

Convolutional Neural Networks (CNN)

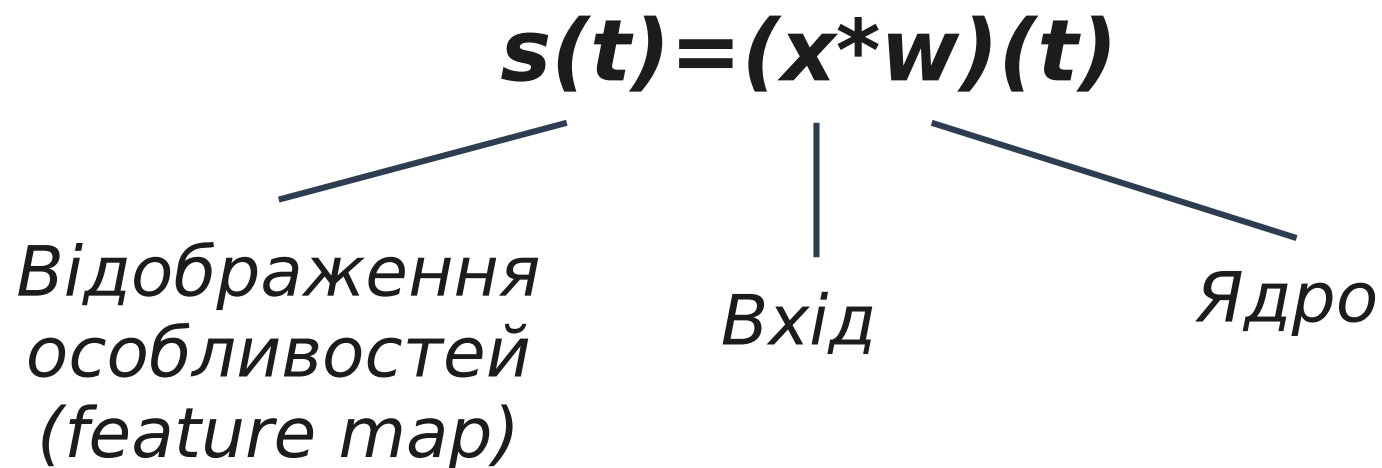
Котенко А.
ФВЕ

Визначення

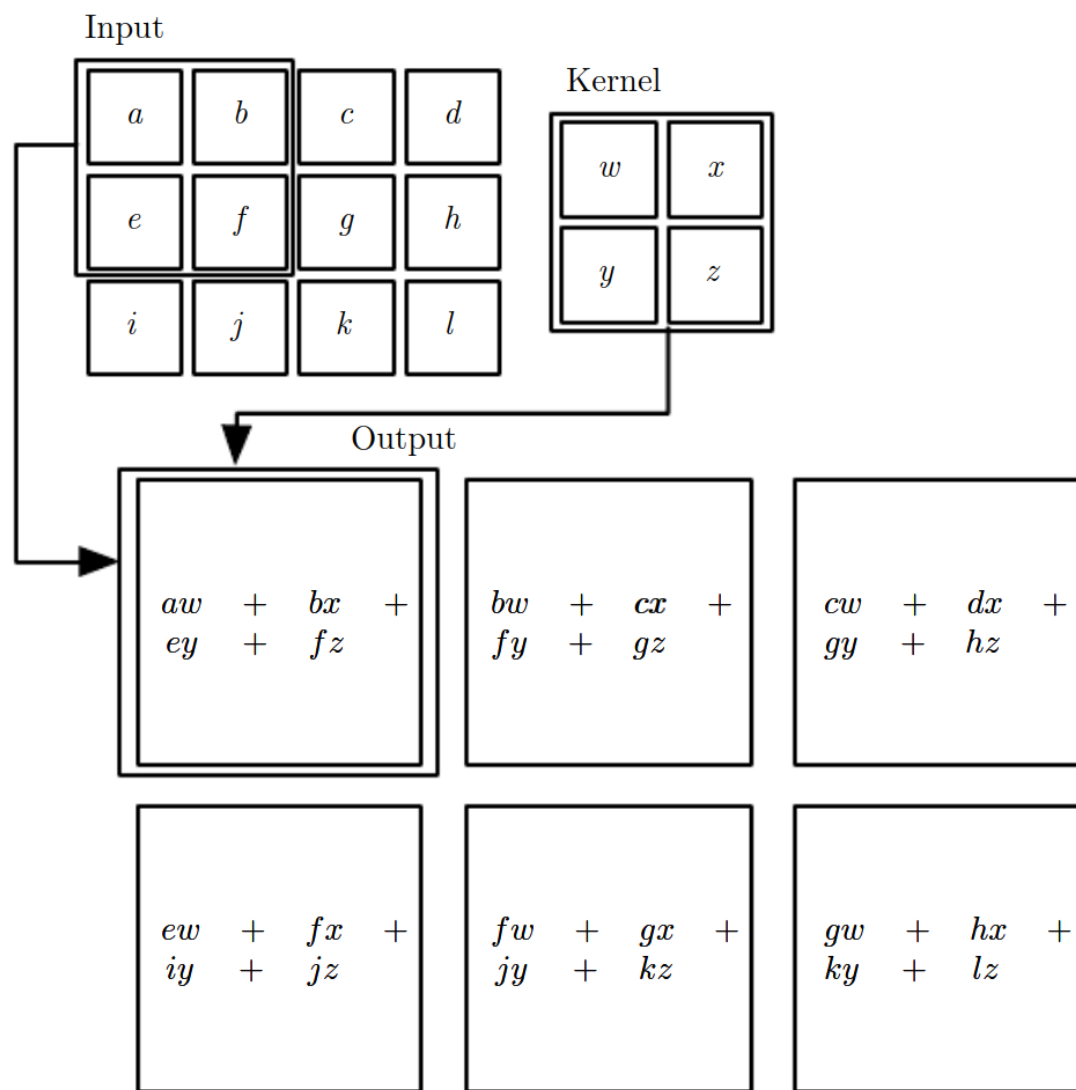
Конволюційні (згорткові) мережі - це нейронні мережі, які використовують згортку замість загальної матриці множення принаймні в одному з їх шарів.

- I. Goodfellow "Deep Learning"

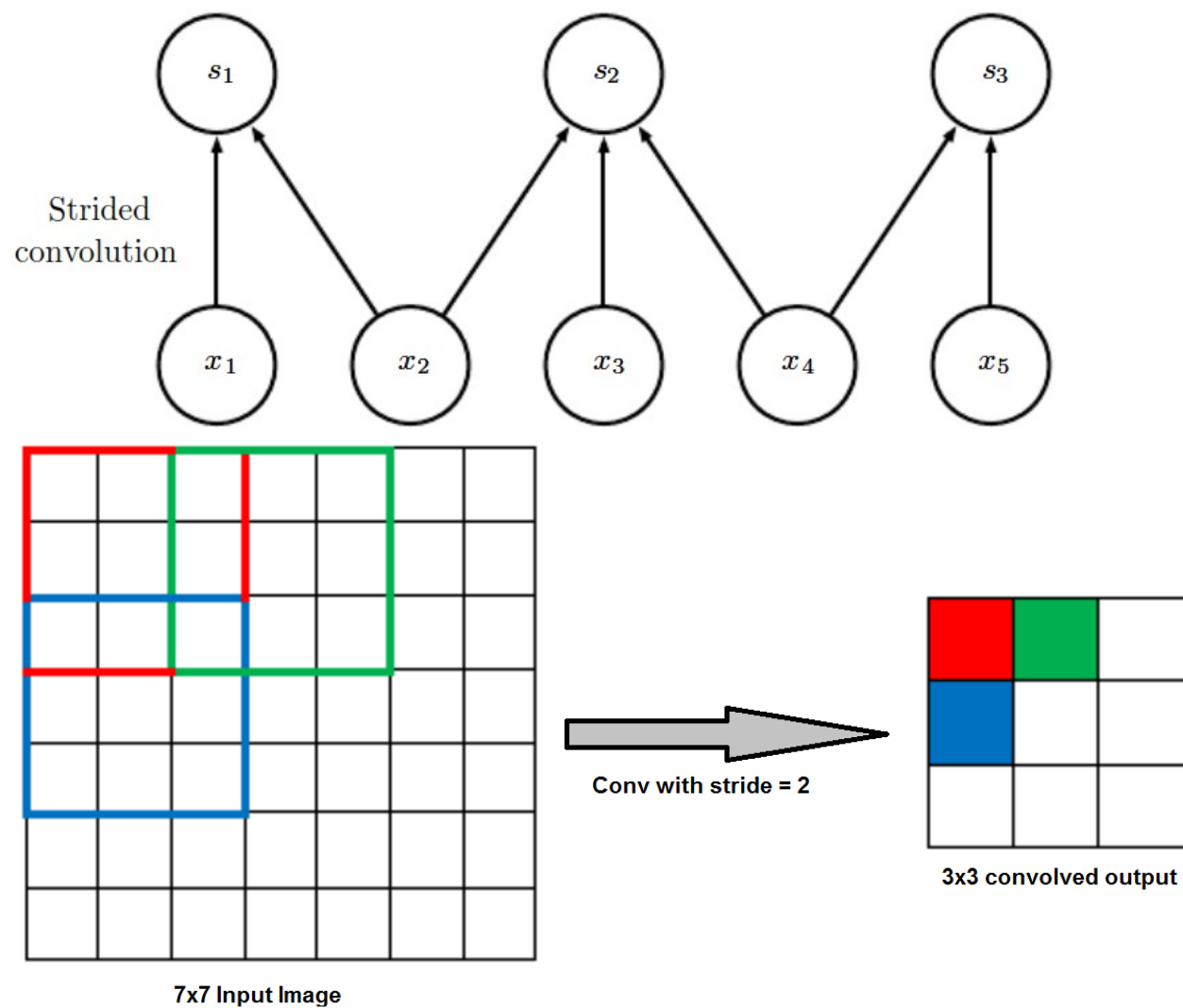
Визначення



Згортка



Згортка



Мотивація

Згорткові мережі вперше набули широкого вжитку при розв'язанні задачі розпізнавання зображень.

Для HD зображень повністю зв'язана нейронна мережа потребувала би $\sim O(10^6)$ параметрів на кожен шар мережи.

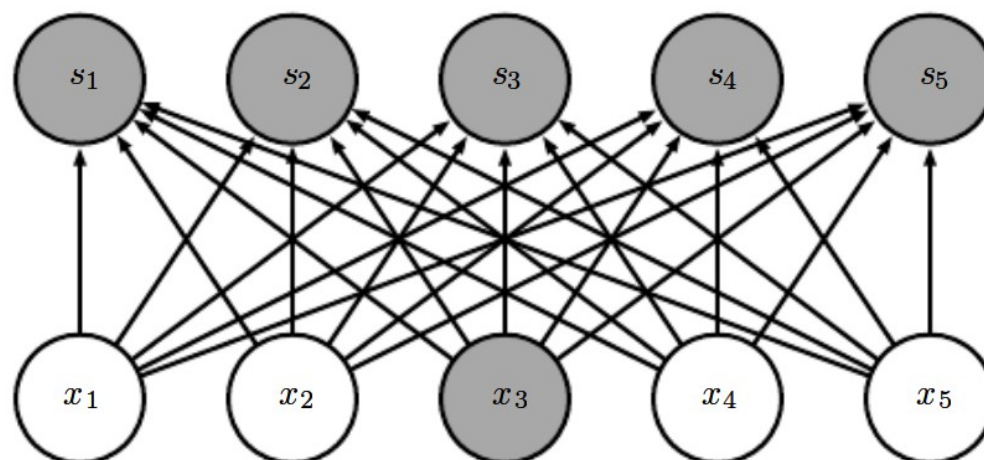
Мотивація

Згорткові мережі втілюють 3 важливі ідеї:

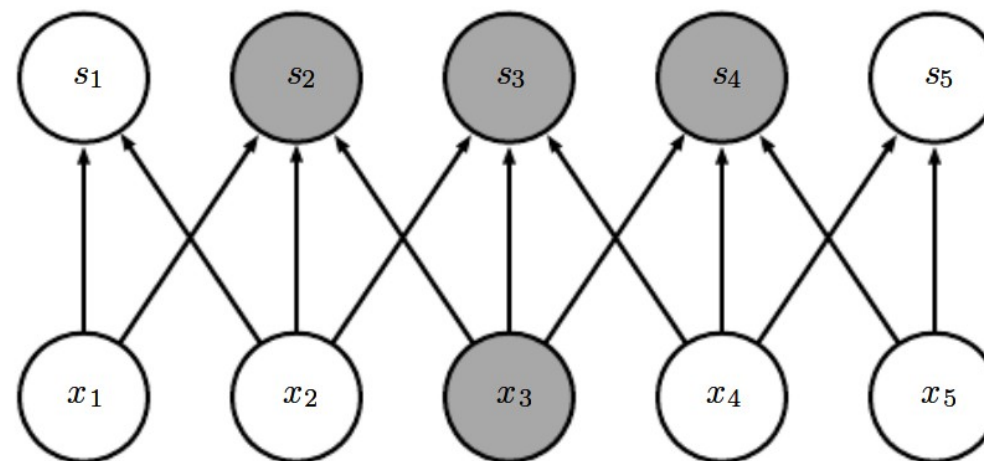
- Розрідженна взаємодія (sparse interactions)
- Спільне використання параметрів (parameter sharing)
- Еквіваріантні представлення (equivariant representations)

Sparse interactions

Матричне множення
(повністю зв'язані
мережі):



Згортка:



Sparse interactions

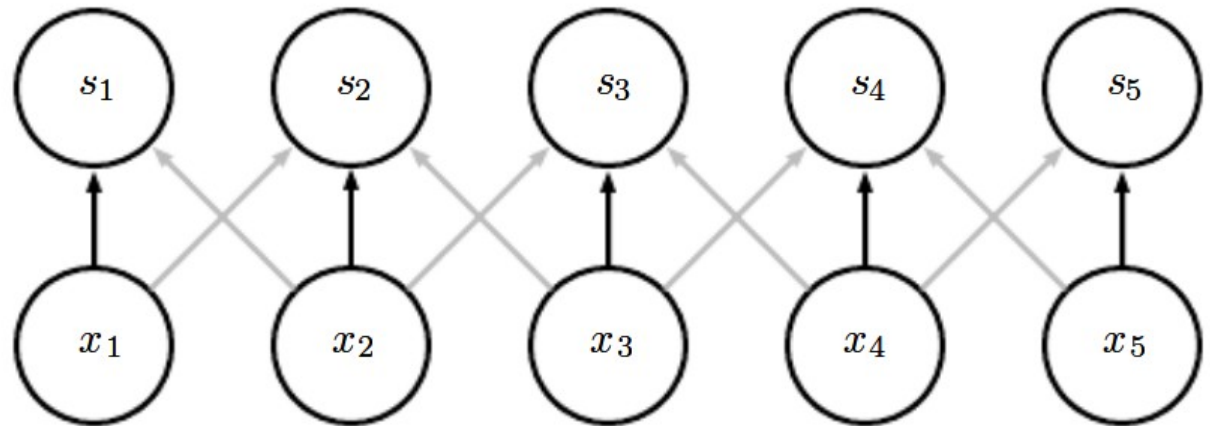
При наявності n входів, та m виходів, матричне множення вимагатиме $(m \times n)$ параметрів. Тому загальний час роботи алгоритму буде $O(m \times n)$.

Зменшивши кількість зв'язків з виходом до k , то такий “розрідженно-зв'язаний” підхід потребуватиме $(k \times n)$ параметрів та відповідно $O(k \times n)$ часу.

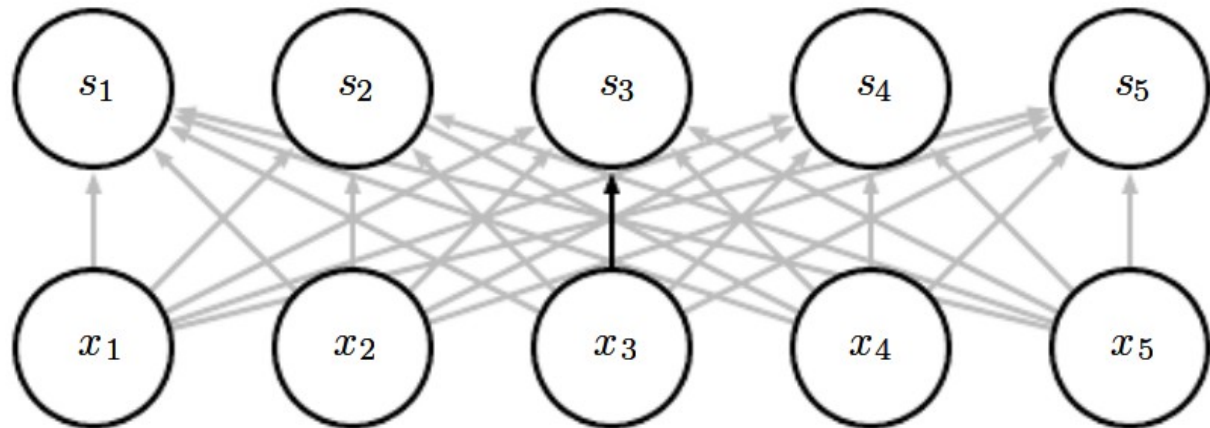
Для багатьох задач можливо отримувати велику ефективність алгоритму з k меншими за m на декілька порядків.

Parameter sharing

Повністю зв'язані мережі:



Згорткові мережі:



Equivariance

Функція $f(x)$ еквіваріантна до перетворення g , якщо $f(g(x)) = g(f(x))$

Конволюція еквіваріантна до операції зсуву.

Конволюція **не** еквіваріантна до операцій масштабування та поворотів.

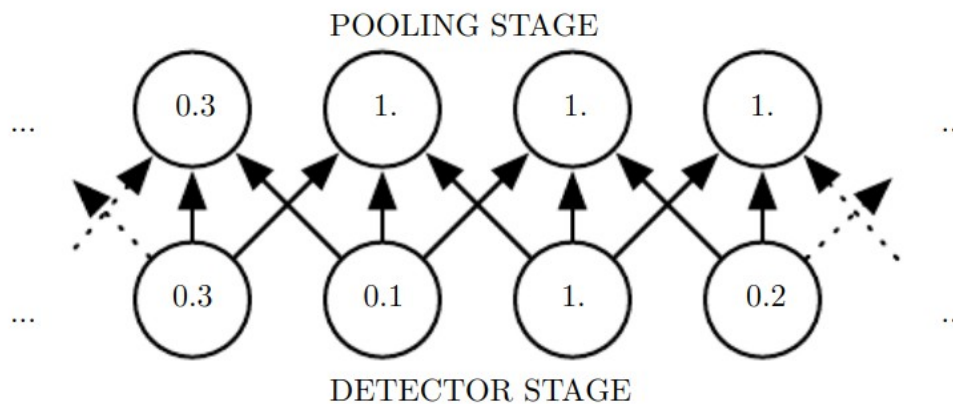
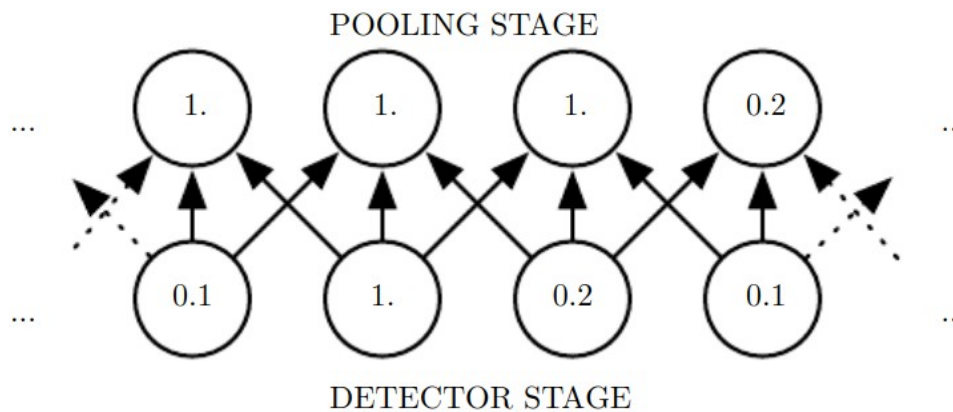
Pooling (Об'єднання)

Pooling шар мережі замінює вихід попереднього шару в певному місці підсумковою статистикою виходів поблизу цього місця.

Найпоширенішим типом є max pooling.

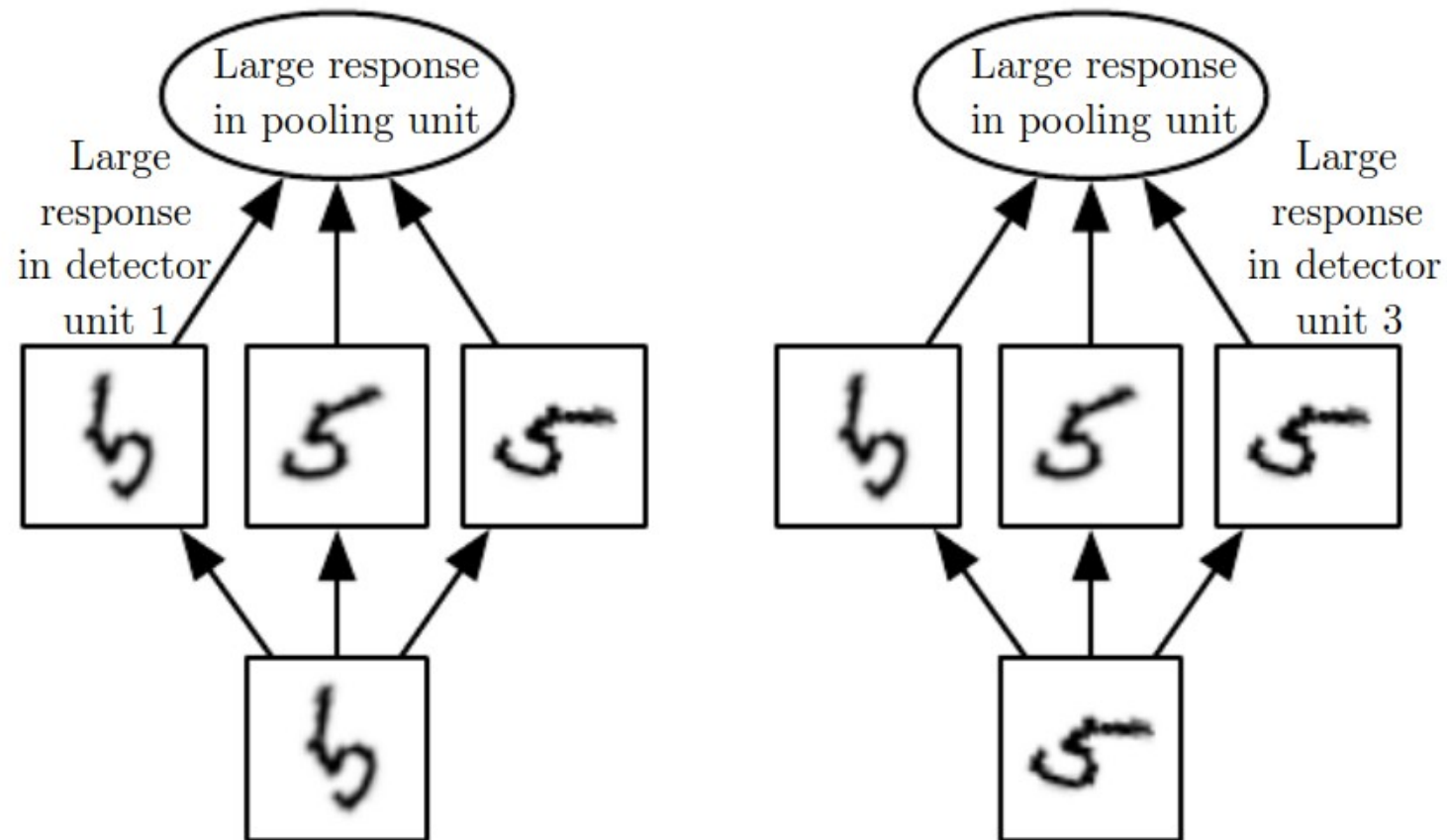
max pooling

Інваріантність до невеликих зсувів:



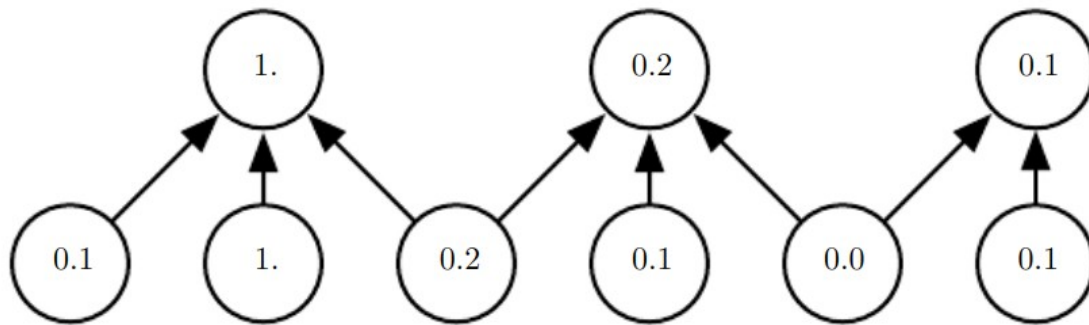
max pooling

Можливість отримання інваріантності до поворотів:

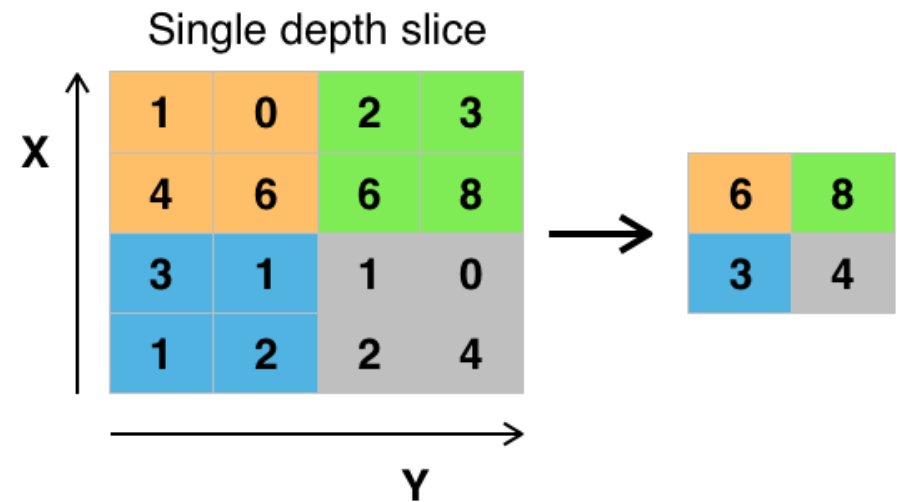


max pooling

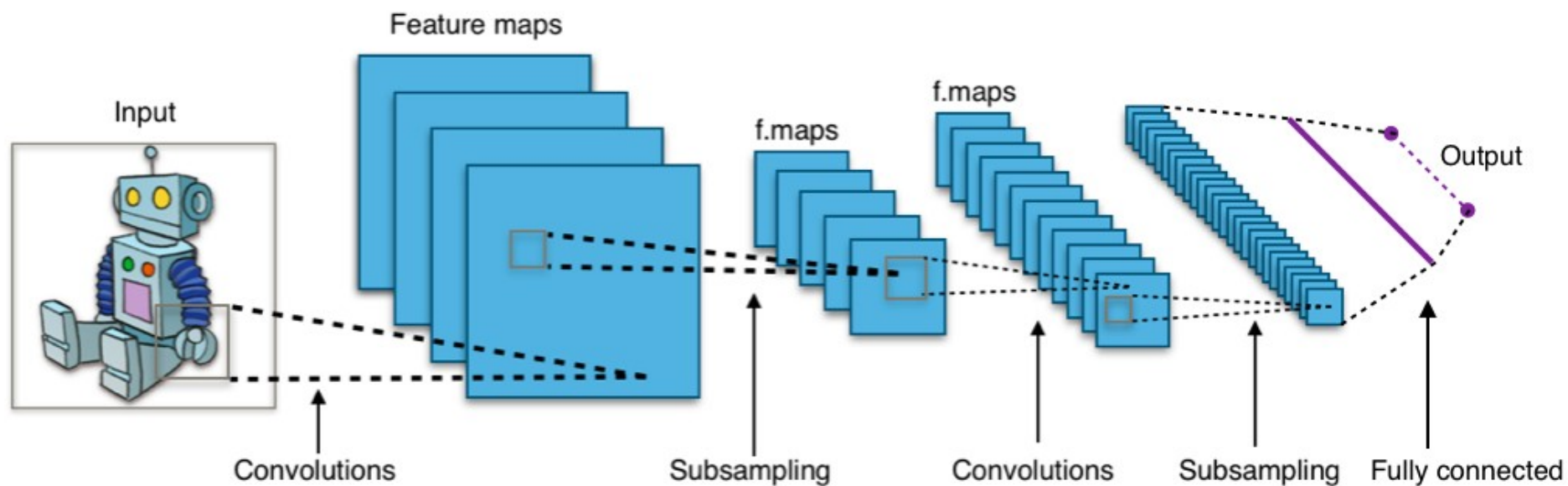
В більшості випадків, роблять pooling шар, що зменшує розмір.



max pooling з кроком 2:



Типова структура CNN



Feature map

Задача розпізнавання 4 знаків на 3x3 зображеннях:

1	-1	-1
-1	1	-1
-1	-1	1

\

1	-1	1
-1	1	-1
1	-1	1

X

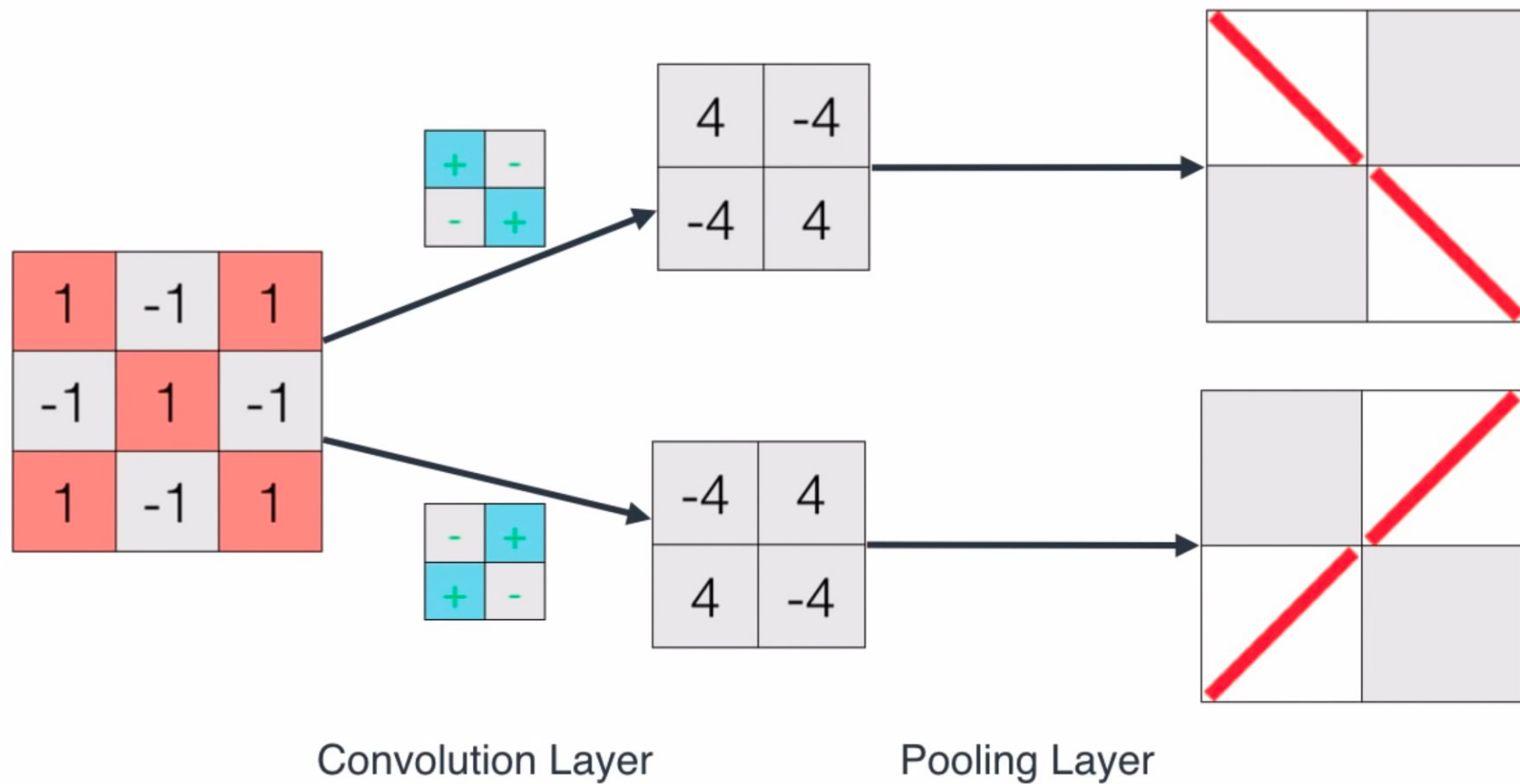
-1	-1	1
-1	1	-1
1	-1	-1

/

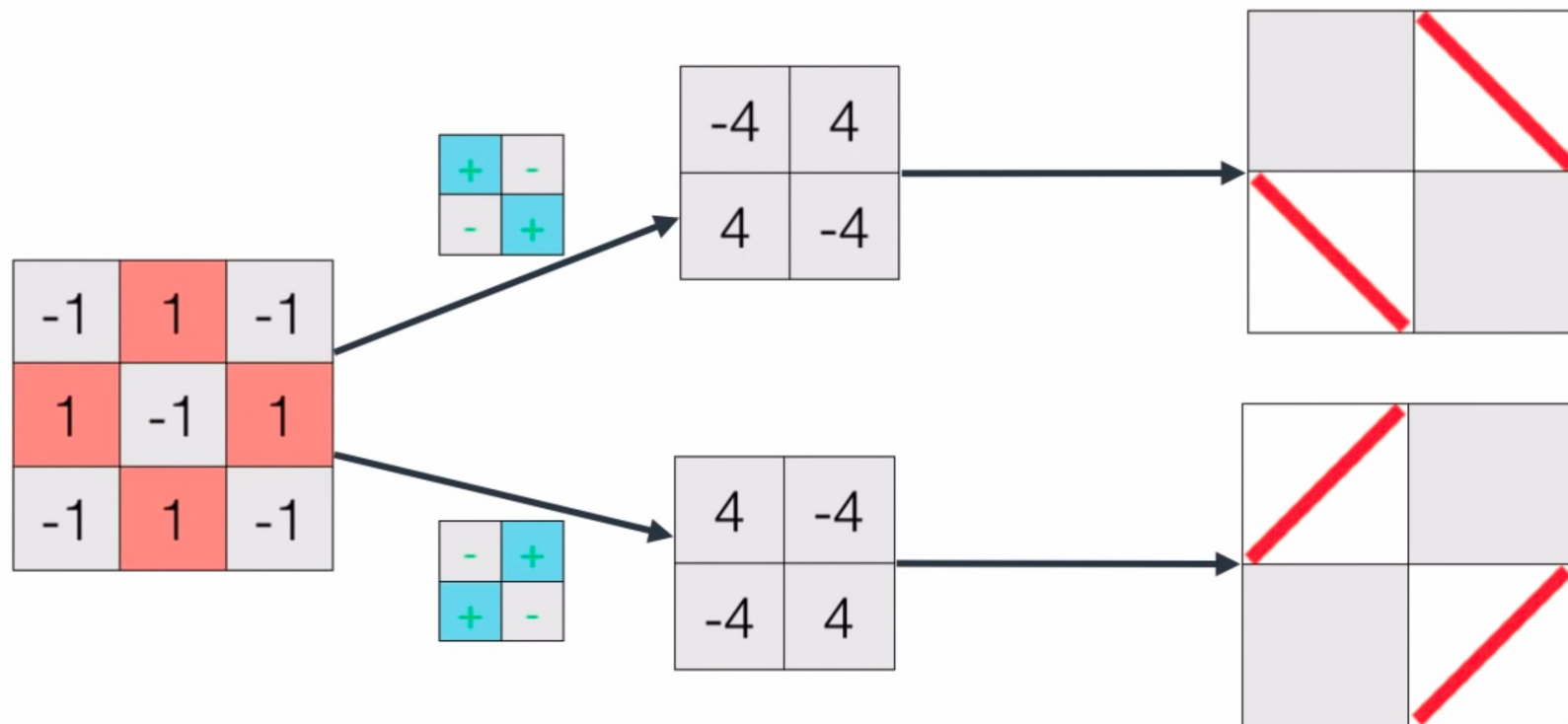
-1	1	-1
1	-1	1
-1	1	-1

O

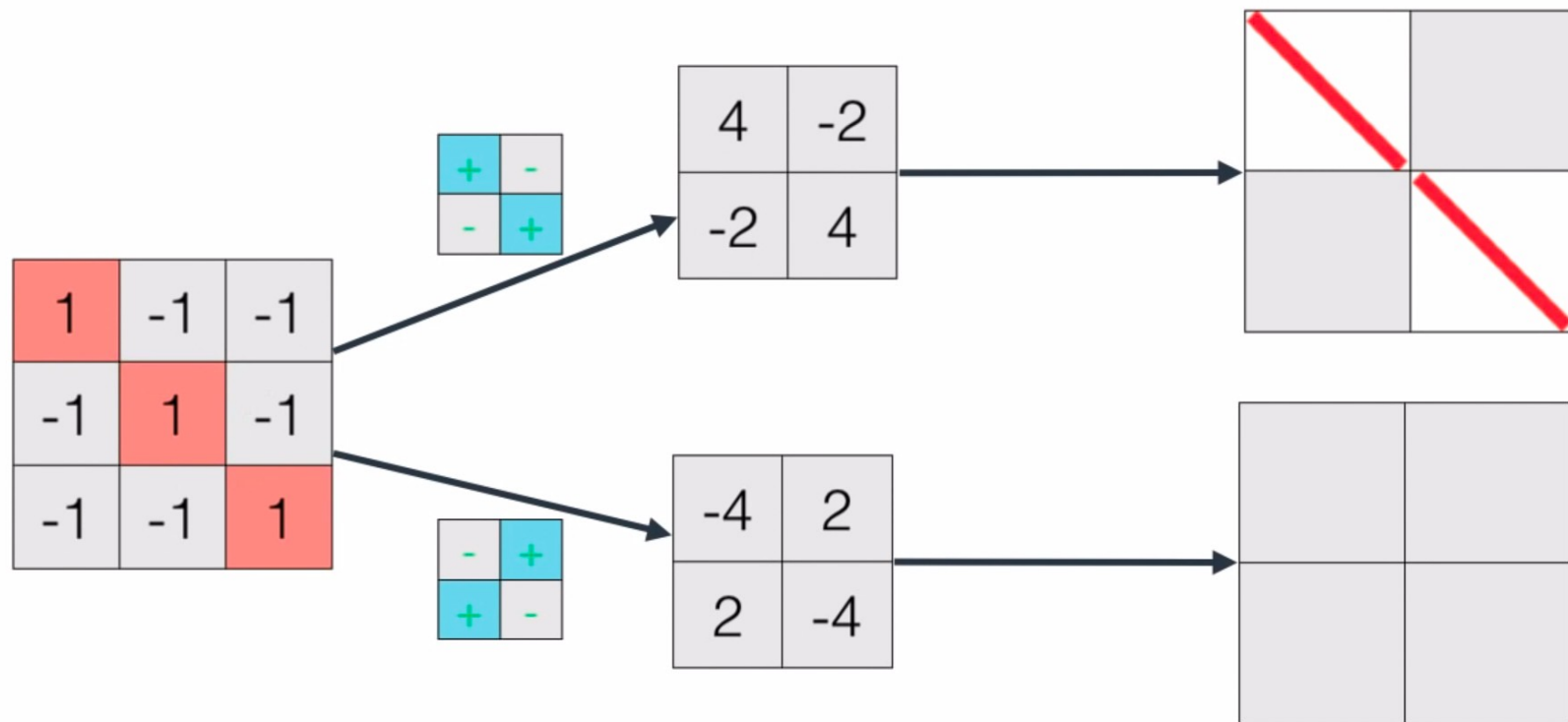
Feature map



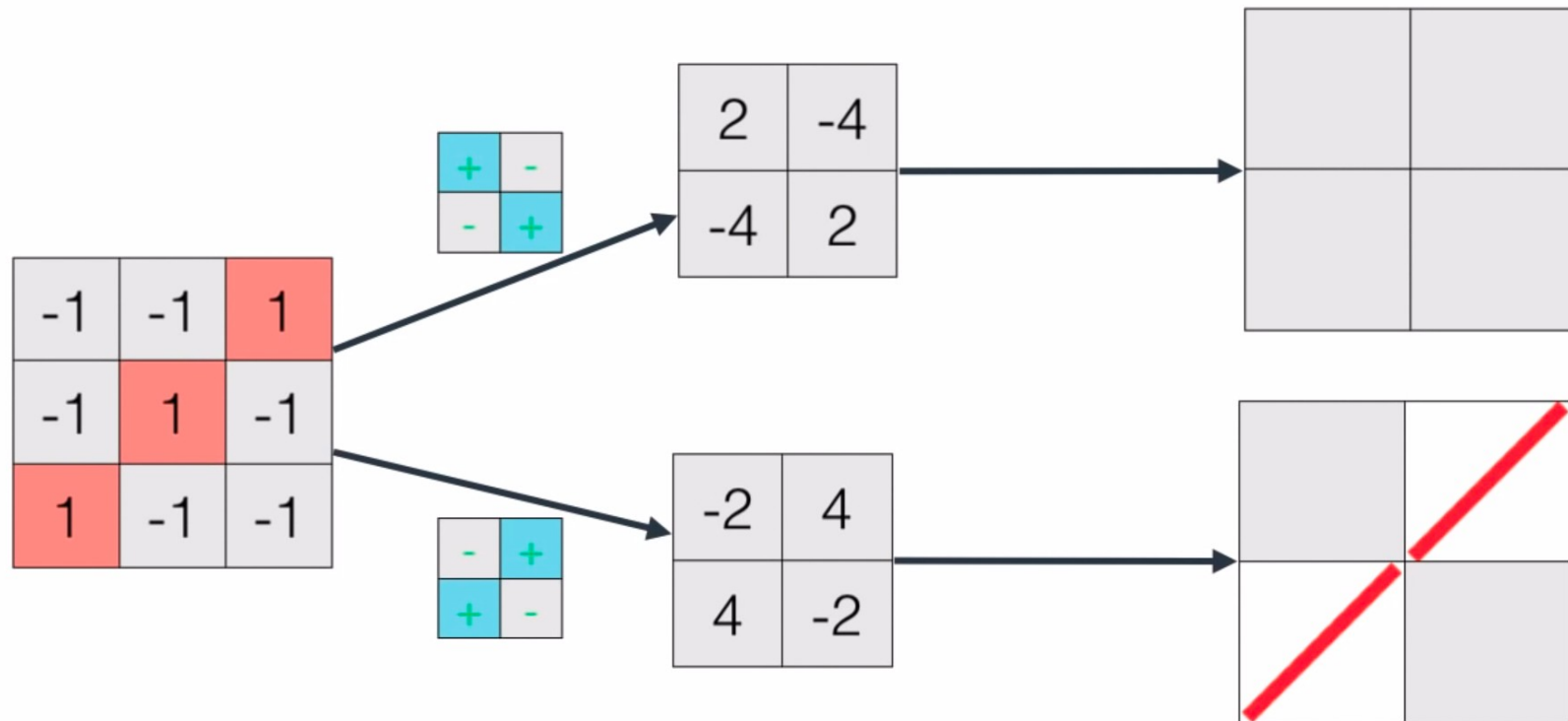
Feature map



Feature map

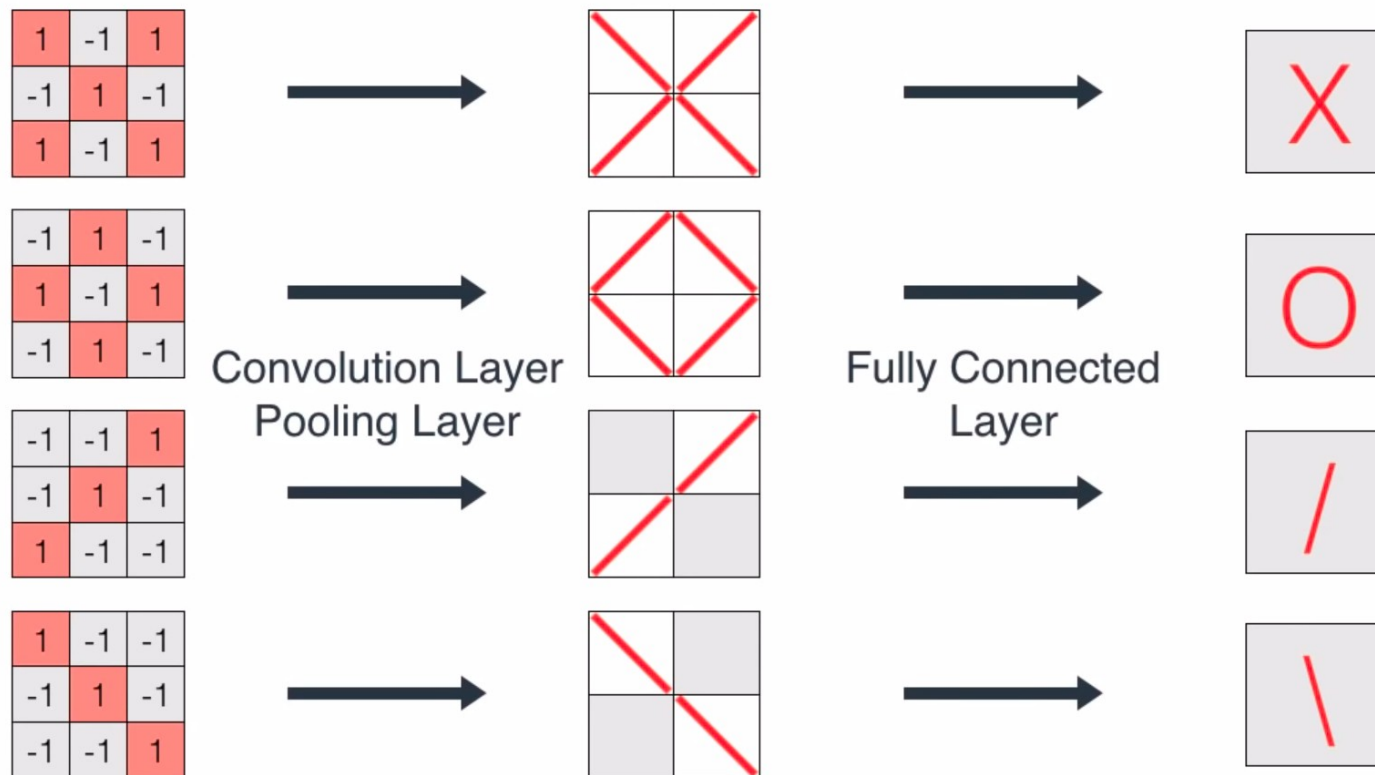


Feature map



Повна схема мережі

Convolutional Neural Network



Спс за увагу!