



Sisteme de Vedere Artificială

Recunoaşterea Obiectelor

Sorin M. Grigorescu



Cuprins

- Rețele neurale convolutive

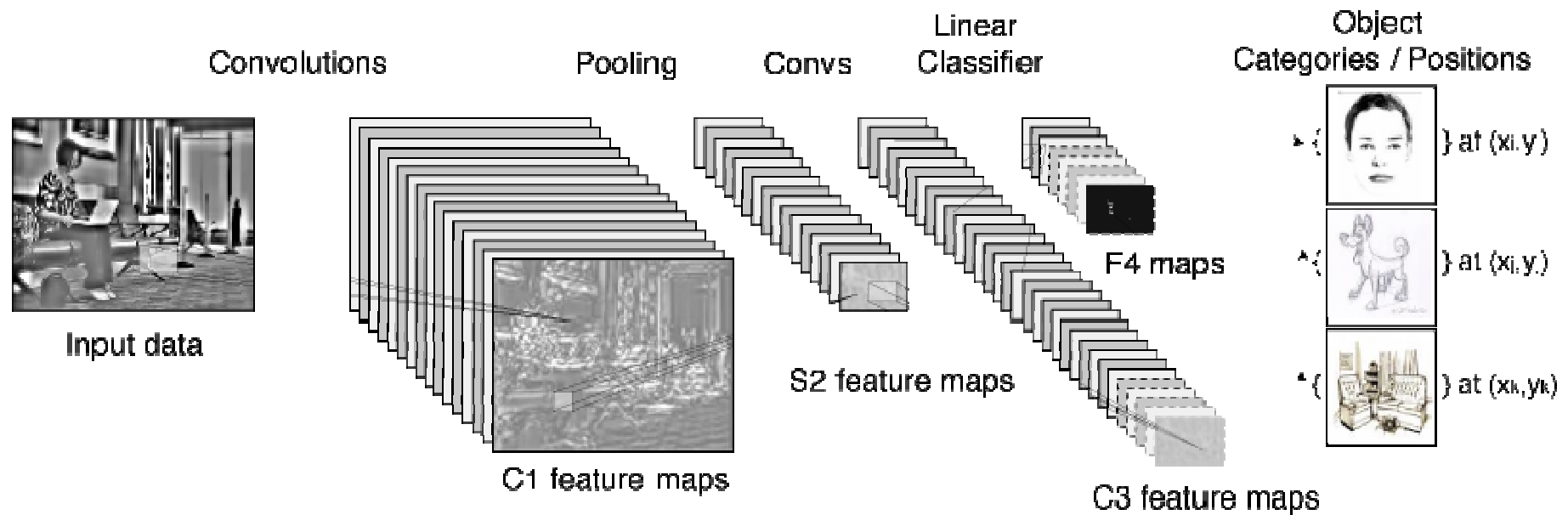


Rețele Neurale Convolutive



Rețea Neurală Convolutivă

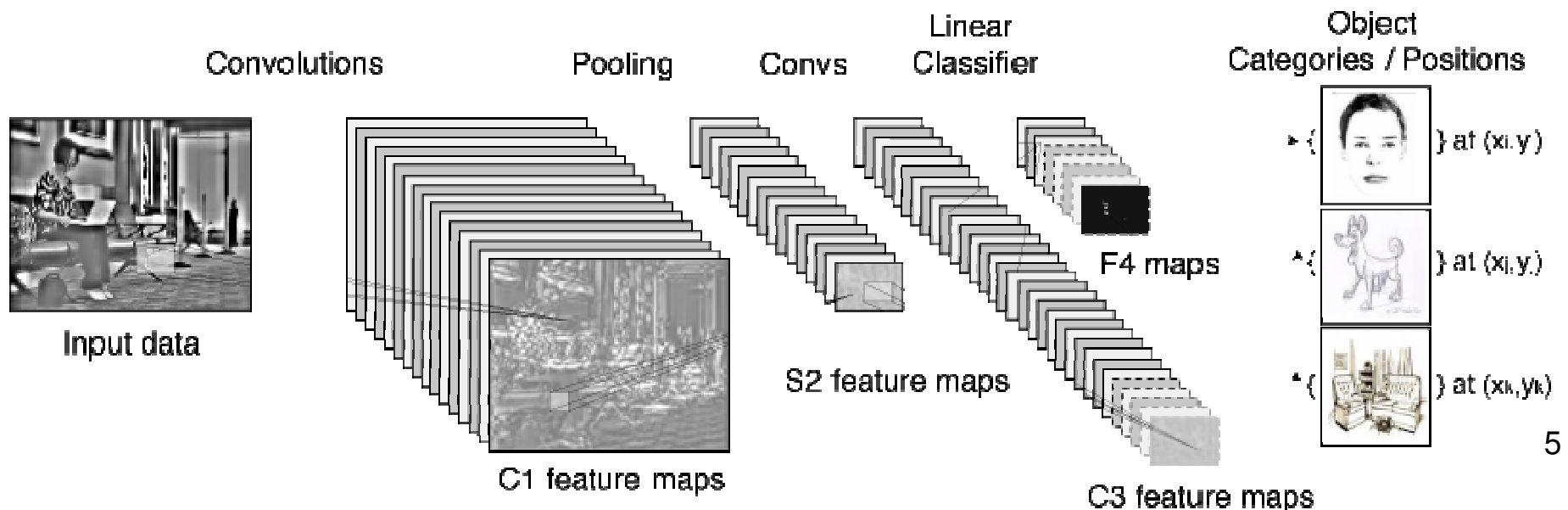
- Caracteristicile (features) sunt activate pe măsură ce datele de intrare străbat straturile rețelei
- Într-un anumit strat, convoluția are loc pe datele calculate în stratul precedent





Rețea Neurală Convolutivă

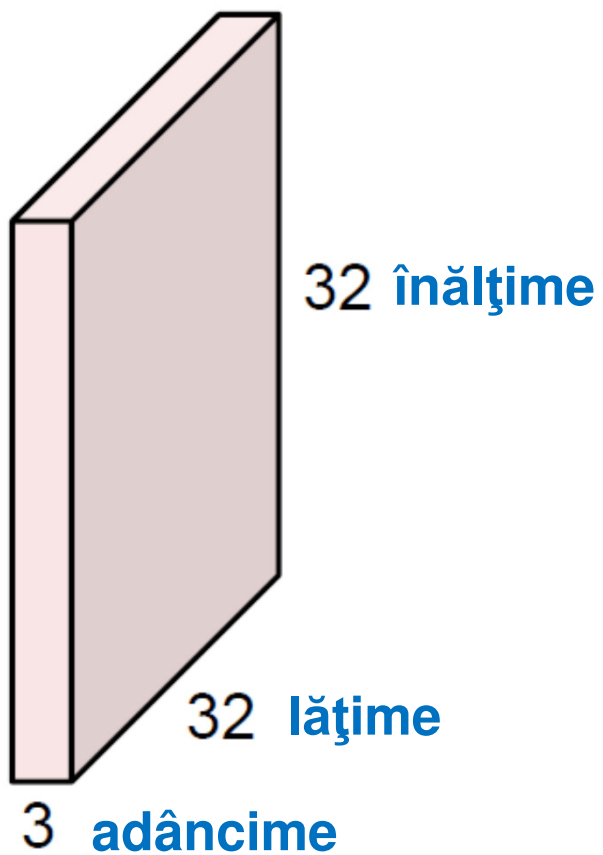
- Există patru tipuri de straturi:
 - Convolutiv
 - Activare printr-o funcție de activare (ex. Sigmoid sau ReLU – Rectified Linear Unit)
 - Pooling (statistică asupra unei regiuni de interes, reducerea cantității de informație)
 - Conectivitate completă (rețea neurală clasică)





Strat convolutiv

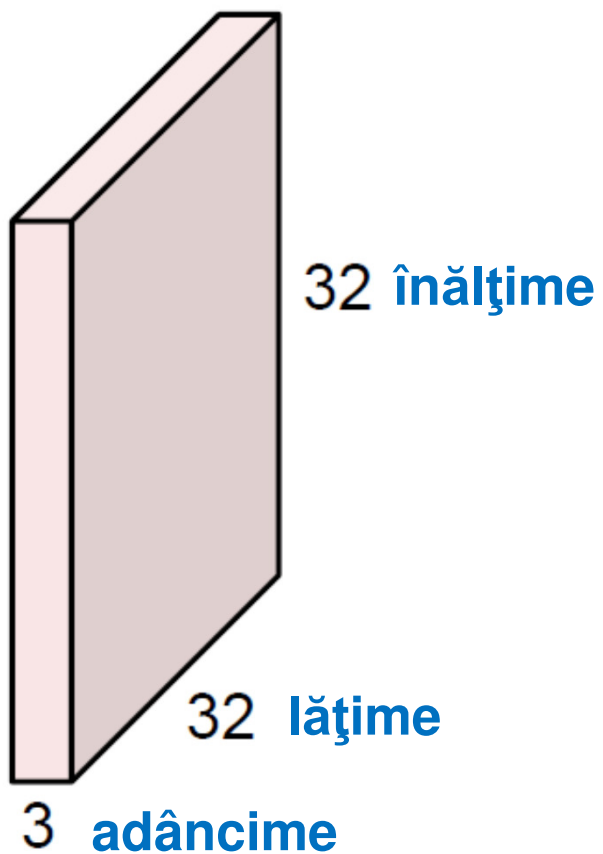
imagine 32x32x3





Strat convolutiv

imagine 32x32x3

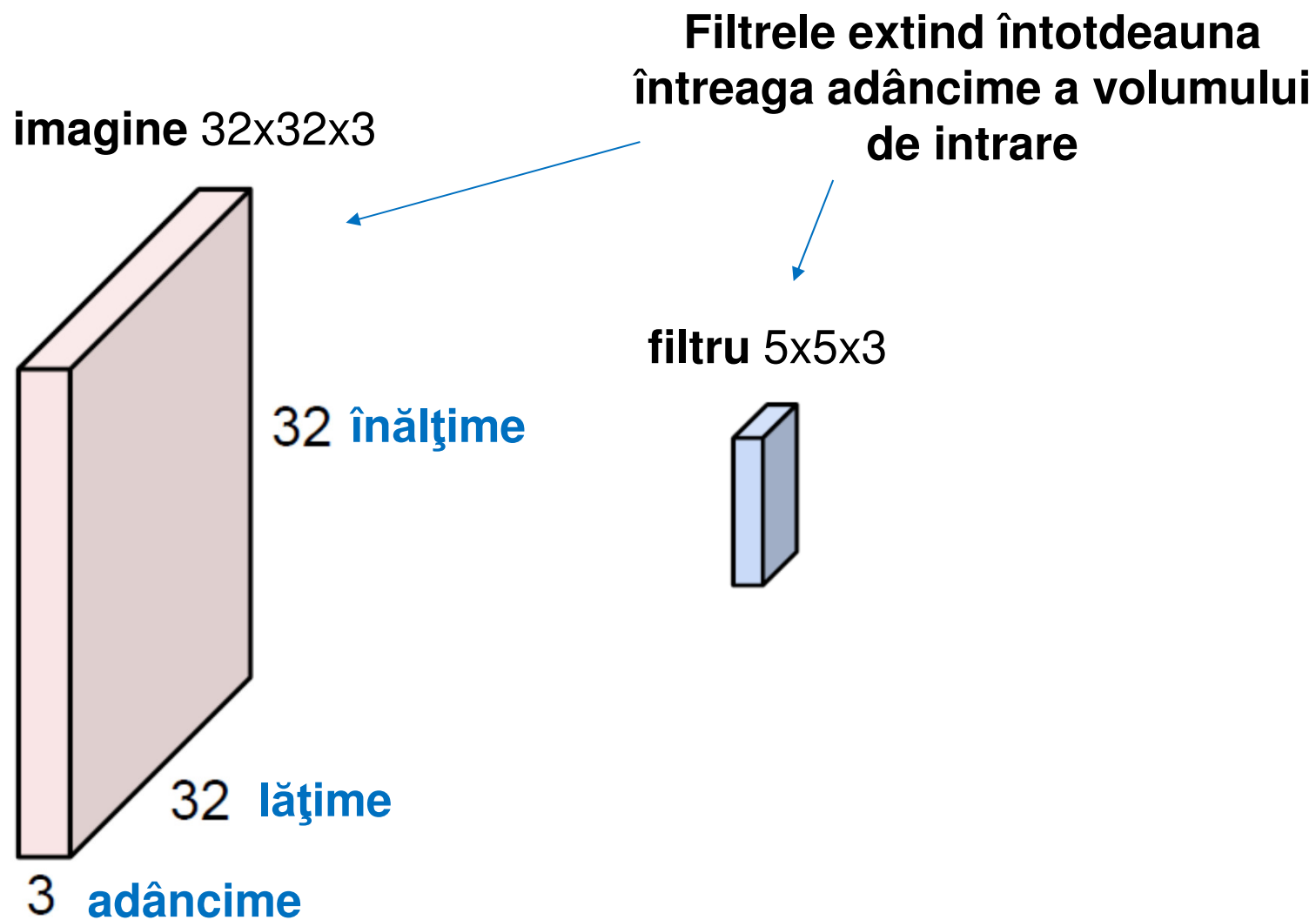


filtru 5x5x3



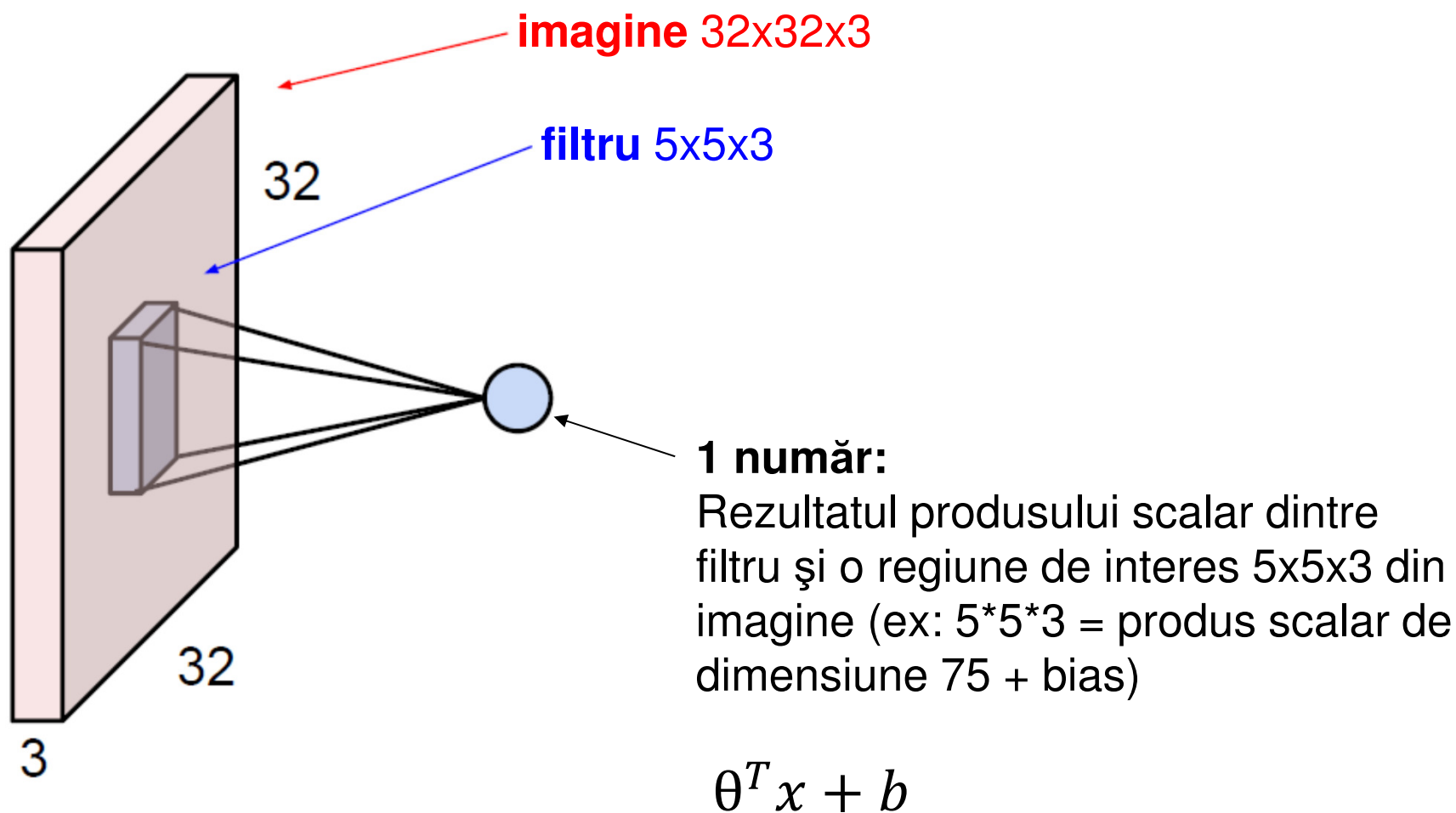


Strat convolutiv





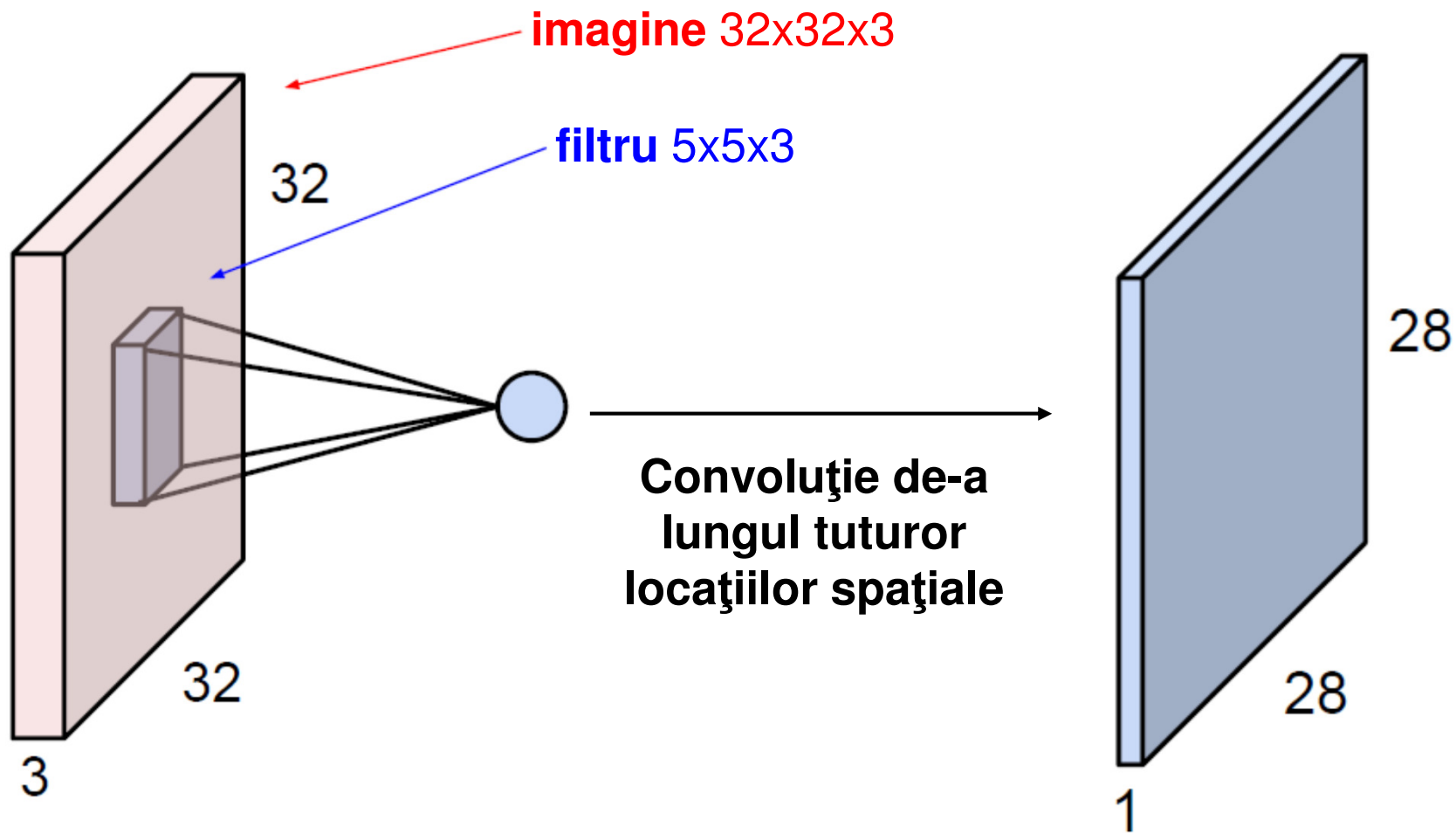
Strat convolutiv





Strat convolutiv

Hartă de activare

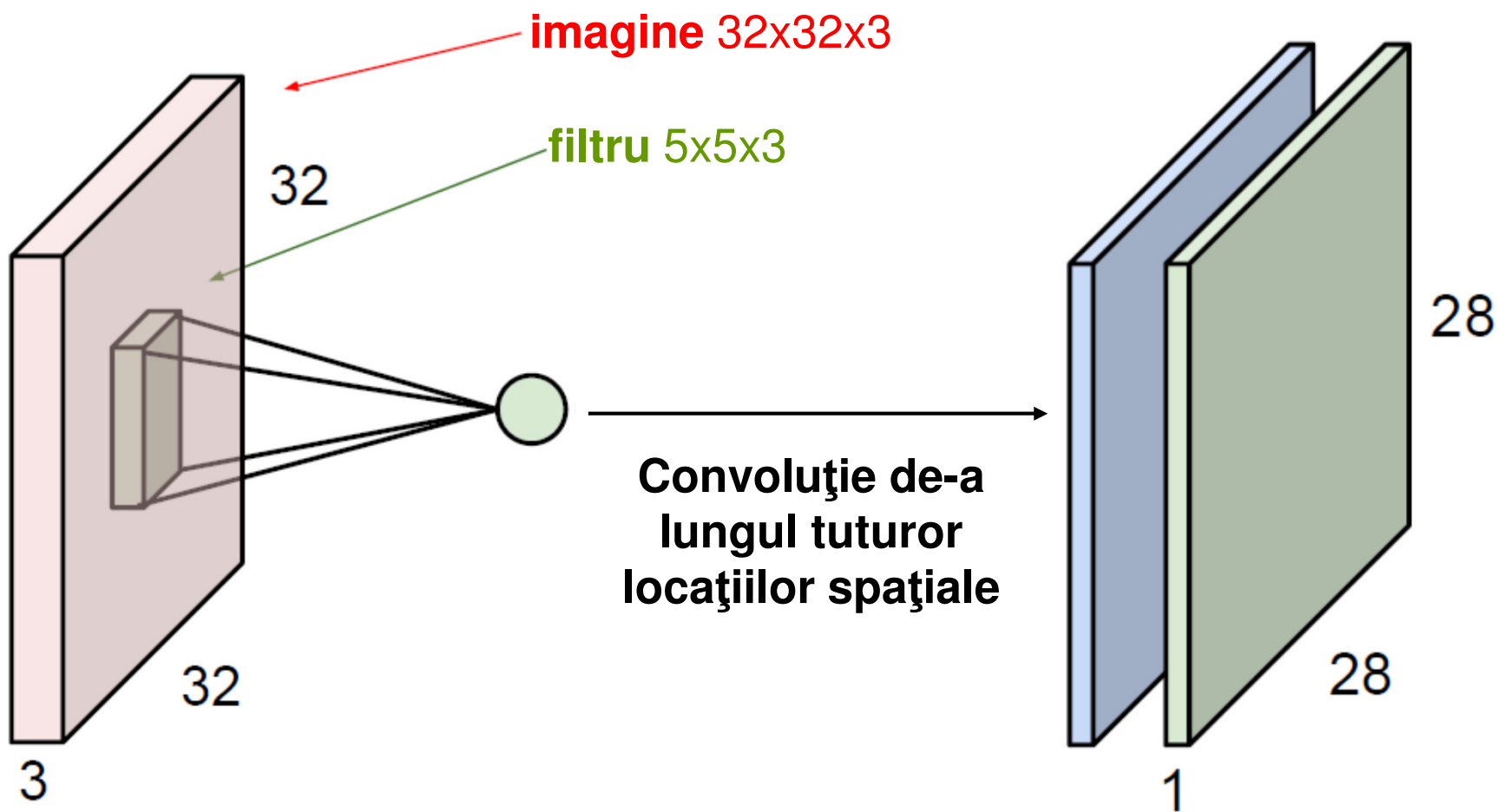


Strat convolutiv

Al doilea filtru (verde)



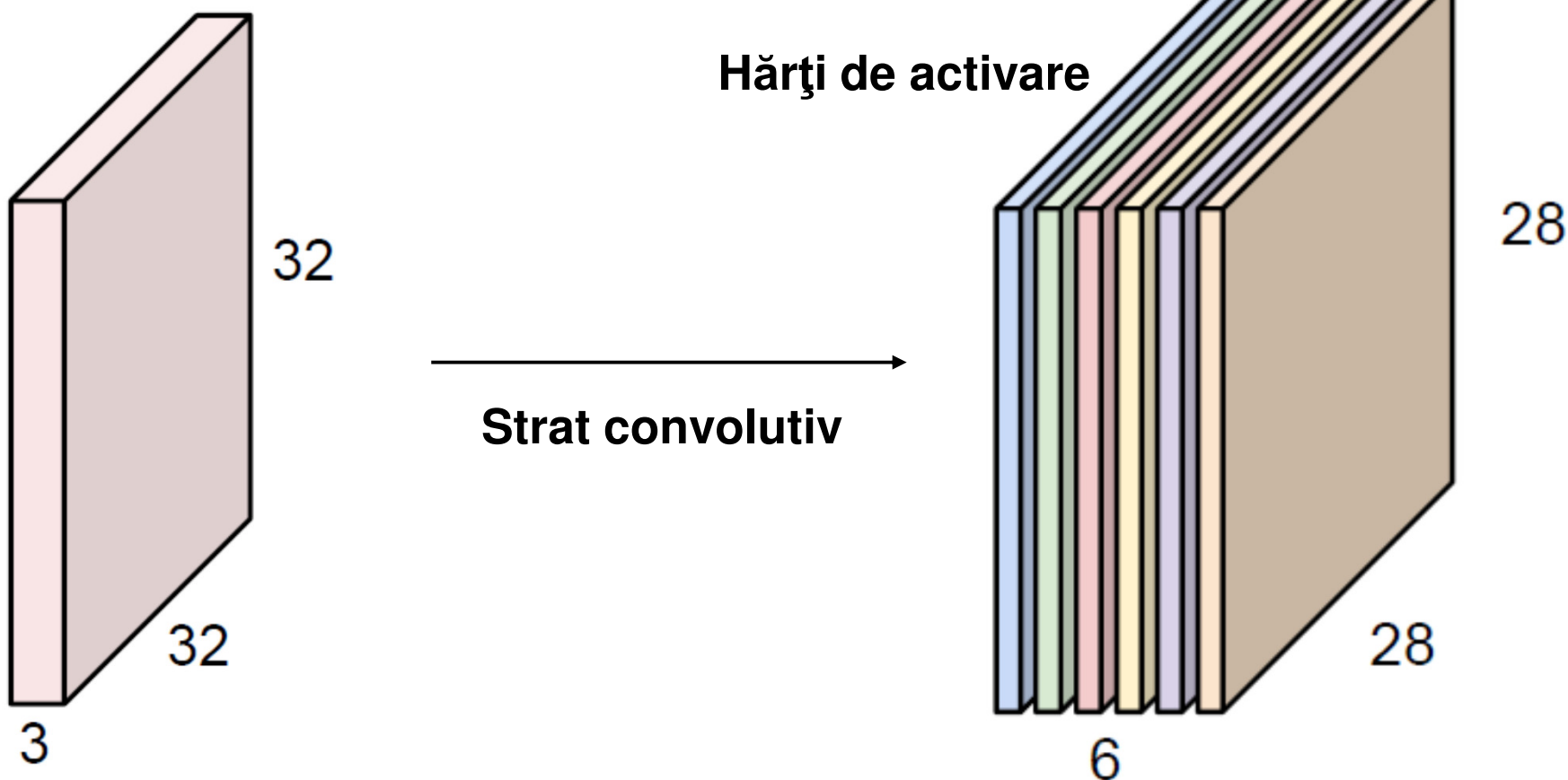
Hărți de activare





Strat convolutiv

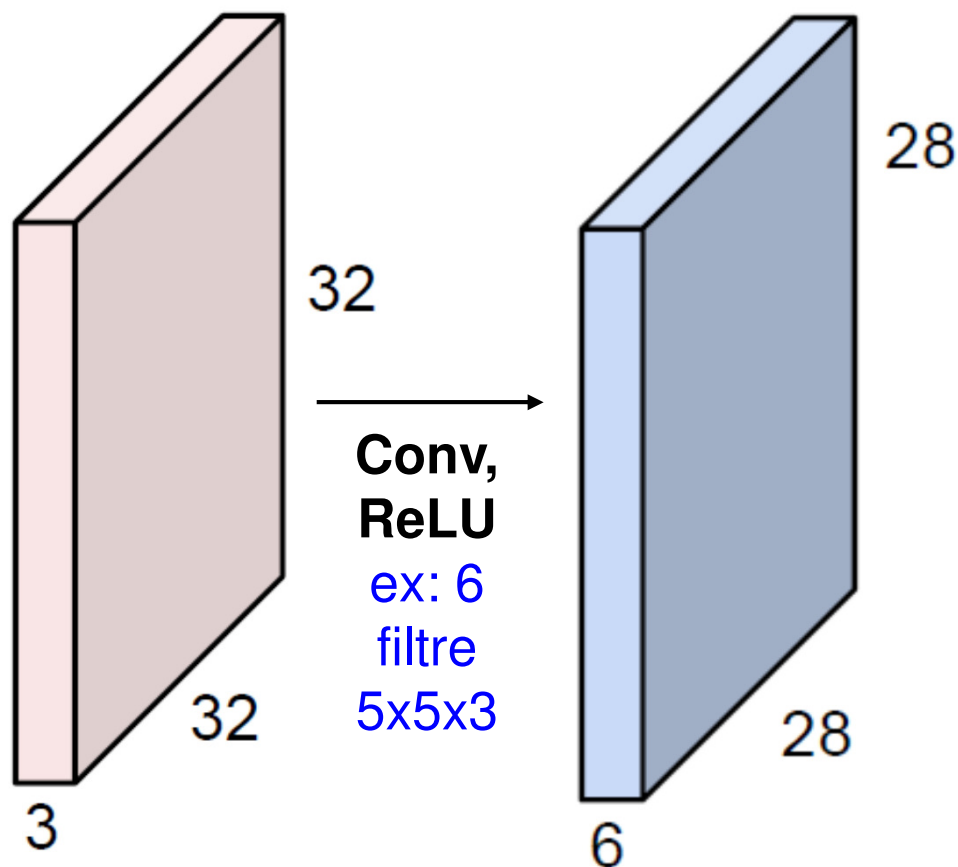
- Ex: în cazul a 6 filtre 5x5, vom obține 6 hărți de activare
- Suprapunând cele 6 hărți de activare, vom obține o noua imagine 28x28x6





Strat convolutiv

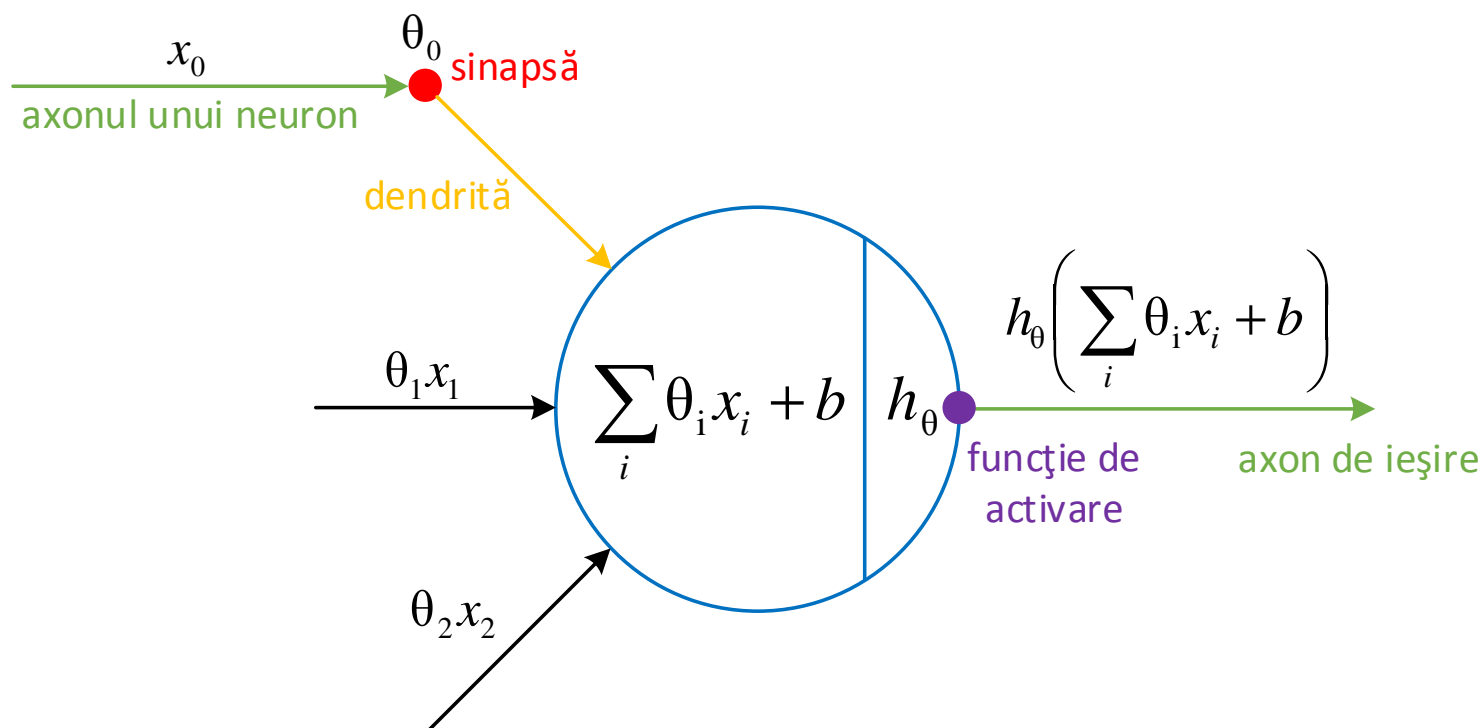
- O rețea neurală convolutivă este o secvență de straturi convolutive, intercalate cu funcții de activare (ex: ReLU sau Sigmoid)





Modelul unui neuron

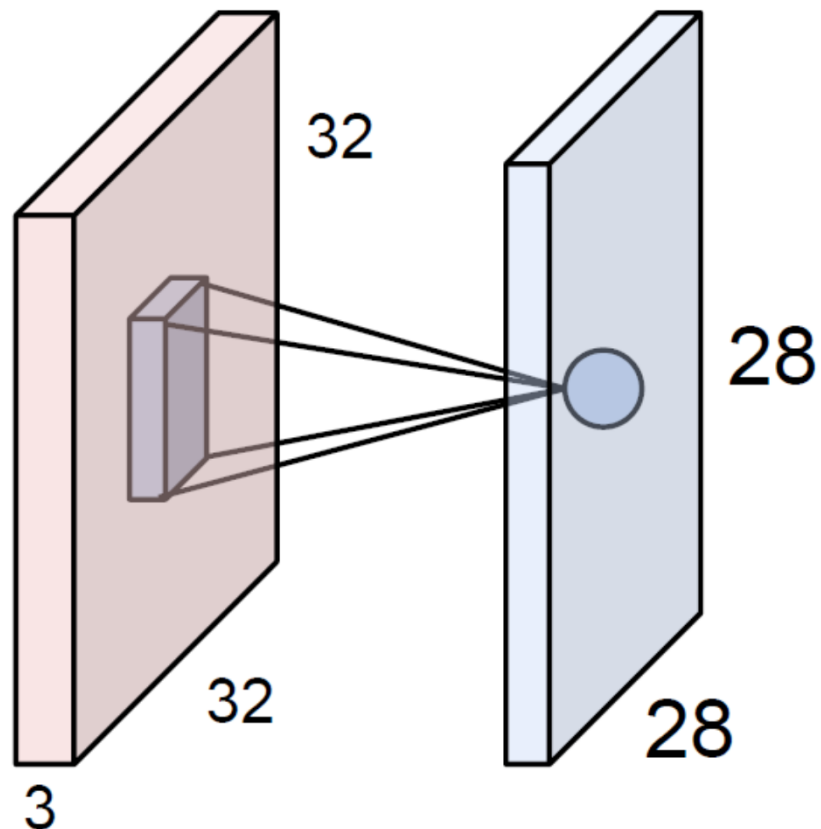
- Neuron cu conectivitate locală



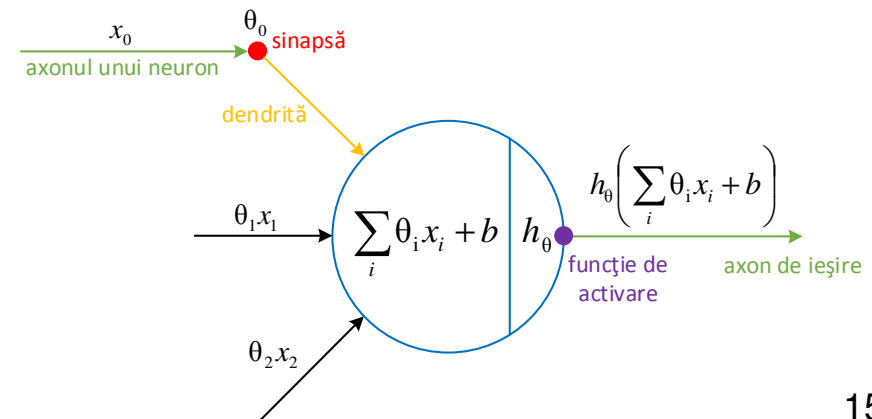


Modelul unui neuron

- O hartă de activare este un strat de neuroni de ieșire, de dimensiune 28x28
- Fiecare neuron este conectat la o mică regiune de interes din volumul de intrare



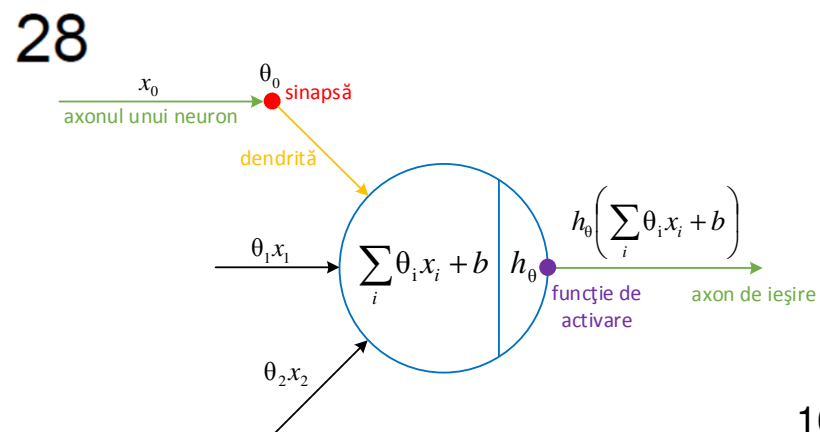
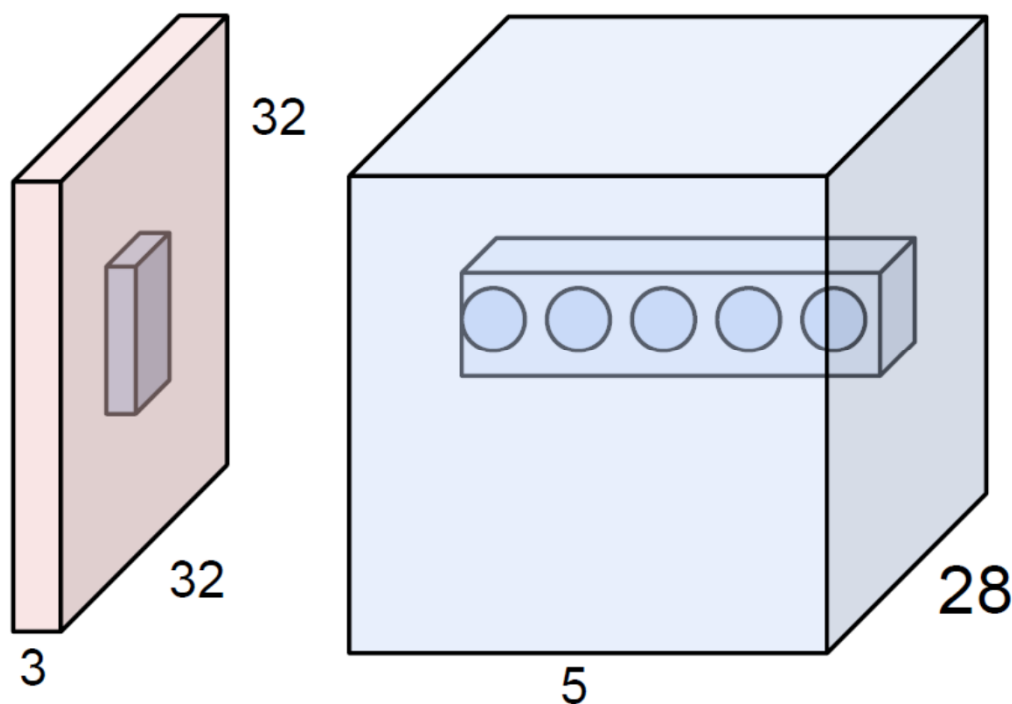
- Toți neuronii împart parametrii θ
- Filtru 5x5: Câmp receptiv de dimensiune 5x5 pentru fiecare neuron





Modelul unui neuron

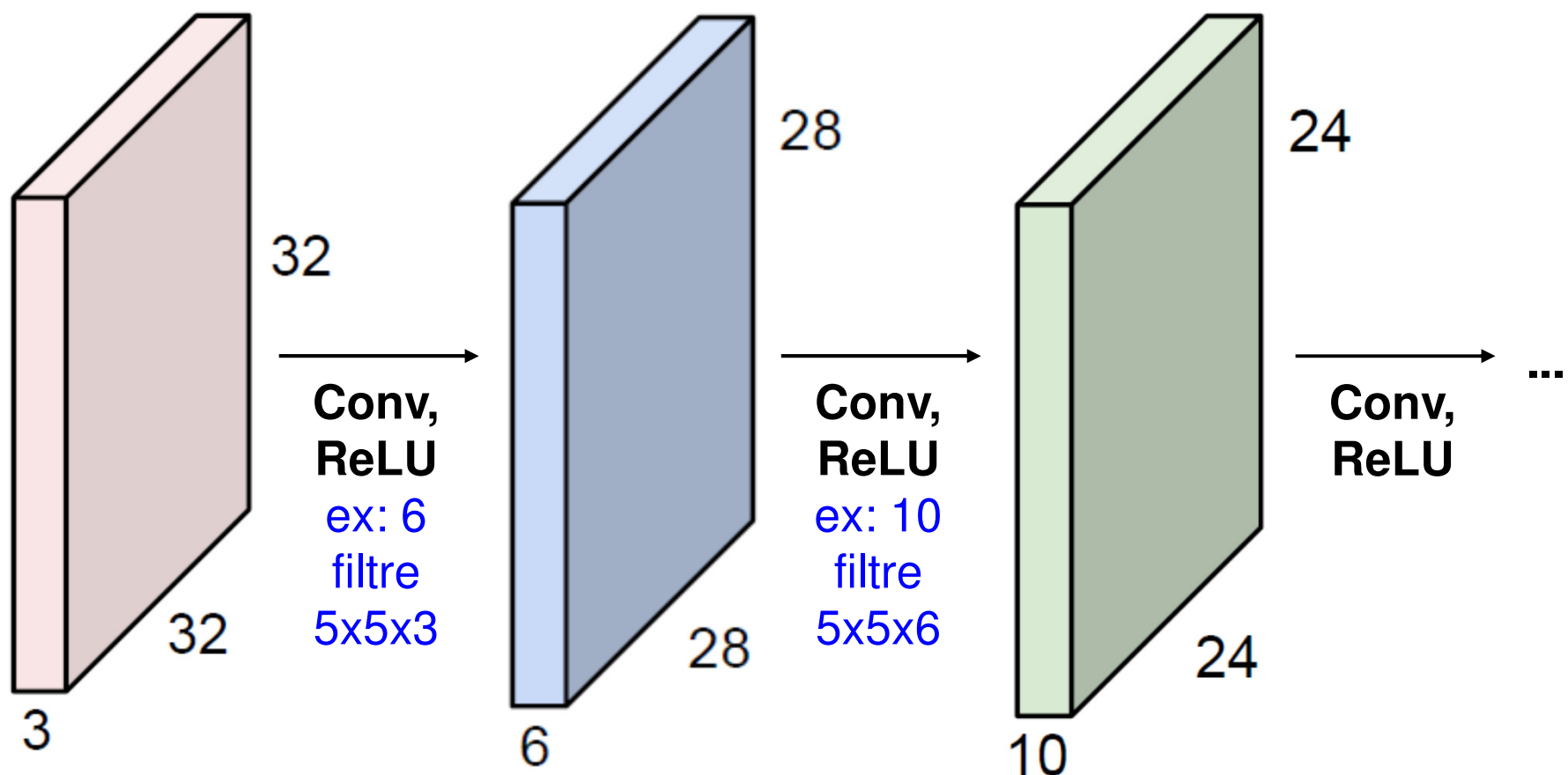
- Ex: având 5 filtre, stratul convolutiv este alcătuit din neuroni ordonați într-un grid 3D de dimensiune 28x28x5
- Vor exista 5 neuroni diferiți, toți analizând aceeași regiune din volumul de intrare





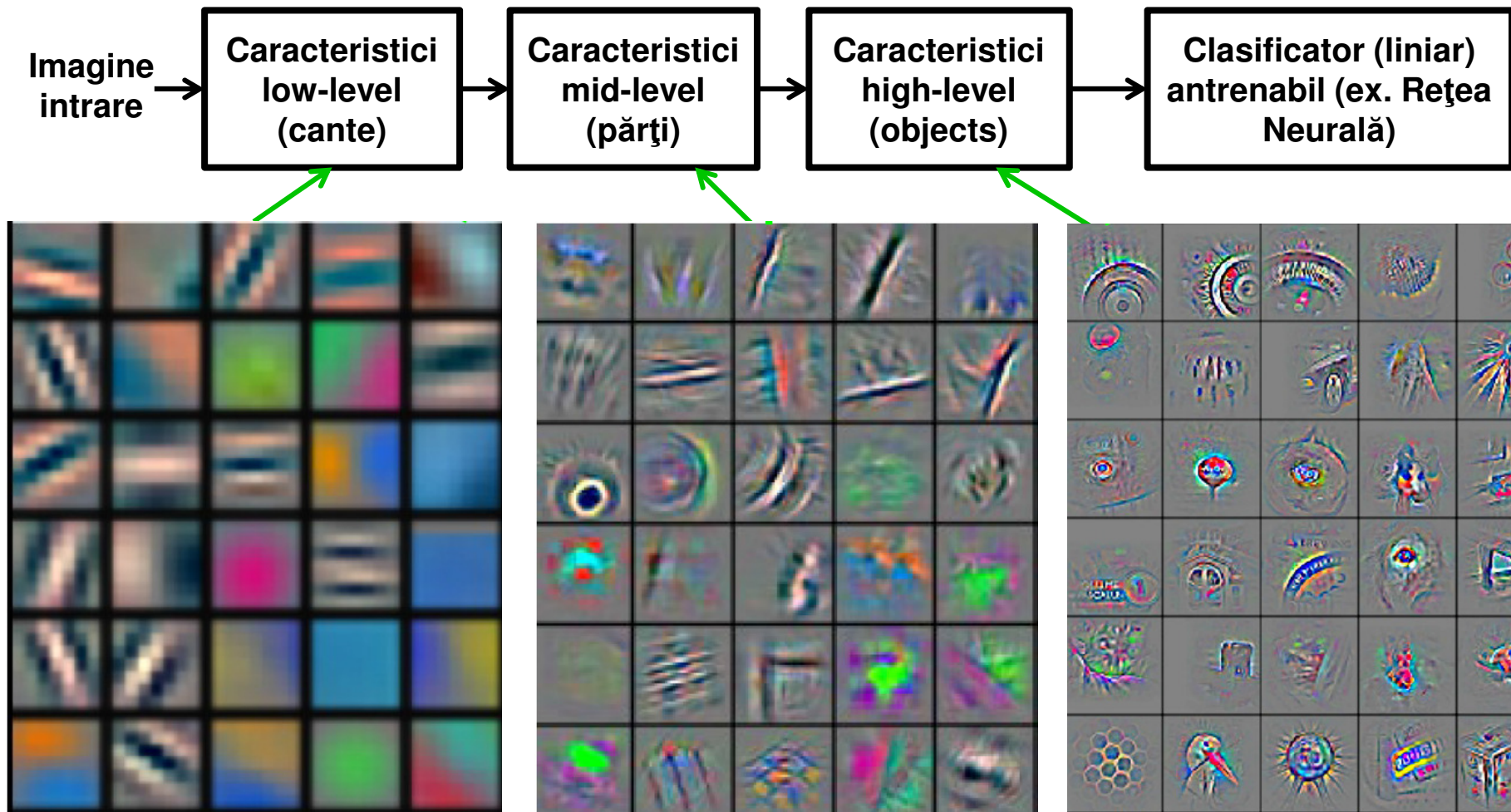
Strat convolutiv

- O rețea neurală convolutivă este o secvență de straturi convolutive, intercalate cu funcții de activare (ex: ReLU sau Sigmoid)





Strat convolutiv



Feature visualization of a convolutional net trained on ImageNet [Zeiler & Fergus 2013]



Strat convolutiv

one filter =>
one activation map

example 5x5 filters
(32 total)

Activations:

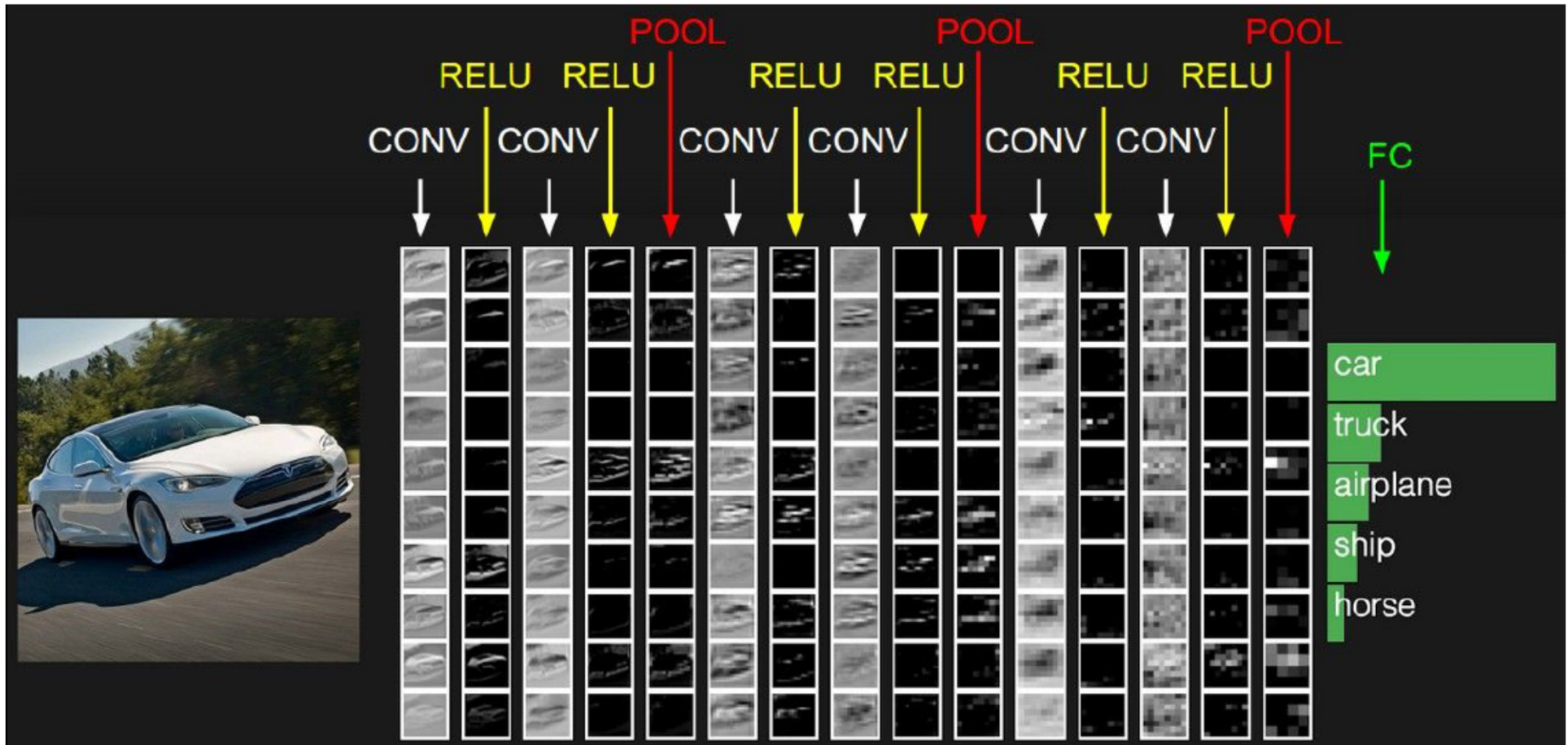
We call the layer convolutional because it is related to convolution of two signals:

$$f[x,y] * g[x,y] = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f[n_1,n_2] \cdot g[x-n_1,y-n_2]$$

↑
elementwise multiplication and sum of a filter and the signal (image)



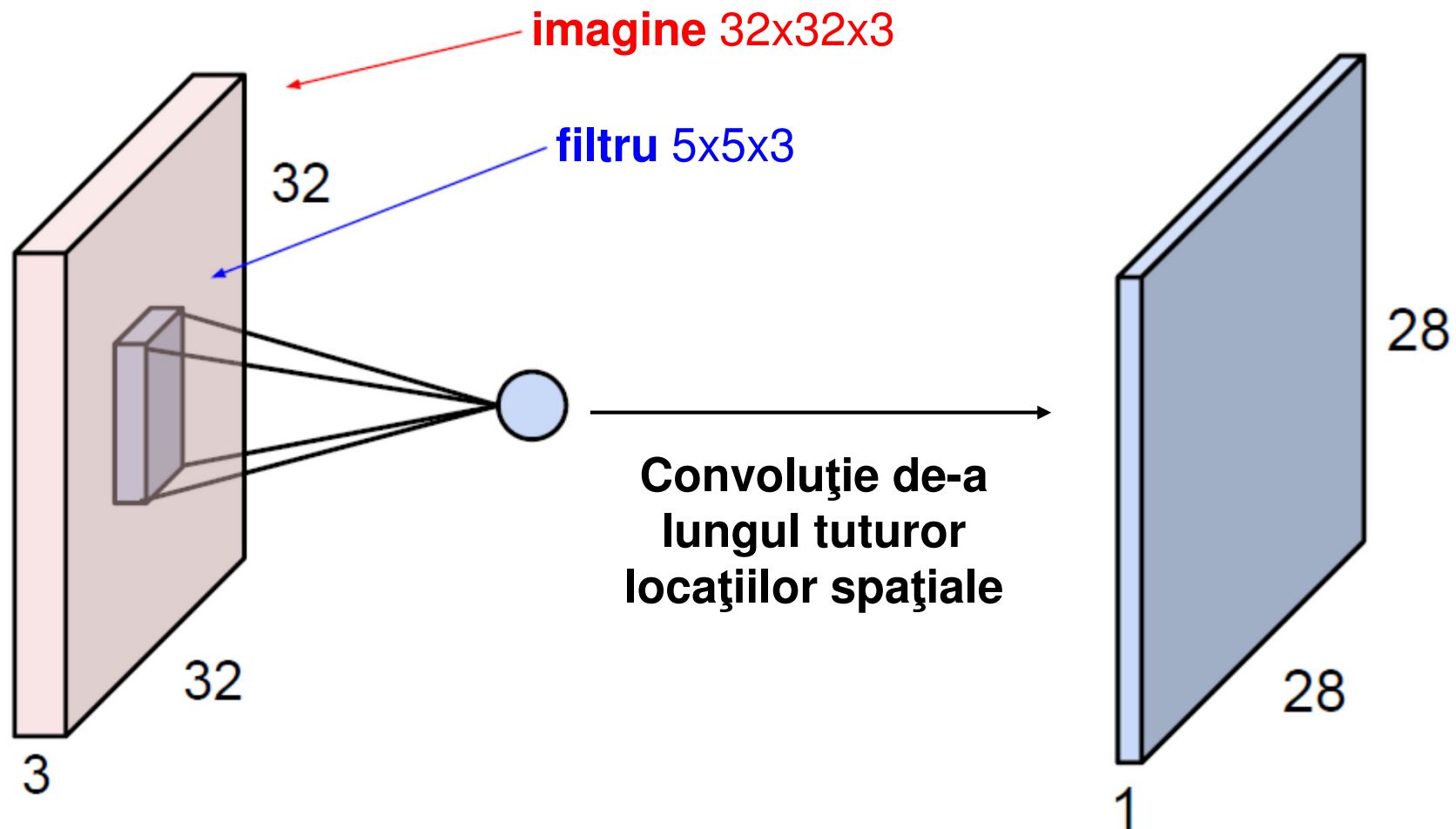
Strat convolutiv





Dimensiuni spațiale

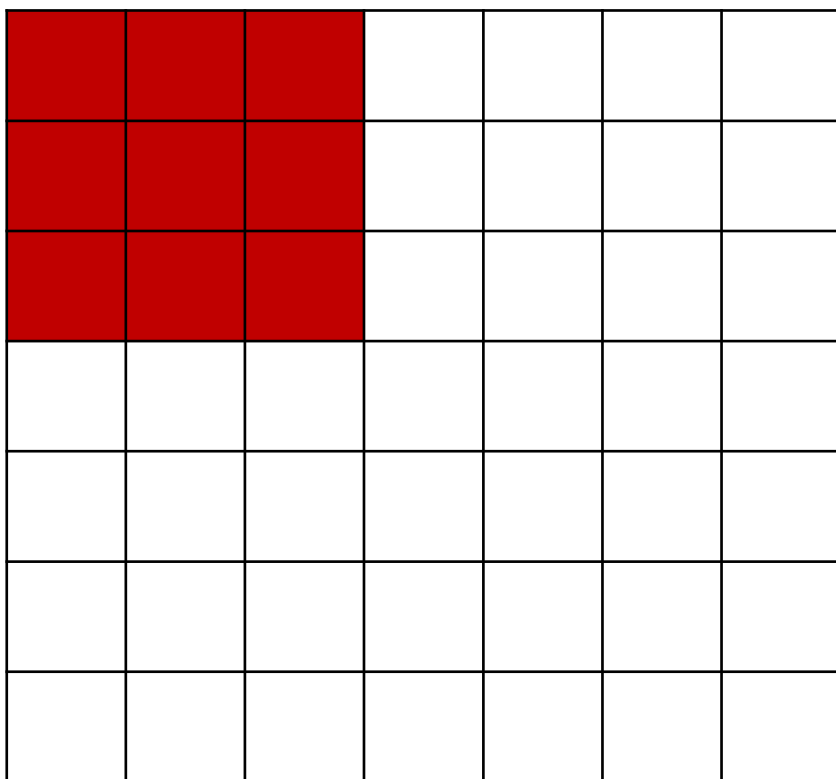
Hartă de activare



Dimensiuni spațiale



7



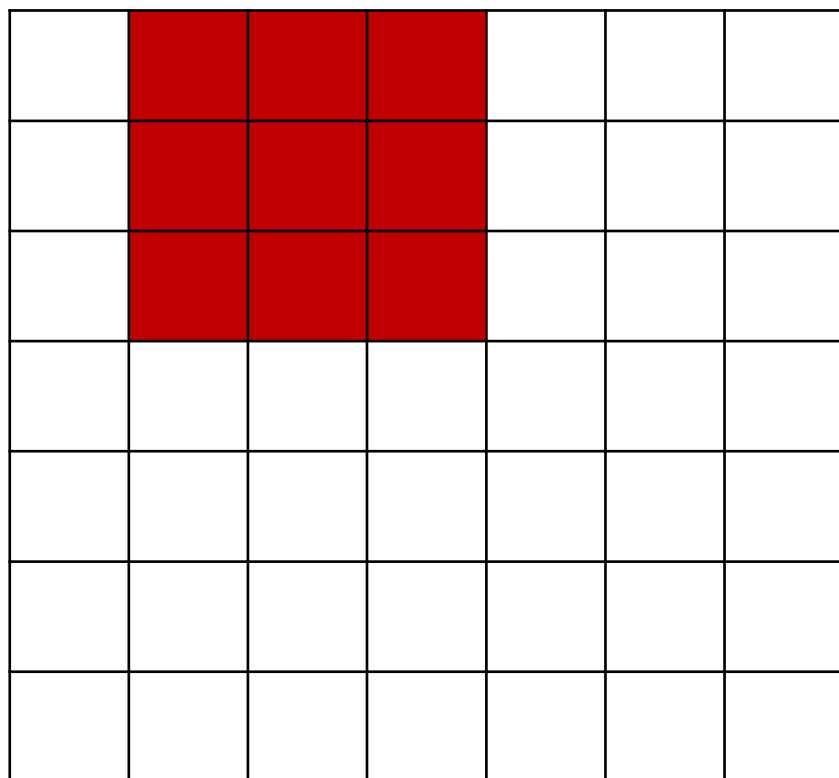
7

Imagine de intrare 7x7
Filtru 3x3



Dimensiuni spațiale

7



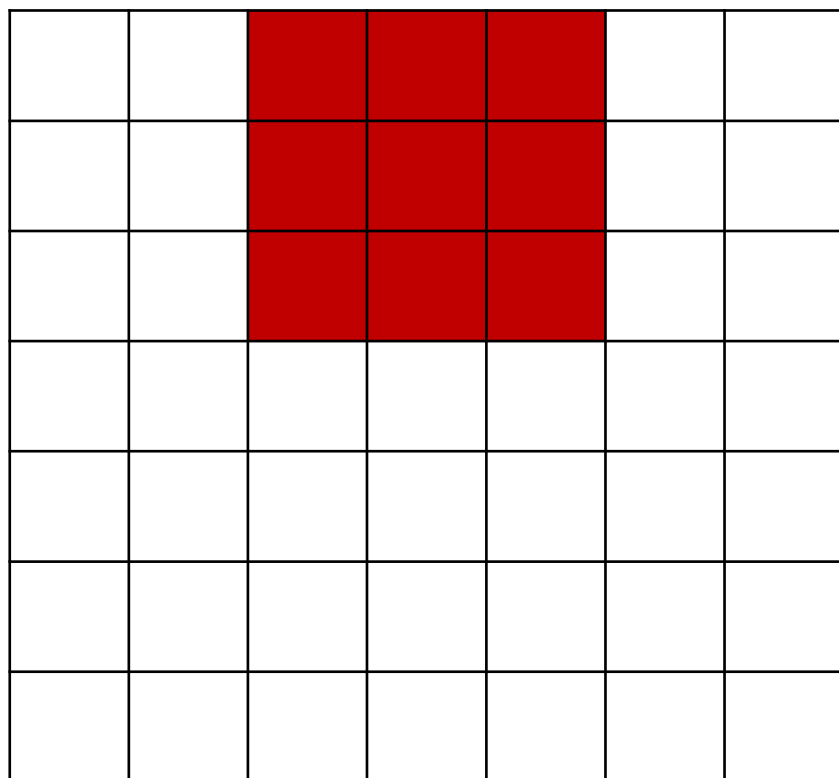
7

Imagine de intrare 7x7
Filtru 3x3



Dimensiuni spațiale

7



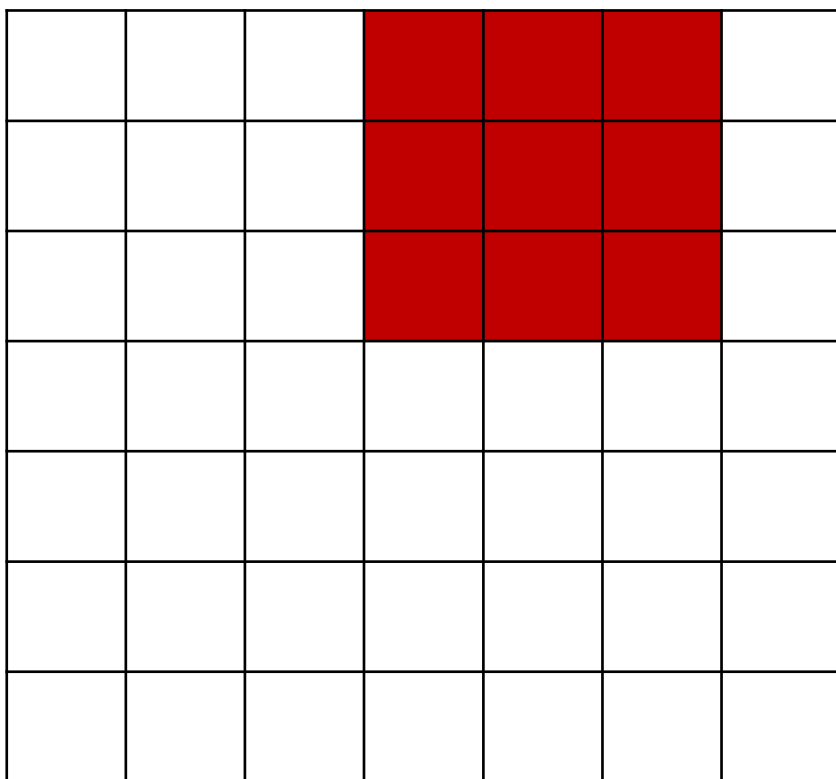
7

Imagine de intrare 7x7
Filtru 3x3

Dimensiuni spațiale



7



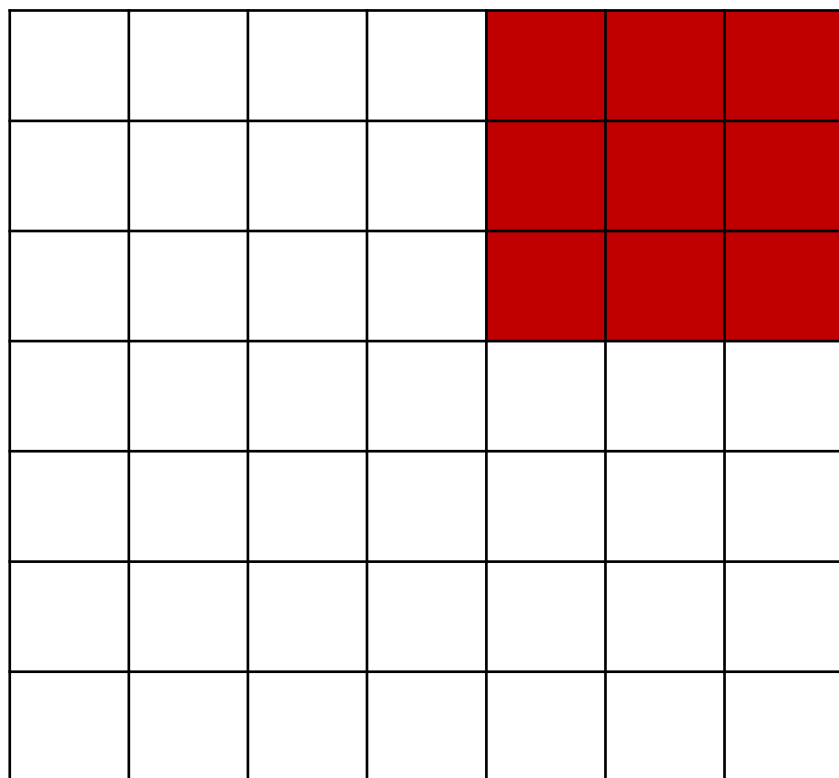
7

Imagine de intrare 7x7
Filtru 3x3



Dimensiuni spațiale

7



7

Imagine de intrare 7x7

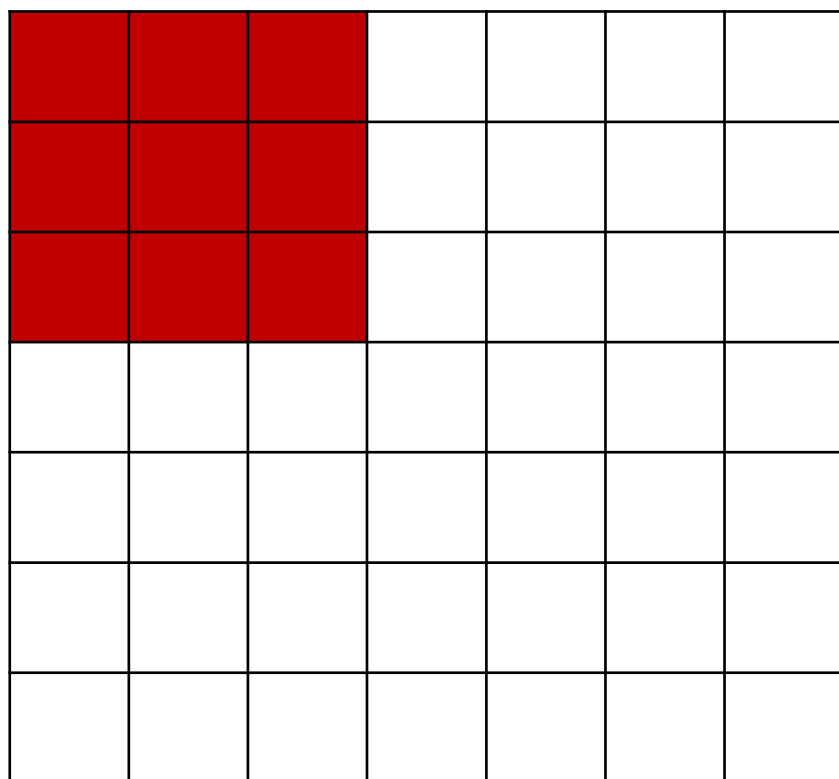
Filtru 3x3

Ieșire 5x5



Dimensiuni spațiale: stride

7



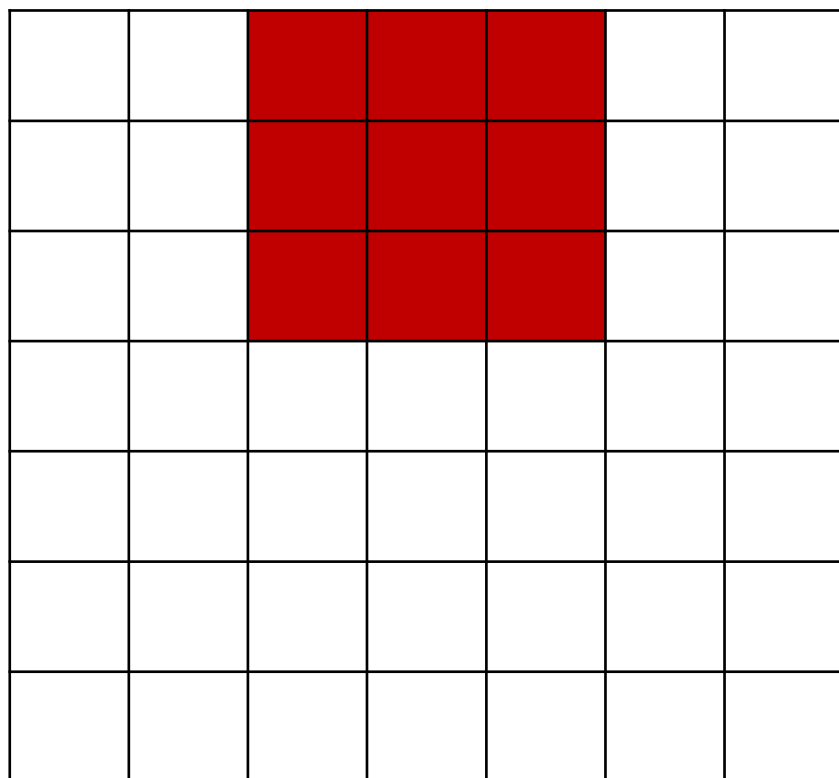
7

Imagine de intrare 7x7
Filtru 3x3
Stride 2



Dimensiuni spațiale: stride

7



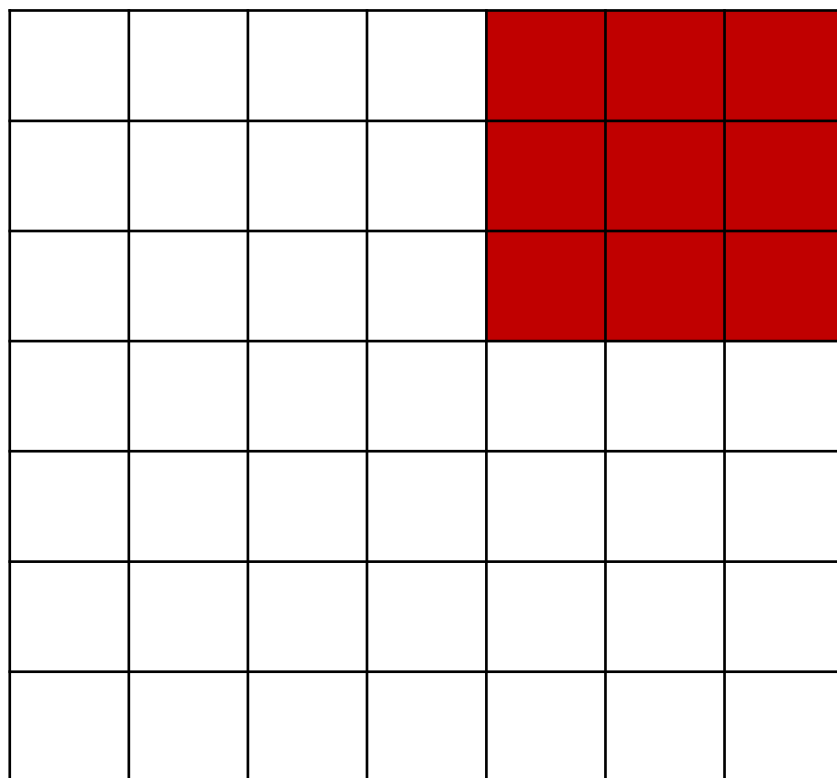
7

Imagine de intrare 7x7
Filtru 3x3
Stride 2



Dimensiuni spațiale: stride

7



7

Imagine de intrare 7x7

Filtru 3x3

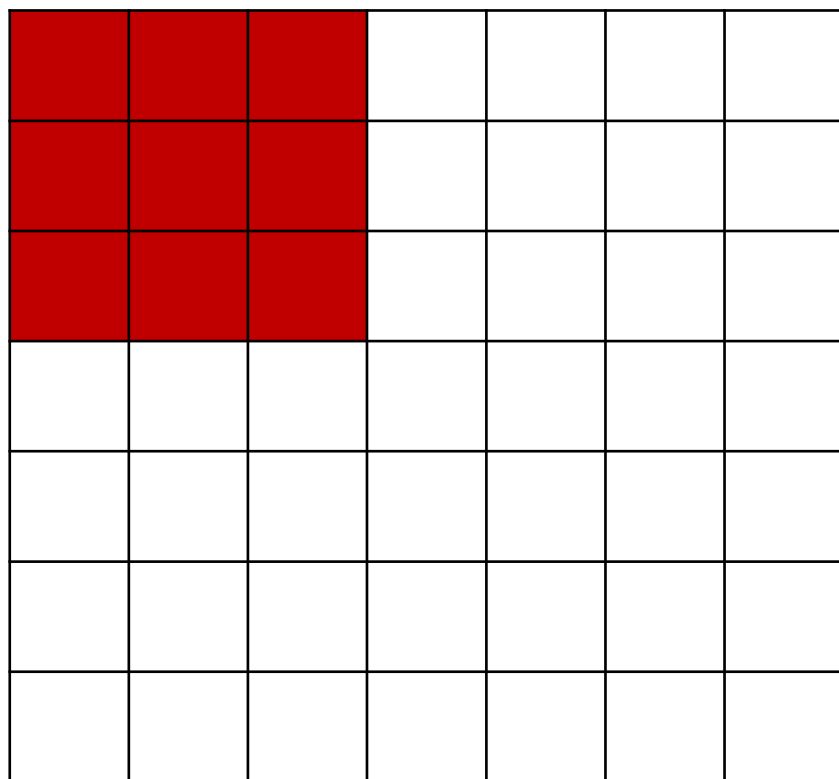
Stride 2

ieșire 3x3



Dimensiuni spațiale: stride

7



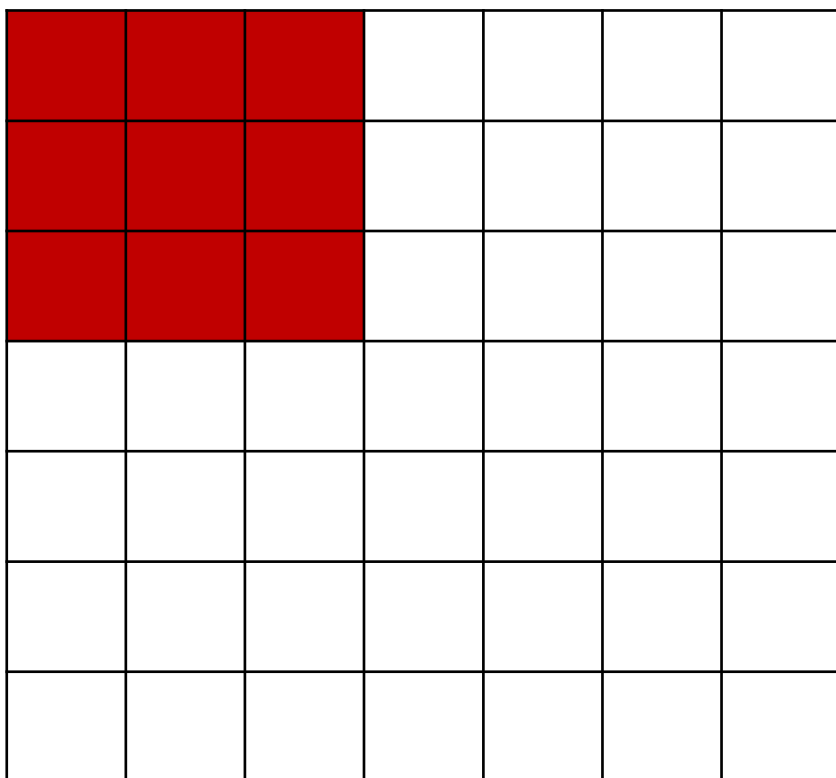
7

Imagine de intrare 7x7
Filtru 3x3
Stride 3



Dimensiuni spațiale: stride

7



7

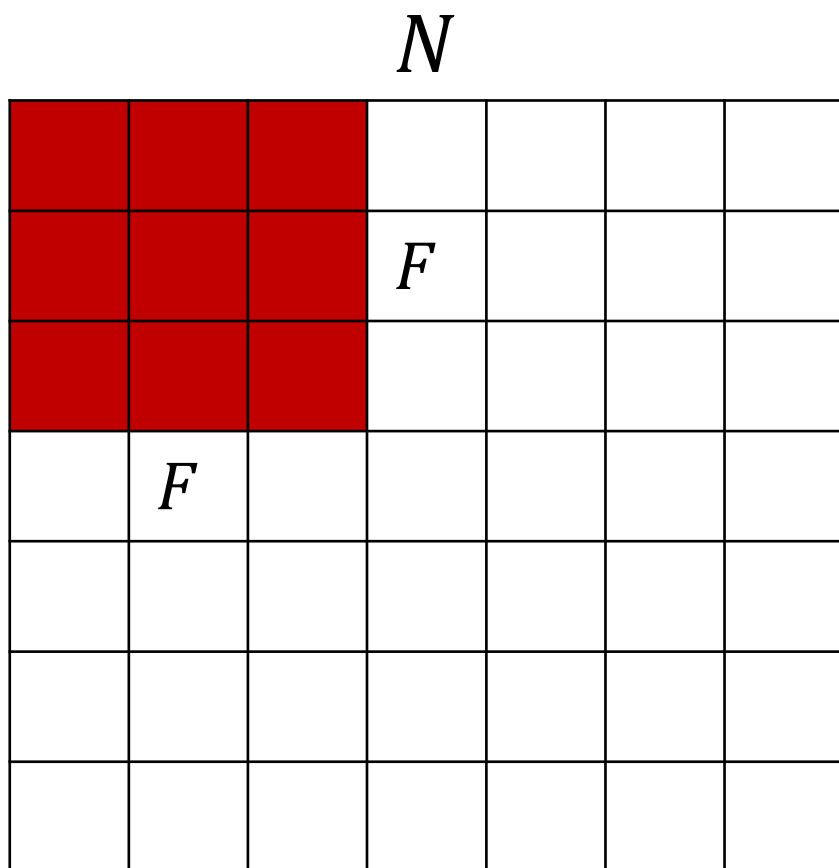
Imagine de intrare 7x7
Filtru 3x3
Stride 2

Dimensiunile nu se
potrivesc

Un filtru 3x3 nu poate fi
aplicat pe imagine 7x7,
utilizând stride 3



Dimensiuni spațiale: stride



Dimensiunea de ieșire:

$$\frac{N - F}{\text{stride}} + 1$$

Ex: $N = 7, F = 3$

Stride 1: $(7-3)/1 + 1 = 5$

Stride 2: $(7-3)/2 + 1 = 3$

Stride 3: $(7-3)/3 + 1 = 2.33$ (eroare)



Margine zero-pad

- Ex: imagine 7x7, filtru 3x3, stride 1. Se poate adauga o margine zero-pad de 1 pixel
- Se folosește des zero-padding de dimensiune $F-1/2$

0	0	0	0	0			
0							
0							
0							
0							

Ex: $F = 3 \rightarrow$ zero-pad 1

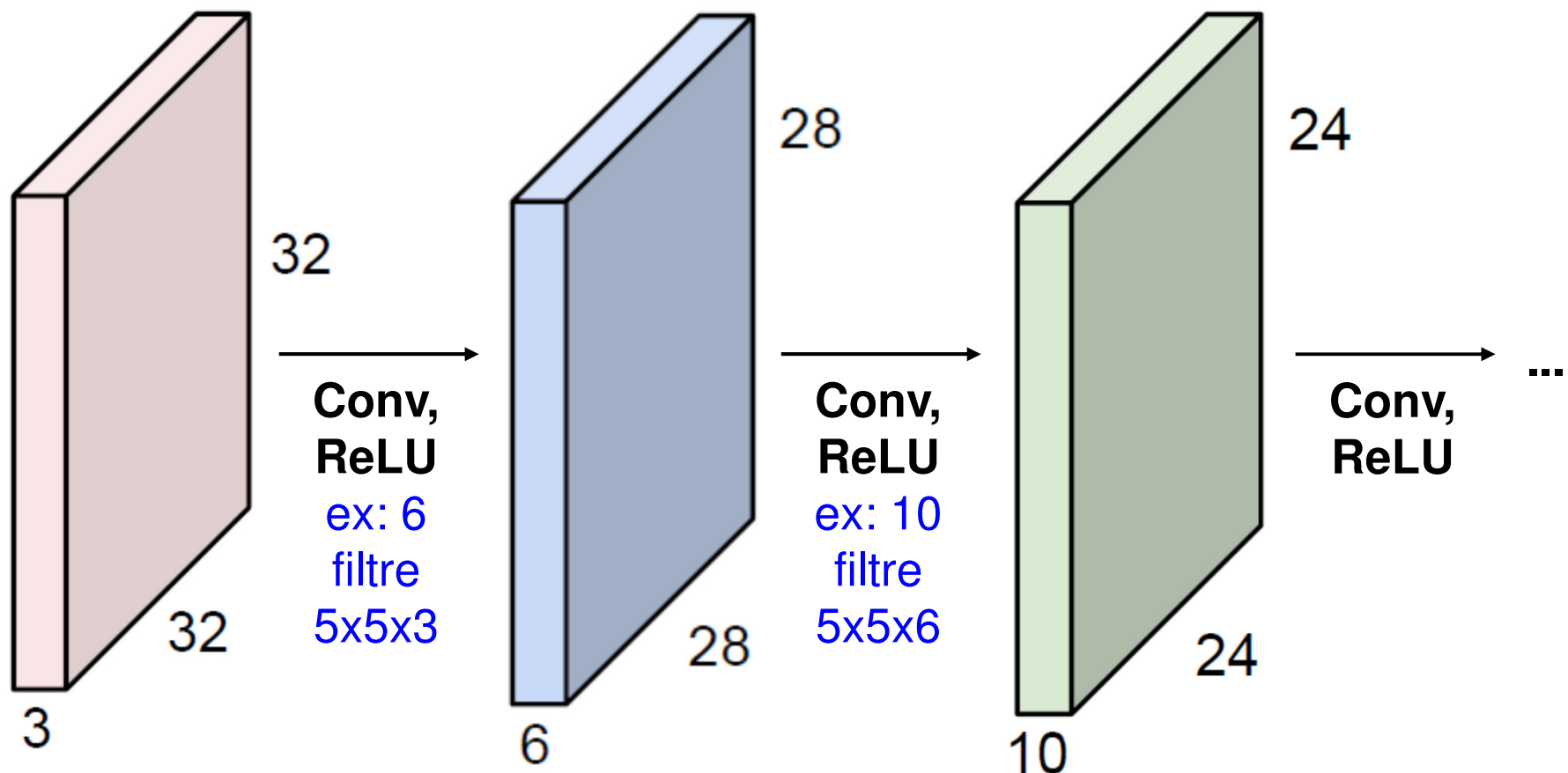
$F = 5 \rightarrow$ zero-pad 2

$F = 7 \rightarrow$ zero-pad 3



Strat convolutiv

- Aplicarea repetată de filtre 5x5 va micșora spațial volumele de intrare (32 → 28 → 24)
- Micșorarea spațială rapidă dăunează performanței rețelei





Strat convolutiv: rezumat

- Prelucreează volume de dimensiune $W_1 \times H_1 \times D_1$
- Necesită patru hiperparametrii:
 - Numărul de filtre K
 - Dimensiunea spațială F a filtrelor
 - Stride S
 - Dimensiunea zero-padding P
- Produce un volum de dimensiune $W_2 \times H_2 \times D_2$, unde:

$$W_2 = \frac{W_1 - F + 2P}{S} + 1$$

$$H_2 = \frac{H_1 - F + 2P}{S} + 1$$

$$D_2 = K$$

(lățimea și înălțimea sunt calculate identic prin simetrie)



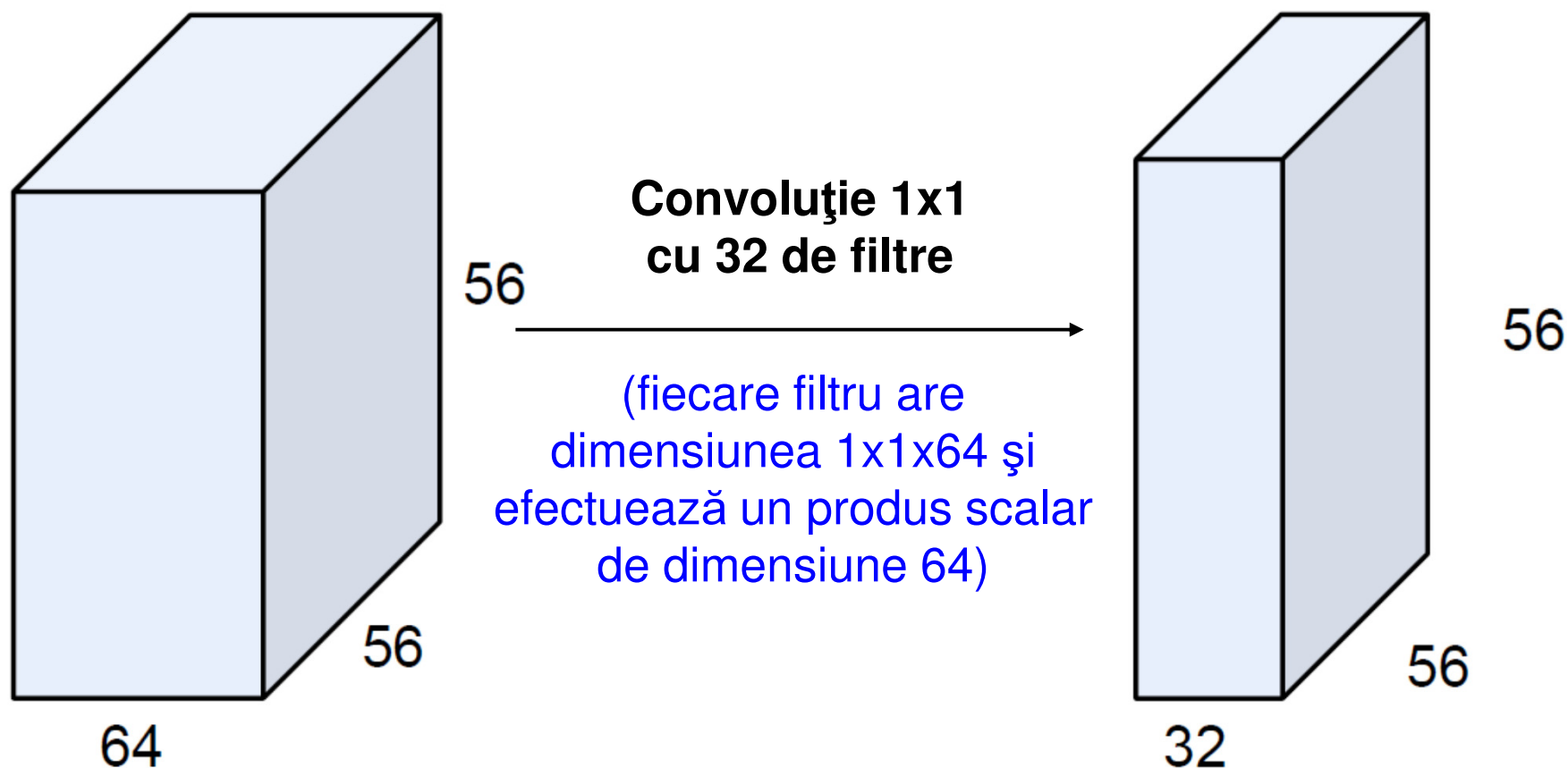
Strat convolutiv: rezumat

- Împărțind parametrii θ ai rețelei, vom obține $F \cdot F \cdot D_1$ parametrii (ponderi) pe filtru, pentru un număr total de $(F \cdot F \cdot D_1) \cdot K$ ponderi și K neuroni de bias
- În volumul de ieșire, harta de activare d (de dimensiune $W_2 \times H_2$) este rezultatul unei convoluții valide al filtrului d de-a lungul volumului de intrare cu un stride S și un offset al bias-ului d



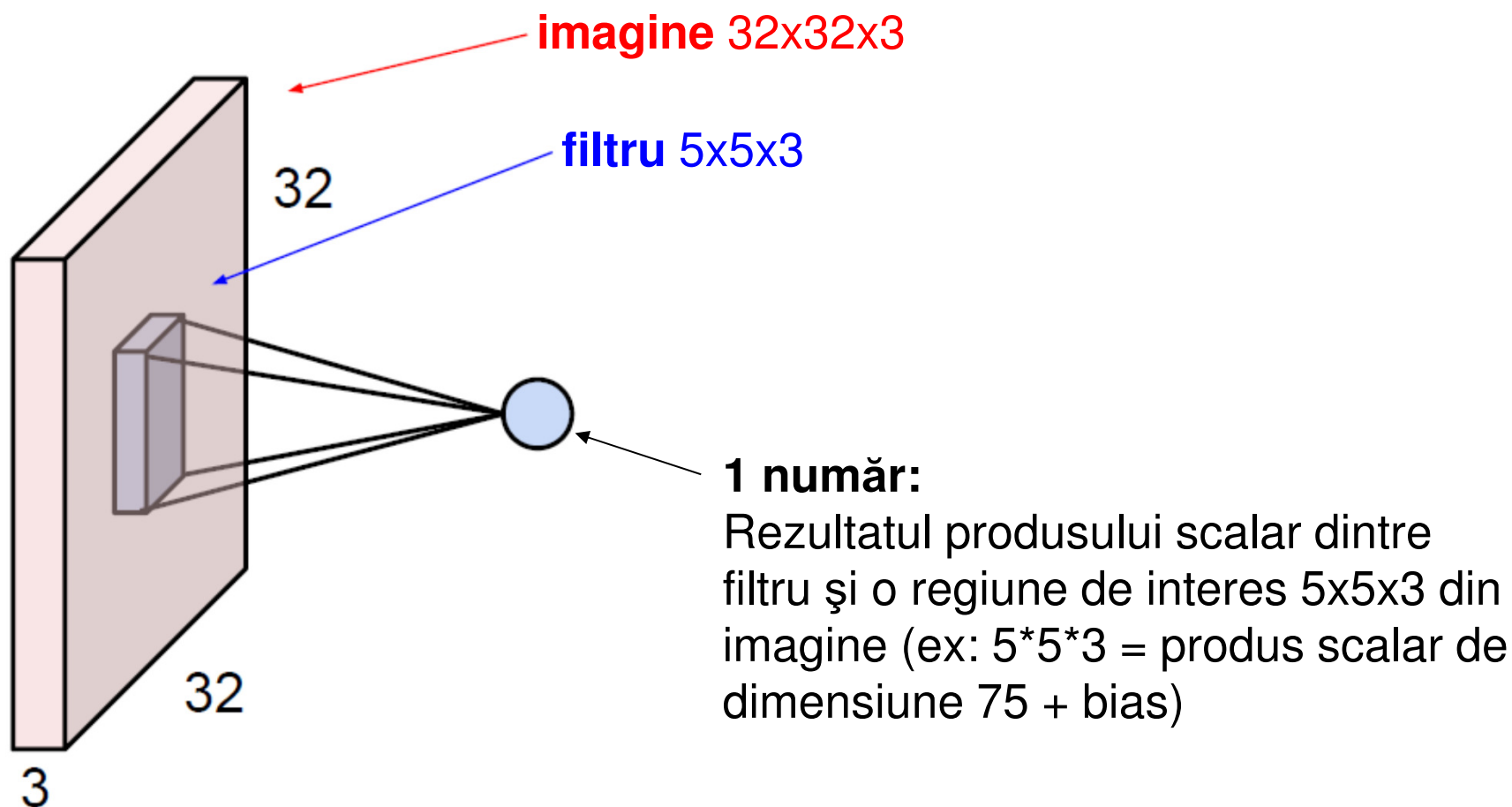
Strat convolutiv

- Este perfect normal să efectuăm convoluții cu filtre de dimensiune 1x1



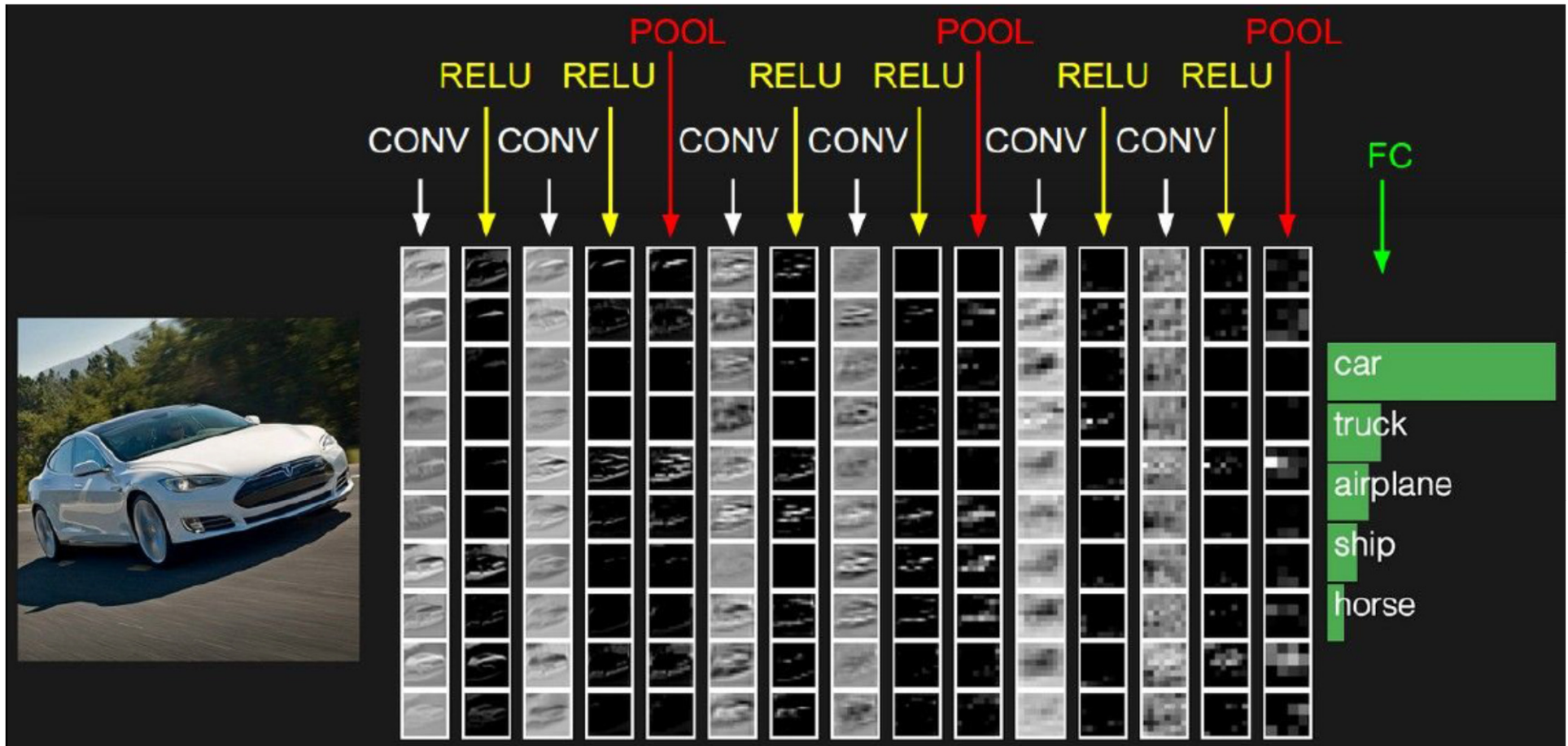


Modelul unui neuron





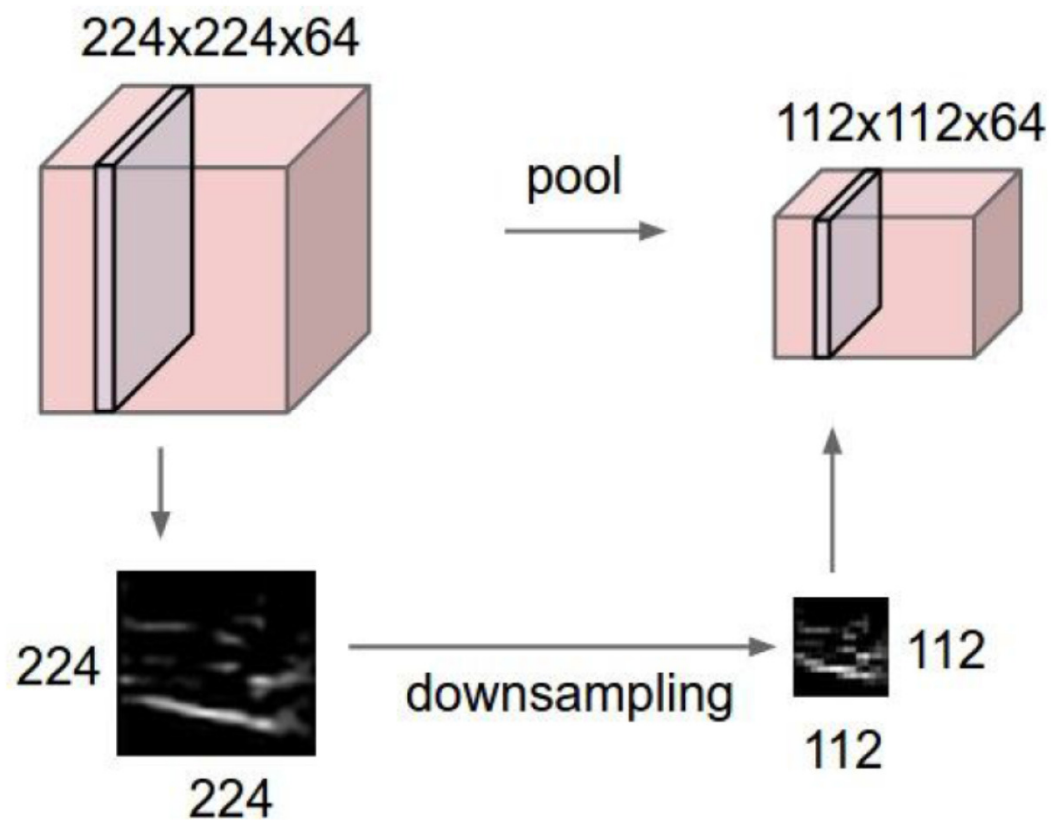
Stratul de pooling





Stratul de pooling

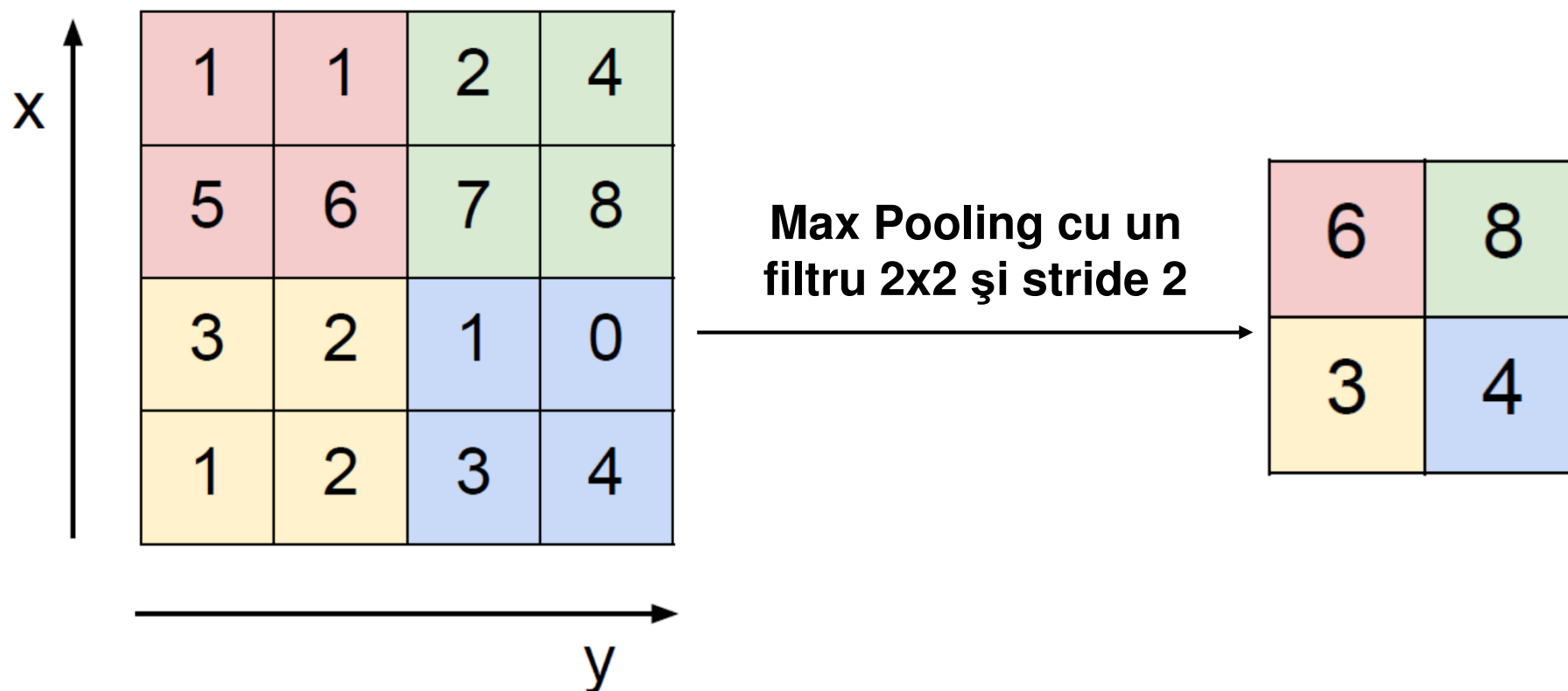
- Modifică volumul într-o reprezentare mai mică și mai flexibilă
- Operează independent de-a lungul fiecărei hărți de activare





Max pooling

- Filtrul de Max Pooling va extrage valoarea maximă din regiunea de interes (ex: 2x2)





Max Pooling: rezumat

- Prelucrează volume de dimensiune $W_1 \times H_1 \times D_1$
- Necesită doi hiperparametrii:
 - Dimensiunea spațială F a filtrelor
 - Stride S
- Produce un volum de dimensiune $W_2 \times H_2 \times D_2$, unde:

$$W_2 = \frac{W_1 - F}{S} + 1$$

$$H_2 = \frac{H_1 - F}{S} + 1$$

$$D_2 = D_1$$

Parametrii comuni:

$$F = 2, S = 2$$

$$F = 3, S = 2$$

- Nu există parametrii (ponderi) θ , din moment ce calculează o funcție fixă a intrării
- Nu este obișnuit să se folosească zero-padding pentru straturi de Max Pooling



Stratul conectat complet (FC)

- FC: Fully Connected
- Conține neuroni ce sunt conectați la întregul volum de intrare, ca și o rețea neurală clasică

