

Redes Neurais e Aprendizagem Profunda

REDES NEURAIS CONVOLUCIONAIS PASSO / PREENCHIMENTO / AGRUPAMENTO

Zenilton K. G. Patrocínio Jr

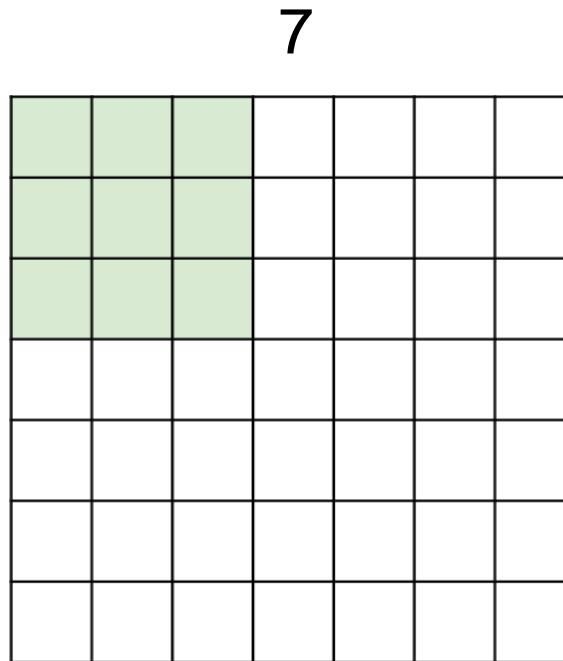
zenilton@pucminas.br

Camada Convolucional



Camada Convolucional – Tamanho do Passo

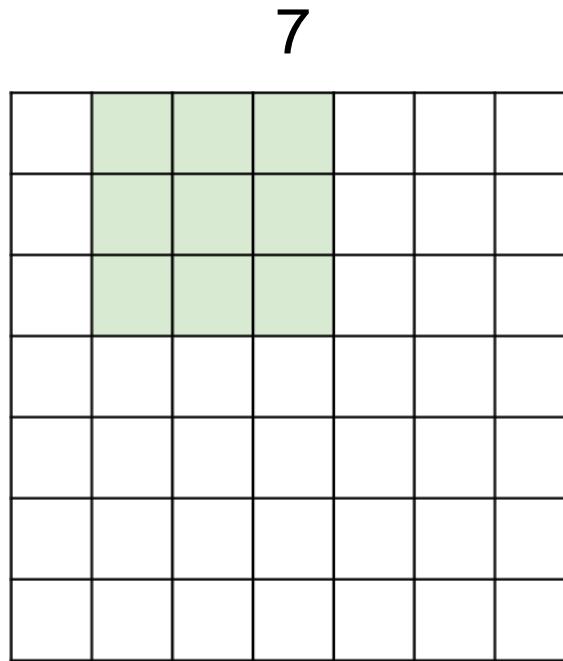
Um olhar mais atento às dimensões espaciais:



Entrada 7×7 (espacialmente)
Assumindo um filtro 3×3 com **passo = 1**

Camada Convolucional – Tamanho do Passo

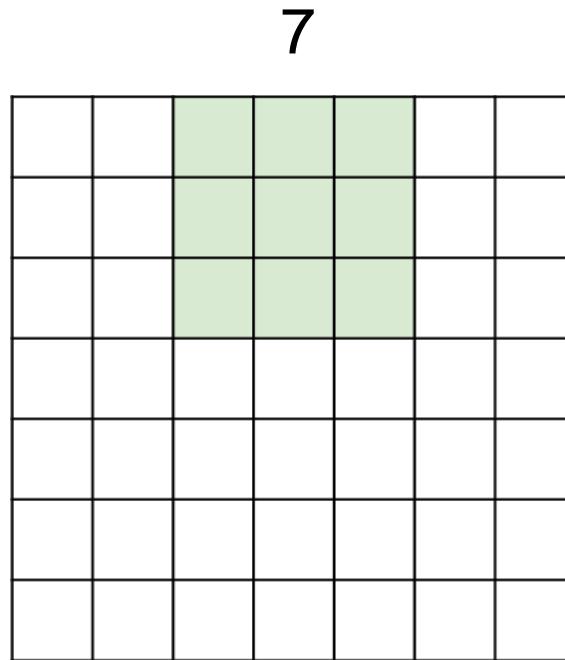
Um olhar mais atento às dimensões espaciais:



Entrada 7×7 (espacialmente)
Assumindo um filtro 3×3 com **passo = 1**

Camada Convolucional – Tamanho do Passo

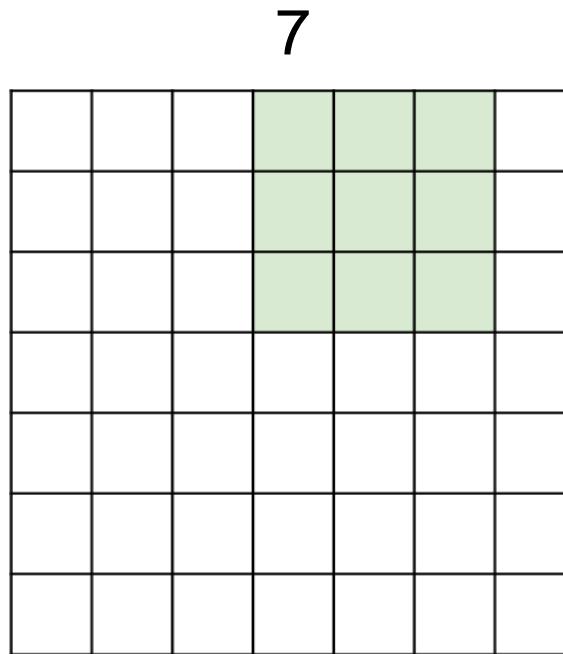
Um olhar mais atento às dimensões espaciais:



Entrada 7×7 (espacialmente)
Assumindo um filtro 3×3 com **passo = 1**

Camada Convolucional – Tamanho do Passo

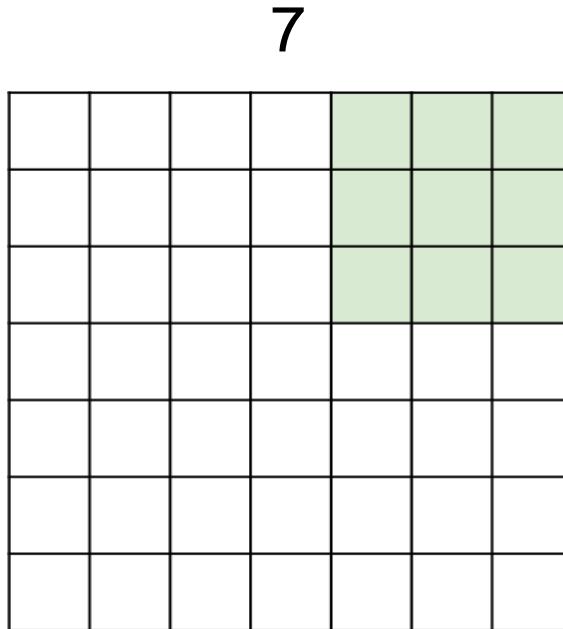
Um olhar mais atento às dimensões espaciais:



Entrada 7×7 (espacialmente)
Assumindo um filtro 3×3 com **passo = 1**

Camada Convolucional – Tamanho do Passo

Um olhar mais atento às dimensões espaciais:

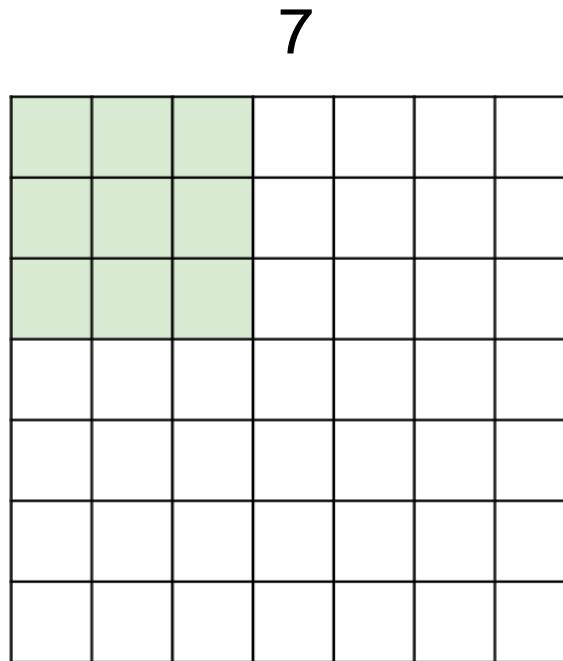


Entrada 7×7 (espacialmente)
Assumindo um filtro 3×3 com **passo = 1**

7 \Rightarrow **Saída 5×5 (para passo = 1)**

Camada Convolucional – Tamanho do Passo

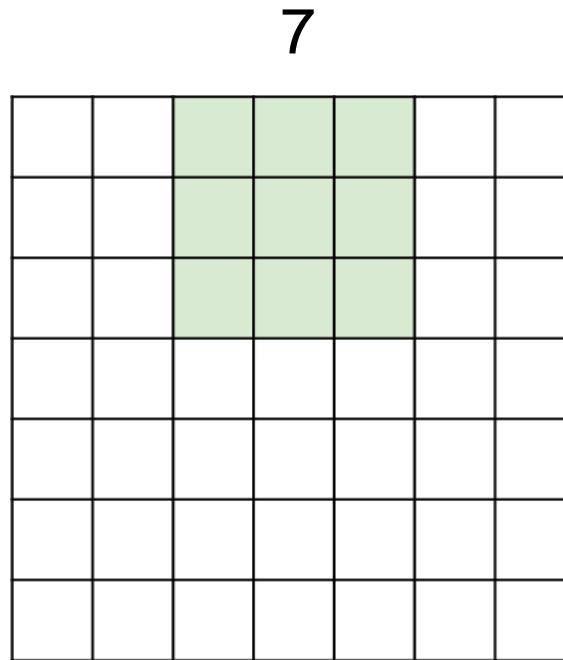
Um olhar mais atento às dimensões espaciais:



Entrada 7×7 (espacialmente)
Assumindo um filtro 3×3 com **passo = 2**

Camada Convolucional – Tamanho do Passo

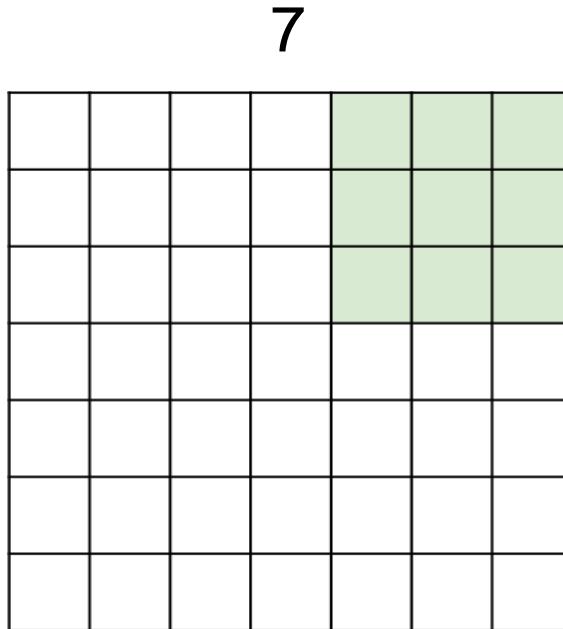
Um olhar mais atento às dimensões espaciais:



Entrada 7×7 (espacialmente)
Assumindo um filtro 3×3 com **passo = 2**

Camada Convolucional – Tamanho do Passo

Um olhar mais atento às dimensões espaciais:

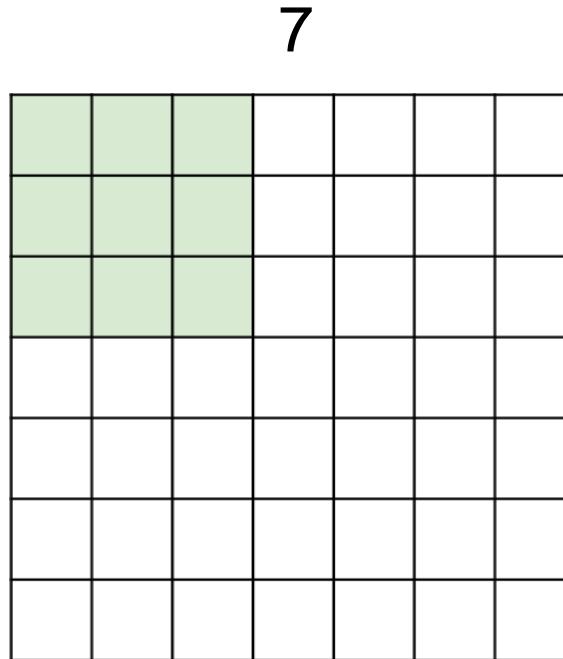


Entrada 7×7 (espacialmente)
Assumindo um filtro 3×3 com **passo = 2**

7 \Rightarrow **Saída 3×3 (para passo = 2)**

Camada Convolucional – Tamanho do Passo

Um olhar mais atento às dimensões espaciais:

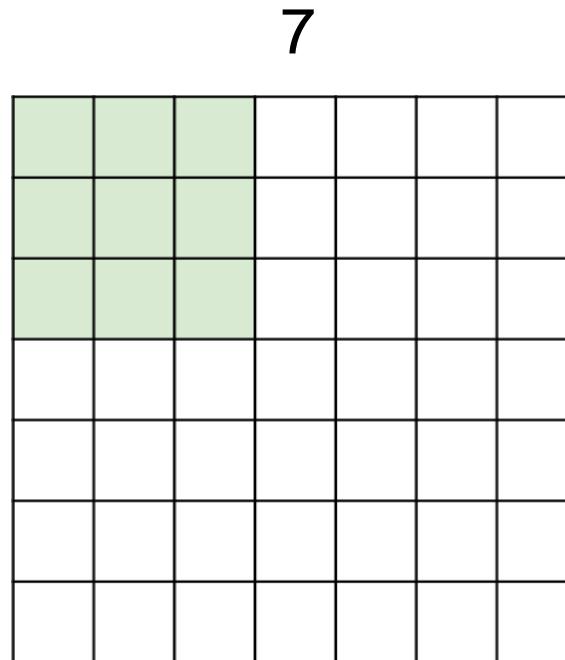


Entrada 7×7 (espacialmente)
Assumindo um filtro 3×3 com **passo = 3**

7 **Qual é a saída?**

Camada Convolucional – Tamanho do Passo

Um olhar mais atento às dimensões espaciais:



7

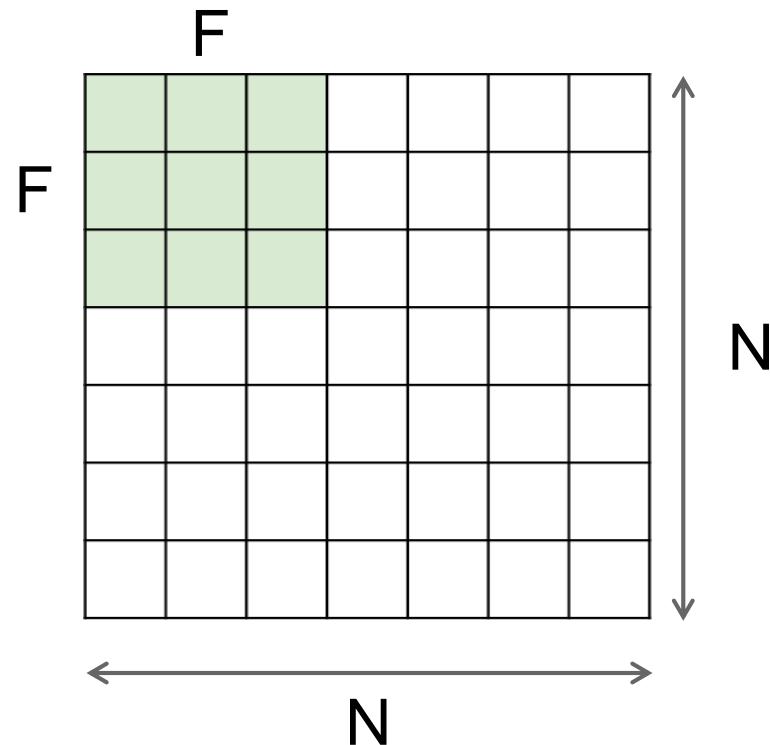
7

Entrada 7×7 (espacialmente)
Assumindo um filtro 3×3 com **passo = 3**

Não se encaixa adequadamente!

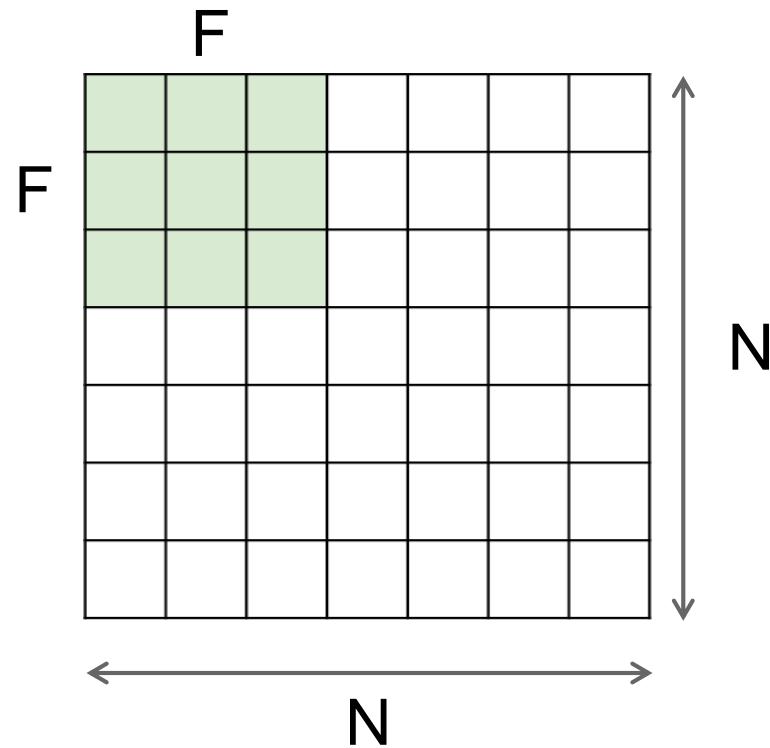
**Não é possível aplicar o filtro 3×3 na
entrada 7×7 com o passo 3**

Camada Convolucional – Tamanho do Passo



Tamanho da saída:
 $(N - F) / \text{passo} + 1$

Camada Convolucional – Tamanho do Passo

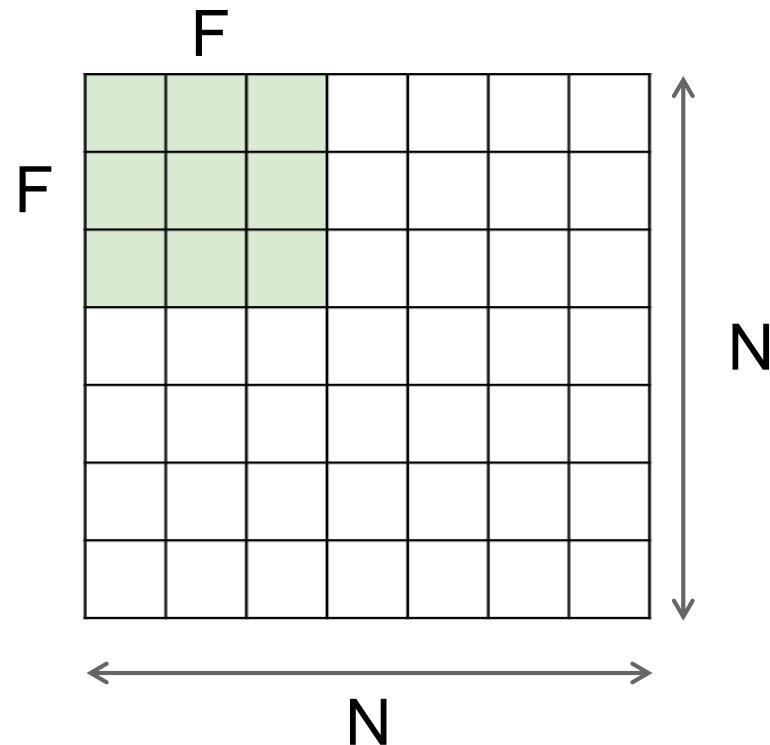


Tamanho da saída:

$$(N - F) / \text{passo} + 1$$

p.ex. $N = 7$, $F = 3$:

Camada Convolucional – Tamanho do Passo



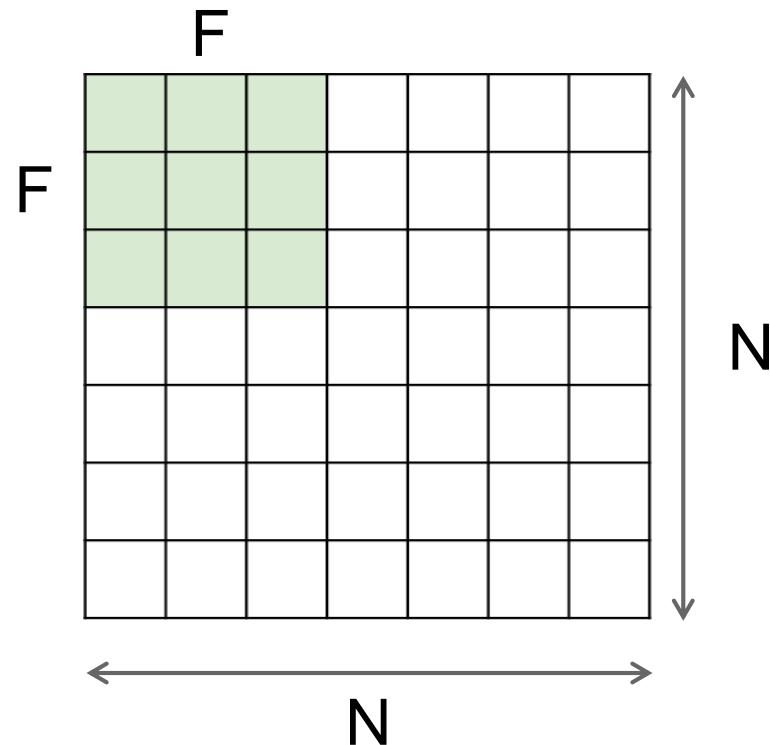
Tamanho da saída:

$$(N - F) / \text{passo} + 1$$

p.ex. $N = 7$, $F = 3$:

- passo 1 $\Rightarrow (7 - 3)/1 + 1 = 5$

Camada Convolucional – Tamanho do Passo



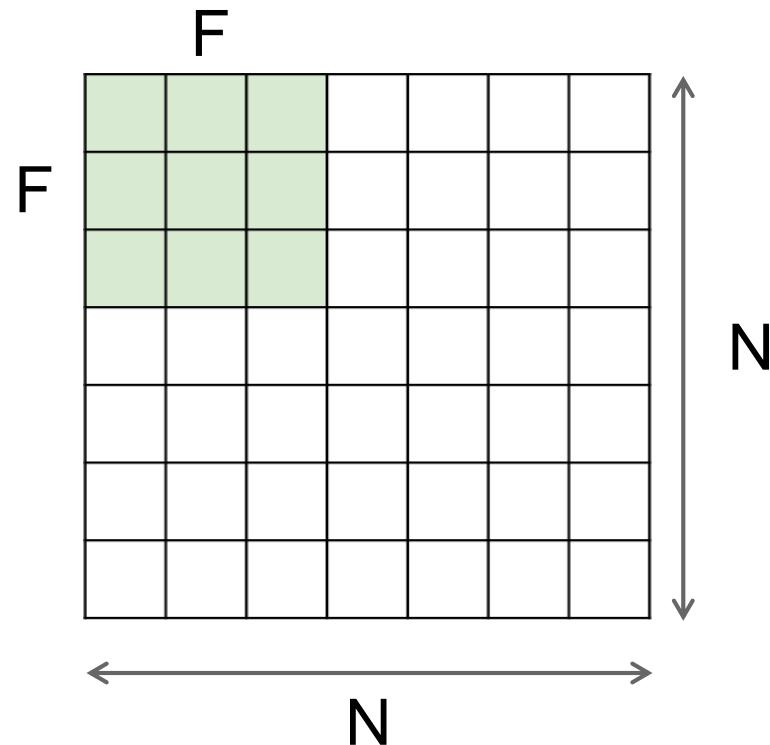
Tamanho da saída:

$$(N - F) / \text{passo} + 1$$

p.ex. $N = 7$, $F = 3$:

- passo 1 $\Rightarrow (7 - 3)/1 + 1 = 5$
- passo 2 $\Rightarrow (7 - 3)/2 + 1 = 3$

Camada Convolucional – Tamanho do Passo



Tamanho da saída:

$$(N - F) / \text{passo} + 1$$

p.ex. $N = 7$, $F = 3$:

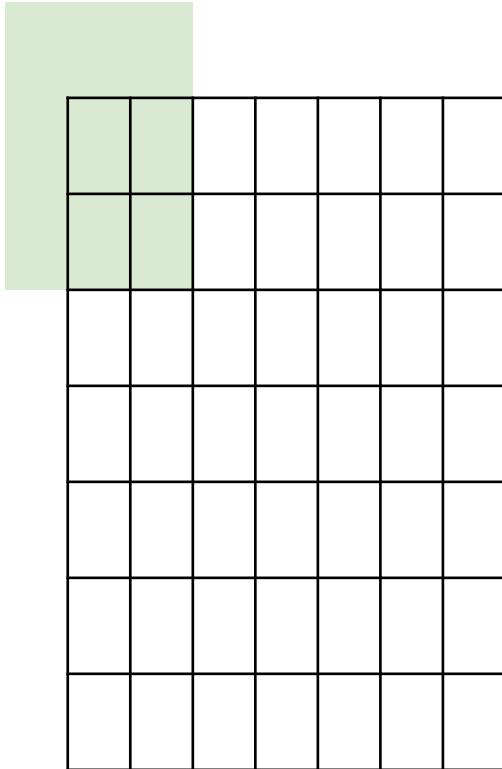
- passo 1 $\Rightarrow (7 - 3)/1 + 1 = 5$

- passo 2 $\Rightarrow (7 - 3)/2 + 1 = 3$

- passo 3 $\Rightarrow (7 - 3)/3 + 1 = \mathbf{2,33}$ ☹

Camada Convolucional – Preenchimento

Problema de cálculo nas bordas



Entrada 7×7

Filtro 3×3 aplicado com **passo = 1**

Camada Convolucional – Preenchimento

Problema de cálculo nas bordas

0	0	0	0	0	0			
0								
0								
0								
0								

Entrada 7×7

Filtro 3×3 aplicado com **passo = 1**

⇒ Na prática, é comum se acrescentar uma **borda de zeros** !

Camada Convolucional – Preenchimento

Problema de cálculo nas bordas

0	0	0	0	0	0			
0								
0								
0								
0								

Entrada 7×7

Filtro 3×3 aplicado com **passo = 1**

Preenchimento (*padding*) com borda de tamanho 1

Qual é a saída?

⇒ Na prática, é comum se acrescentar uma **borda de zeros** !

Camada Convolucional – Preenchimento

Problema de cálculo nas bordas

0	0	0	0	0	0			
0								
0								
0								
0								

Entrada 7×7

Filtro 3×3 aplicado com **passo = 1**

Preenchimento (*padding*) com borda de tamanho 1

Qual é a saída? \Rightarrow **Saída 7×7**

\Rightarrow Na prática, é comum se acrescentar uma **borda de zeros** !

Camada Convolucional – Preenchimento

Problema de cálculo nas bordas

0	0	0	0	0	0			
0								
0								
0								
0								

Entrada 7×7

Filtro 3×3 aplicado com **passo = 1**

Preenchimento (*padding*) com borda de tamanho 1

Qual é a saída? \Rightarrow **Saída 7×7**

É comum, camadas convolucionais com passo 1, filtros de tamanho F e preenchimento **pad = $(F - 1)/2$** zeros (para preservar as dimensões espaciais)

\Rightarrow Na prática, é comum se acrescentar uma **borda de zeros** !

Camada Convolucional – Preenchimento

Problema de cálculo nas bordas

0	0	0	0	0	0			
0								
0								
0								
0								

Entrada 7×7

Filtro 3×3 aplicado com **passo = 1**

Preenchimento (*padding*) com borda de tamanho 1

Qual é a saída? \Rightarrow **Saída 7×7**

É comum, camadas convolucionais com passo 1, filtros de tamanho F e preenchimento **pad = $(F - 1)/2$** zeros (para preservar as dimensões espaciais)

p.ex. $F = 3 \Rightarrow$ preenchimento com **1 zero**

$F = 5 \Rightarrow$ preenchimento com **2 zeros**

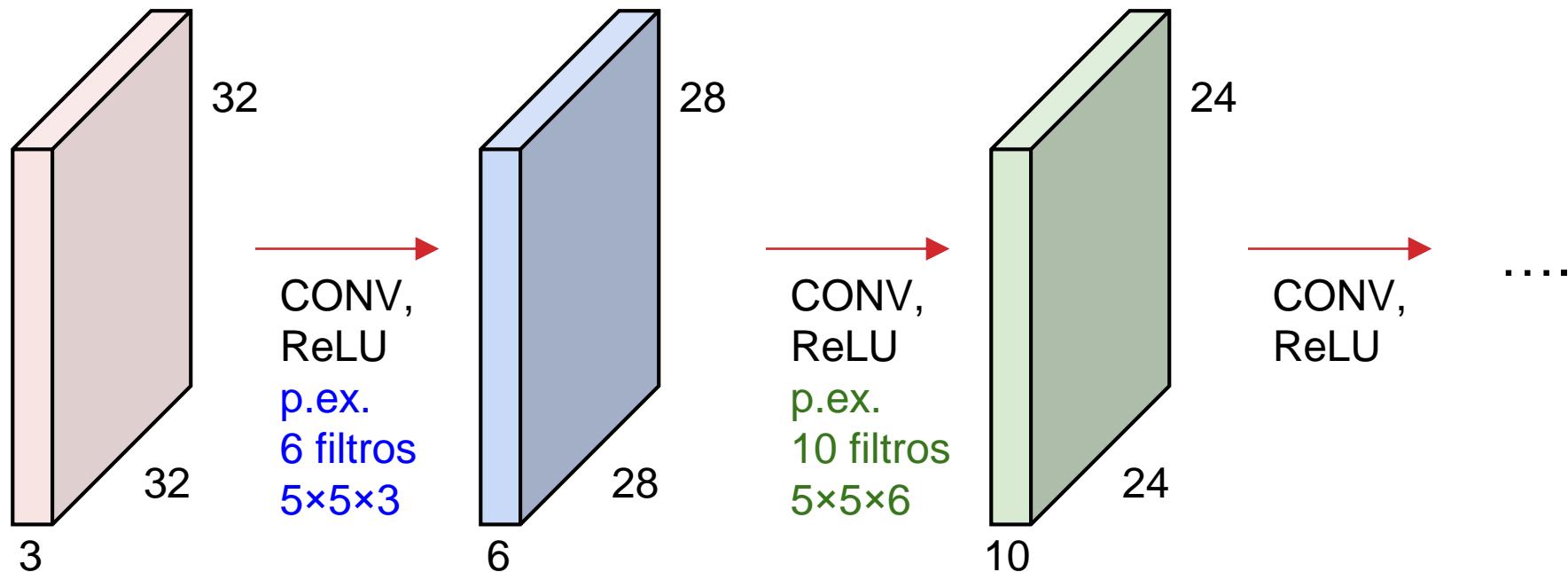
$F = 7 \Rightarrow$ preenchimento com **3 zeros**

\Rightarrow Na prática, é comum se acrescentar uma **borda de zeros** !

Camada Convolucional – Preenchimento

Lembre-se ...

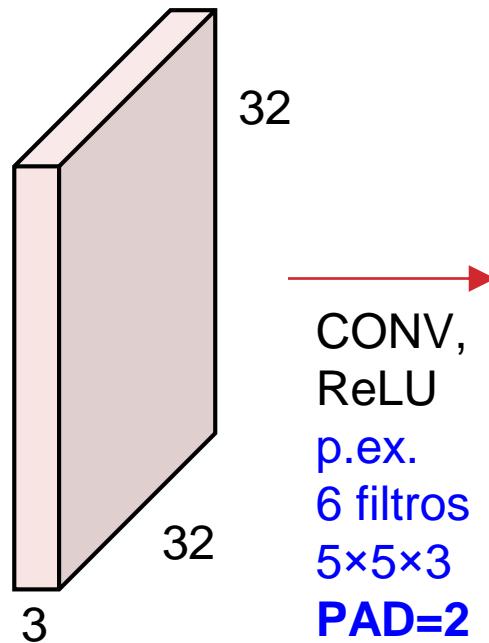
Por exemplo, uma entrada 32x32 convoluída repetidamente com filtros 5x5 reduz o volume espacialmente!
(32 → 28 → 24 ...) Encolher muito rápido não é bom e não funciona bem!



Camada Convolucional – Preenchimento

Lembre-se ...

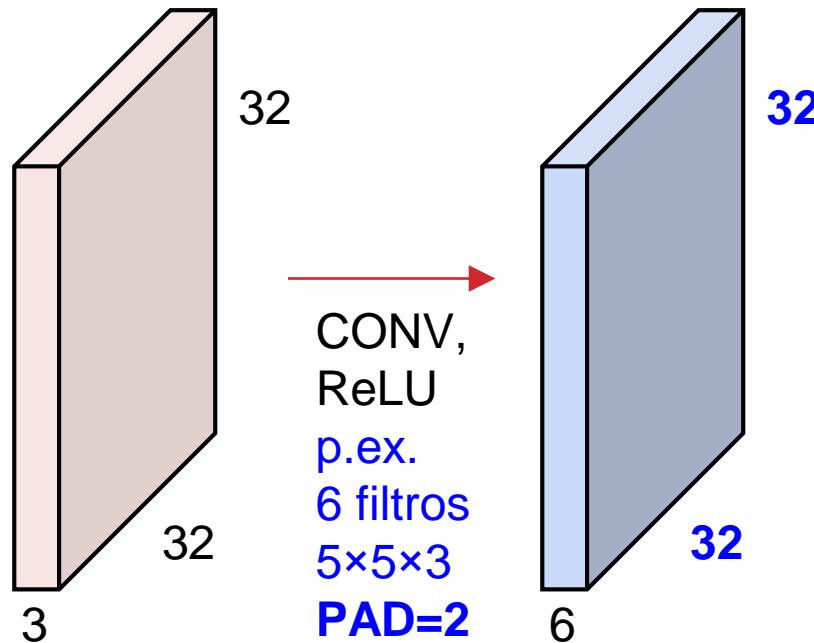
Por exemplo, uma entrada 32x32 convoluída repetidamente com filtros 5x5 reduz o volume espacialmente!
(32 → 28 → 24 ...) Encolher muito rápido não é bom e não funciona bem!



Camada Convolucional – Preenchimento

Lembre-se ...

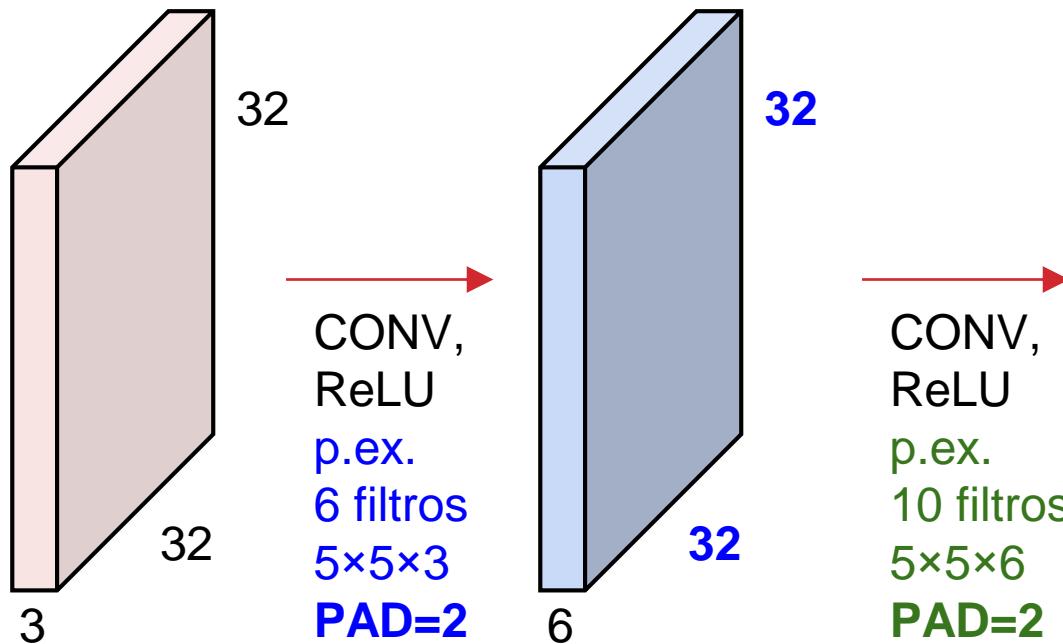
Por exemplo, uma entrada 32x32 convoluída repetidamente com filtros 5x5 reduz o volume espacialmente!
(32 → 28 → 24 ...) Encolher muito rápido não é bom e não funciona bem!



Camada Convolucional – Preenchimento

Lembre-se ...

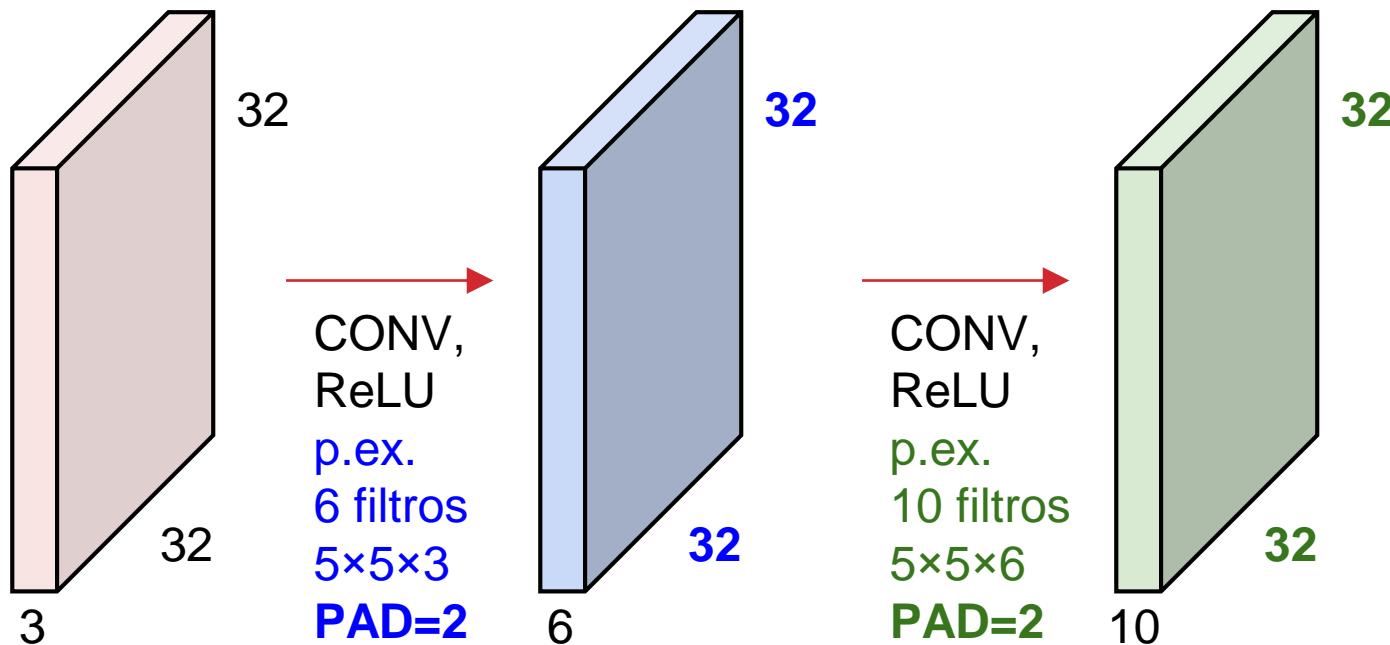
Por exemplo, uma entrada 32x32 convoluída repetidamente com filtros 5x5 reduz o volume espacialmente!
(32 → 28 → 24 ...) Encolher muito rápido não é bom e não funciona bem!



Camada Convolucional – Preenchimento

Lembre-se ...

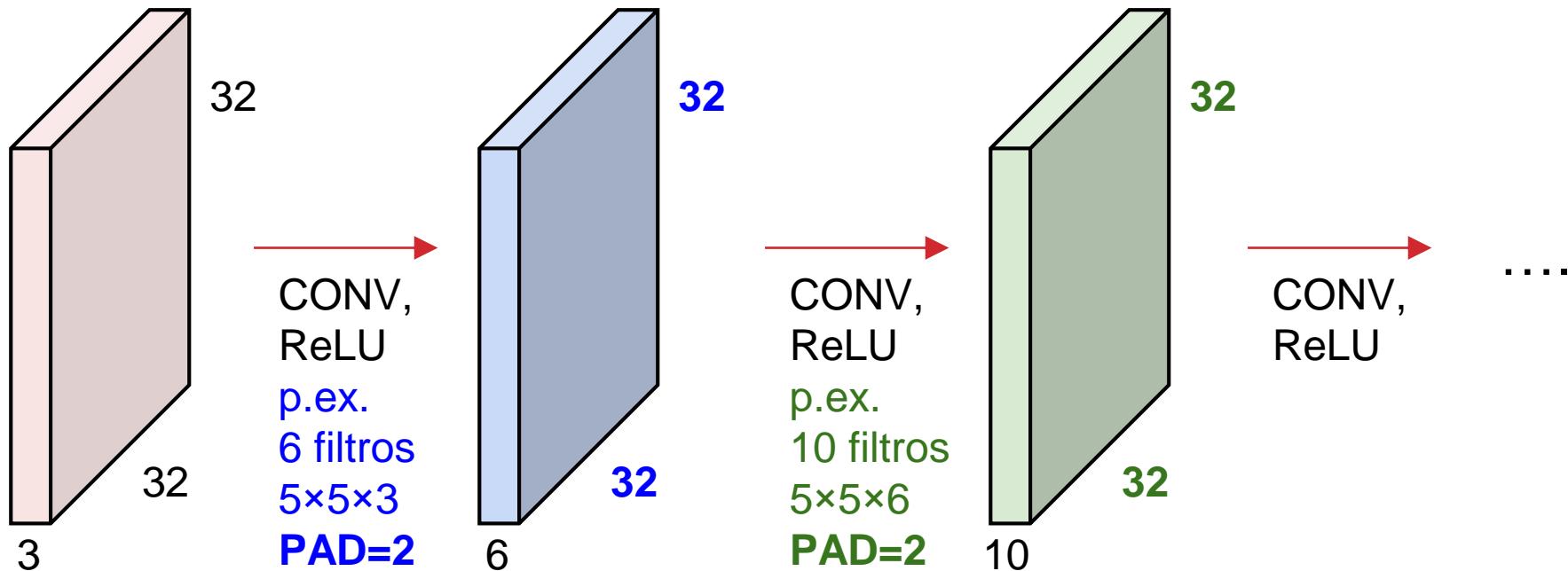
Por exemplo, uma entrada 32x32 convoluída repetidamente com filtros 5x5 reduz o volume espacialmente!
(32 → 28 → 24 ...) Encolher muito rápido não é bom e não funciona bem!



Camada Convolucional – Preenchimento

Lembre-se ...

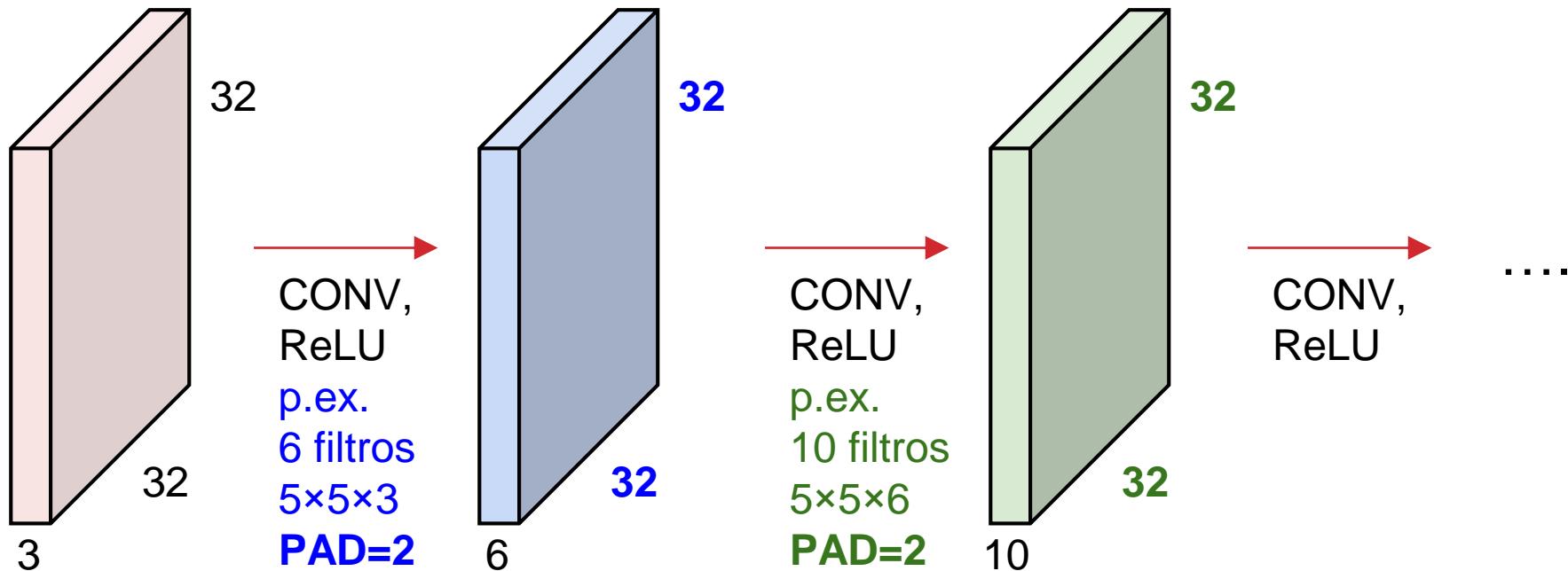
Por exemplo, uma entrada 32x32 convoluída repetidamente com filtros 5x5 reduz o volume espacialmente!
(32 → 28 → 24 ...) Encolher muito rápido não é bom e não funciona bem!



Camada Convolucional – Preenchimento

Lembre-se ...

Por exemplo, uma entrada 32×32 convoluída repetidamente com filtros 5×5 reduz o volume espacialmente!
($32 \rightarrow 28 \rightarrow 24 \dots$) Encolher muito rápido não é bom e não funciona bem!



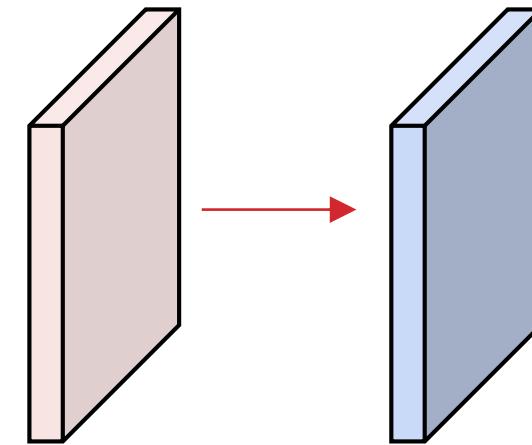
Preenchimento ajuda a minimizar o efeito da redução rápida de dimensionalidade espacial

Camada Convolucional – Detalhes

Volume de entrada: **32×32×3**

10 filtros 5×5 com passo = 1, preenchimento = 2

Tamanho do volume de saída: ?



Camada Convolucional – Detalhes

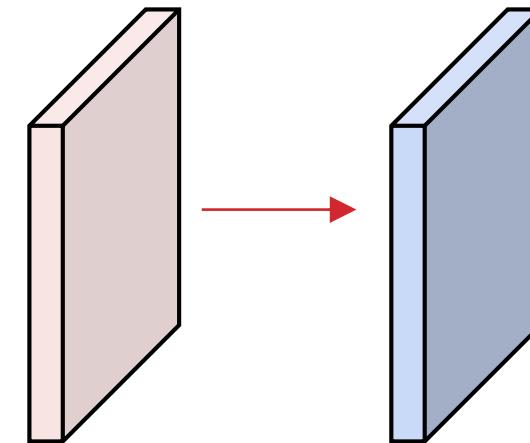
Volume de entrada: $32 \times 32 \times 3$

10 filtros 5×5 com passo = **1**, preenchimento = **2**

Tamanho do volume de saída:

$$(32 + 2 \times 2 - 5) / 1 + 1 = 32 \text{ espacialmente,}$$

Portanto o volume de saída é **$32 \times 32 \times 10$**

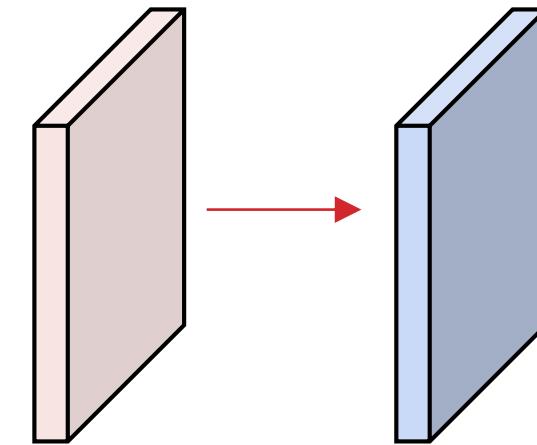


Camada Convolucional – Detalhes

Volume de entrada: **32×32×3**

10 filtros 5×5 com passo = 1, preenchimento = 2

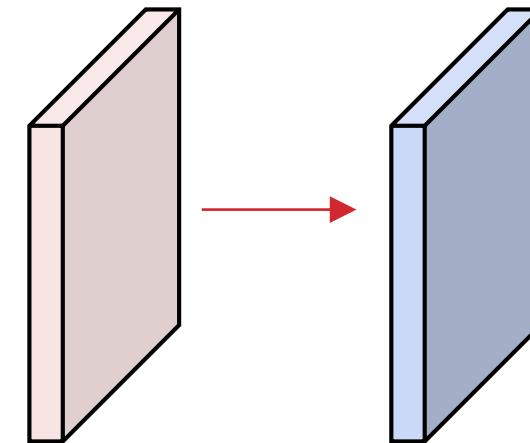
Número de parâmetros nessa camada: ?



Camada Convolucional – Detalhes

Volume de entrada: **32×32×3**

10 filtros **5×5** com passo = 1, preenchimento = 2



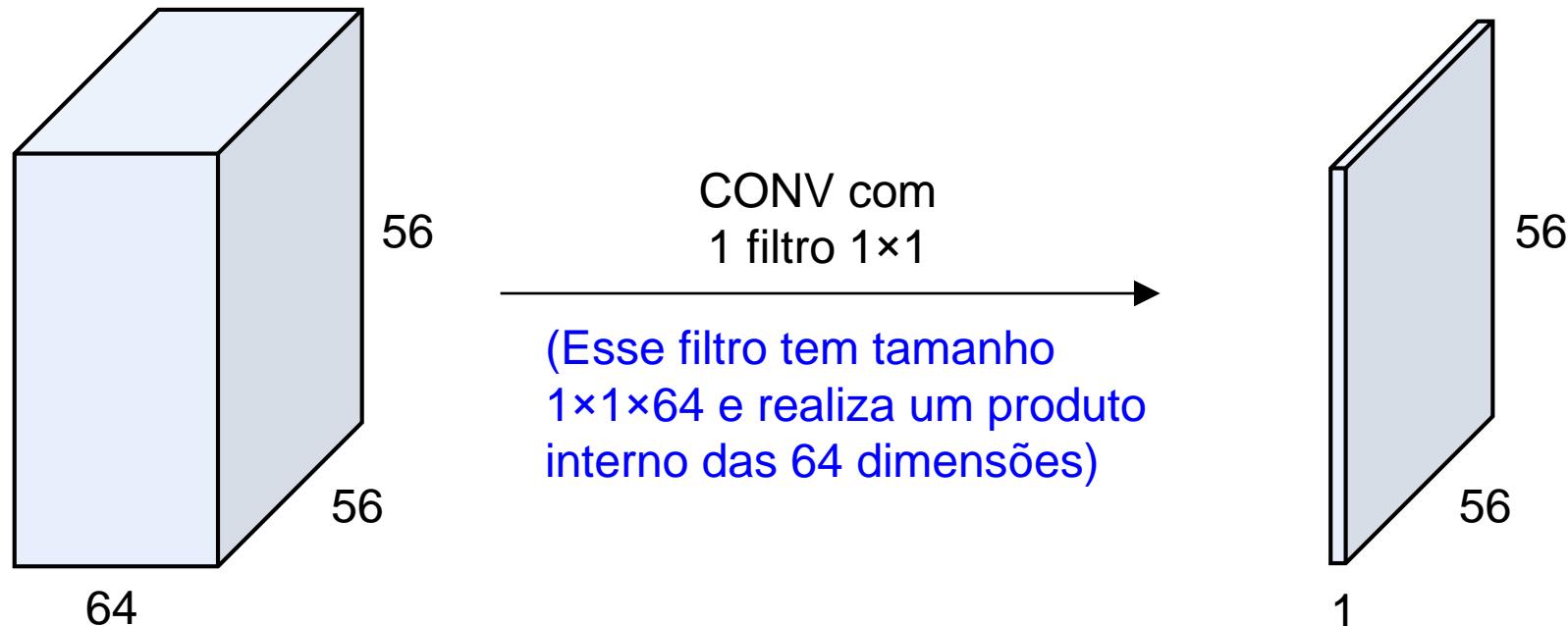
Número de parâmetros nessa camada:

Cada filtro possui **5×5×3 + 1 = 76** params (+1 para o viés)

Como são 10 filtros $\Rightarrow 76 \times 10 = 760$ params

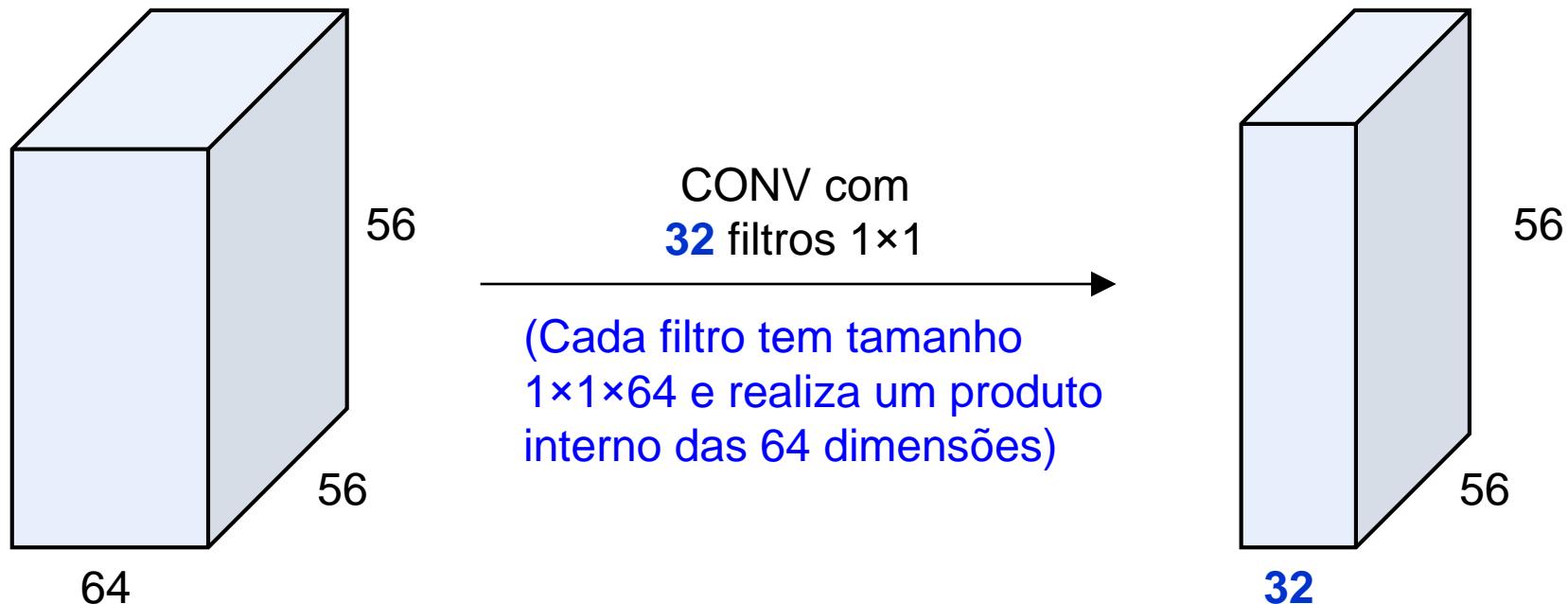
Camada Convolucional

OBS: Camadas convolucionais 1×1 também fazem sentido!

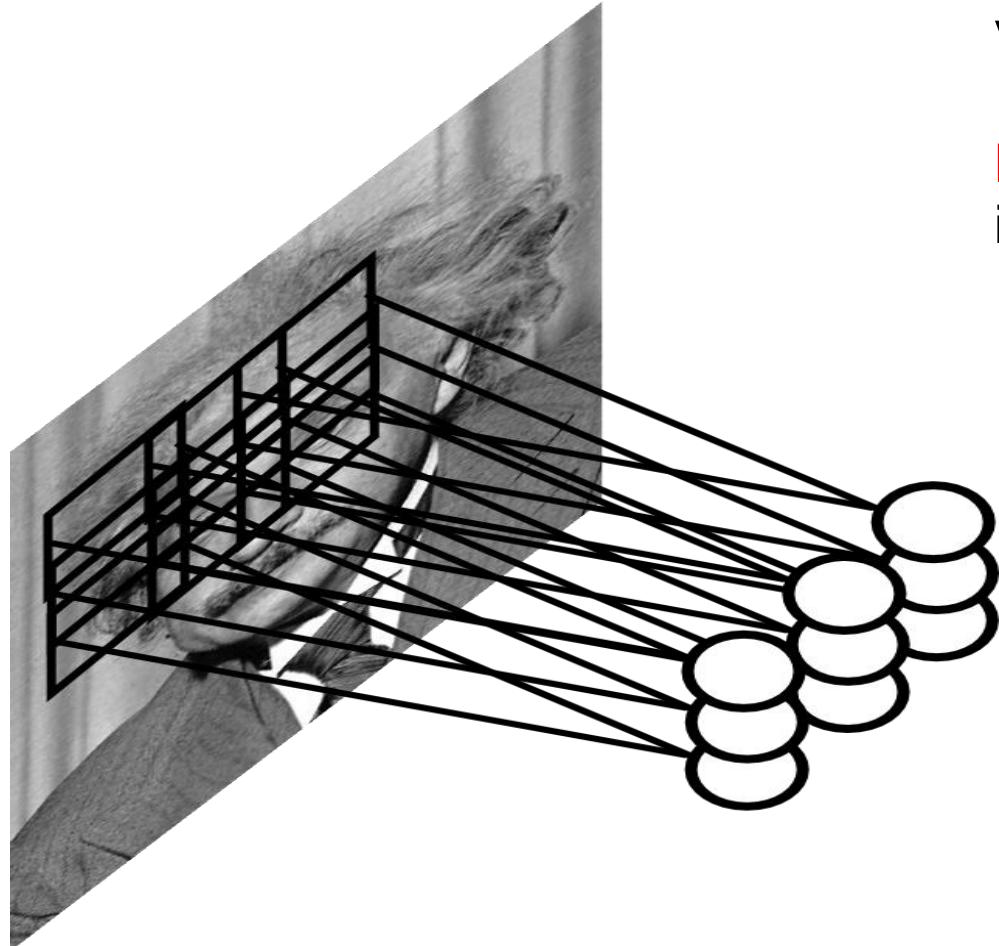


Camada Convolucional

OBS: Camadas convolucionais 1×1 também fazem sentido!



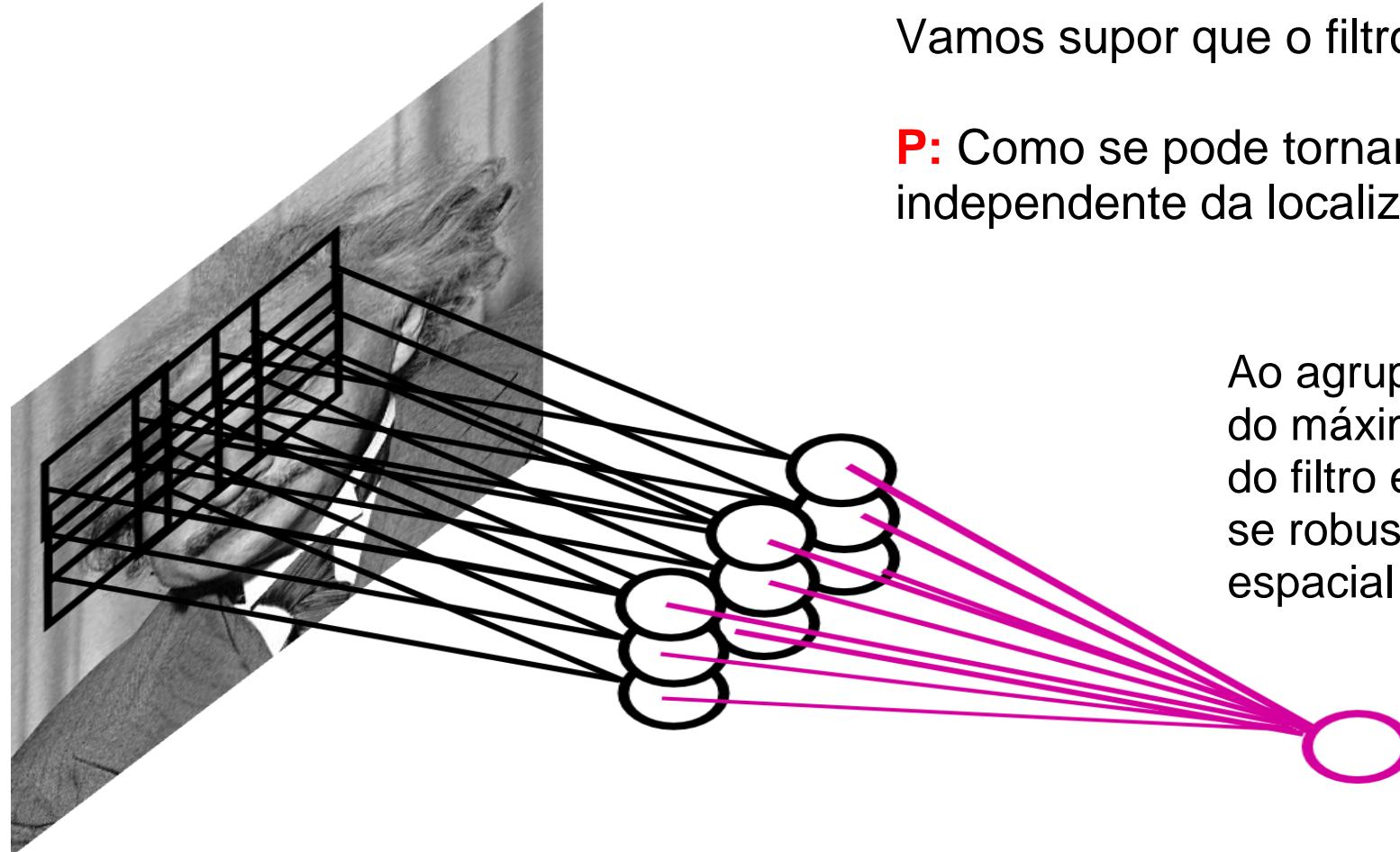
Camada de Agrupamento (*Pooling*)



Vamos supor que o filtro seja um detector de “olhos”

P: Como se pode tornar a detecção mais robusta e independente da localização exata do olho?

Camada de Agrupamento (*Pooling*)



Vamos supor que o filtro seja um detector de “olhos”

P: Como se pode tornar a detecção mais robusta e independente da localização exata do olho?

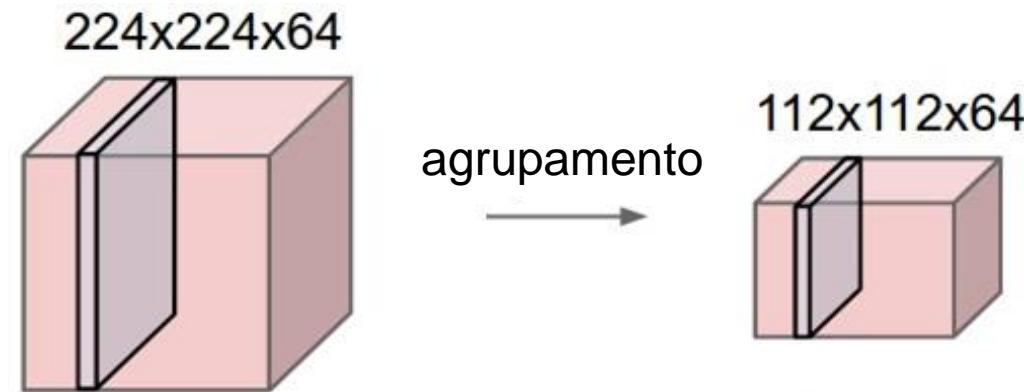
Ao agrupar (por exemplo, por meio do máximo ou média) as respostas do filtro em diferentes locais, ganha-se robustez em relação a localização espacial exata de uma característica

Camada de Agrupamento (*Pooling*)

- Torna as representações menores e mais gerenciáveis

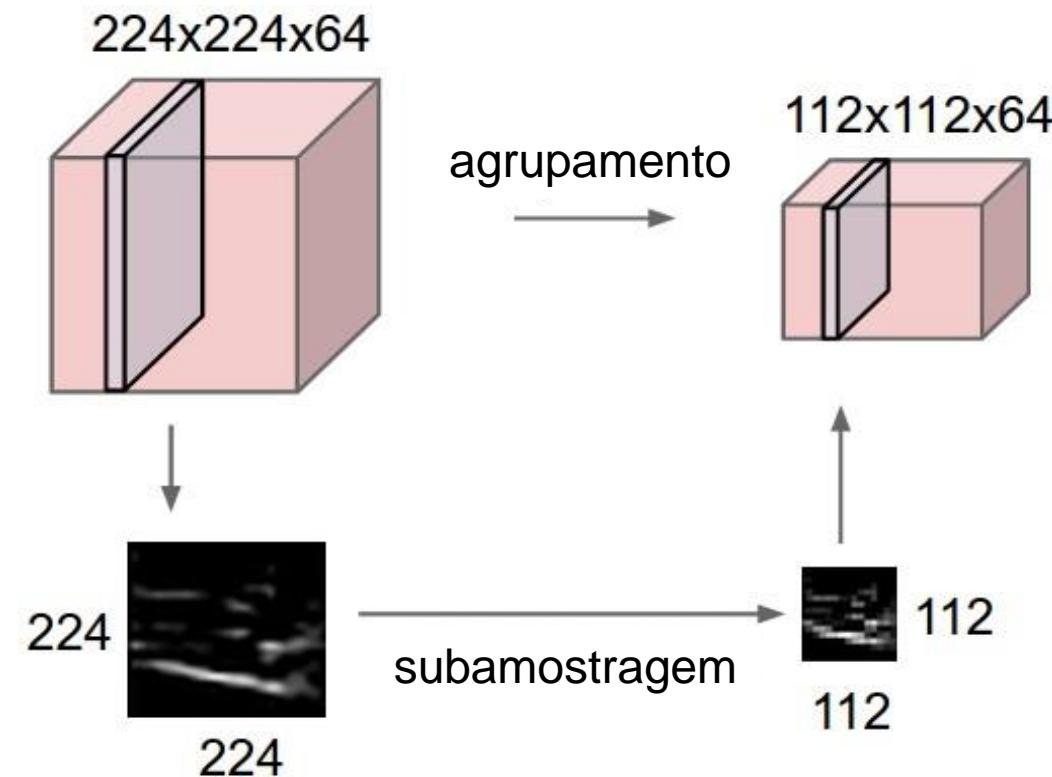
Camada de Agrupamento (*Pooling*)

- Torna as representações menores e mais gerenciáveis
- Opera sobre cada mapa de ativação de forma independente



Camada de Agrupamento (*Pooling*)

- Torna as representações menores e mais gerenciáveis
- Opera sobre cada mapa de ativação de forma independente



Camada de Agrupamento (*Pooling*)

Agrupamento utiliza uma operação para condensar as informações:

- Máximo
- Média
- L^2

Camada de Agrupamento (*Pooling*)

Agrupamento utiliza uma operação para condensar as informações:

- Máximo
- Média
- L^2

Exemplo de agrupamento pelo máximo (*max pooling*)

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

Um único mapa

Agrupamento pelo max
com filtros 2×2 e passo 2

The diagram illustrates the max pooling process. On the left is a 4x4 input grid labeled "Um único mapa". An arrow points from this grid to a 2x2 output grid on the right. The input grid has values: Row 1: 1, 1, 2, 4; Row 2: 5, 6, 7, 8; Row 3: 3, 2, 1, 0; Row 4: 1, 2, 3, 4. The output grid has four cells, each containing the maximum value from its corresponding 2x2 receptive field in the input grid. The output values are: Top-left cell: 5 (max of 1, 1, 5, 6); Top-right cell: 8 (max of 2, 4, 7, 8); Bottom-left cell: 3 (max of 3, 2, 1, 0); Bottom-right cell: 4 (max of 1, 2, 3, 4).

Camada de Agrupamento (*Pooling*)

Agrupamento utiliza uma operação para condensar as informações:

- Máximo
- Média
- L^2

Exemplo de agrupamento pelo máximo (*max pooling*)

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

Um único mapa

Agrupamento pelo max
com filtros 2×2 e passo 2

The diagram illustrates a max pooling operation. On the left, a 4x4 input map is shown with values: Row 1: 1, 1, 2, 4; Row 2: 5, 6, 7, 8; Row 3: 3, 2, 1, 0; Row 4: 1, 2, 3, 4. An arrow points from this input to a 2x2 output map on the right. The output map has values: Top-left: 5 (the maximum of 1, 1, 2, 4); Top-right: 8 (the maximum of 5, 6, 7, 8); Bottom-left: 3 (the maximum of 3, 2, 1, 0); Bottom-right: 4 (the maximum of 1, 2, 3, 4). This represents a 2x2 pooling window with a stride of 2.

5	
3	

Camada de Agrupamento (*Pooling*)

Agrupamento utiliza uma operação para condensar as informações:

- Máximo
- Média
- L^2

Exemplo de agrupamento pelo máximo (*max pooling*)

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

Um único mapa

Agrupamento pelo max
com filtros 2×2 e passo 2

6	

Camada de Agrupamento (*Pooling*)

Agrupamento utiliza uma operação para condensar as informações:

- Máximo
- Média
- L^2

Exemplo de agrupamento pelo máximo (*max pooling*)

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

Um único mapa

Agrupamento pelo max
com filtros 2×2 e passo 2

6	8

Camada de Agrupamento (*Pooling*)

Agrupamento utiliza uma operação para condensar as informações:

- Máximo
- Média
- L^2

Exemplo de agrupamento pelo máximo (*max pooling*)

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

Um único mapa

Agrupamento pelo max
com filtros 2×2 e passo 2

The diagram illustrates the max pooling process. On the left, a 4x4 input map is shown with values: Row 1: 1, 1, 2, 4; Row 2: 5, 6, 7, 8; Row 3: 3, 2, 1, 0; Row 4: 1, 2, 3, 4. An arrow points from this input to a smaller 2x2 output map on the right. The output map has values: Top-left cell (row 1, col 1) is 6 (the maximum of the 2x2 receptive field in the first row); Top-right cell (row 1, col 2) is 8 (the maximum of the 2x2 receptive field in the second row); Bottom-left cell (row 2, col 1) is 3 (the maximum of the 2x2 receptive field in the third row); Bottom-right cell (row 2, col 2) is empty.

6	8
3	

Camada de Agrupamento (*Pooling*)

Agrupamento utiliza uma operação para condensar as informações:

- Máximo
- Média
- L^2

Exemplo de agrupamento pelo máximo (*max pooling*)

1	1	2	4
5	6	7	8
3	2	1	0
1	2	3	4

Um único mapa

Agrupamento pelo max
com filtros 2×2 e passo 2

The diagram illustrates the max pooling process. On the left, a 4x4 input map is shown with values: Row 1: 1, 1, 2, 4; Row 2: 5, 6, 7, 8; Row 3: 3, 2, 1, 0; Row 4: 1, 2, 3, 4. An arrow points from this input to a smaller 2x2 output map on the right. The output map has values: Top-left cell (row 1, col 1) is 6 (the maximum of the 2x2 receptive field in the first row); Top-right cell (row 1, col 2) is 8 (the maximum of the 2x2 receptive field in the second row); Bottom-left cell (row 2, col 1) is 3 (the maximum of the 2x2 receptive field in the third row); Bottom-right cell (row 2, col 2) is 4 (the maximum of the 2x2 receptive field in the fourth row).

6	8
3	4