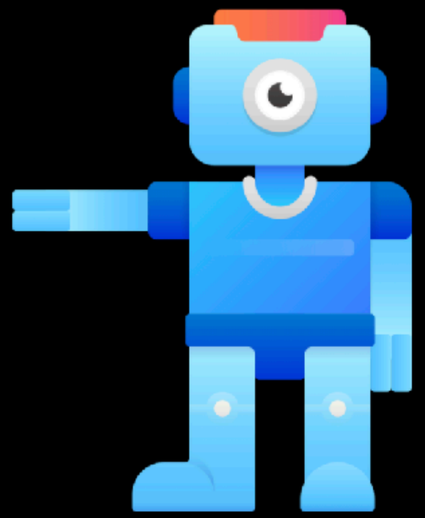


MODERN COMPUTER VISION

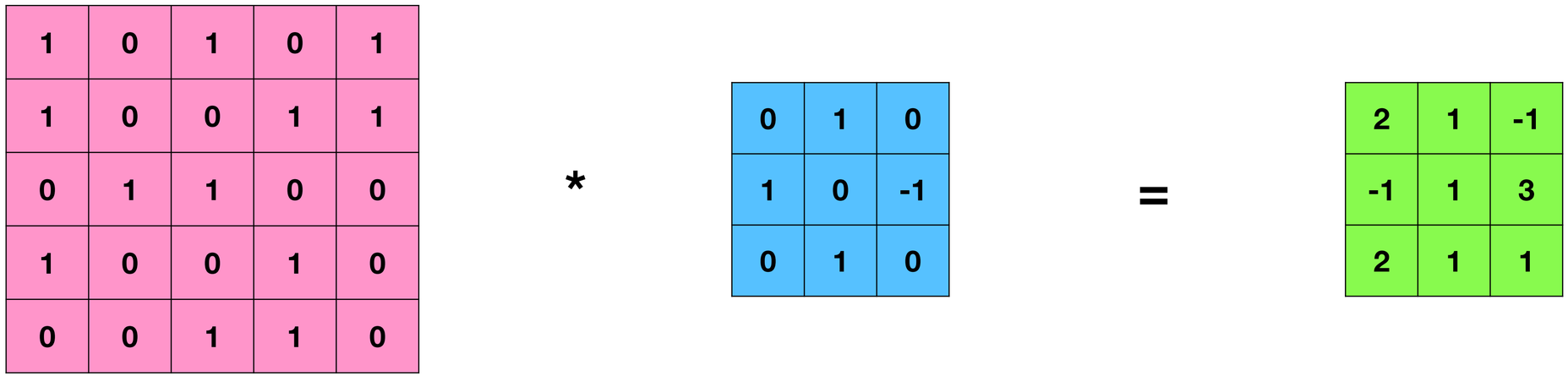
BY RAJEEV RATAN

Padding

Manipulating the input size

Padding

Notice How Conv Filters Produce an Output Smaller than the Input



5 x 5

$n \times n$

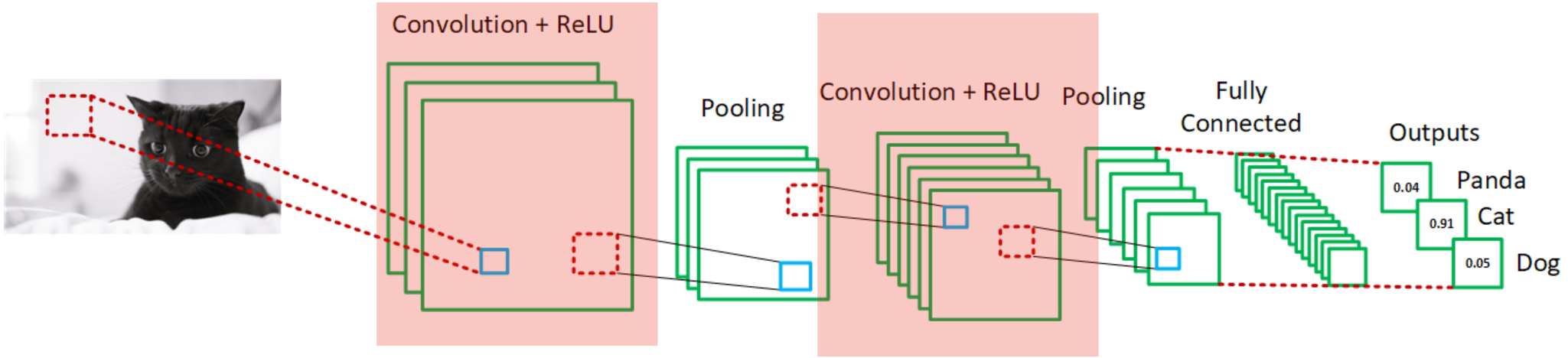
3 x 3

$f \times f$

3 x 3

$m \times m$

CNNs can have several sequences of Convolution layers



Consecutive Conv layers would keep shrinking the output

Can we preserve our image size?

We've added a 1 pixel pad of zeros (zero padding) around our input

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

*

0	1	0
1	0	-1
0	1	0

=

7 x 7

$n \times n$

3 x 3

$f \times f$

Padding

Let’s Perform our Convolution with the Padding

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

*

0	1	0
1	0	-1
0	1	0

=

7 x 7

$n \times n$

3 x 3

$f \times f$

Padding

Let’s Perform our Convolution with the Padding

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

*

0	1	0
1	0	-1
0	1	0

=

7 x 7

$n \times n$

3 x 3

$f \times f$

Padding

Let’s Perform our Convolution with the Padding

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

7 x 7

$n \times n$

*

0	1	0
1	0	-1
0	1	0

3 x 3

$f \times f$

=

Padding

Let’s Perform our Convolution with the Padding

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

7 x 7

$n \times n$

*

0	1	0
1	0	-1
0	1	0

3 x 3

$f \times f$

=

Padding

Let’s Perform our Convolution with the Padding

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

*

0	1	0
1	0	-1
0	1	0

=

7 x 7

$n \times n$

3 x 3

$f \times f$

Padding

Let’s Perform our Convolution with the Padding

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

7 x 7

$n \times n$

*

0	1	0
1	0	-1
0	1	0

3 x 3

$f \times f$

=

Padding

Let’s Perform our Convolution with the Padding

Feature Map Size = $n - f + 1 = m$

Feature Map Size = $7 - 3 + 1 = 5$

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

7 x 7

$n \times n$

*

0	1	0
1	0	-1
0	1	0

3 x 3

$f \times f$

=

2	1	-1	2	2
-1	1	3	2	1
2	1	1	1	2
1	1	1	0	2
2	0	2	3	1

5 x 5

$m \times m$

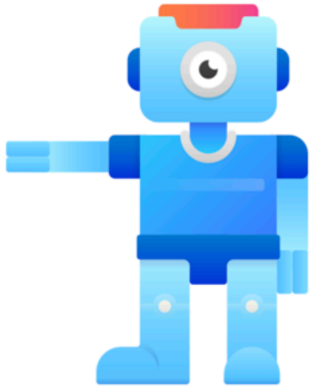
Why Use Padding?

- For very deep networks we don’t want to keep reducing the size
- Pixels at the edges contribute less to the output Feature Maps, thus we’re throwing away information from them

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

Without padding, our top left pixel is only touched by the Conv Filter once

Whereas, our centre pixel is passed over numerous times



MODERN COMPUTER VISION

BY RAJEEV RATAN

Next...

Stride