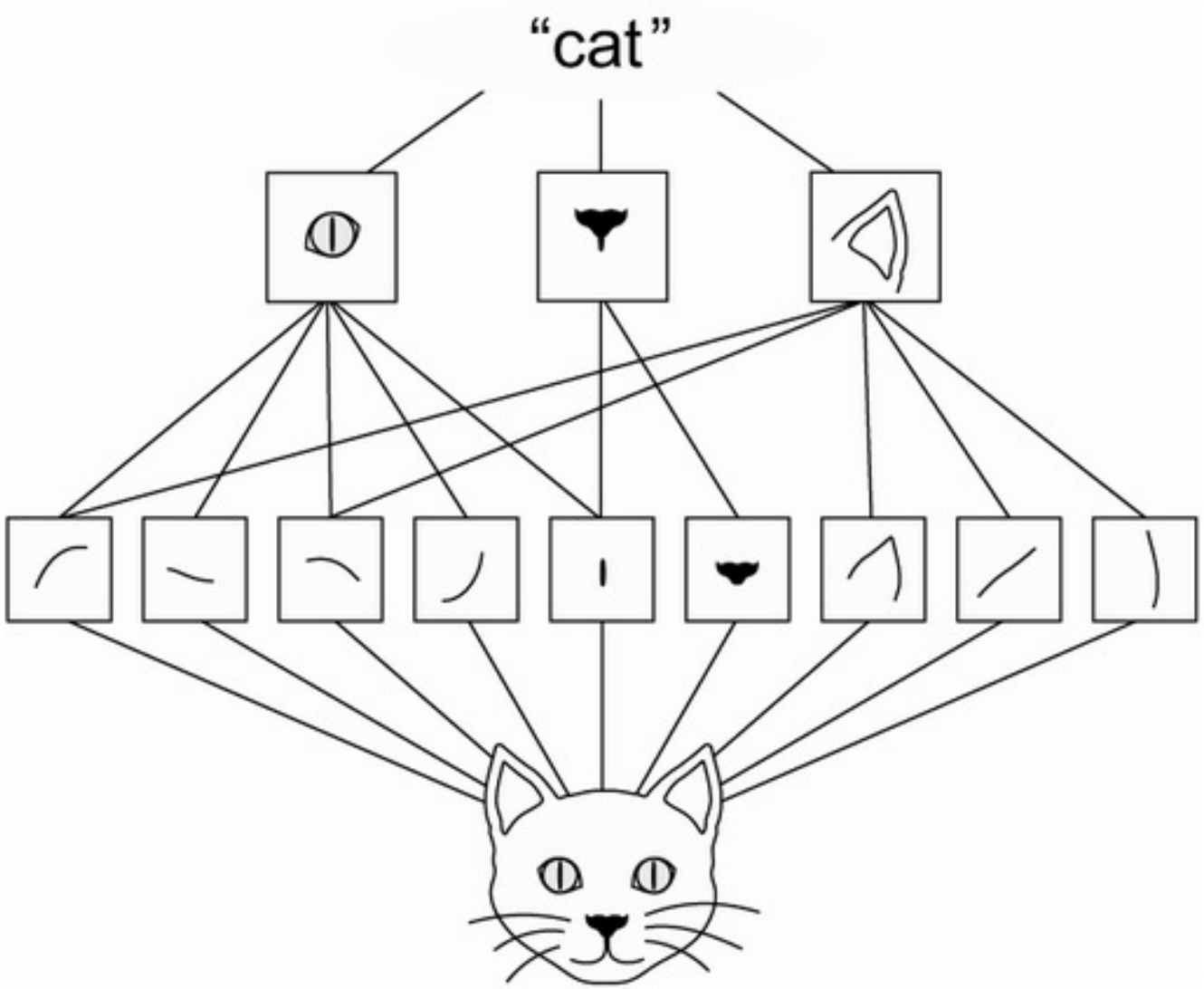
Hierarchical setup



Source: Goodfellow

BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

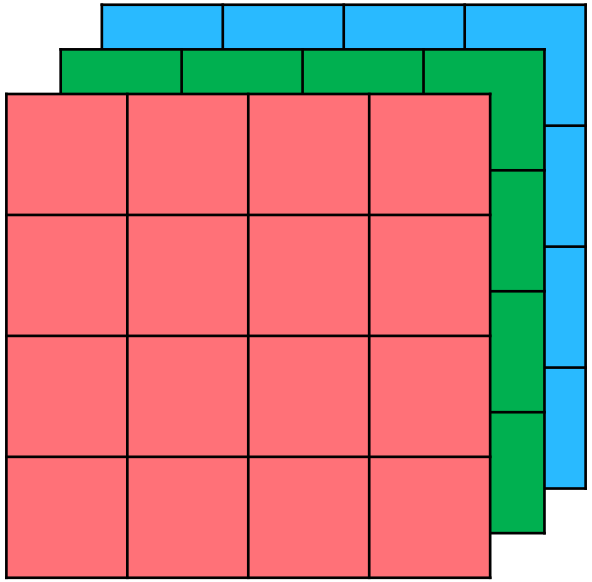
Hierarchical structure supports a spatial hierarchy of learning



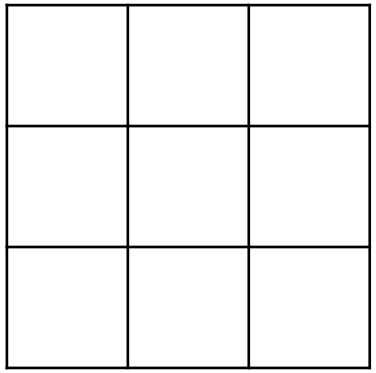
Source: Chollet

From convolution to CNNs

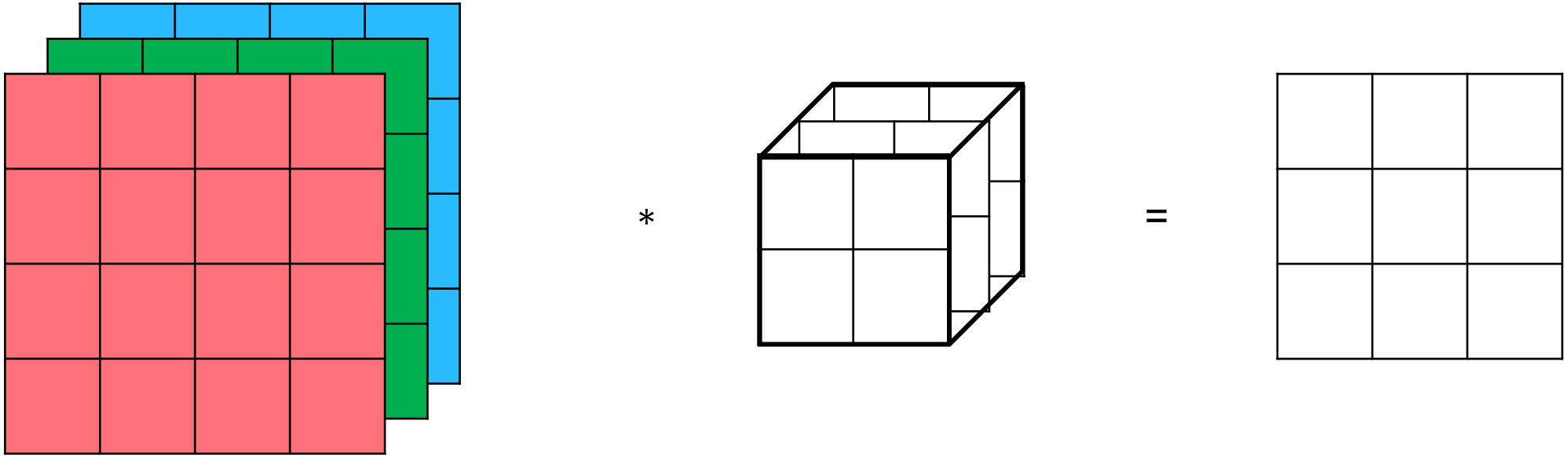
Convolution on a 3D array



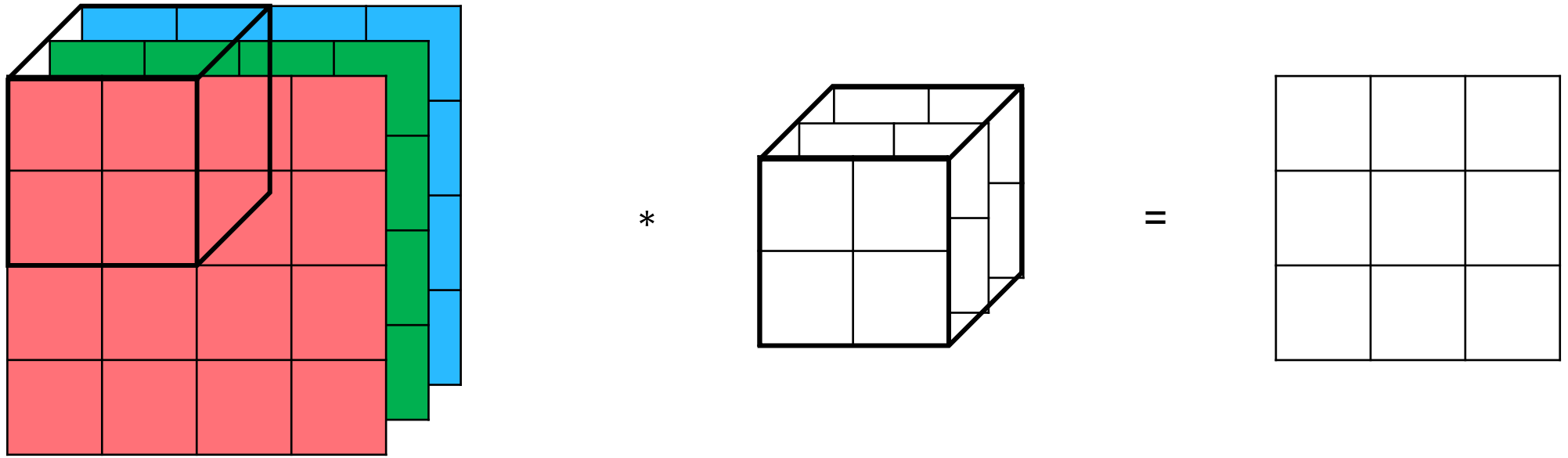
=



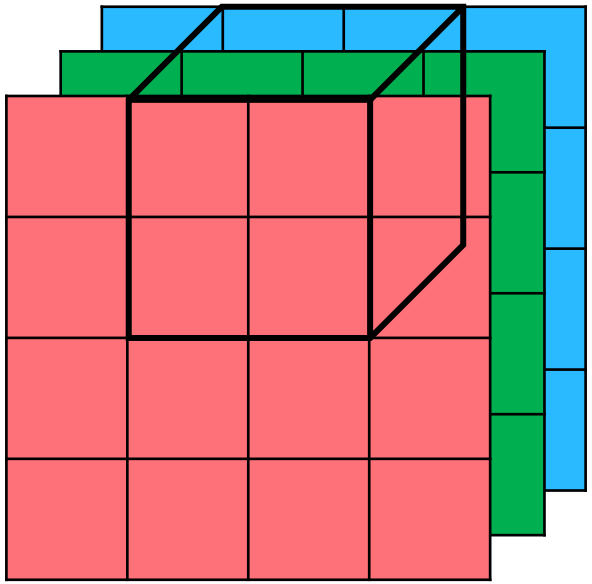
Convolution on a 3D array



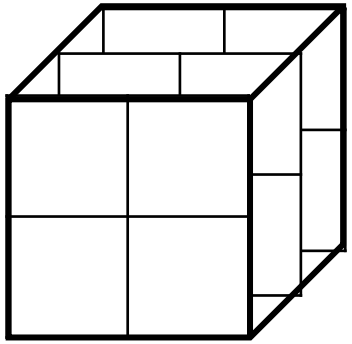
Convolution on a 3D array



Convolution on a 3D array



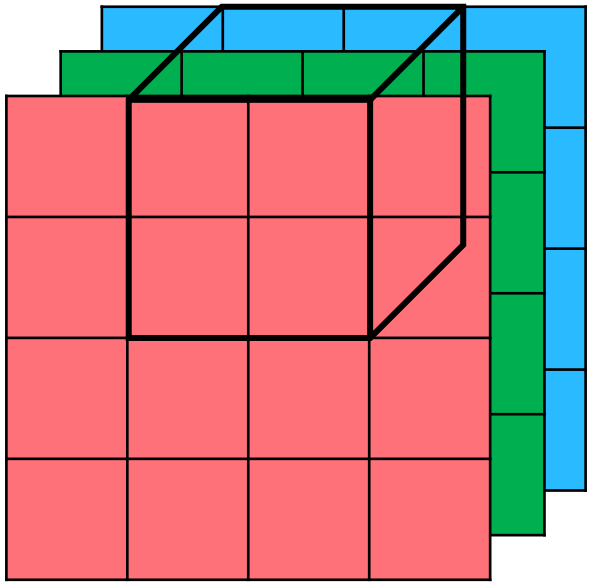
*



=

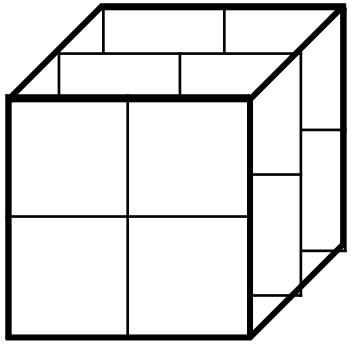
20		

Convolution on a 3D array



4x4x3

*



2x2x3

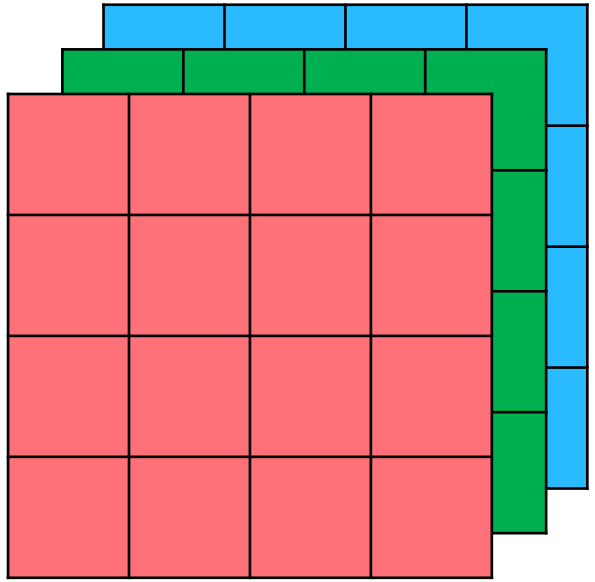
=

20		

3x3

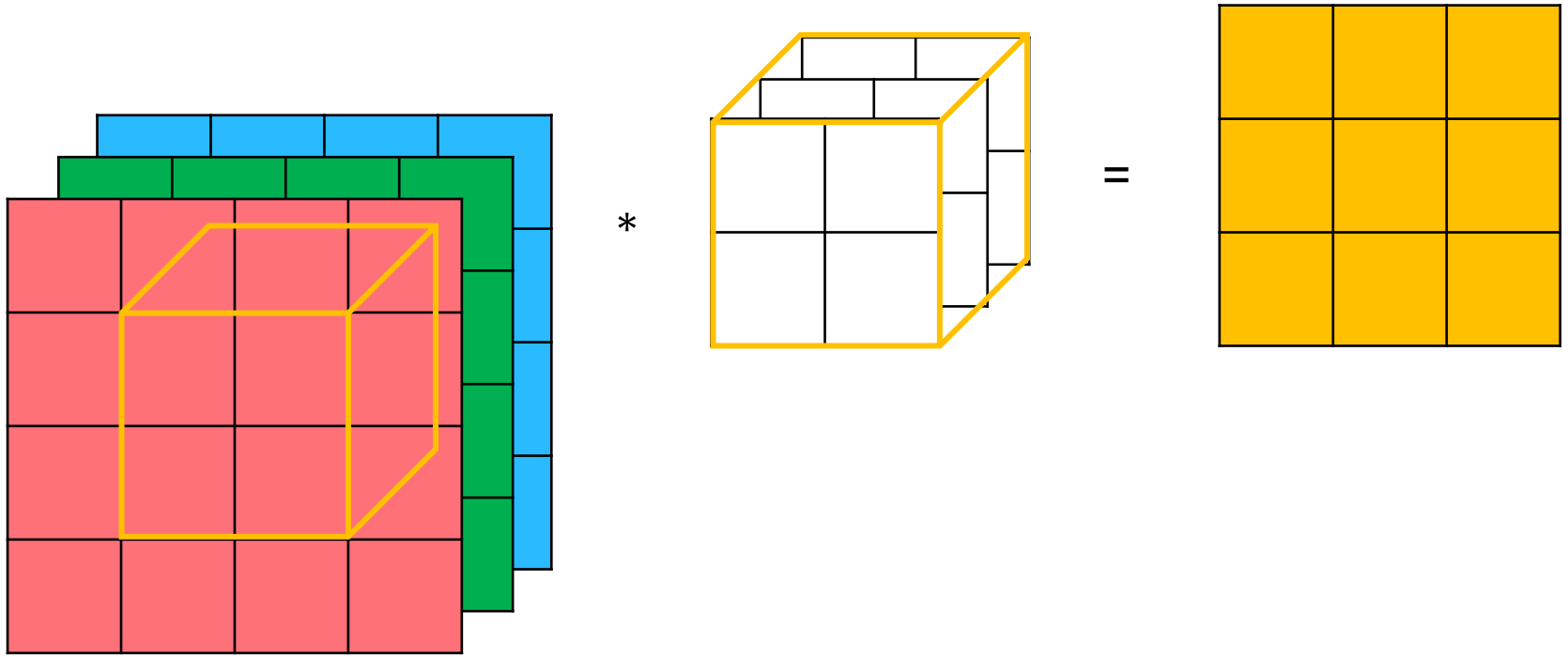
BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

Multiple 3D convolutions

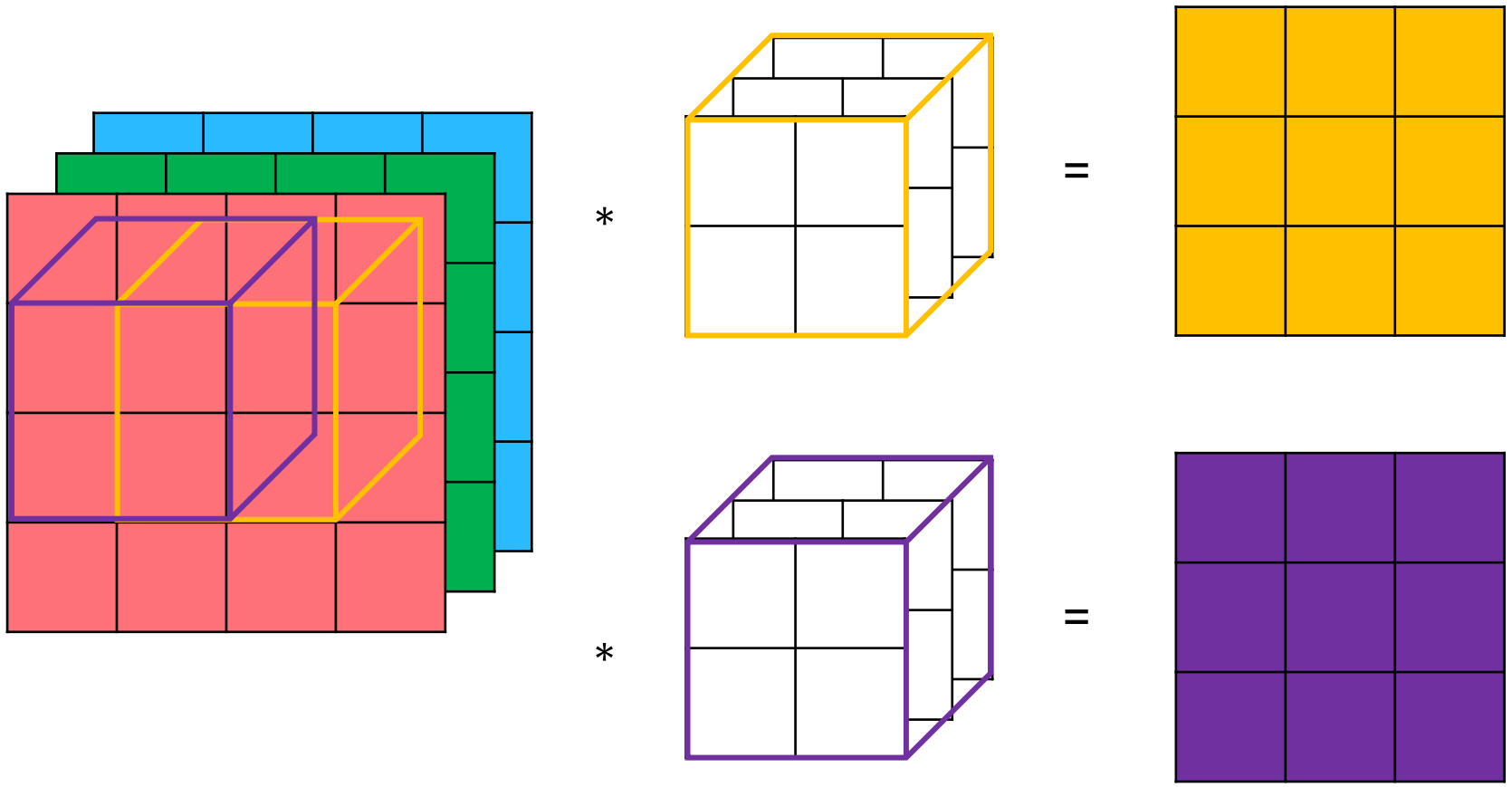


BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

Multiple 3D convolutions

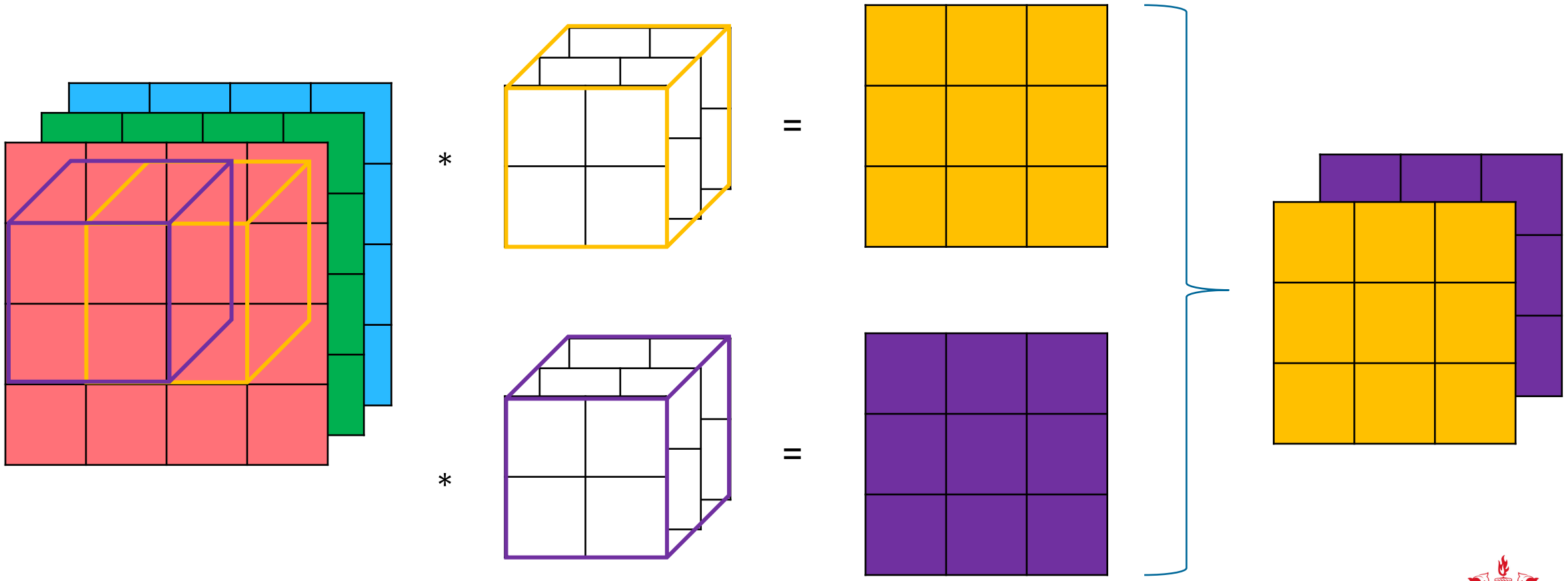


Multiple 3D convolutions



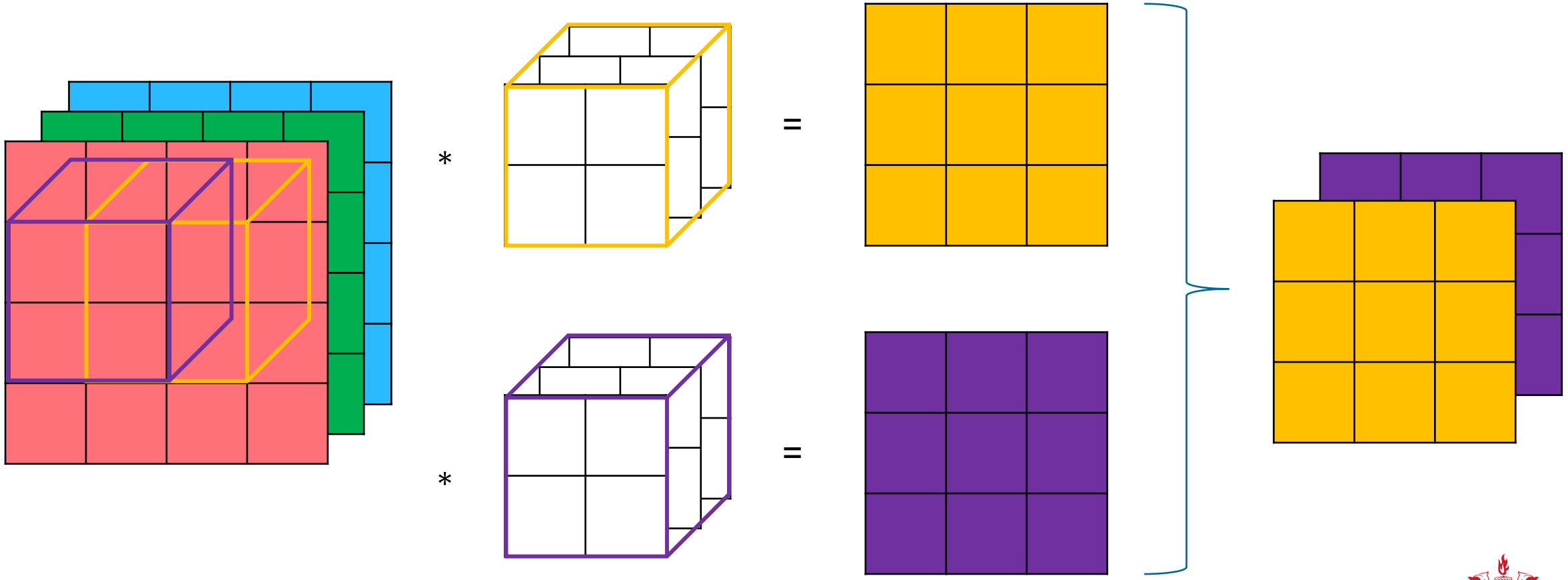
BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

Multiple 3D convolutions



BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

Multiple 3D convolutions



4x4x3

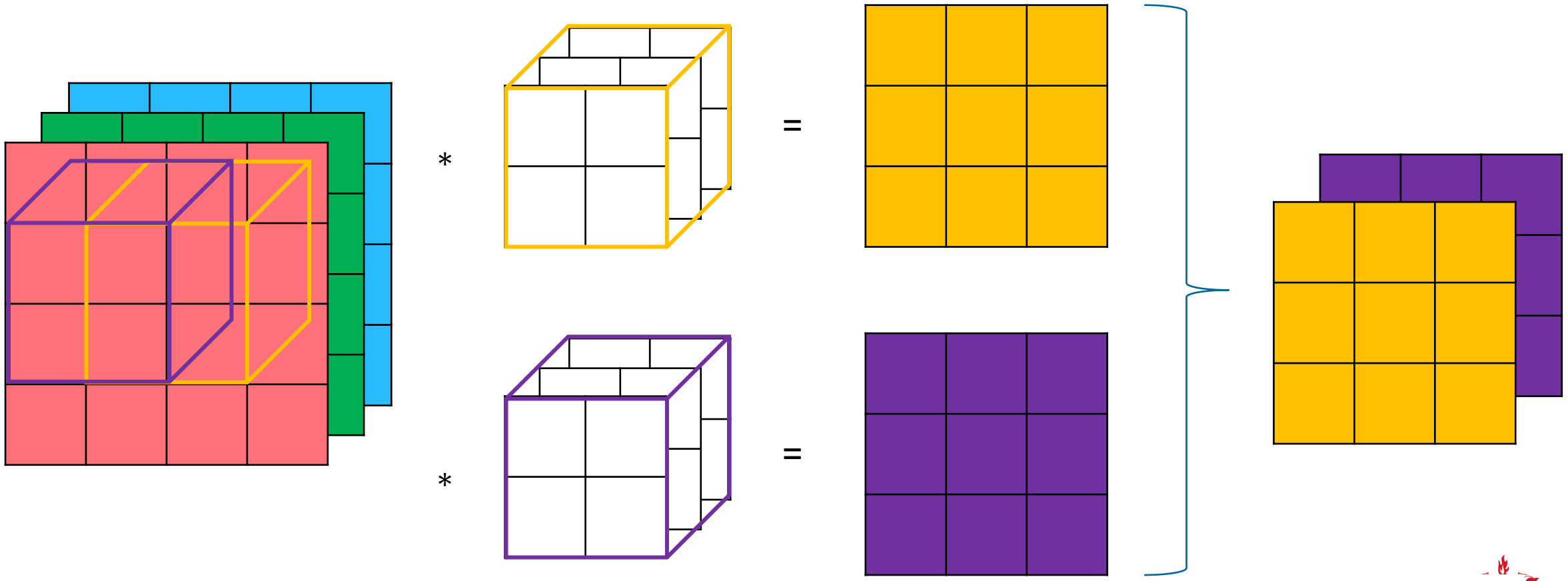
$(n_H^{[0]}, n_W^{[0]}, \mathrm{n}_\mathrm{C}^{[0]})$

2x2x3

$(f_H, f_W, \mathrm{n}_\mathrm{C}^{[0]})$

$n_C^{[1]}$= number of filters

Multiple 3D convolutions



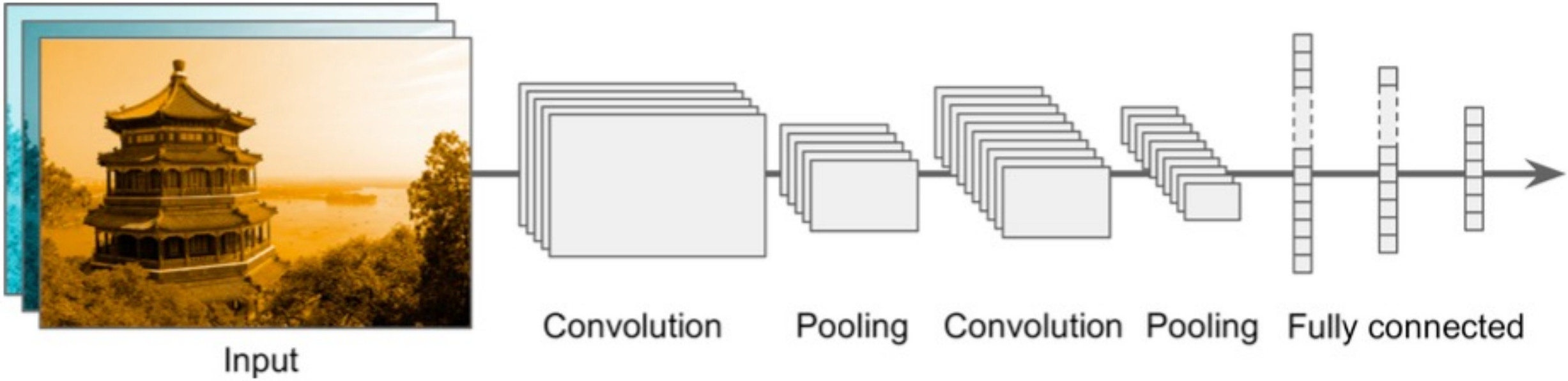
4x4x3
$(n_H^{[0]}, n_W^{[0]}, \mathrm{n}_\mathrm{C}^{[0]})$

2x2x3
$(f_H, f_W, \mathrm{n}_\mathrm{C}^{[0]})$

$n_C^{[1]}$= number of filters

3x3x2
$(n_H^{[1]}, n_W^{[1]}, \mathrm{n}_\mathrm{C}^{[1]})$

Typical architecture

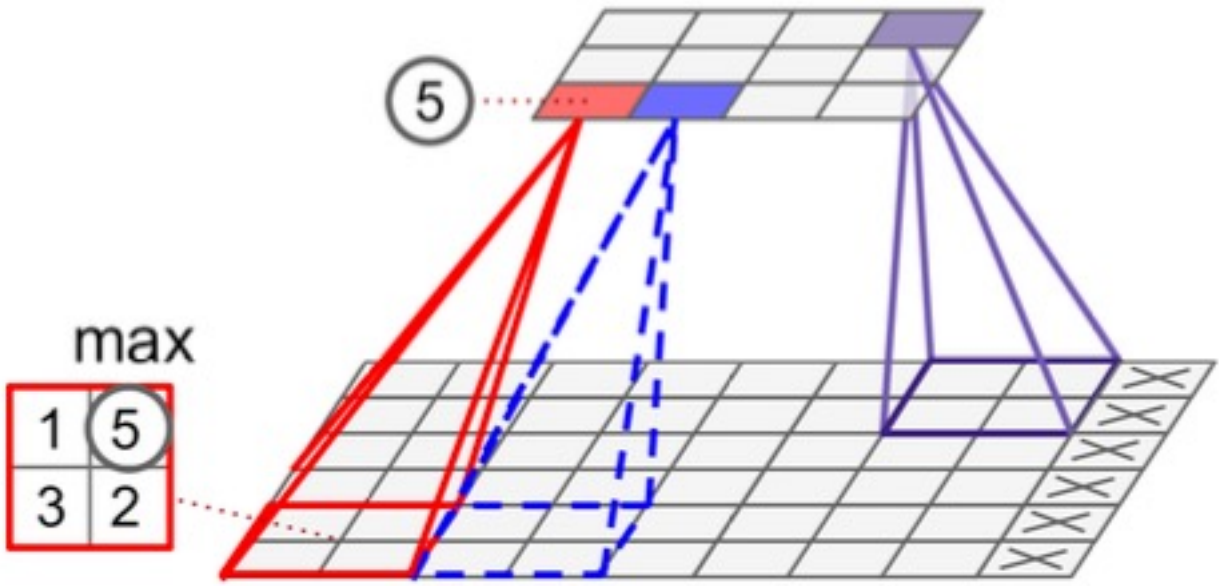


Source: Géron

Pooling layers

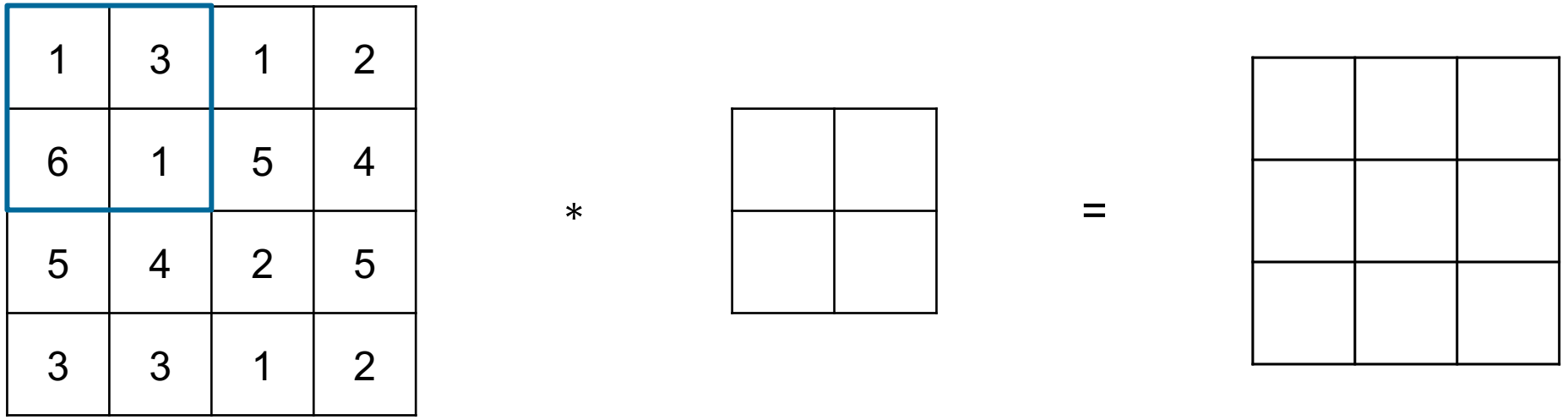
- Subsample (i.e., summarize) the input
  - Reduced computational load
  - Reduced memory usage
  - Fewer parameters (and, thus, less overfitting)

Max pooling



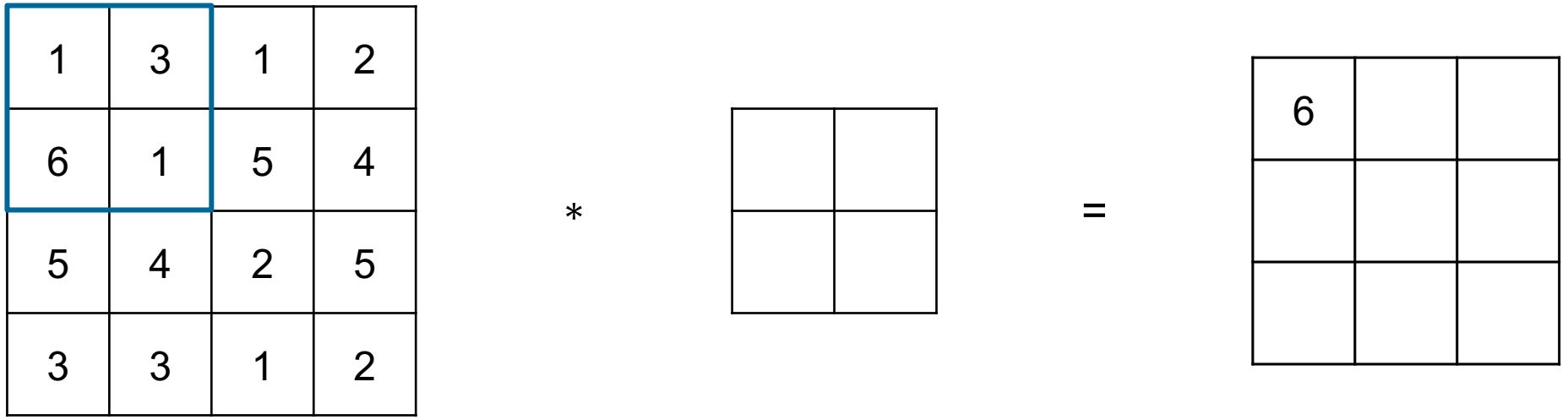
Source: Géron

Max pooling in practice



Source: Géron

Max pooling in practice



Source: Géron

Max pooling in practice

1	3	1	2
6	1	5	4
5	4	2	5
3	3	1	2

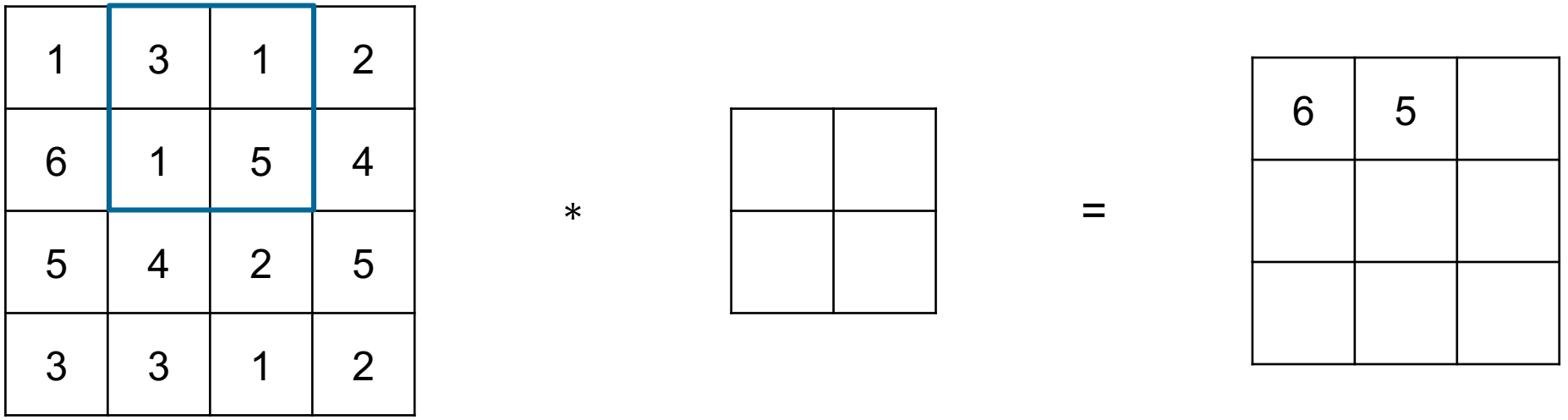
*

=

6		

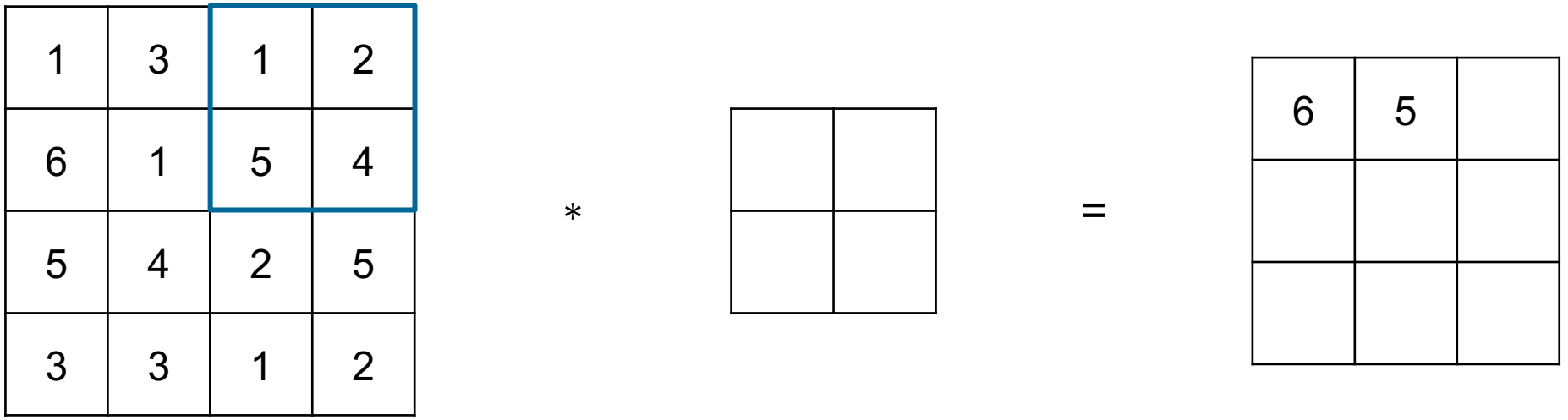
Source: Géron

Max pooling in practice



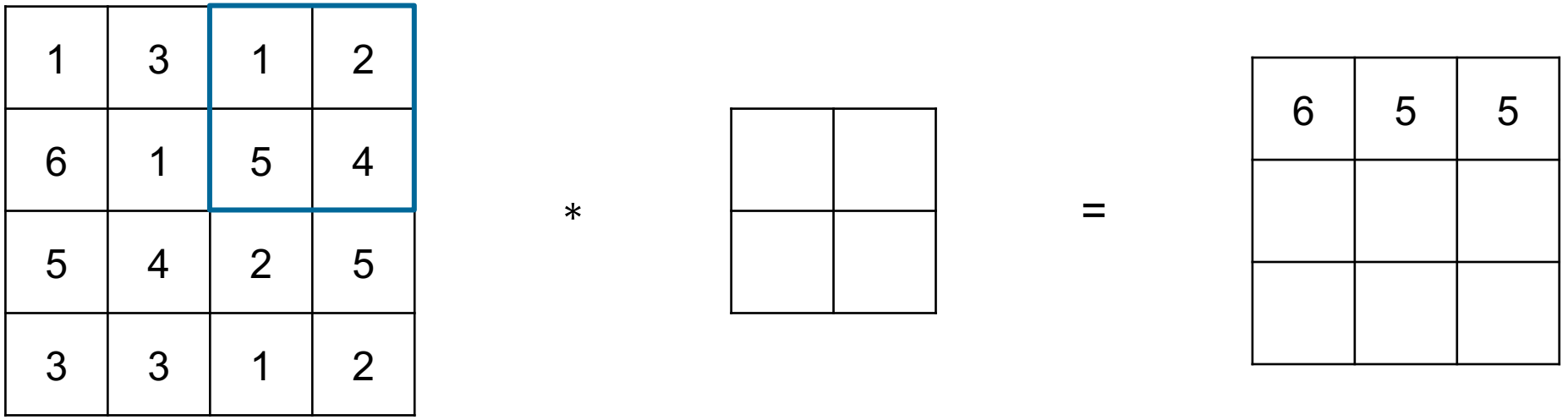
Source: Géron

Max pooling in practice



Source: Géron

Max pooling in practice



Source: Géron

Max pooling in practice

1	3	1	2
6	1	5	4
5	4	2	5
3	3	1	2

*

=

6	5	5
6	5	5
5	4	5

Source: Géron

Timer permitting – additional details on convolutions

With convolution, the center matters more

1	3	1	2
6	1	5	4
5	4	2	5
3	3	1	2

*

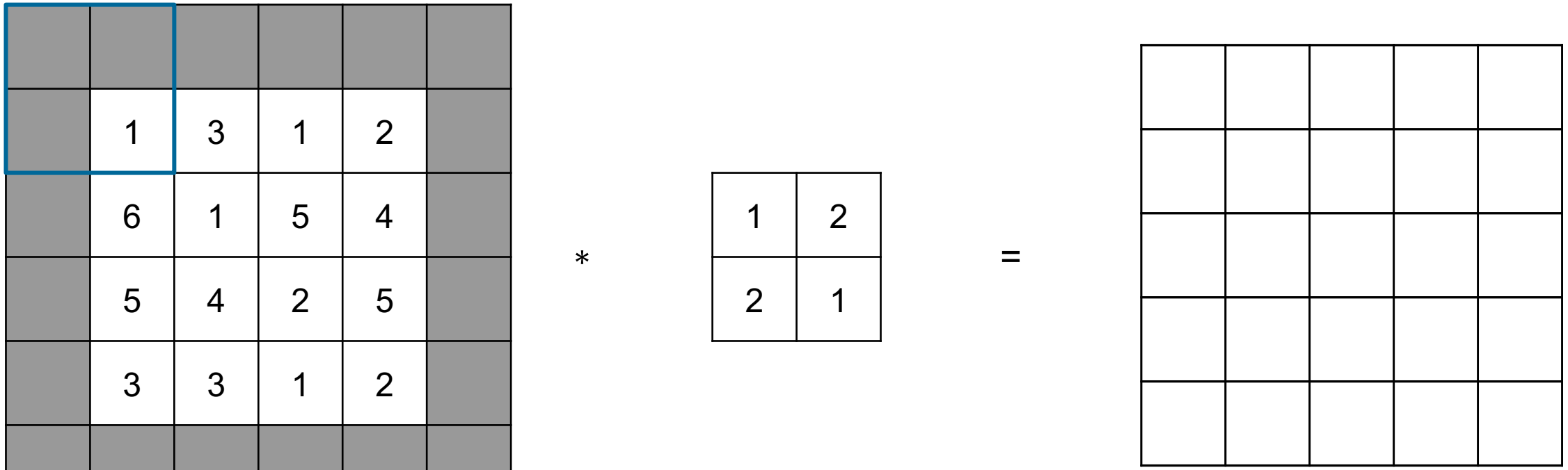
1	2
2	1

=

20	12	19
22	21	22
22	15	16

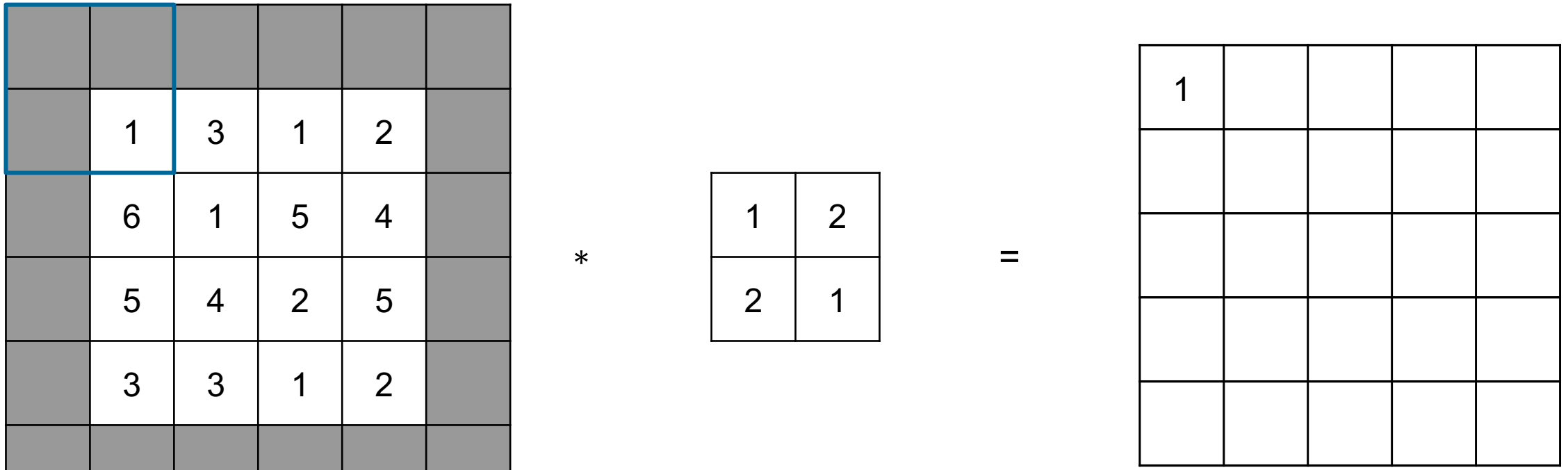
BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding



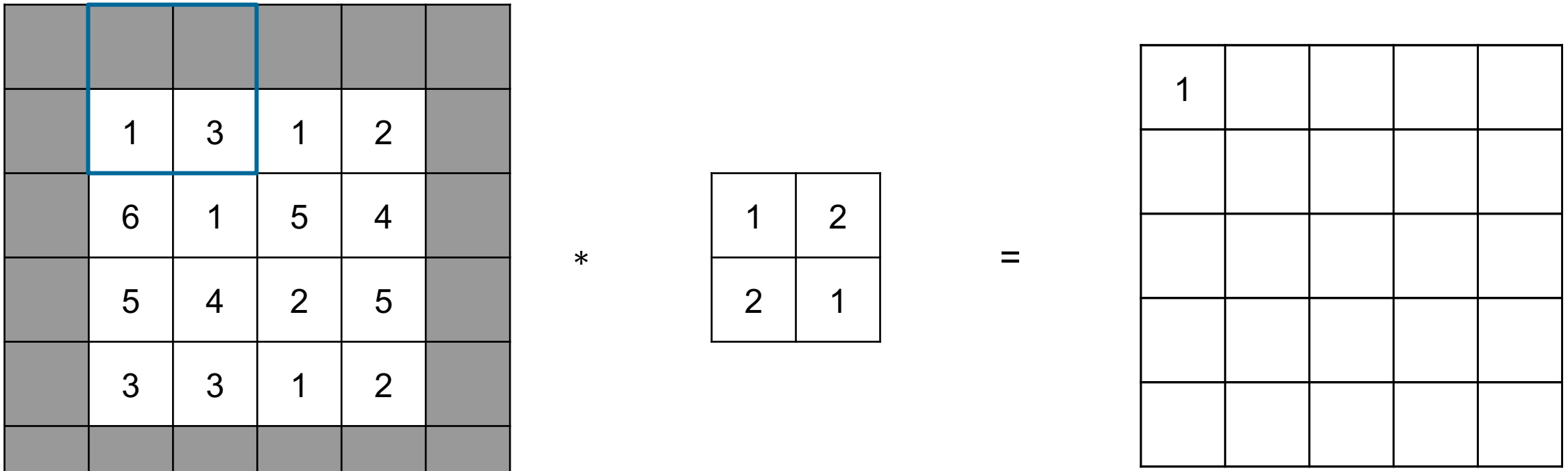
BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding



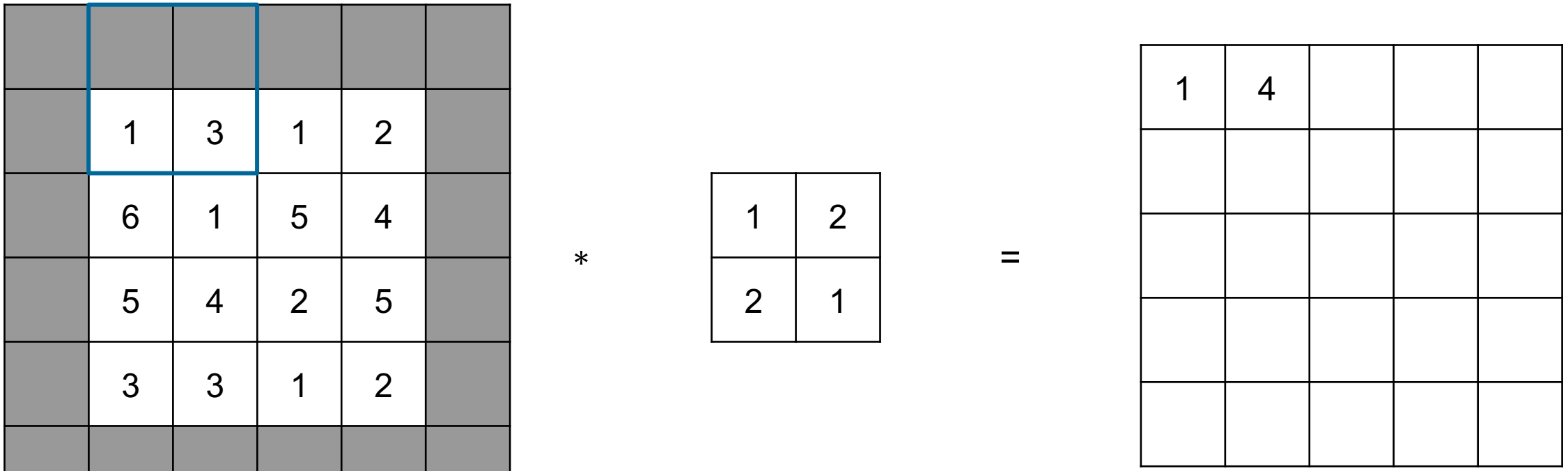
BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding



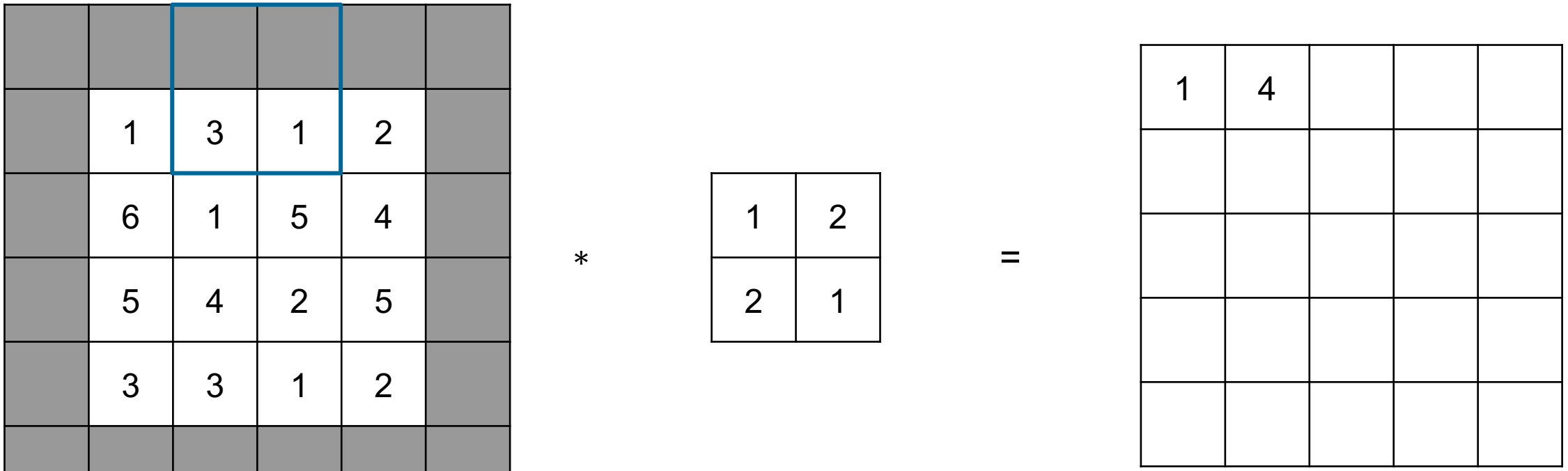
BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding



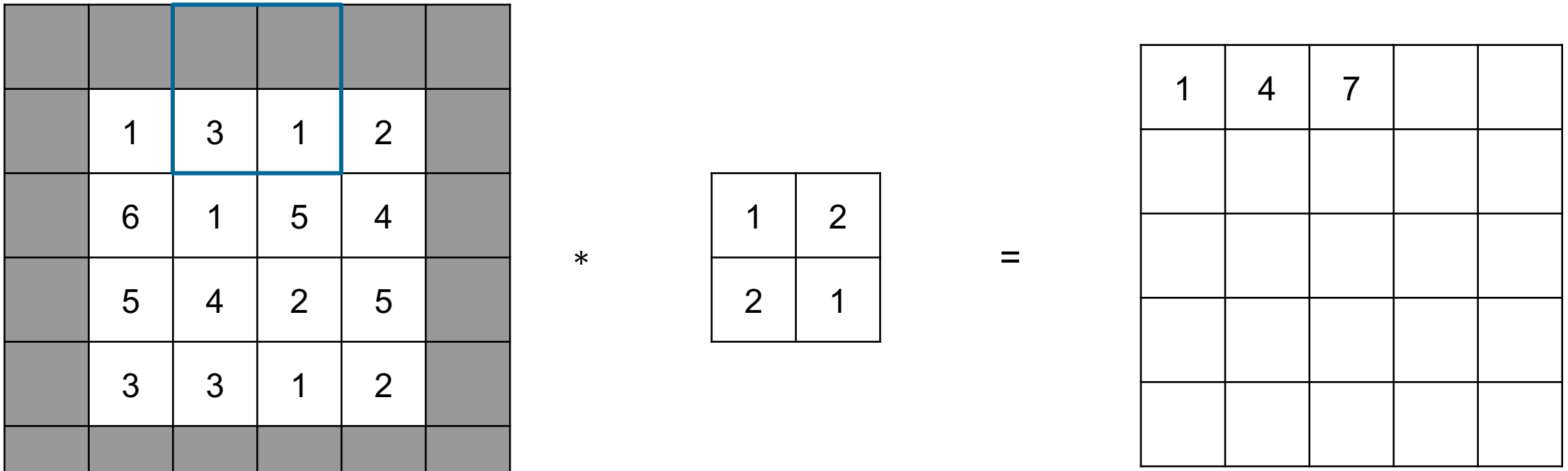
BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding



BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding



BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding

	1	3	1	2	
	6	1	5	4	
	5	4	2	5	
	3	3	1	2	

*

1	2
2	1

=

1	4	7	4	4
8	20	12	19	10
17	22	21	22	14
13	22	15	16	9
6	9	5	5	2

The convolution operator with padding

	1	3	1	2	
	6	1	5	4	
	5	4	2	5	
	3	3	1	2	

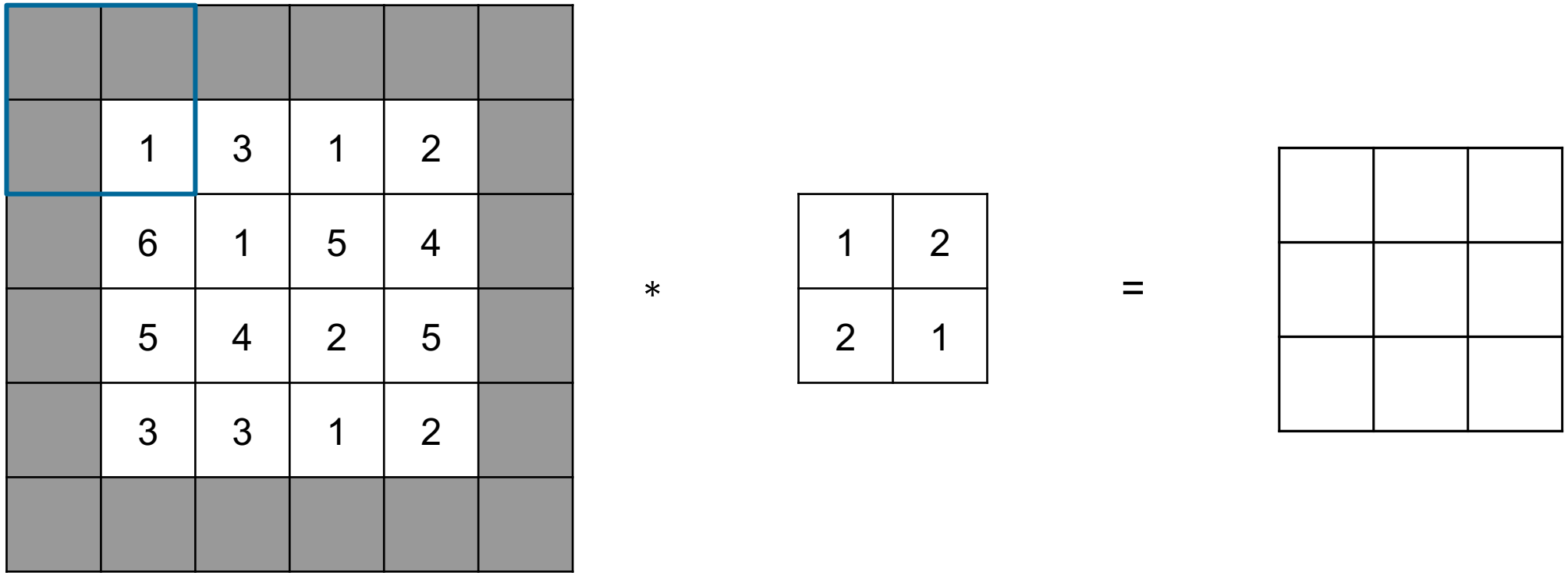
*

1	2
2	1

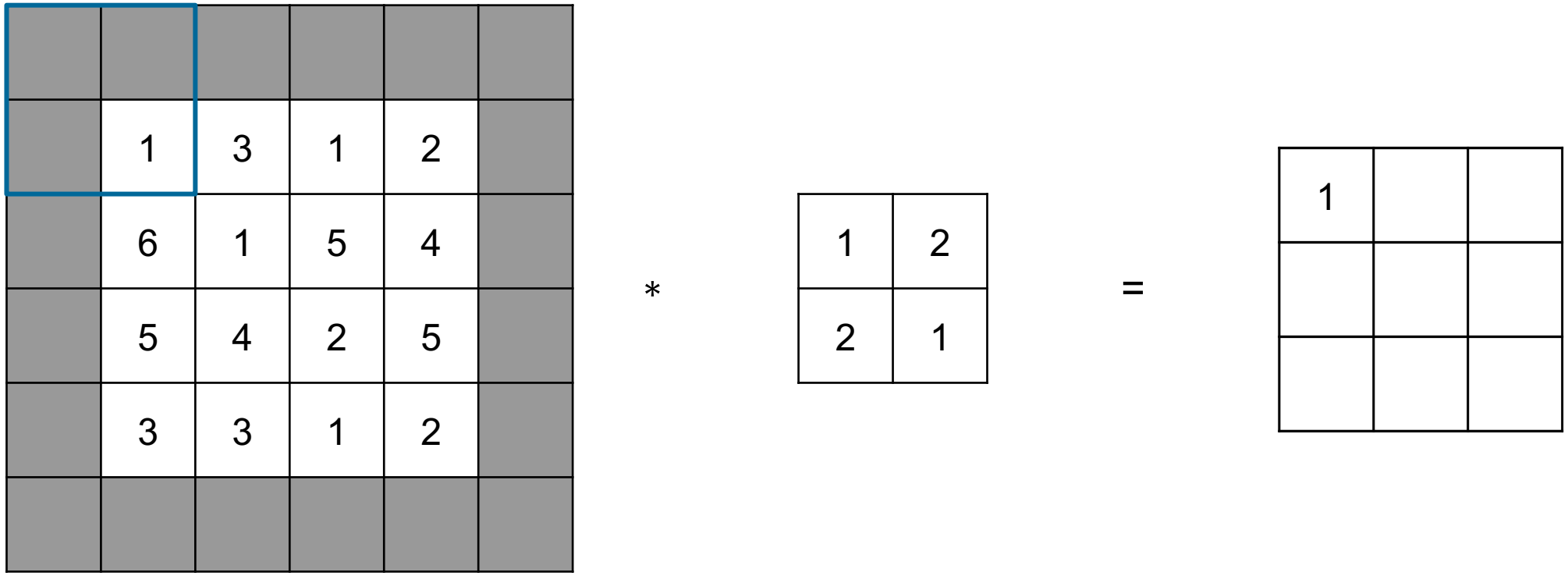
=

1	4	7	4	4
8	20	12	19	10
17	22	21	22	14
13	22	15	16	9
6	9	5	5	2

The convolution operator with padding and stride

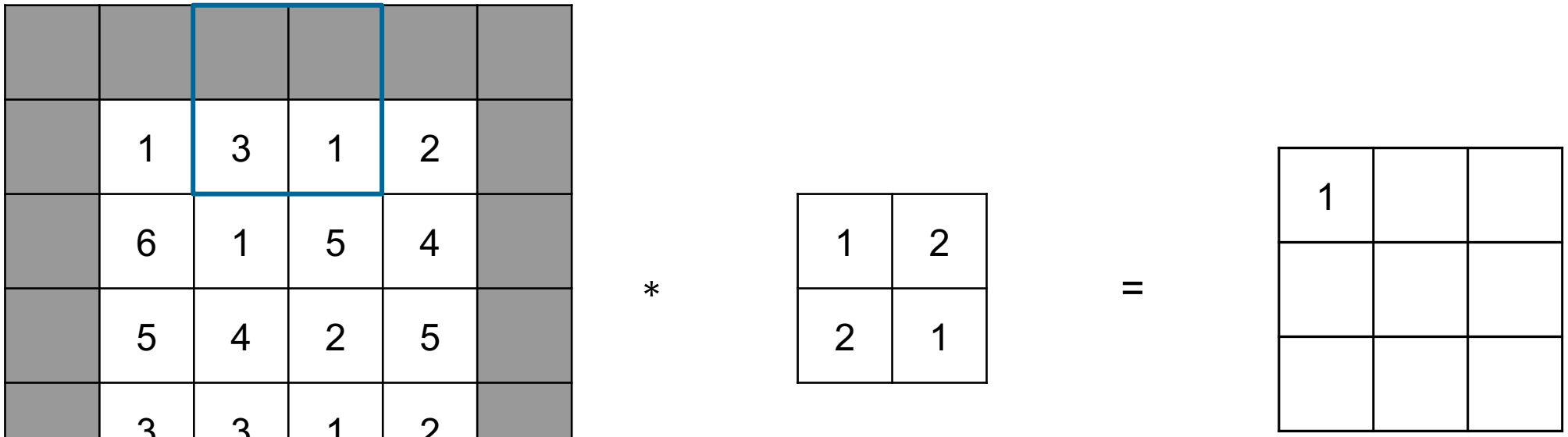


The convolution operator with padding and stride



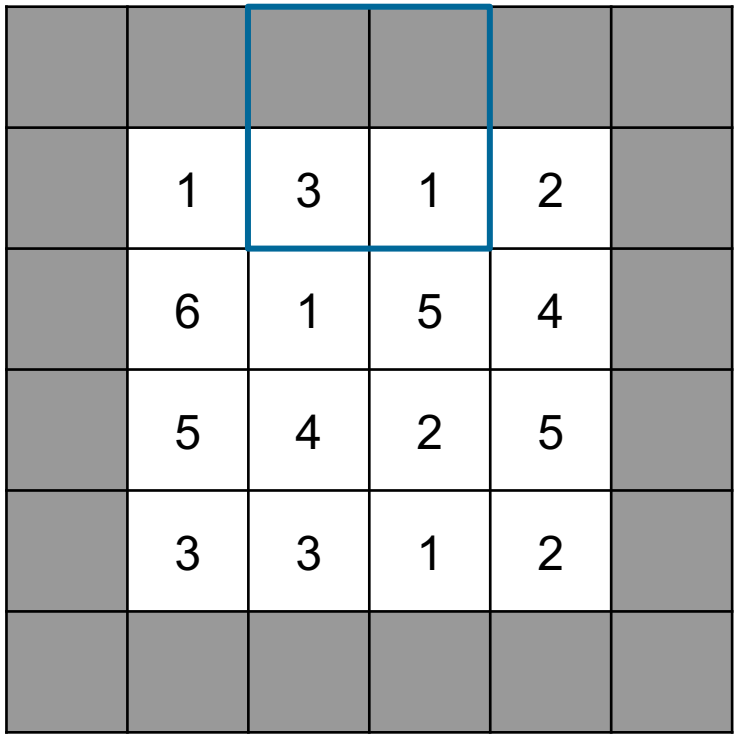
BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding and stride



BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding and stride



*

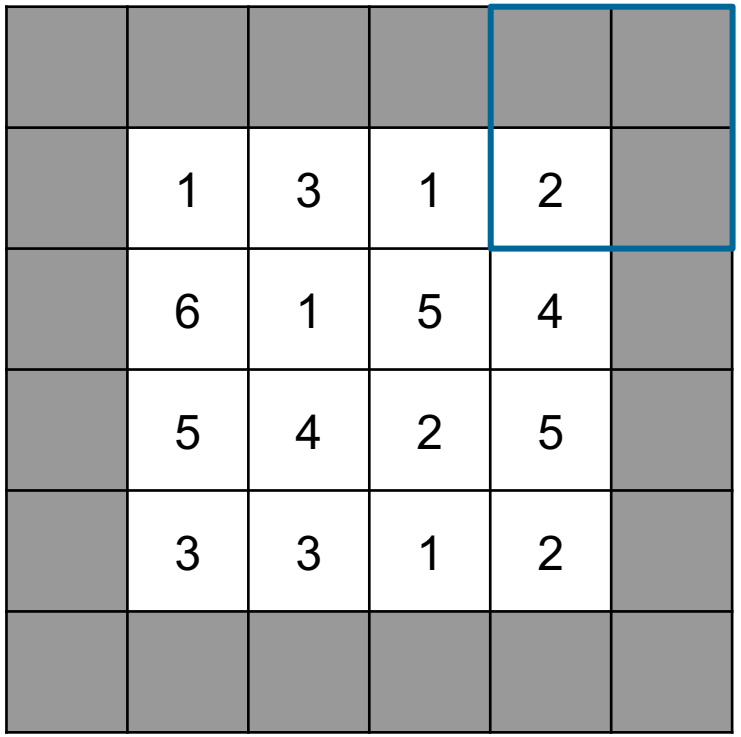
1	2
2	1

=

1	7	

BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding and stride



*

1	2
2	1

=

1	7	

BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding and stride

	1	3	1	2	
	6	1	5	4	
	5	4	2	5	
	3	3	1	2	

*

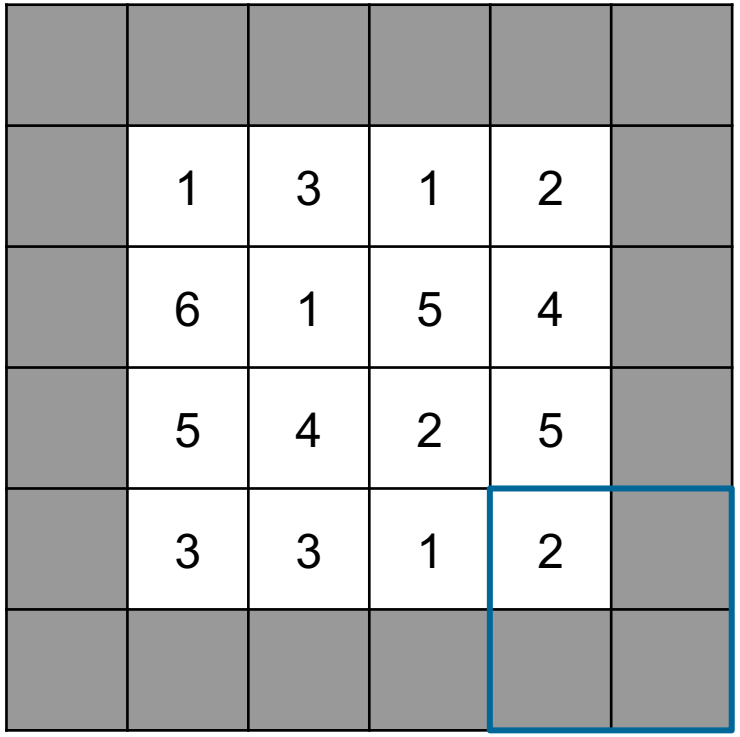
1	2
2	1

=

1	7	4

BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding and stride



*

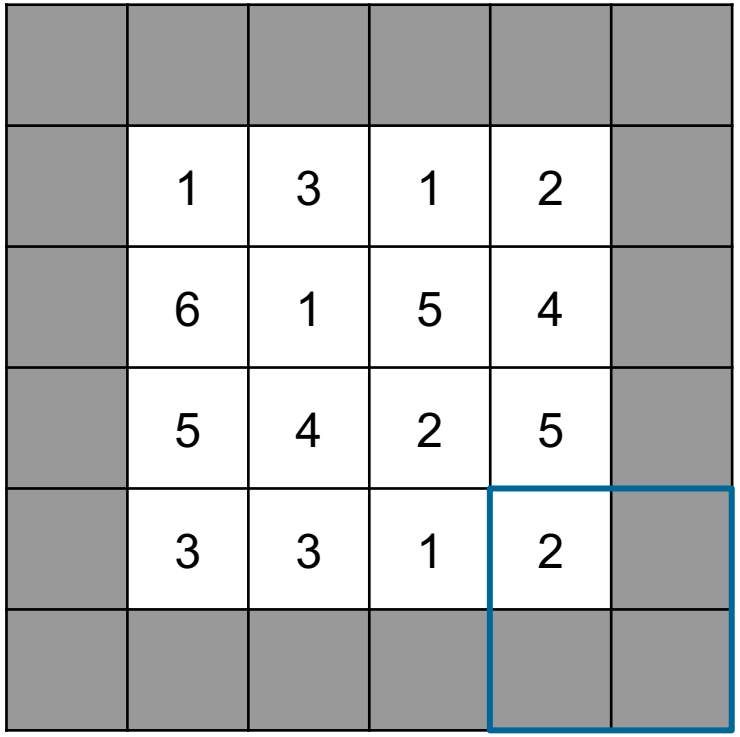
1	2
2	1

=

1	7	4
17	21	14
6	5	2

BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding and stride



*

1	2
2	1

=

1	7	4
17	21	14
6	5	2

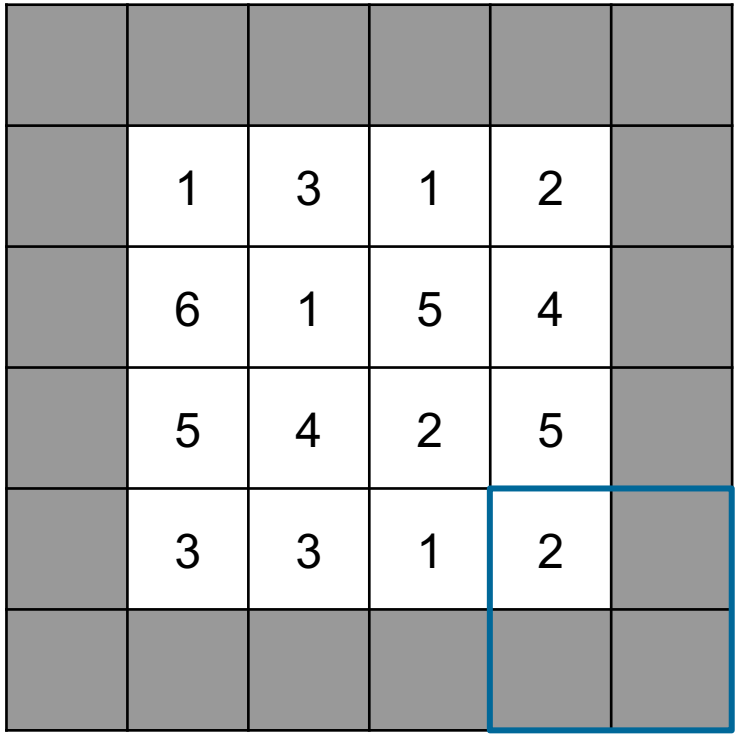
4x4
$(n_H^{[0]}, n_W^{[0]})$

2x2 with padding 1x1 and stride 2x2
(f_H, f_W) with (p_H, p_W) and (s_H, s_W)

3x3
$(n_H^{[1]}, n_W^{[1]})$

BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

The convolution operator with padding and stride



*

1	2
2	1

=

1	7	4
17	21	14
6	5	2

4x4
$(n_H^{[0]}, n_W^{[0]})$

2x2 with padding 1x1 and stride 2x2
(f_H, f_W) with (p_H, p_W) and (s_H, s_W)

3x3
$(n_H^{[1]}, n_W^{[1]})$

$$n_H^{[1]} = \left\lfloor \frac{n_H^{[0]} + 2p_H - f_H}{s_H} + 1 \right\rfloor \quad n_W^{[1]} = \left\lfloor \frac{n_W^{[0]} + 2p_W - f_W}{s_W} + 1 \right\rfloor$$

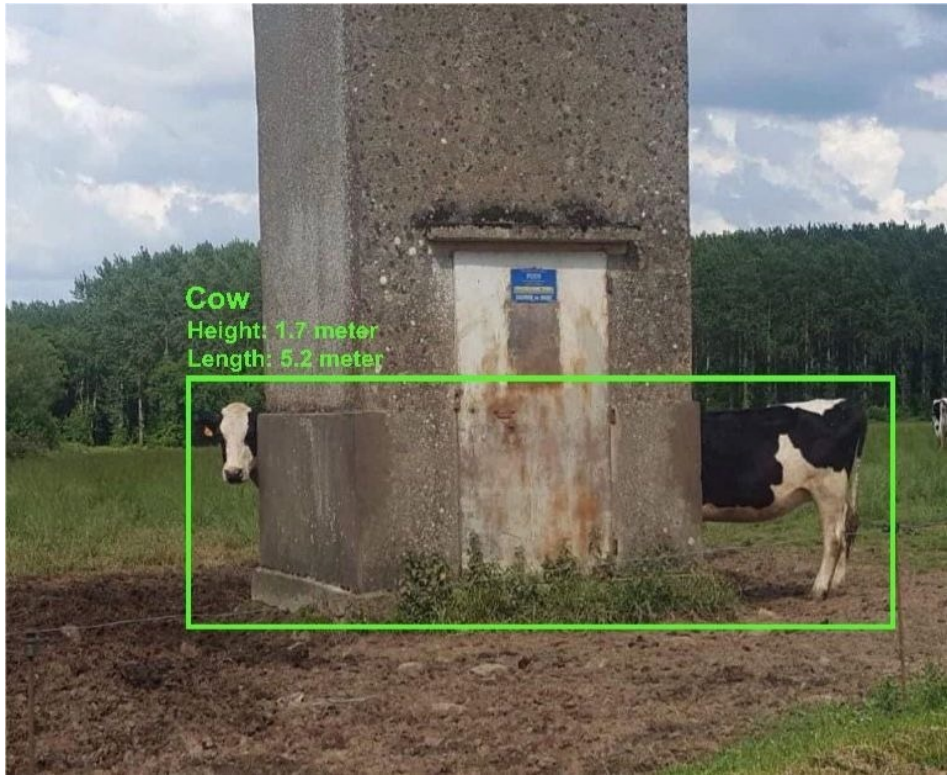
BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

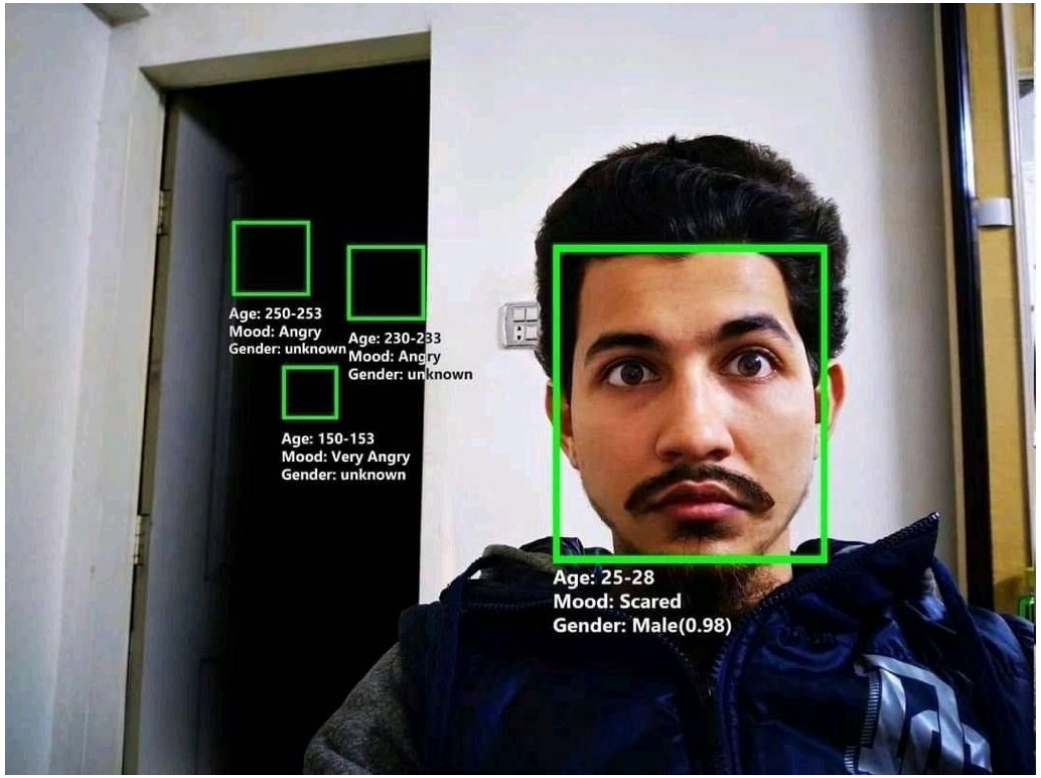
Image classification as a business

Use cases and challenges

Beyond image classification

Object detection





BAYES
BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON

Semantic segmentation

Source: Towards AI, 2021

Image generation

DALL·E History Collections

Edit the detailed description

Surprise me Upload

A lecture theatre full of students in business analytics

Generate




Source: OpenAI

BAYES BUSINESS SCHOOL
CITY, UNIVERSITY OF LONDON


See you next week!

Sources

- Alla & Kalyan Adari, 2019, Beginning Anomaly Detection Using Python-Based Deep Learning, Apress
- Bhaskhar, 2021, Introduction to Deep Learning: https://cs229.stanford.edu/syllabus.html
- Chollet, 2021, Deep Learning with Python, 2nd edition
- DeepLearning.AI, n.d.: deeplearning.ai
- Dieleman, 2020, Lecture 3: Convolutional Neural Networks: https://storage.googleapis.com/deepmind-media/UCLxDeepMind_2020/L3%20-%20UUCLxDeepMind%20DL2020.pdf
- Géron, 2019, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow
- Goodfellow, Bengio, Courville, 2016, The Deep Learning Book: http://www.deeplearningbook.org
- Liang, 2016, Introduction to Deep Learning: https://www.cs.princeton.edu/courses/archive/spring16/cos495/
- Lin et al., 2014, Microsoft COCO: Common Objects in Context: https://arxiv.org/pdf/1405.0312.pdf
- OpenAI et al., n.d., DALL-E 2: https://openai.com/dall-e-2/
- Towards AI, 2021, Semantic Segmentation: A Complete Guide: https://towardsai.net/p/l/machine-learning-7