Deep Learning

Vazgen Mikayelyan

YSU, Krisp

November 5, 2020

Outline

Data Augmentation

- Images:

Data Augmentation

- Images:
  - horizontal flips,

Data Augmentation

- Images:
  - horizontal flips,
  - crops,

Data Augmentation

- Images:
  - horizontal flips,
  - crops,
  - change contrast of colours,

Data Augmentation

- Images:
  - horizontal flips,
  - crops,
  - change contrast of colours,
  - add noise.
- Audios:

Data Augmentation

- Images:
  - horizontal flips,
  - crops,
  - change contrast of colours,
  - add noise.
- Audios:
  - change speed,

Data Augmentation

- Images:
  - horizontal flips,
  - crops,
  - change contrast of colours,
  - add noise.
- Audios:
  - change speed,
  - change pitch,

Data Augmentation

- Images:
  - horizontal flips,
  - crops,
  - change contrast of colours,
  - add noise.
- Audios:
  - change speed,
  - change pitch,
  - add noise.

Outline

What is convolution?

Definition 1

Convolution of the functions $f, g : \mathbb{R} \to \mathbb{R}$ is defined as the integral of the product of the two functions after one is reversed and shifted:

$$(f * g)(t) =: \int_{-\infty}^{+\infty} f(x)\, g(t - x)\, dx.$$

What is convolution?

Definition 1

Convolution of the functions $f, g : \mathbb{R} \to \mathbb{R}$ is defined as the integral of the product of the two functions after one is reversed and shifted:

$$(f * g)(t) =: \int_{-\infty}^{+\infty} f(x)\, g(t - x)\, dx.$$

It easy to see that $f * g = g * f$.

What is convolution?

Definition 1

Convolution of the functions $f, g : \mathbb{R} \to \mathbb{R}$ is defined as the integral of the product of the two functions after one is reversed and shifted:

$$(f * g)(t) =: \int_{-\infty}^{+\infty} f(x)\, g(t - x)\, dx.$$

It easy to see that $f * g = g * f$.

Definition 2

Convolution of the sequences of real numbers $\{f_n\}_{n=-\infty}^{+\infty}, \{g_n\}_{n=-\infty}^{+\infty}$ is the following sequence:

$$z_n =: \sum_{k=-\infty}^{+\infty} f_k g_{n-k}.$$

2D Convolution

Definition 3

Convolution of the functions $f, g : \mathbb{R}^2 \to \mathbb{R}^2$ *is the following function:*

$$(f * g)(t, \tau) =: \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y)\, g(t - x, \tau - y)\, dxdy.$$

2D Convolution

Definition 3

Convolution of the functions $f, g : \mathbb{R}^2 \to \mathbb{R}^2$ is the following function:

$$(f * g)(t, \tau) =: \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y)\, g(t - x, \tau - y)\, dxdy.$$

It easy to see that $f * g = g * f$.

2D Convolution

Definition 3

Convolution of the functions $f, g : \mathbb{R}^2 \to \mathbb{R}^2$ is the following function:

$$(f * g)(t, \tau) =: \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y)\, g(t - x, \tau - y)\, dxdy.$$

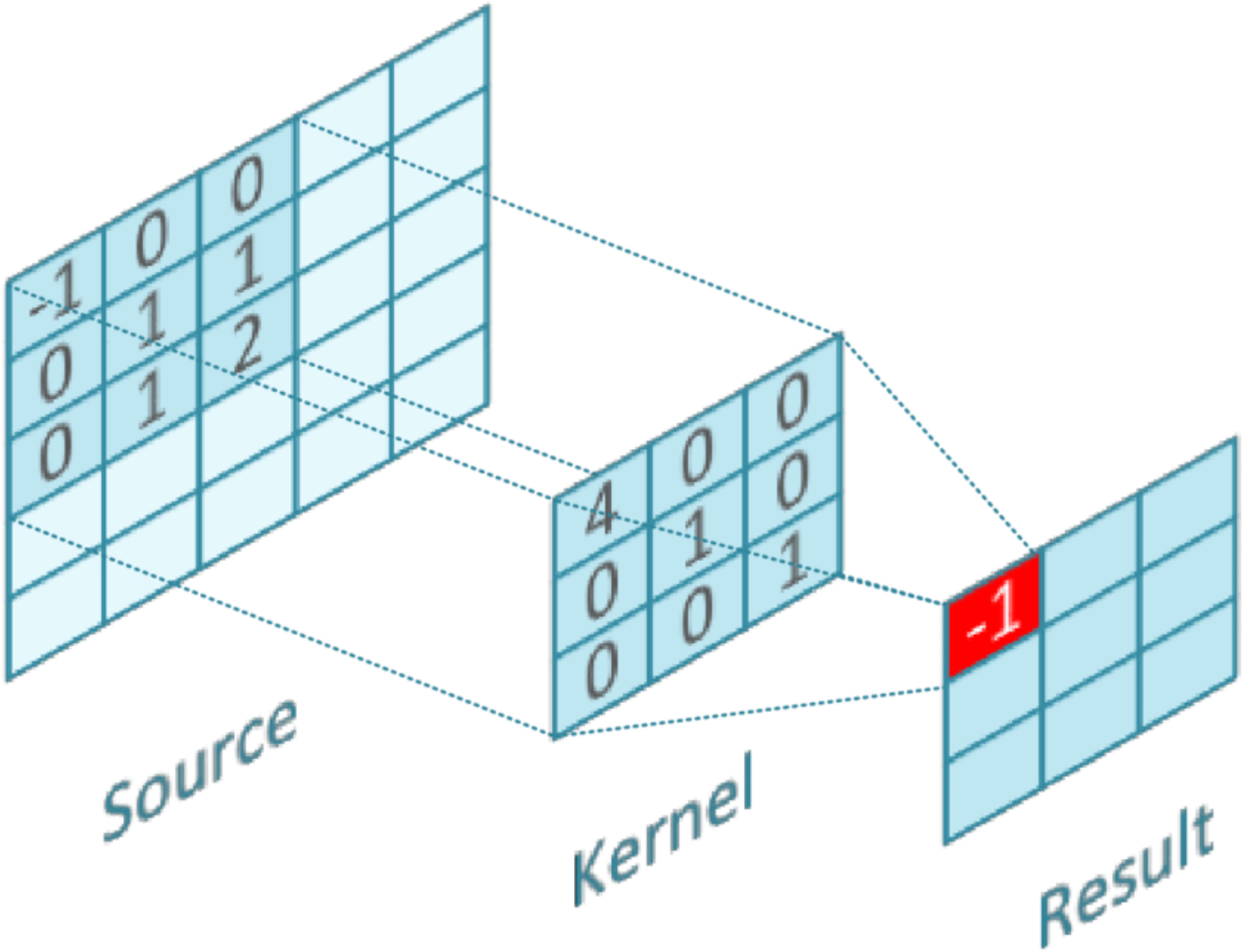
It easy to see that $f * g = g * f$.

Definition 4

Let $f(x, y)$ is an image and $w(s, t)$ is a kernel where $s \in [-a, a], t \in [-b, b]$, $x, y, s, t, a, b, c, d \in \mathbb{Z}$. The convolution between kernel w and image f is the following function

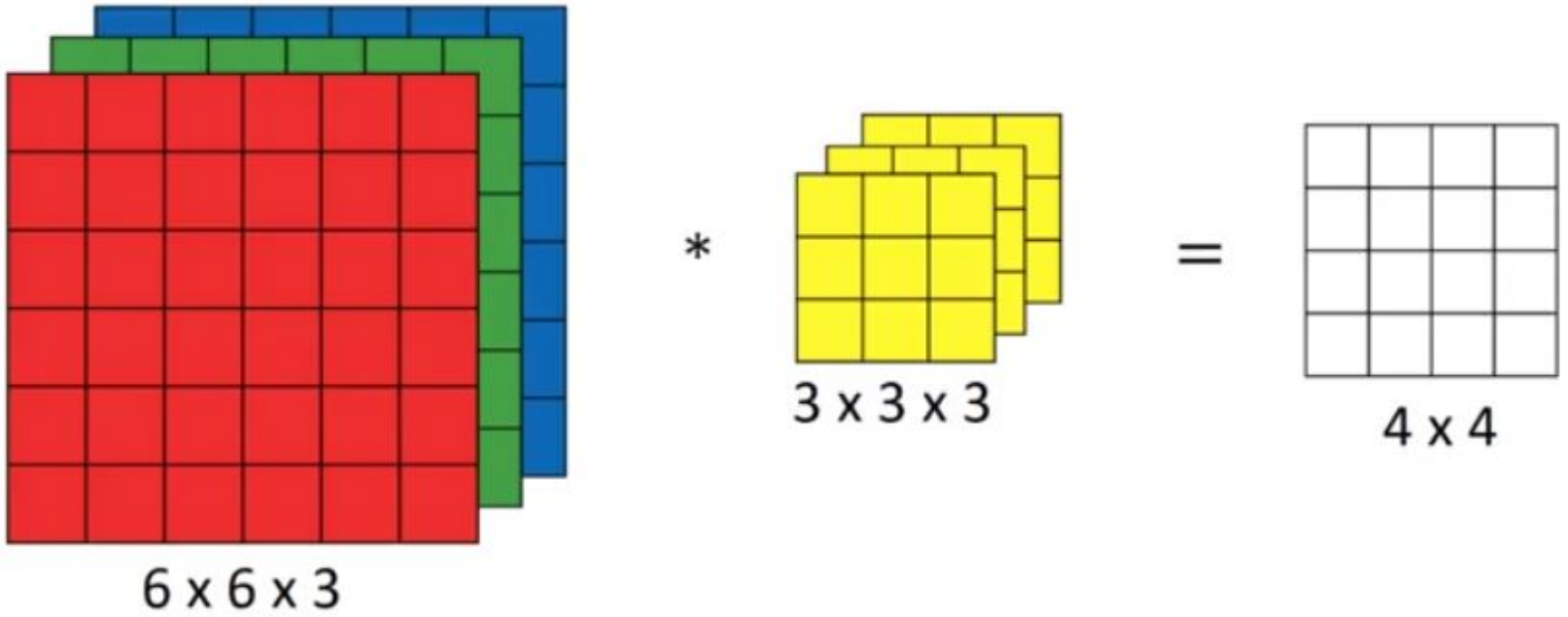
$$(w * f)(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s, t)\, f(x - s, y - t)$$

2D Convolution

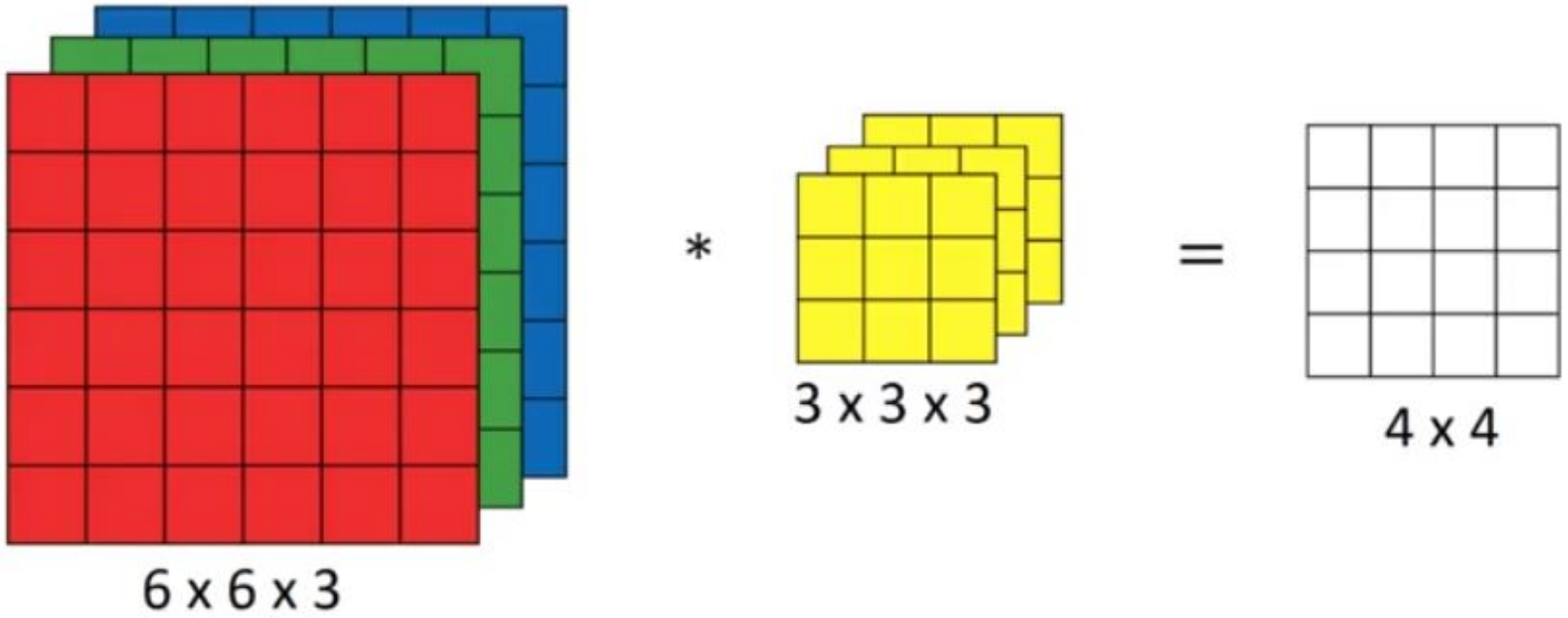


Convolution on RGB Images

Convolution on RGB Images

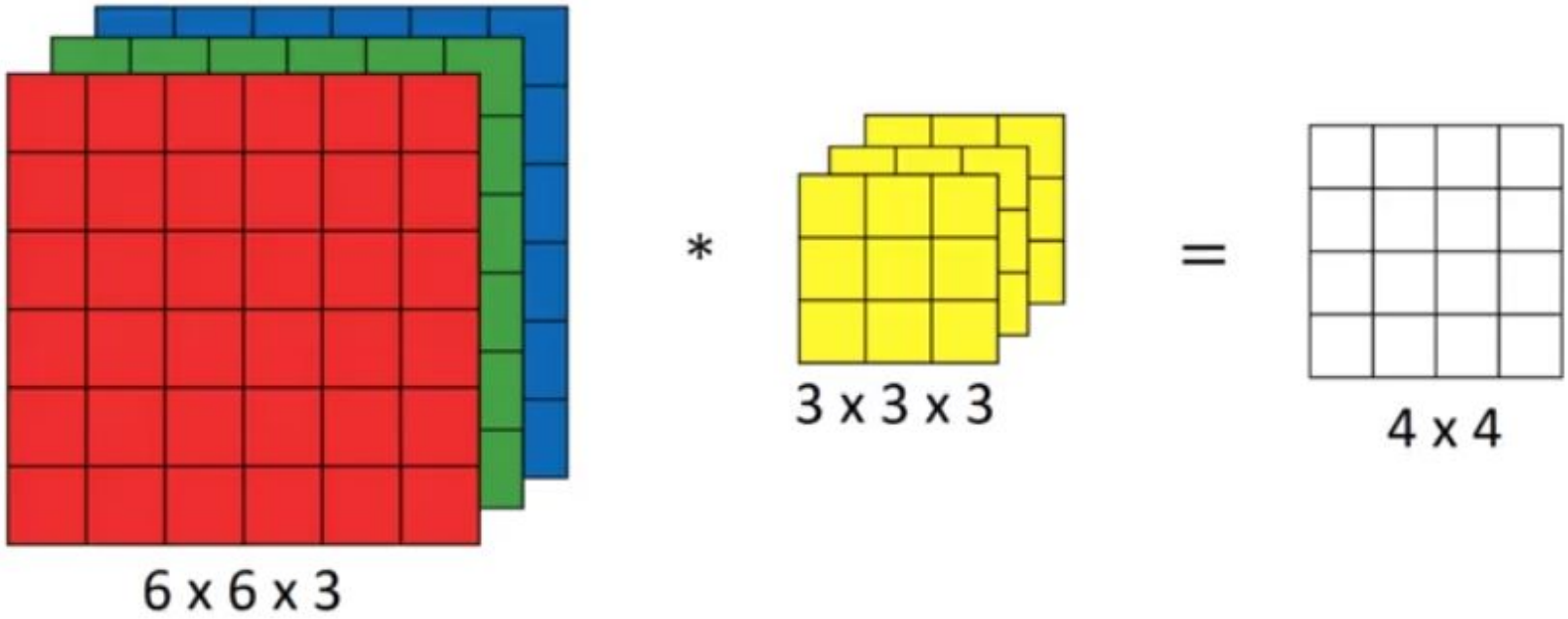


Convolution on RGB Images



- How many dimensions has convolution kernel in general?

Convolution on RGB Images



- How many dimensions has convolution kernel in general?
- What does 1×1 convolution do?

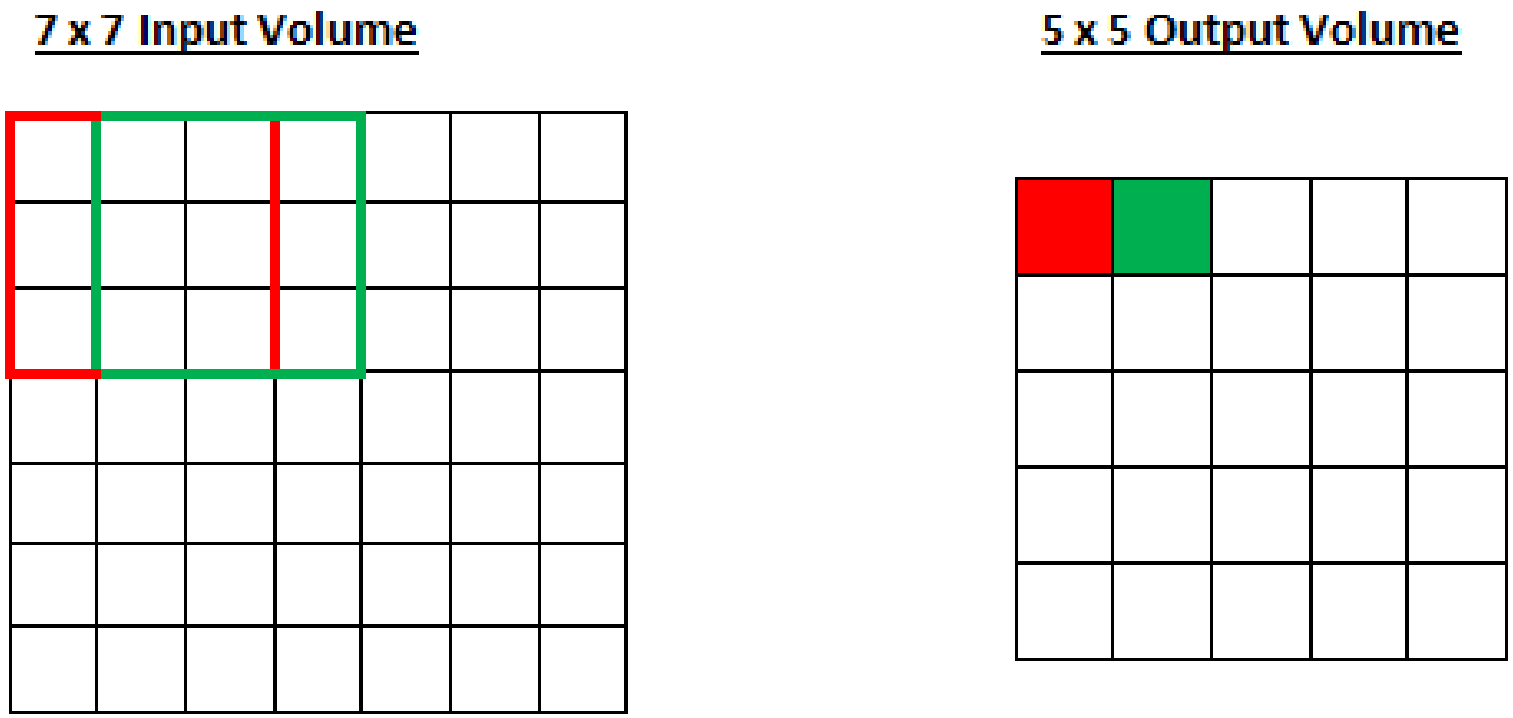
Valid and Same Convolution

- Padding = Same: means the input image ought to have zero padding so that the output in convolution doesn't differ in size as input.

Valid and Same Convolution

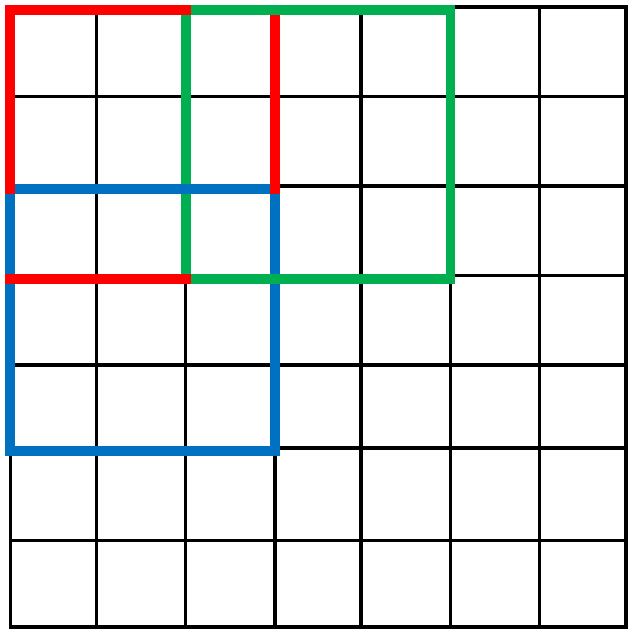
- Padding = Same: means the input image ought to have zero padding so that the output in convolution doesn't differ in size as input.
- Padding = Valid: means we don't add the zero pixel padding around the input matrix, and its like saying, we are ready to loose some information.

Strided Convolution

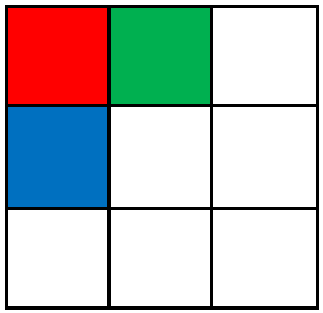


Strided Convolution

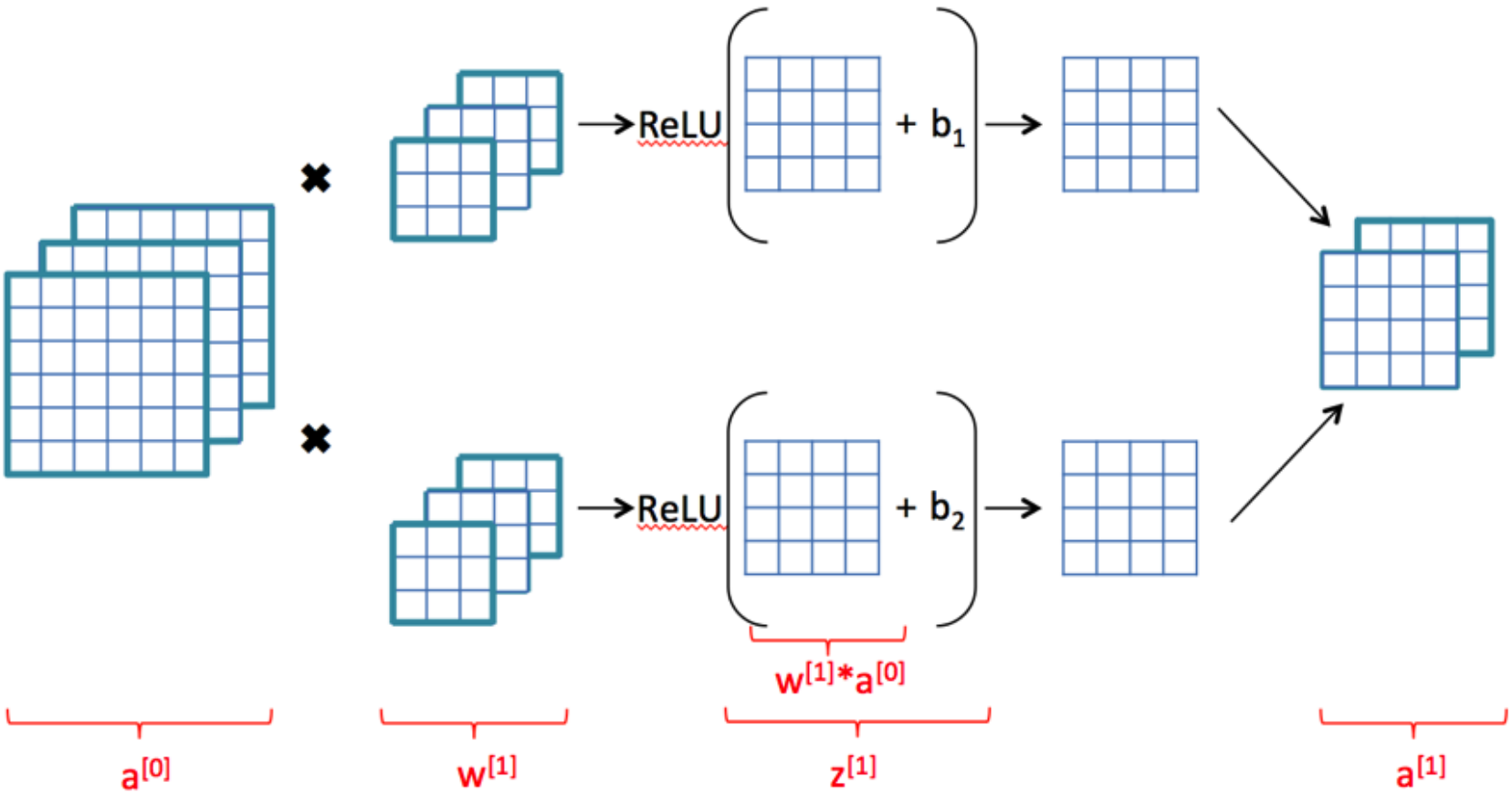
7 x 7 Input Volume



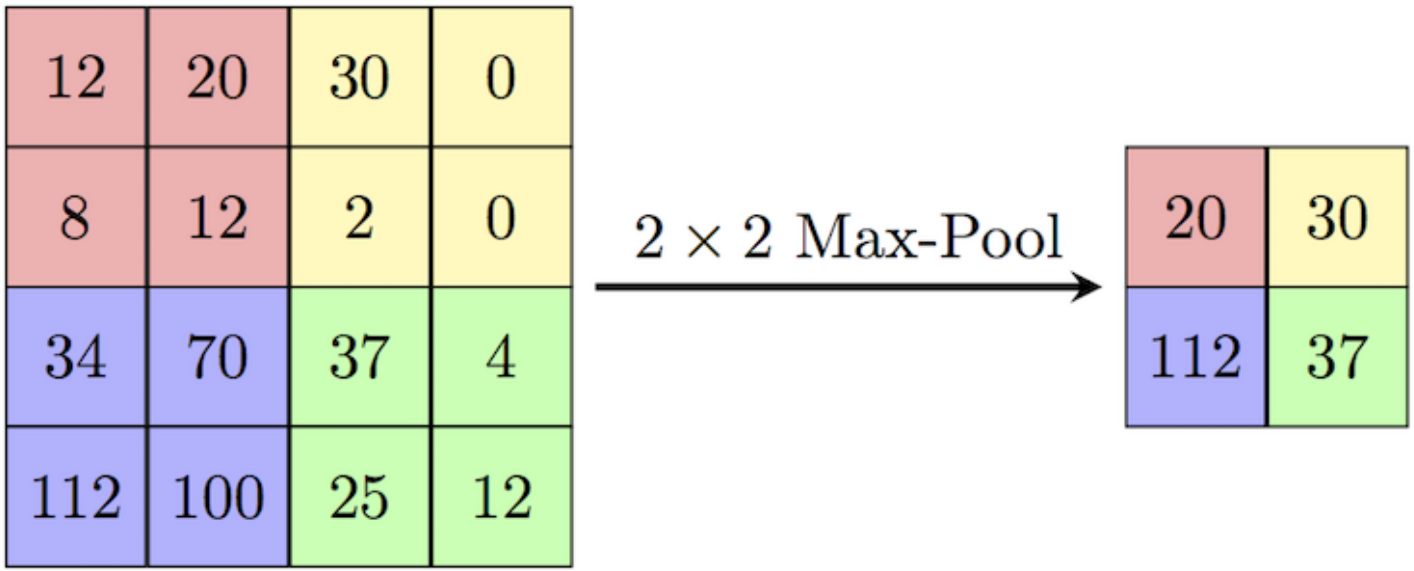
3 x 3 Output Volume



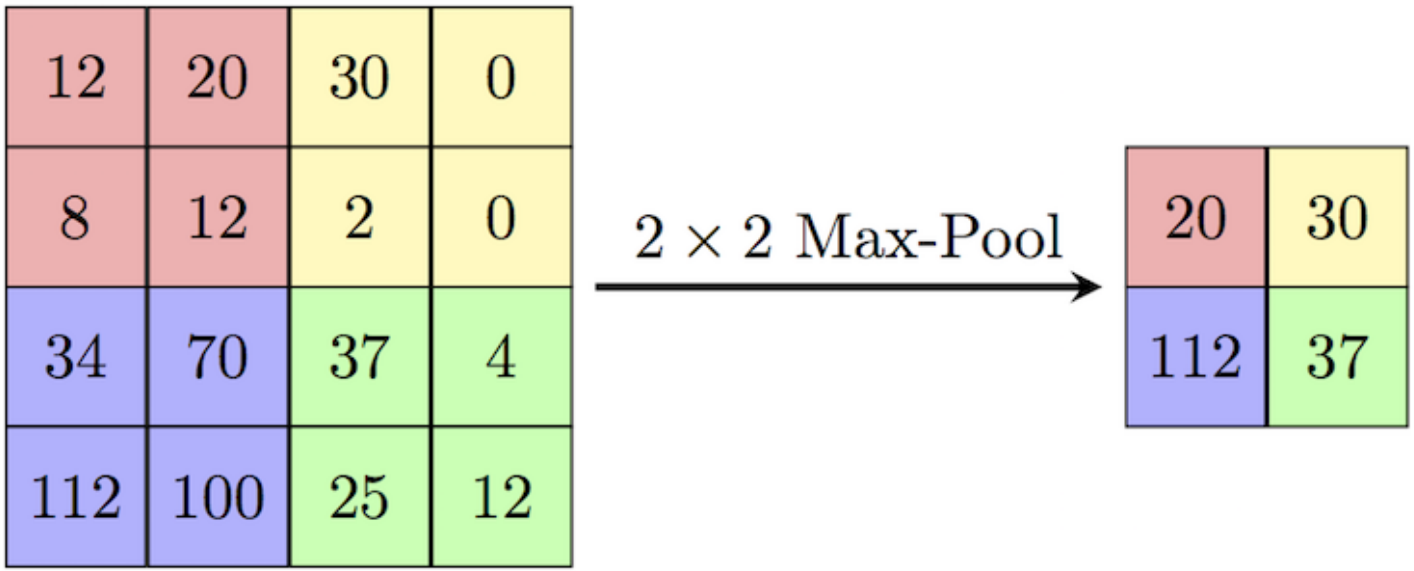
One Layer of CNN



Pooling Layers

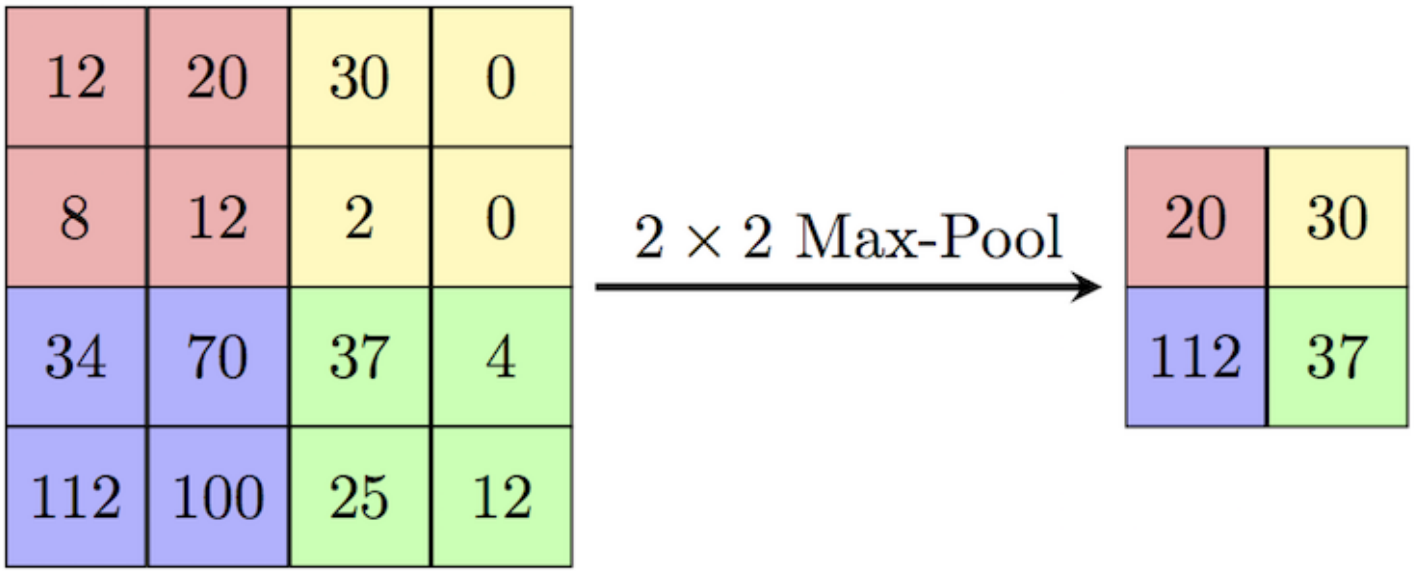


Pooling Layers



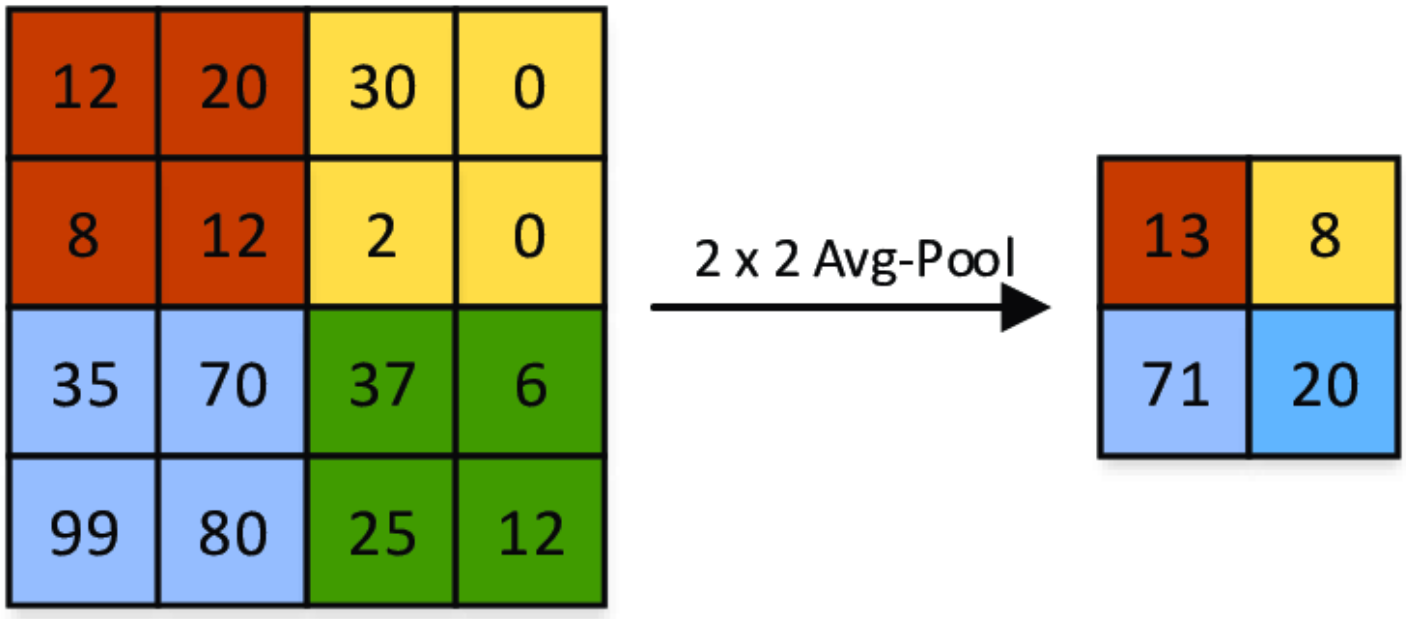
- Number of channels is the same after pooling layer.

Pooling Layers



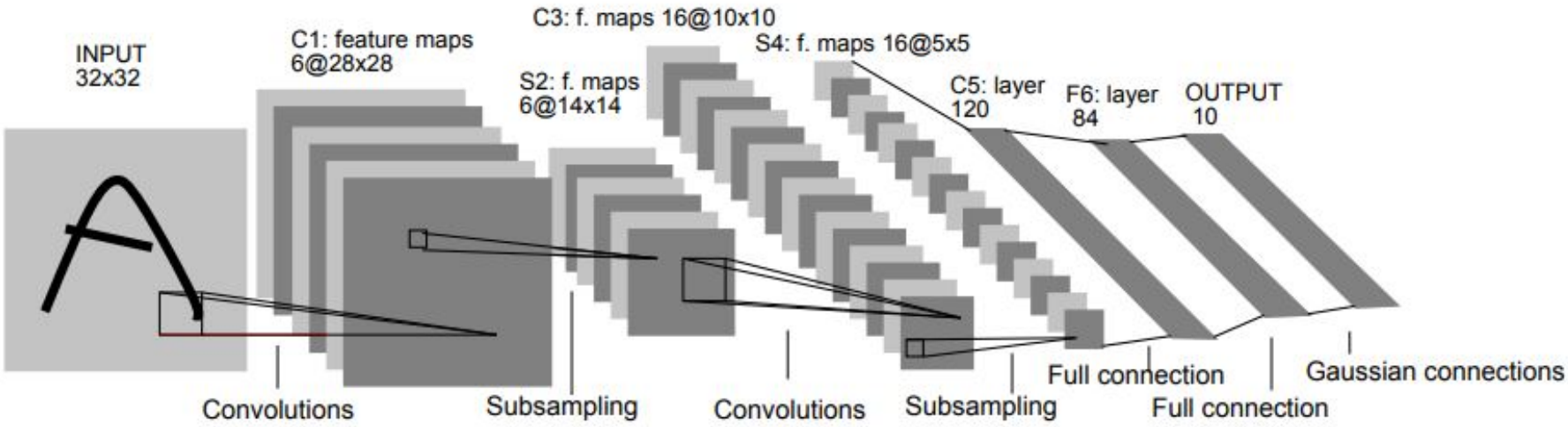
- Number of channels is the same after pooling layer.
- There are not trainable parameters in this layer.

Pooling Layers

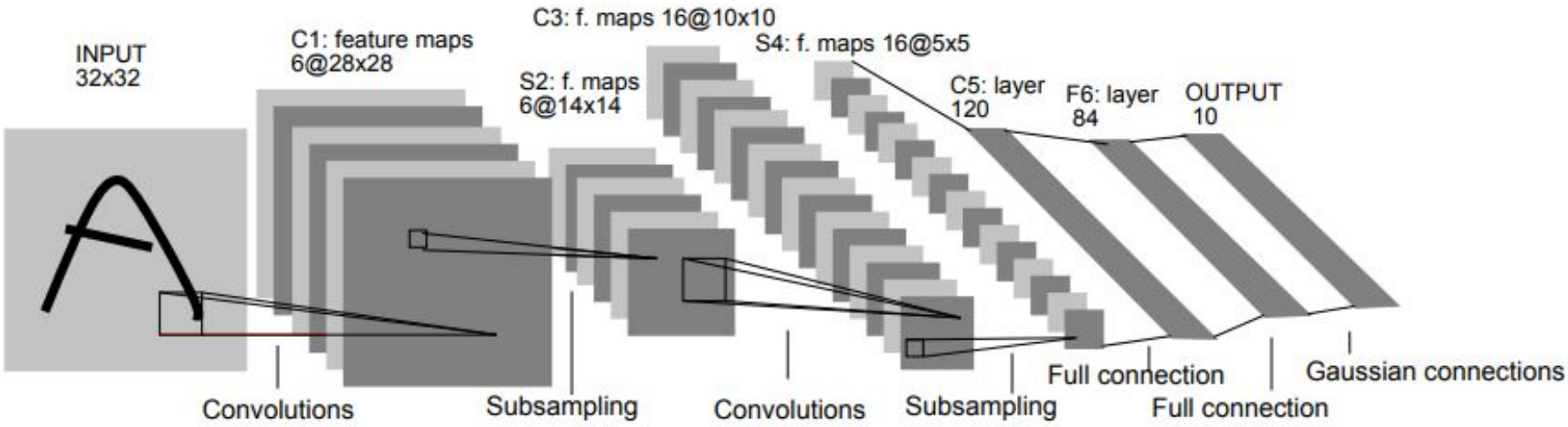


Outline

LeNet-5 (1998)

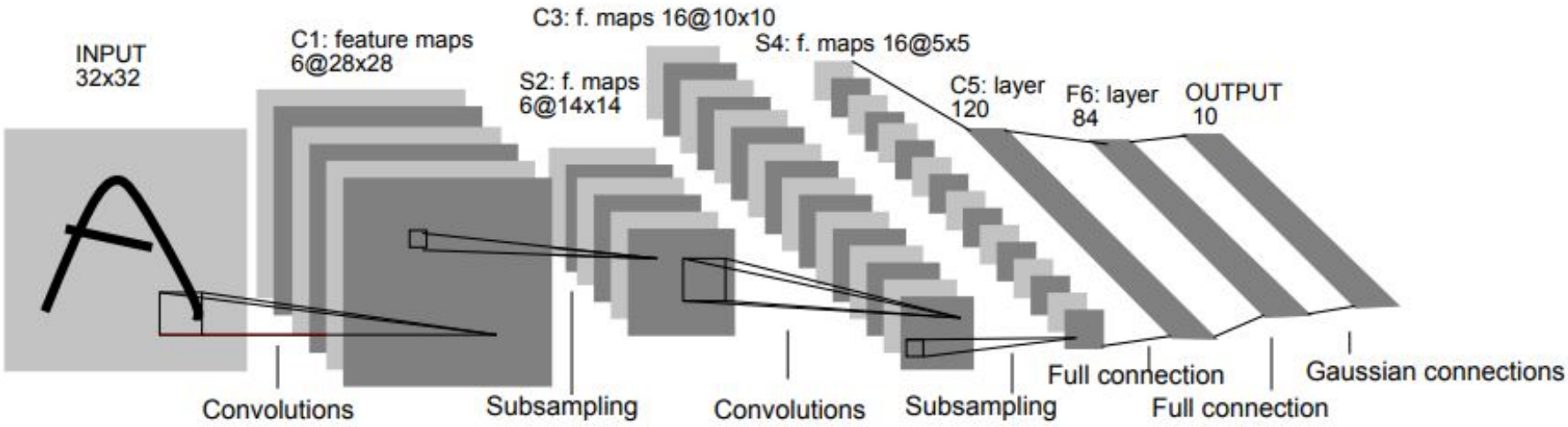

INPUT
32x32
C1: feature maps
6@28x28
S2: f. maps
6@14x14
C3: f. maps 16@10x10
S4: f. maps 16@5x5
C5: layer
120
F6: layer
84
OUTPUT
10
Convolutions
Subsampling
Convolutions
Subsampling
Full connection
Full connection
Gaussian connections

LeNet-5 (1998)



- Activation functions are sigmoids and hyperbolic tangents.

LeNet-5 (1998)



- Activation functions are sigmoids and hyperbolic tangents.
- LeNet-5 has approximately 60k parameters.