Convolutional neural networks

Lecture 7

Changho Suh

September 29, 2021

Convolutional neural networks (CNNs) and history

Recap: Data organization & evaluation

m is a deciding factor for the **ratio** of data split:

4 regimes: Small, middle, large and ultra-large

val set dist. ~ test set dist. ~ target dist.

Cross validation for evaluation:

Choose a hyperparameter that minimizes the average validation loss:

$$\text{val loss} = \frac{\text{val}_1 + \text{val}_2 + \text{val}_3 + \text{val}_4}{4}$$

Recap: Techniques for DNNs

Weight initialization: Xavier's, He's

Training stability: Adam optimizer, BN

Hyperparameter search:

L of layers, # $n^{[\ell]}$ of hidden neurons, activation

Recap: DNNs

Model complexity is very high.

This often leads to the overfitting problem.

This motivates the use of generalization techniques:

Regularization; data augmentation;

early stopping; drop out

Recap: Tensorflow coding

Train-val-test data split:

```
from sklearn.model_selection import train_test_split
```

BN, dropout, activation, early stopping, lr decay:

```
from keras.layers import BatchNormalization, Dropout, ReLU
```

```
from keras.callbacks import EarlyStopping, LearningRateScheduler
```

What is next?

One important question:

Can DNNs be specialized?

CNNs: Image data

RNNs: Text/audio data (language) and any sequential data

Outline of today's lectures

Focus on CNNs.

Specifically we will:

1. Investigate how CNNs were developed;
2. Study the two key building blocks;

   Conv layer

   Pooling layer
3. Discuss two popular CNN architectures.

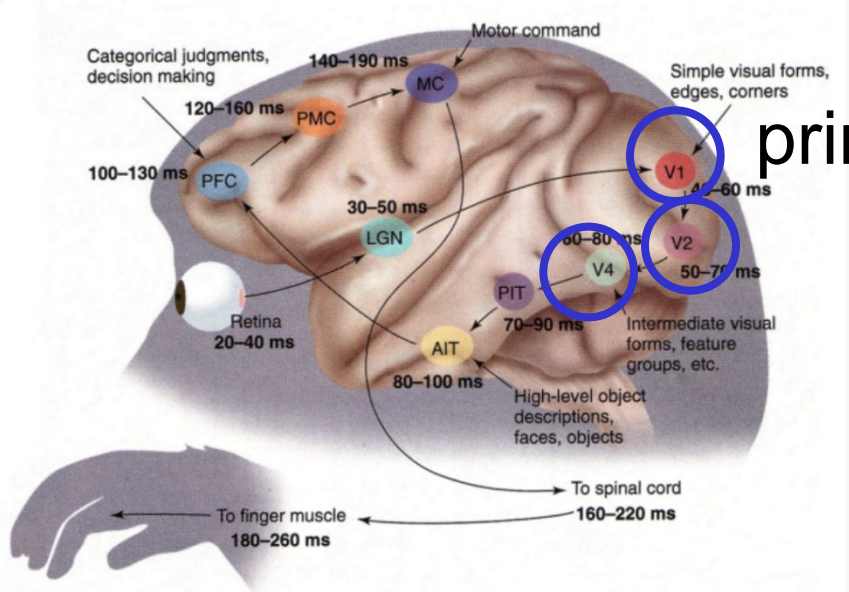
Focus of Lecture 7

Focus on CNNs.

Specifically we will:

1. Investigate how CNNs were developed;
2. Study the two key building blocks;
   - Conv layer
   - Pooling layer
3. Discuss two popular CNN architectures.

Visual cortex (시각피질)



primary visual cortex

Hubel & Wiesel ’58/’59 observed:

1. *React only to* a limited region (*receptive field)*
2. Higher-level neurons are based on lower-level ones.

*Won them the **Nobel Prize** in Physiology or Medicine in **‘81**.

Led to the birth of CNNs

Inspired Fukushima to propose the first CNN in 1980: **Neocognitron**

Kunihiko Fukushima
1980

Developed another CNN in 1988:
LeNet-5

Commercialized for handwritten check numbers recognition.

Yann LeCun
1988

Two building blocks of CNNs

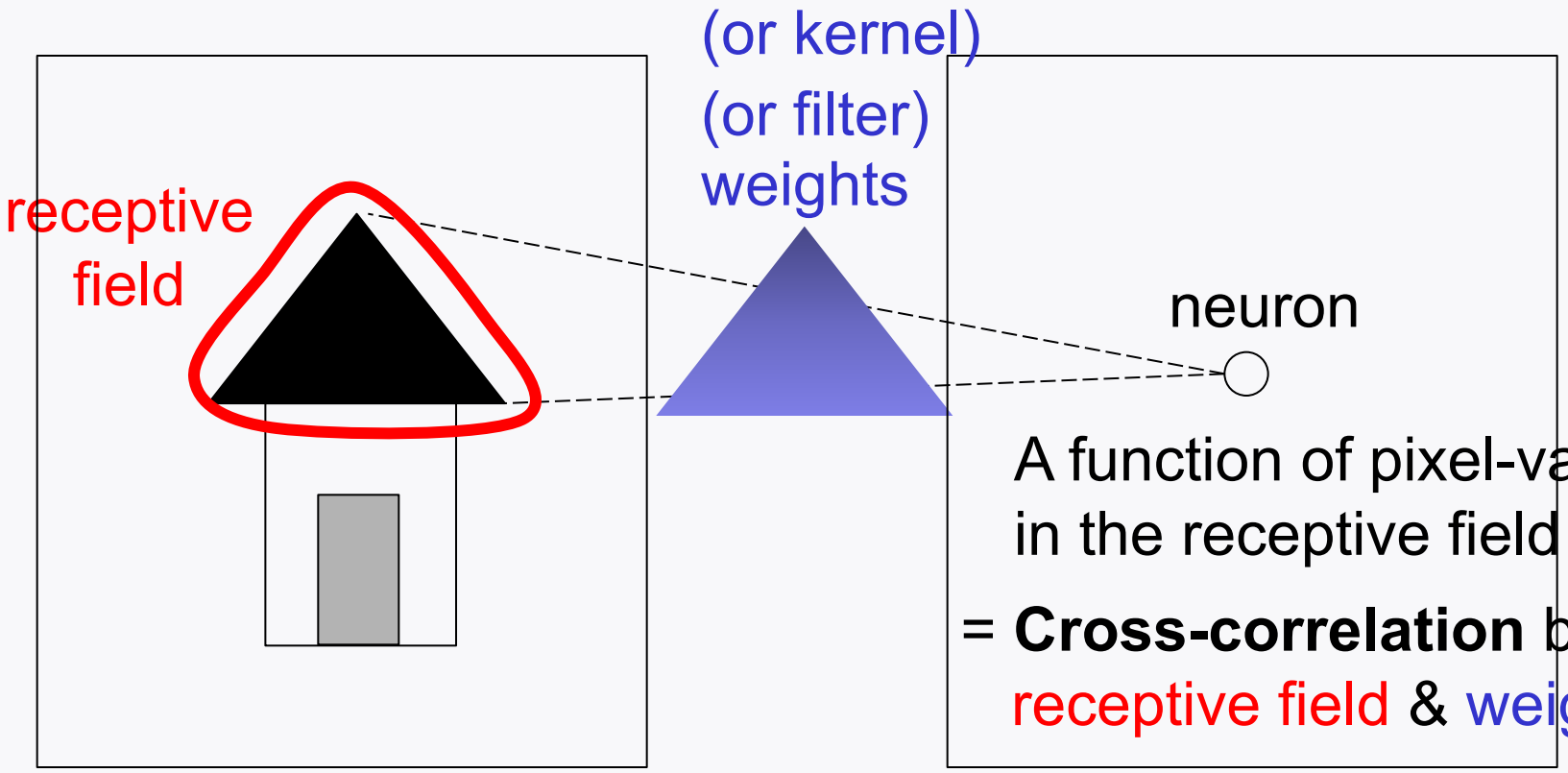
1. Convolutional layer (Conv layer)

Role: Mimick neurons' behaviors:
Reacting only to receptive fields.

2. Pooling layer

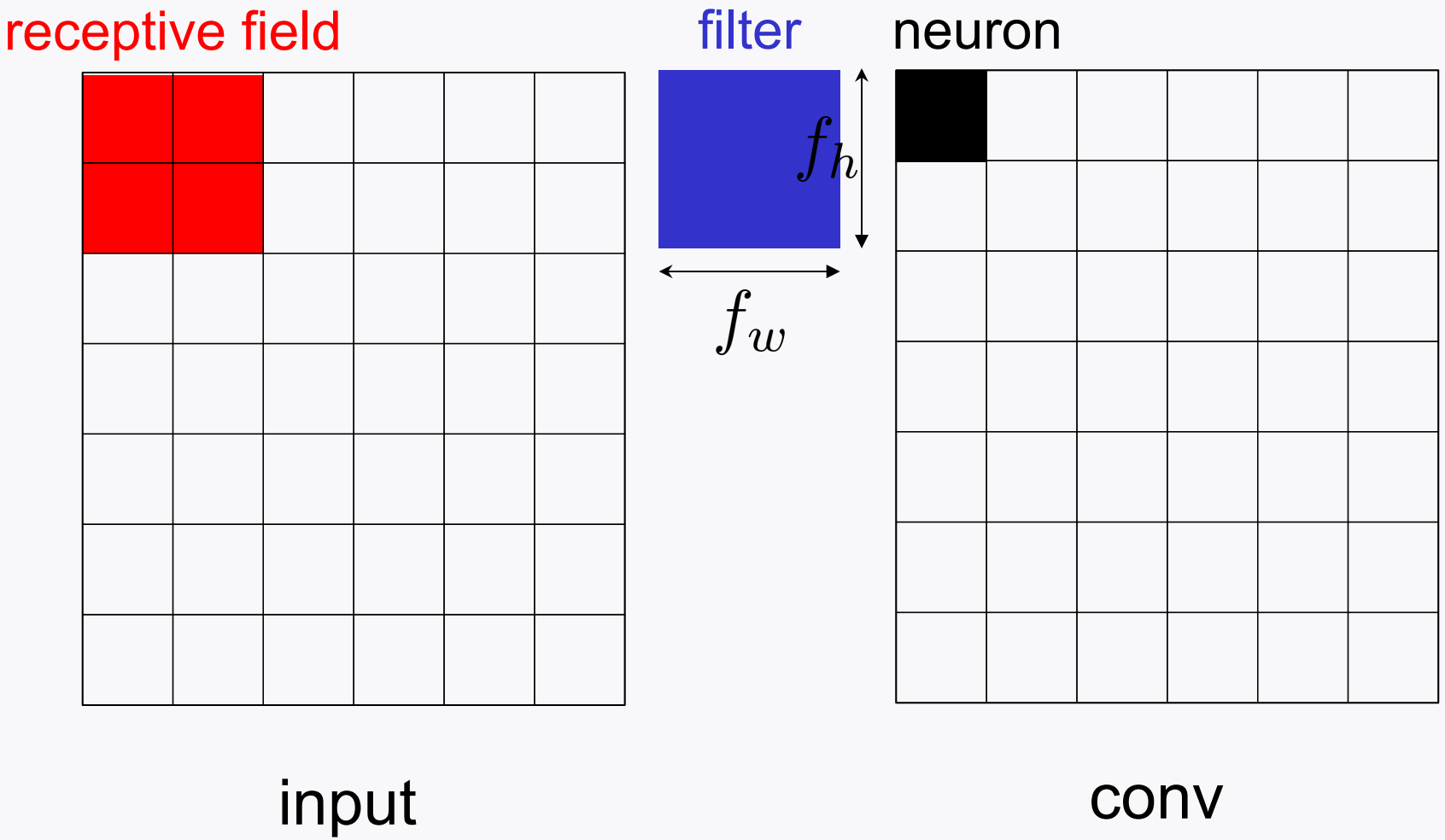
Role: Downsample to reduce complexity
(# parameters & memory size).

Conv layer

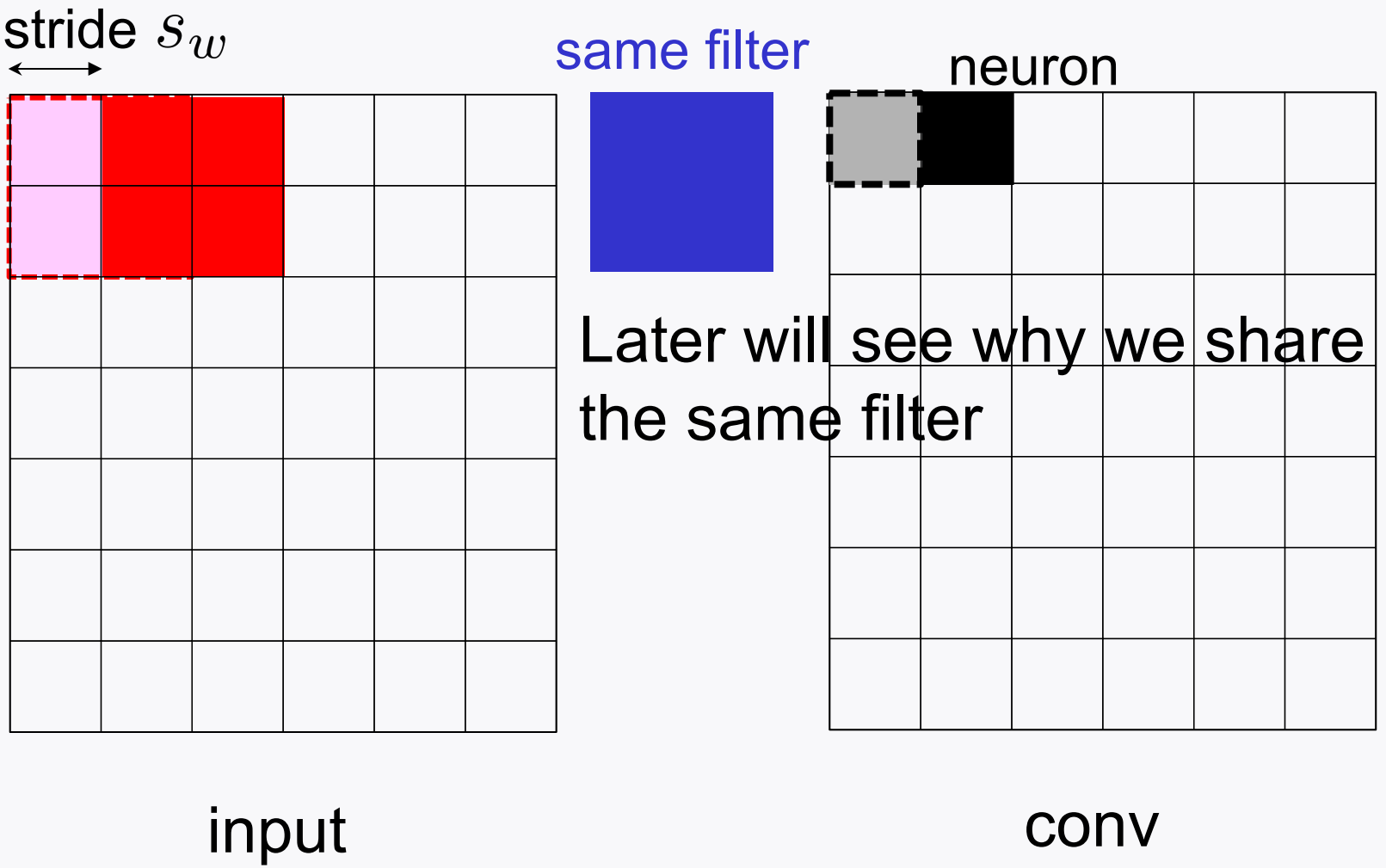


*Convolution operation is very similar to cross-correlation.

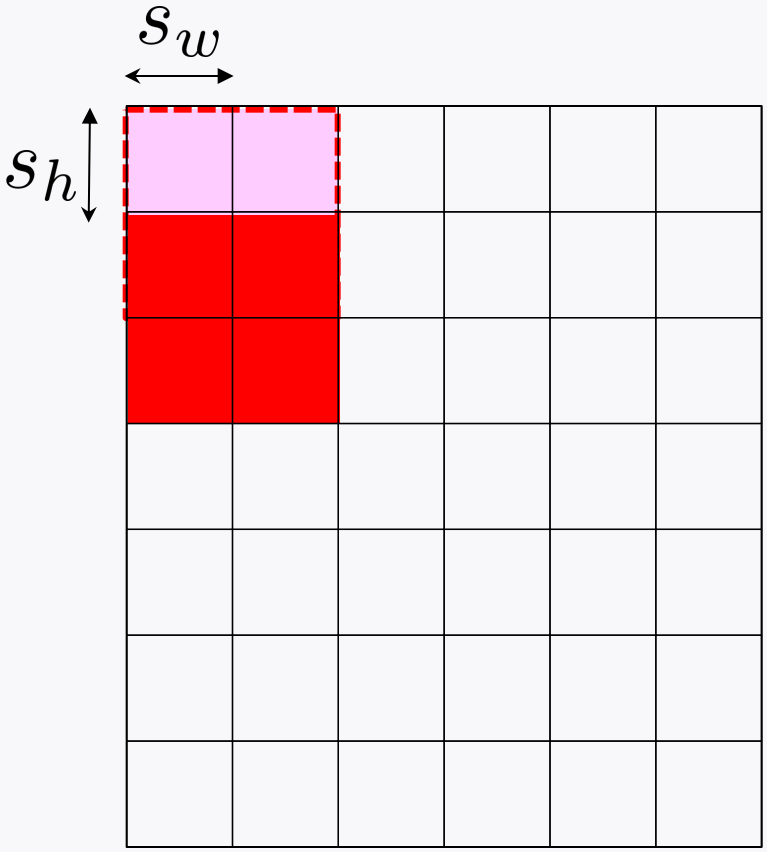
Abstraction via grids



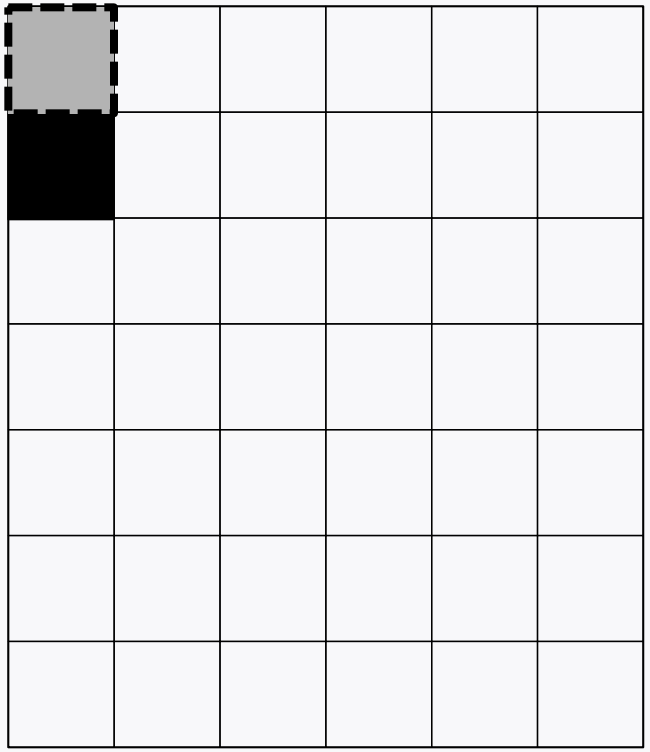
What about for next *right* neuron?



What about for next *below* neuron?

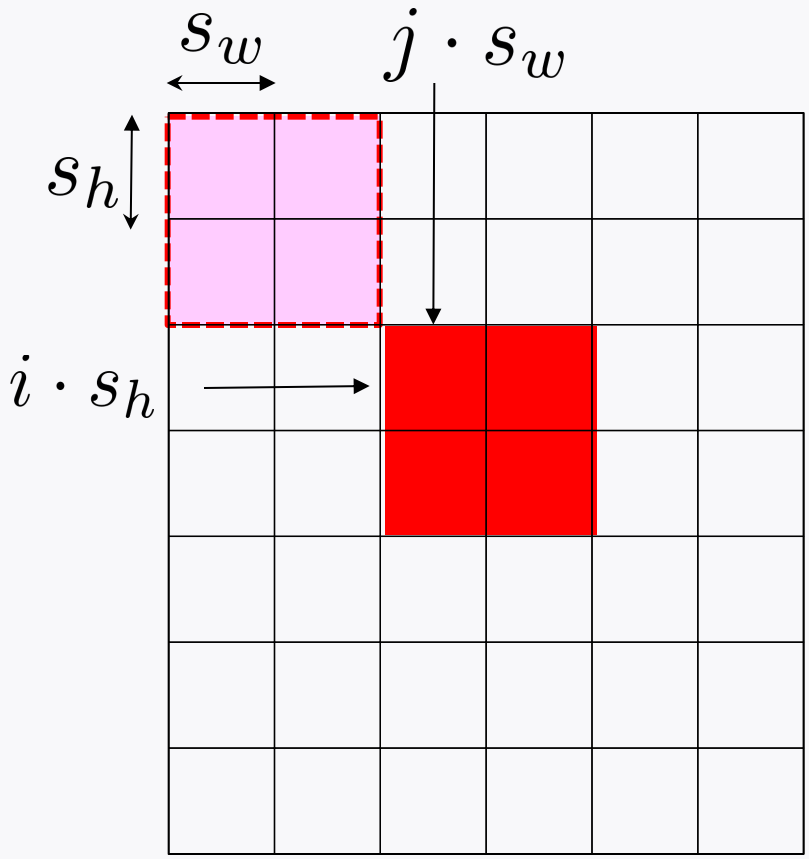


input

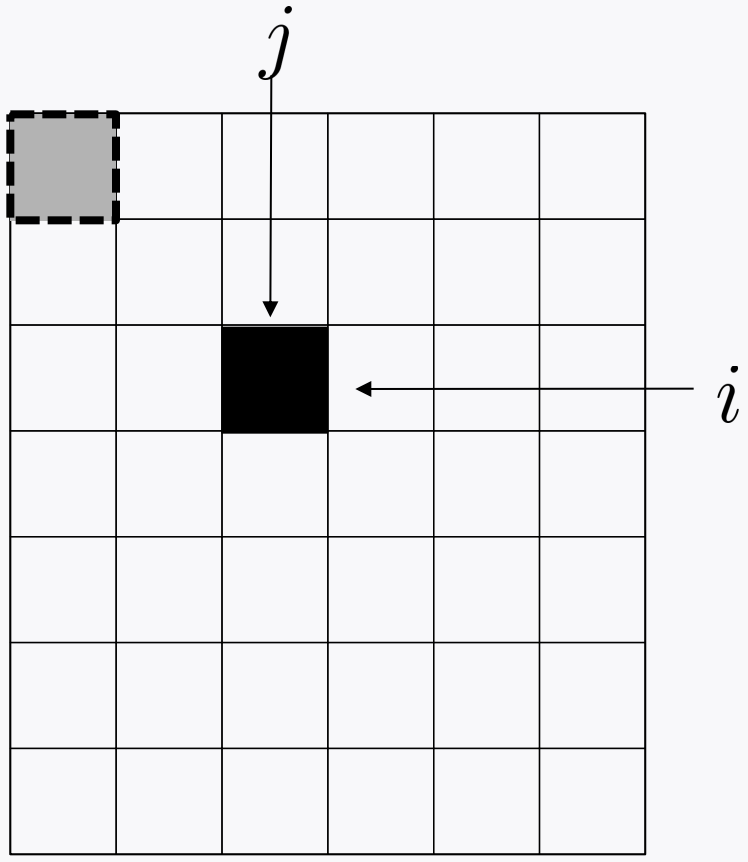


conv

In general …

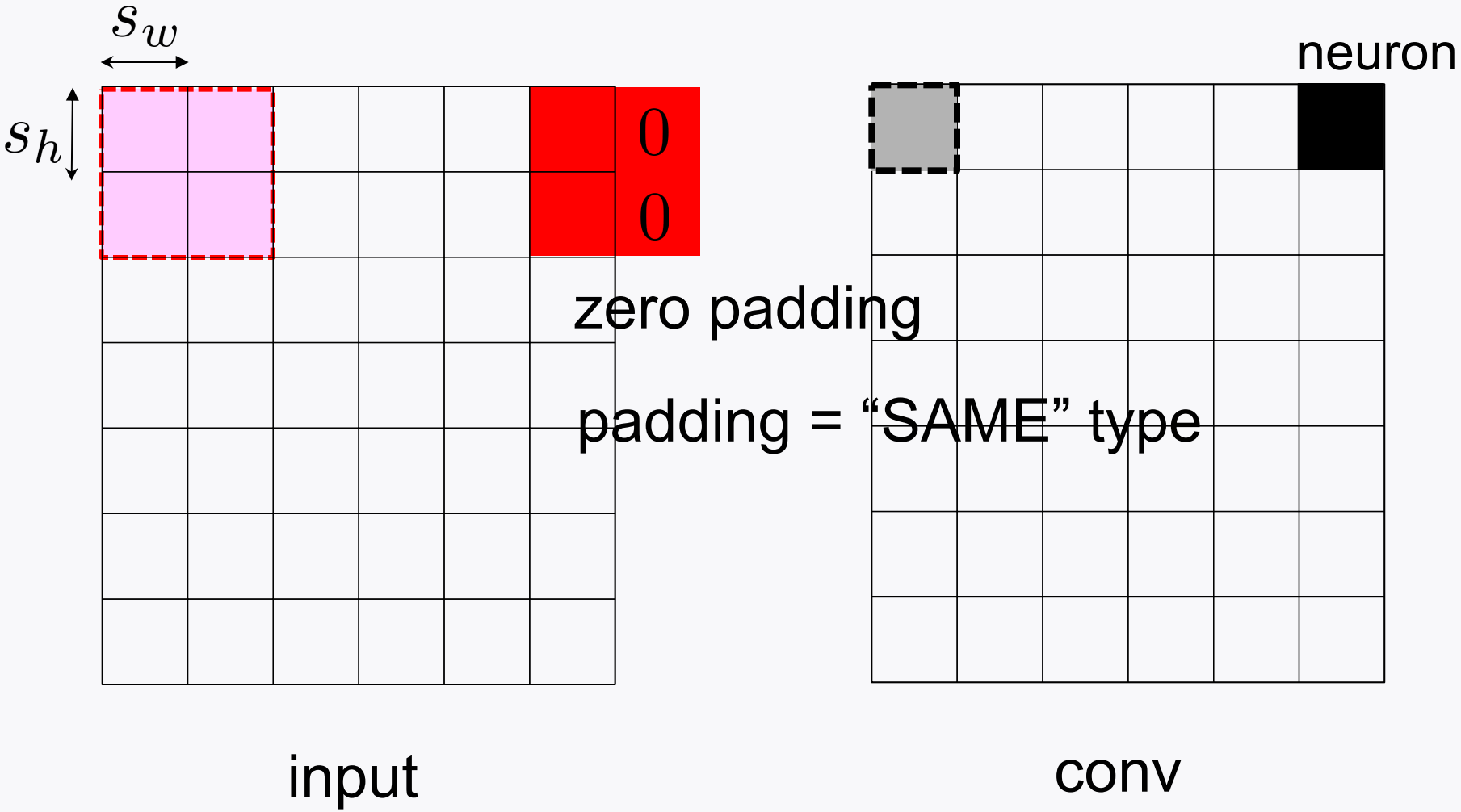


input

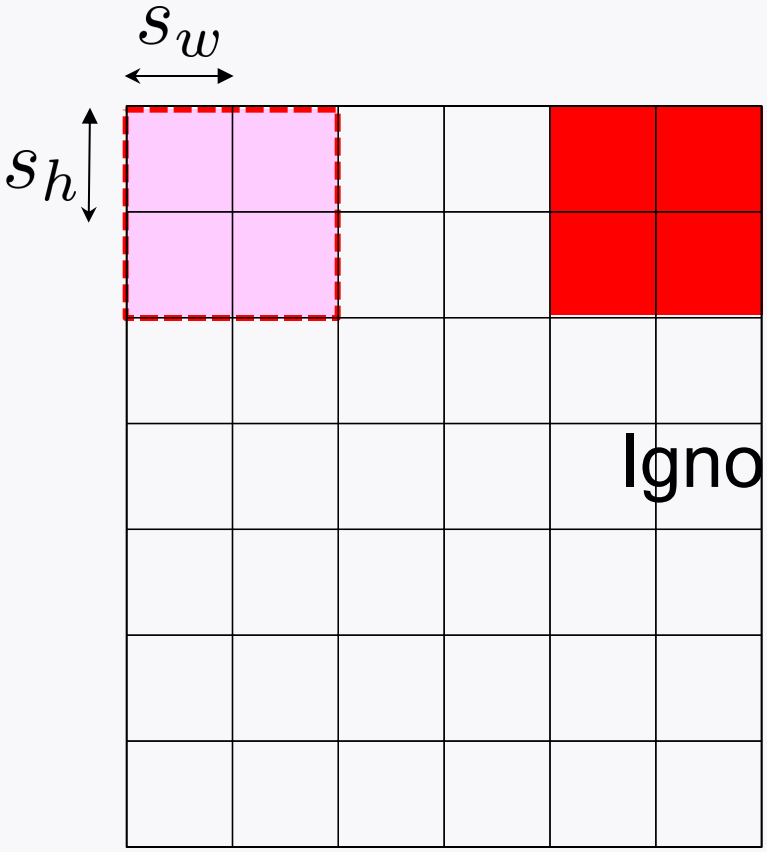


conv

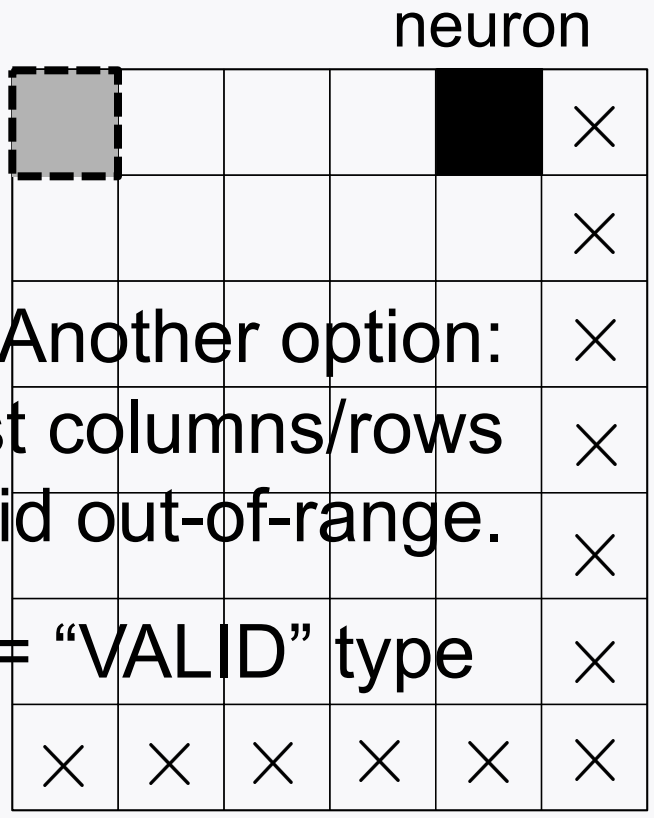
What if a receptive field is out of range?



What if a receptive field is out of range?



input



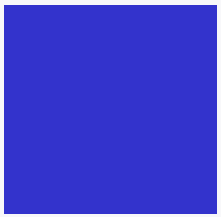
conv

Another option:
Ignore the last columns/rows
to avoid out-of-range.

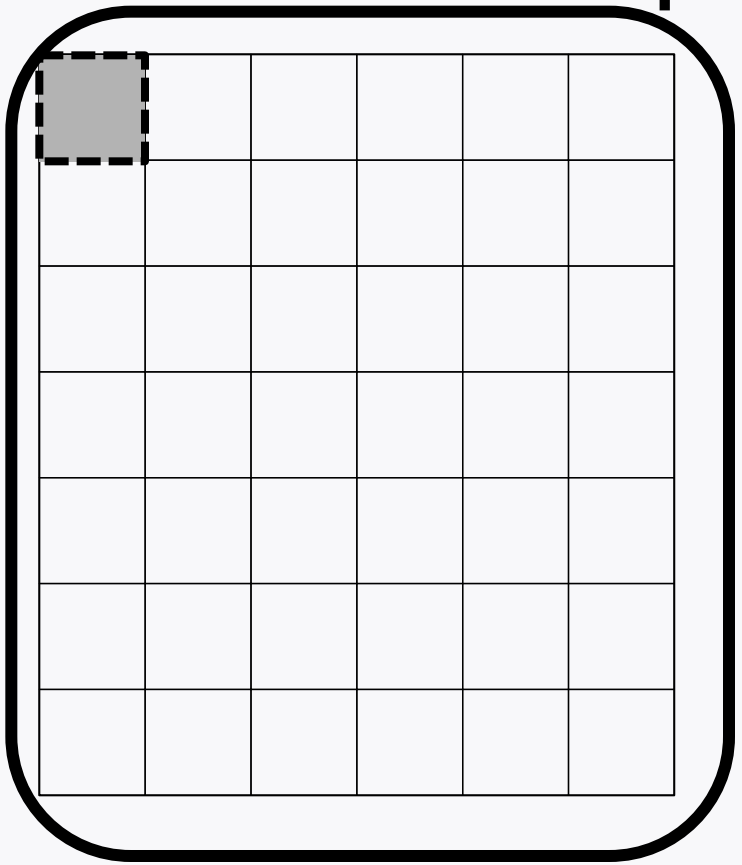
padding = “VALID” type

Feature map

filter



Called: **feature map**



Role: Detect a certain pattern (feature).

Note: Filter type depends on a target pattern that we wish to detect.

This is why we *share the same filter* for computations of other neurons *within the same feature map*.

Look ahead

1. Will study further on “feature map”.
2. Will study 2nd building block: **Pooling** layer