LECTURER: Nghia Duong-Trung

NEURAL NETS AND DEEP LEARNING

TOPIC OUTLINE

Introduction to Neural Networks	1
Feed-forward Networks	2
Overtraining Avoidance	3
Convolutional Neural Networks (Part 1)	4
Convolutional Neural Networks (Part 2)	5
Recurrent Neural Networks	6

- Course book: DLBDSNNDL01_Neural Nets and Deep Learning, provided by IU, myStudies
- Reading list DLBDSNNDL01, provided by IU, myStudies
- This slide is a summarization of important contents in the course book.
- Additional teaching materials:

https://github.com/duongtrung/IU-DLBDSNNDL01_Neural_Nets_and_Deep_Learning

DISCLAIMER

- This is the modified version of the IU slides.
- I used it for my lectures at IU only.

UNIT 4

CONVOLUTIONAL NEURAL NETWORKS (Part 1)

STUDY GOALS

On completion of this session, you will be able to ...

- understand the inspiration behind convolutional neural networks
- learn the purpose and importance of the architectural design of a convolutional neural network.
- explain the role of different types of layers in a convolutional neural network.

EXPLAIN SIMPLY

- What inspired the development of convolutional neural networks?
- Why is the architectural design of a convolutional neural network crucial for its performance and effectiveness?
- How do different types of layers contribute to the overall functioning of a convolutional neural network?

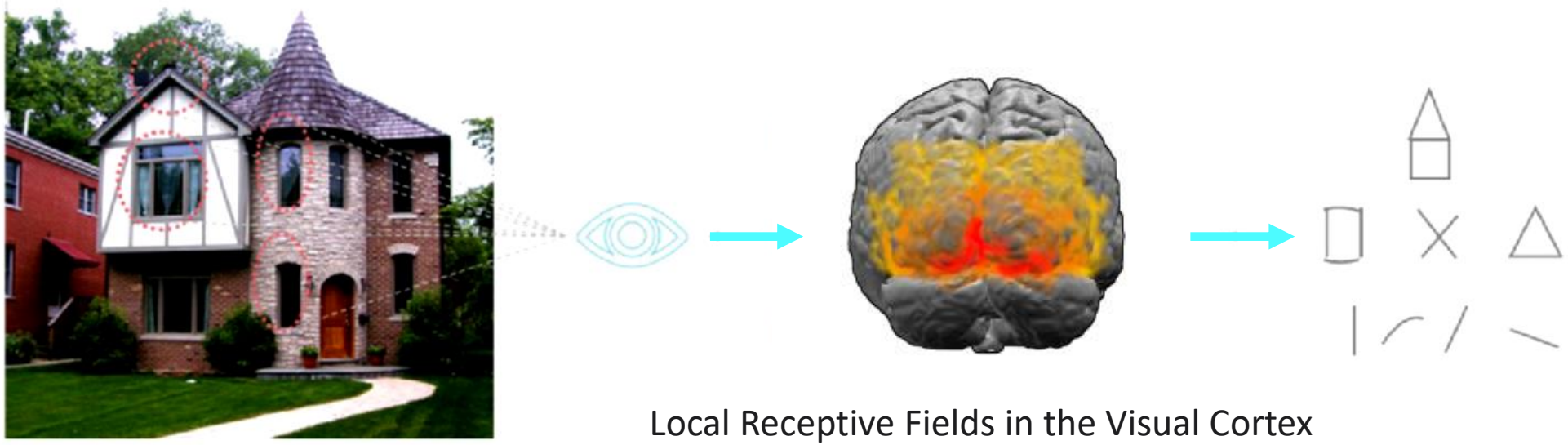
Convolutional neural networks (CNN)...

- are **deep neural networks** commonly used to identify and recognize objects in **image** data.
- are inspired from the study of **brain's** visual cortex.
- are effective in terms of computational **time**, model **size** and prediction **quality**.
- detect objects by progressing from basic **shapes** to recognizing specific **objects**.
- learn **feature** extraction effectively by assigning importance to image aspects.
- need **less preprocessing** of the input data compared to other algorithms.
- have been used extensively and with remarkable results in may **complex task**.

Source of the text: Zöller, 2023
Source of the image: Own Creation

Convolution in brain's visual cortex

- Visual cortex can detect **complex** patterns in the visual field by detecting and combining **lower-level** patterns.
- **Object recognition** involves extracting **features** from visual information, building upon them, and constructing layers of **abstractions**



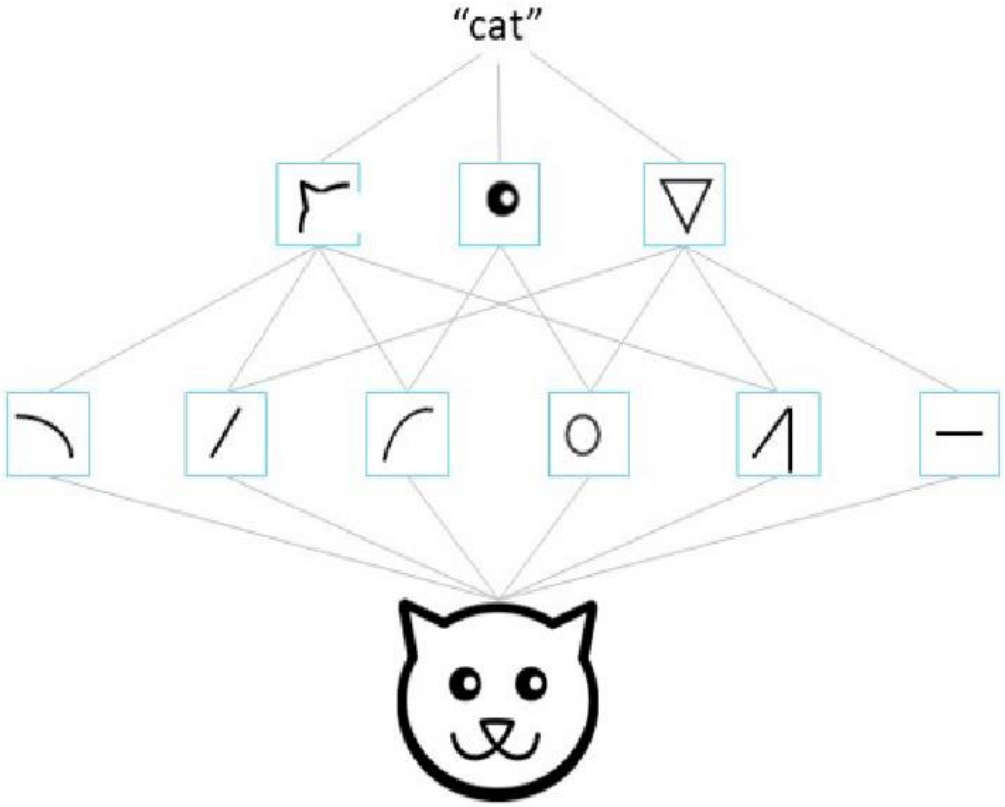
Local Receptive Fields in the Visual Cortex

Source of the text : Zöller (2023)
Source of the image: File:Brodmann_areas_17_18_19.png

Convolution in Image Processing

CNN considers the object as a **combination** of its parts .

- A convolution layer detects low-level **patterns** (lines, edges, or shapes).
- At higher-level layers, some **parts** of the object (ears, nose, or eyes) are identified.
- These parts are assembled into larger abstractions until the **whole object** (cat) is finally recognized.

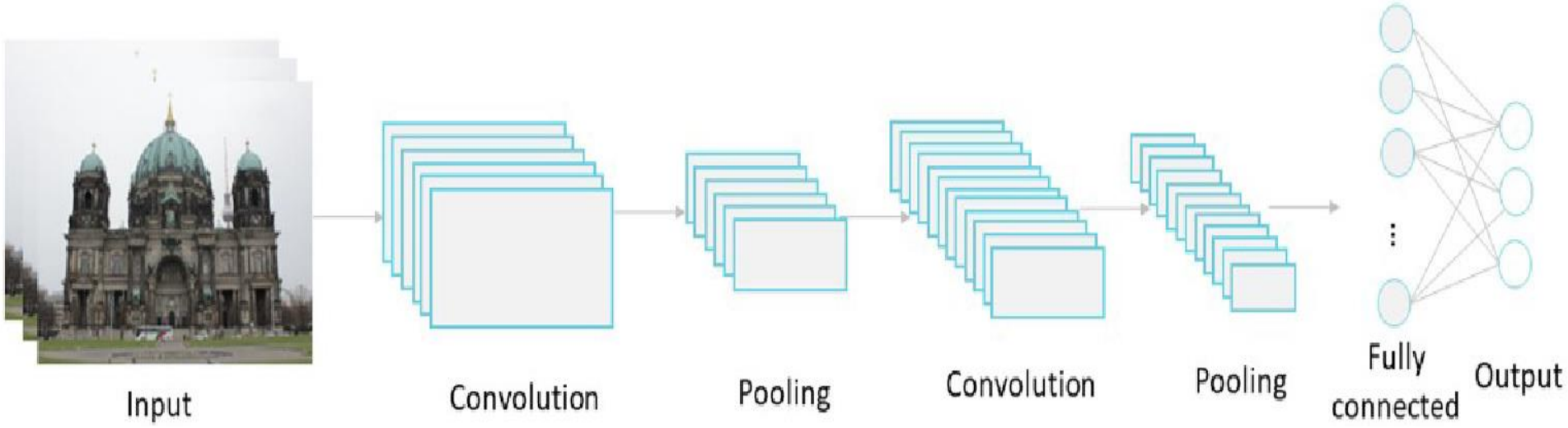


Visual Spatial Hierarchy

Source of the text and the image: Zöller (2023)

CNN Structure

CNNs have three main types of layers, the **convolutional** layer, **pooling** layer, and **fully connected** layer

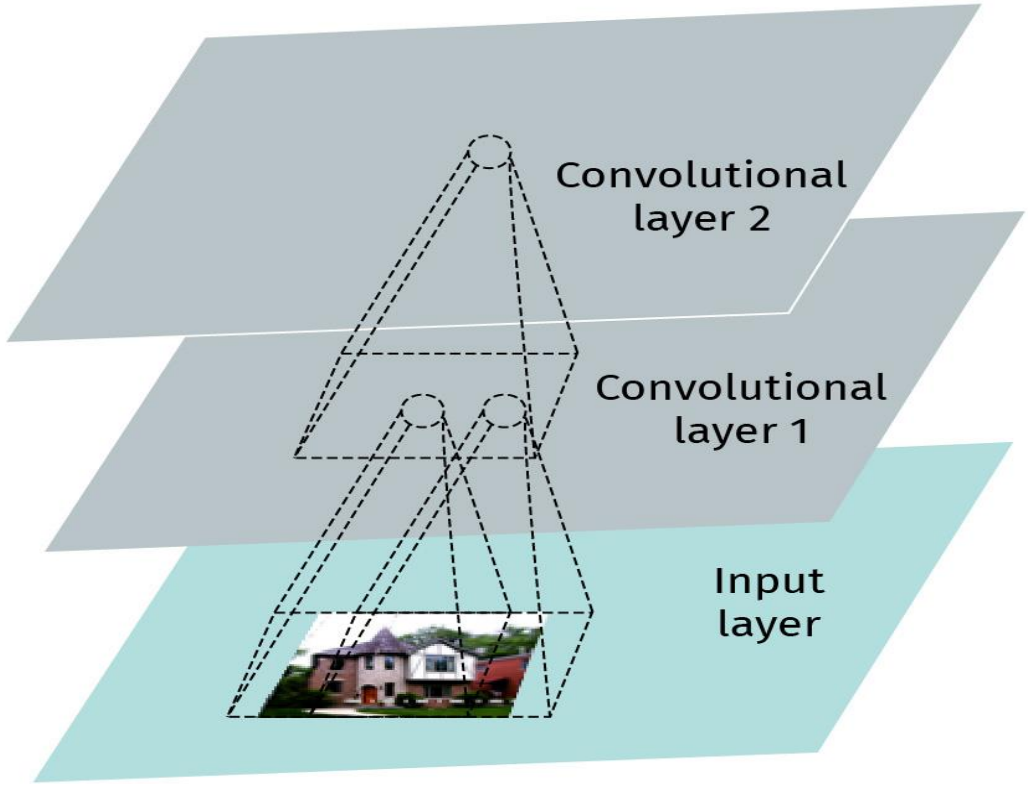


Structure of a typical CNN

Source of the text and the image: Zöller (2023)

Convolutional layer ...

- is a **main** building block of a CNN.
- relies on **input** data, a **filter**, and a **feature map**.
- neurons connect to a small **receptive area** from the previous layer.
- allows the network to focus on **low-level** features and assemble them into more **general** features.



CNN Layers With Rectangular Local Receptive Fields

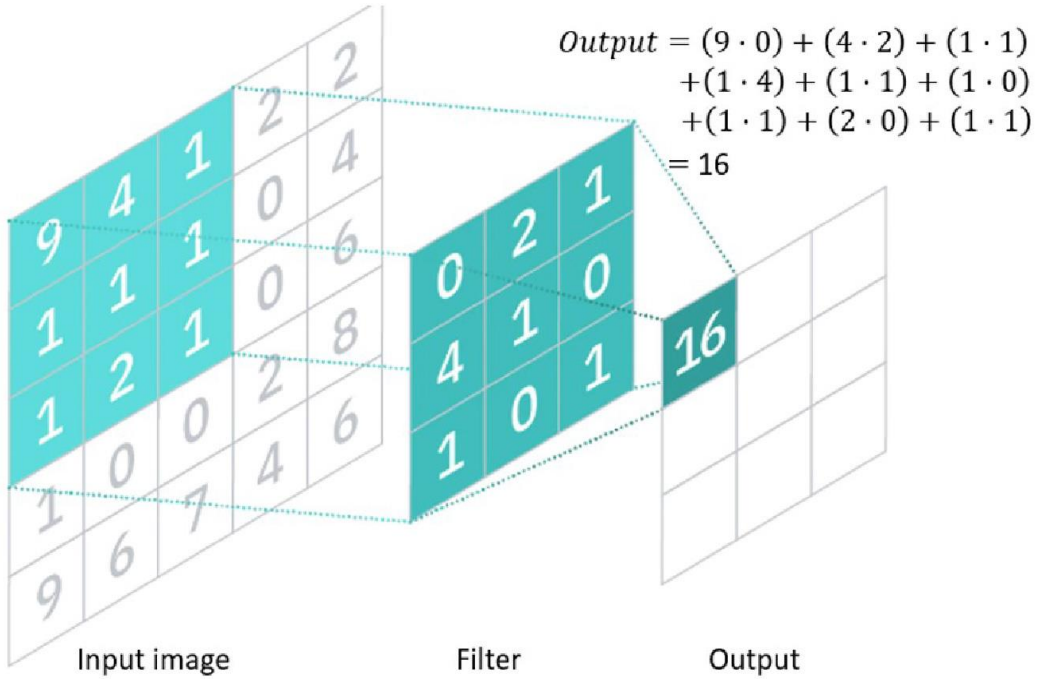
Source of the text and the image: Zöller (2023)

Filter

- Filters are **feature detectors**.
- Convolution is the process of **moving** a filter across the image to detect features.
- The result of filtering is a **feature map**.
- The output connects only to the **receptive field**, not to every pixel in the image.

Spatial arrangement

- Three hyperparameters that determine the size of the output volume:
  - **Depth**: the number of filters to be used
  - **Stride**: the step size when sliding a filter.
  - **Padding**: the number of pixels added to an image.

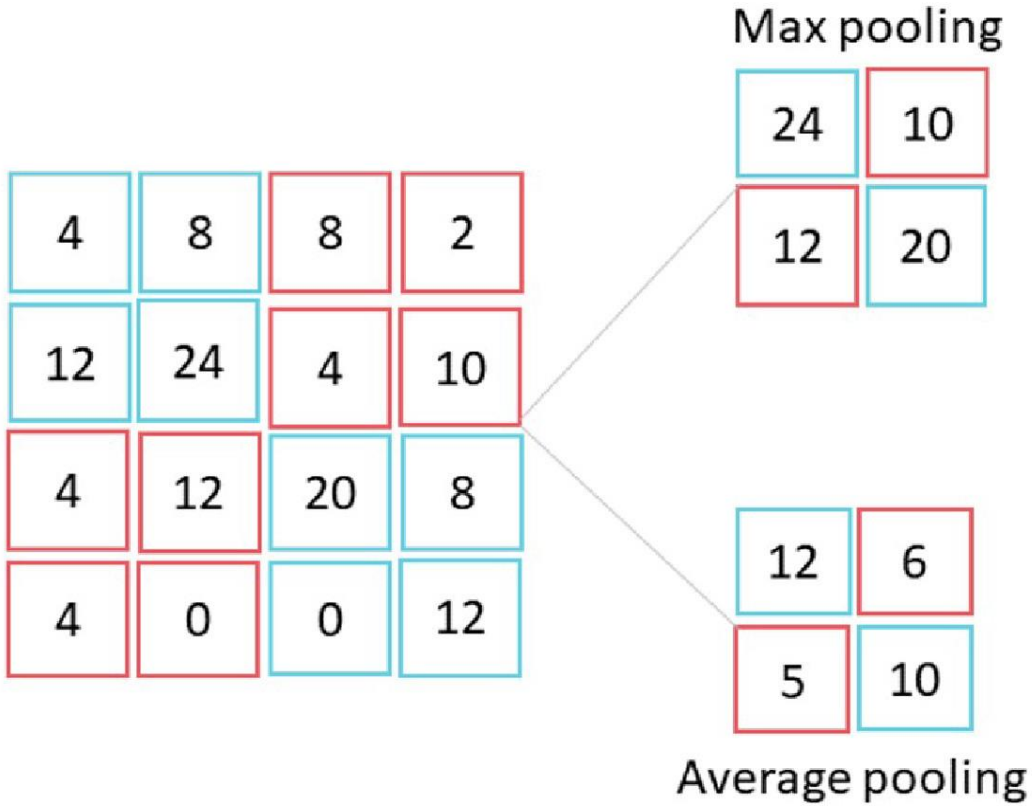


CNN Layers With Rectangular Local Receptive Fields

Source of the text and the image: Zöller (2023)

Pooling layer

- objective is to reduce the spatial **dimensions** and number of **parameters**.
- decreases **computational time**.
- helps to reduce the **complexity** of the network and control **overfitting**

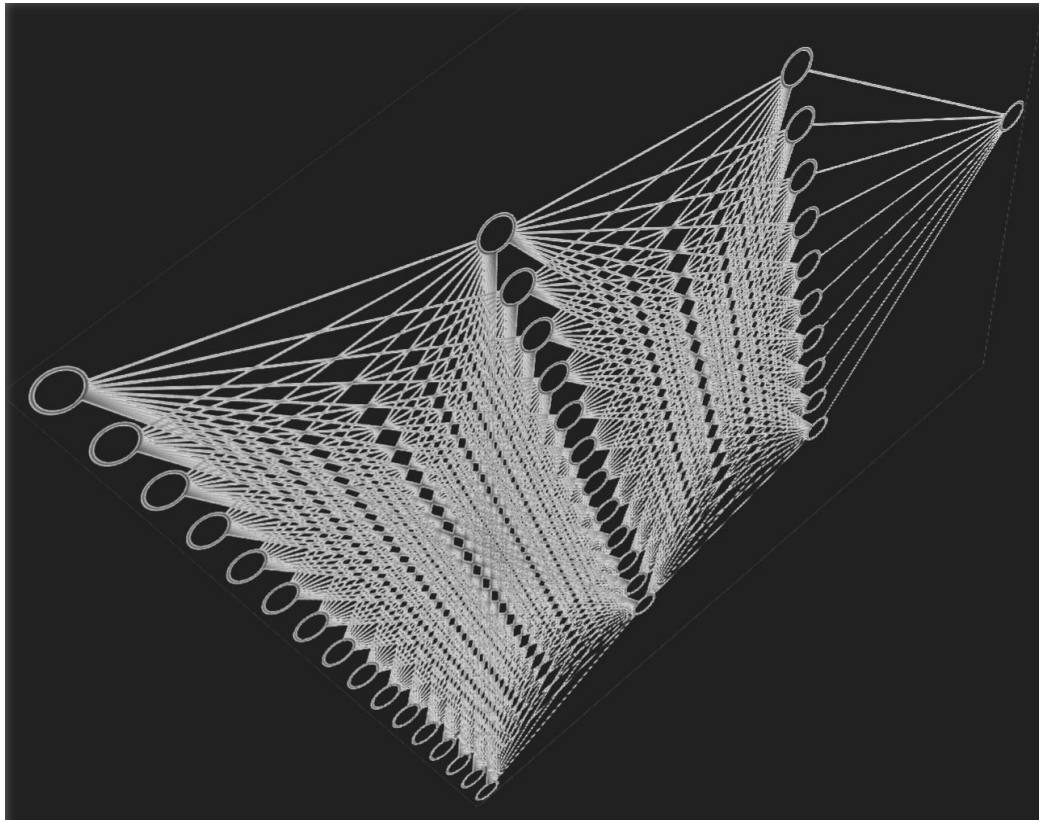


Types of Pooling

Source of the text and the image: Zöller (2023)

Fully connected layer …

- **compiles** the data extracted from the previous layers to form the final **output**.
- neurons have **full connections** to the neurons of the previous layers.
- uses a "**softmax**" activation function to assign an estimated probability (0→1).



Fully connected neurons

Source of the text: Zöller (2023)
Source of the image: File:Neural_network_with_dark_background.png.

REVIEW STUDY GOALS

- understand the inspiration behind convolutional neural networks
- learn the purpose and importance of the architectural design of a convolutional neural network.
- explain the role of different types of layers in a convolutional neural network.

SESSION 4

CONVOLUTIONAL NEURAL NETWORKS (Part 1)

TRANSFER TASK
PRESENTATION OF THE RESULTS

Please present your results.

The results will be discussed in plenary.

Task:

Given :

- the input volume has dimensions of 32x32x16,
- you apply a convolution process with a stride of 2 and a filter of size 2,

What is the output volume ?

Sample solution:

The output volume is:
output_size = (input_size - filter_size) / stride + 1

For the width dimension: output_width = (32 - 2) / 2 + 1 = 15 + 1 = 16
For the height dimension: output_height = (32 - 2) / 2 + 1 = 15 + 1 = 16
For the depth dimension: output_depth = input_depth = 16
So, the output volume would be 16x16x16.

1. Which of the following statements is true about CNNs? Select one.
   a) The quality of a convolutional neural network is significantly reduced when the number of parameters is diminished.
   b) Convolutional neural networks are used widely in image recognition tasks but cannot be used effectively in other machine learning domains such as face recognition, medical analysis, or self-driving cars.
   c) Convolutional neural networks generally require more preprocessing of the input data compared to other classification algorithms.
   d) A convolutional neural network is a deep learning framework commonly used to identify and recognize objects in image data.

2. Which of the following statements is true about CNN layers? Select one
   a) The filter in a convolutional layer is a kind of a feature detector which moves across the image data checking if the feature is present in the respective local receptive field.
   b) A feature map is a mapping mechanism which is applied to the filter of a CNN in the input layers of the network.
   c) The objective of the pooling layer in a CNN is to augment the dimensions of the image input data so that it can improve the generalization and prediction accuracy of the model.
   d) The role of the fully connected layer is to empower the neural network to learn direct linear relationships that map input features into the desired target.

3. Which of the following is the correct order for a convolutional neural network operation? Select one.
   a) Pooling → convolution → flattening → full connection.
   b) Full connection → convolution → pooling → flattening
   c) Convolution → pooling → flattening → full connection

Answers

1. d)
2. a)
3. c)

LIST OF SOURCES

Text

Zöller, T. (2023). Neural Nets and Deep Learning Course Book. IU International University of Applied Sciences.

Images

Zöller (2023).
File: Brodmann_areas_17_18_19.png. Visual cortext (2023, October). *Wikipedia Commons.* *https://en.wikipedia.org/wiki/Visual_cortex*
File:Neural_network_with_dark_background.png. Convolutional neural networks (2023, September). *Wikipedia Commons. https://en.wikipedia.org/wiki/Convolutional_neural_network.*