Neural networks in computer vision

Zalán Bodó

Faculty of Mathematics and Computer Science
Babeş–Bolyai University, Cluj-Napoca

Windesheim University of Applied Sciences, March 25–28, 2024

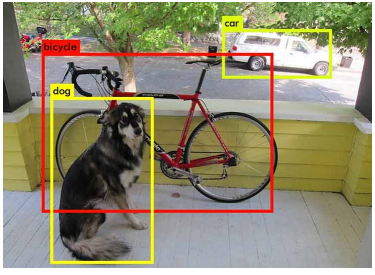
Presentation outline

1. Computer vision
2. A very short history of artificial neural networks
3. Convolutional neural networks (CNNs)
4. Interpretability in machine learning & computer vision

Computer vision (CV):

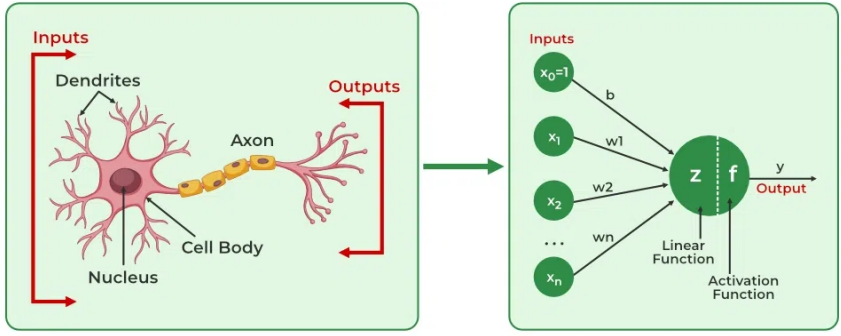
- subfield of AI dealing with the processing, analysis, understanding, etc. of images
- tasks of CV:
  - image classification
  - object detection
  - semantic segmentation
  - object tracking

  …



Artificial neural networks:

- modelling the structure, behavior of (natural) neurons



(From https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/)

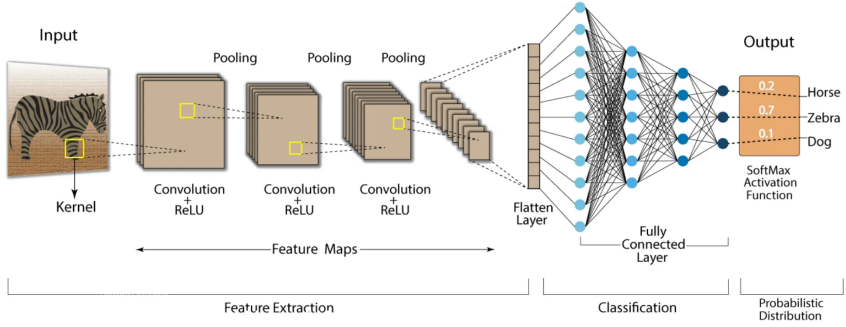
- deep neural networks (DNN): a large number of layers with a large number of neurons
- neurons of the human brain $\approx$ 86 billion (10^9)
  number of neural connections (synapses) $\approx$ ***60–100 trillion*** (10^{12})
- OpenAI:
  - GPT-2 (*Generative Pre-trained Transformer*, 2019): 1.5 billion parameters (= synapses)
  - GPT-3/3.5 (2020): 175 billion
  - GPT-4 (2023): ***1.76 trillion***

A very short history of artificial neural networks

- 1943: McCulloch and Pitts describing a model of artificial neurons with binary threshold activation function capable of simulating logical gates
- 1949: Hebb's rule on synaptic plasticity
- 1958: Rosenblatt's perceptron algorithm
- 1969: *AI winter* caused by Minsky and Papert
- 1986: backpropagation algorithm
- 1995: CNNs (Yann Lecun: LeNet)
- 1997: LSTM
- 2017–2018: transformers (e.g. BERT)
- 2018: GPT
- ...

Convolutional neural networks (CNNs)

- = biologically inspired specialized neural networks – the structure and function of the visual cortex of the brain is similar to this
- neurons have a so-called *receptive field* – each neuron has a receptive field, which can overlap and together cover the entire visual field

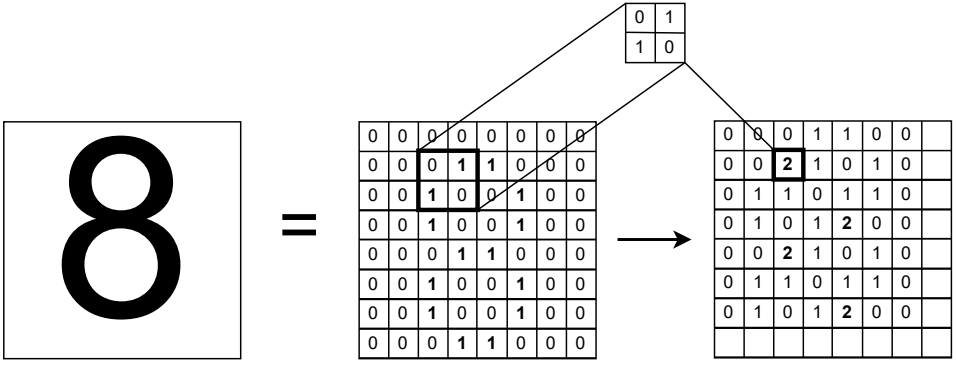


(From https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5)

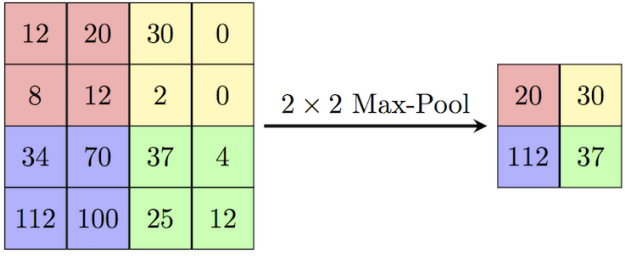
Basic components of a CNN:

- convolutional layers $\leftarrow$
- pooling layers $\leftarrow$
- non-linear activations (ReLU)
- classification layers (fully connected layers)

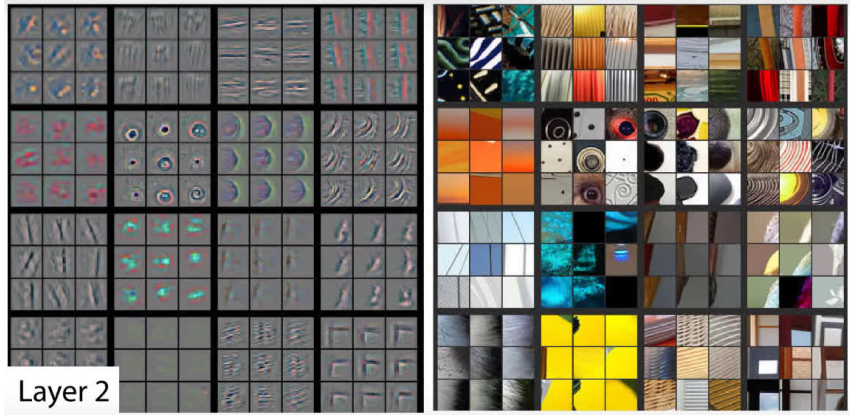
Convolutional layers = feature extraction



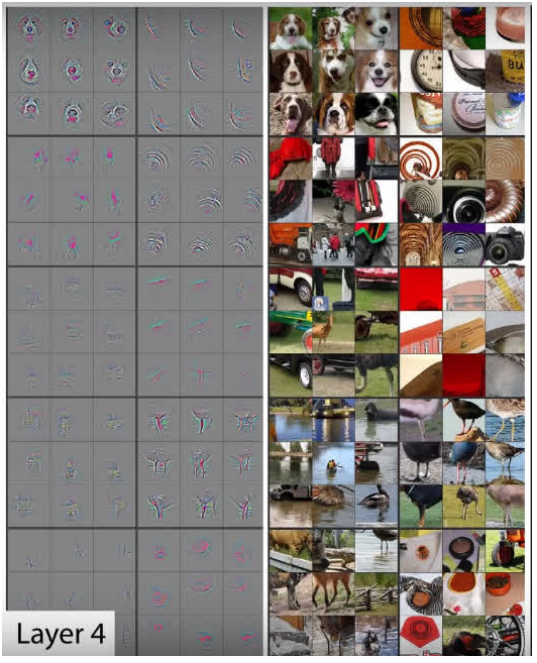
Pooling = finding higher-level features



(From https://paperswithcode.com/method/max-pooling)



(From https://deeplizard.com/learn/video/YRhxdVk_sIs)



(From https://deeplizard.com/learn/video/YRhxdVk_sIs)

Interpretability in machine learning & computer vision

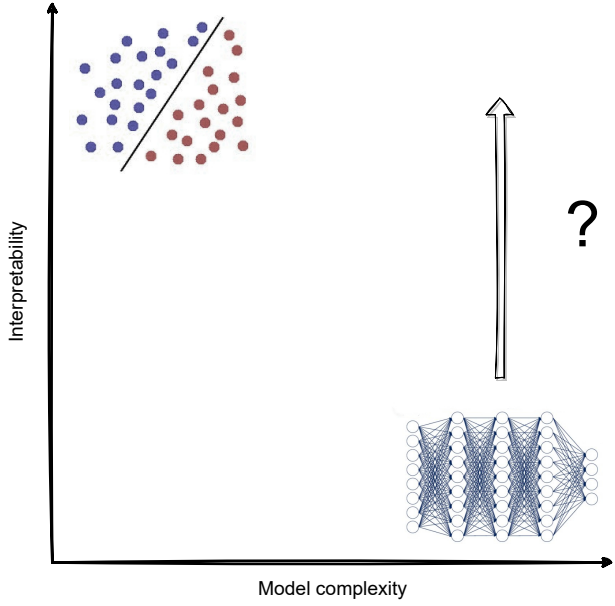
Explainability, interpretability = explaining the predictions/outputs of the method / why?

- simple for linear models (SVM, naive Bayes, logistic regression, decision trees, etc.)

$$f(\boldsymbol{x}) = \boldsymbol{w}^T \boldsymbol{x} = \sum_i w_i \, x_i$$

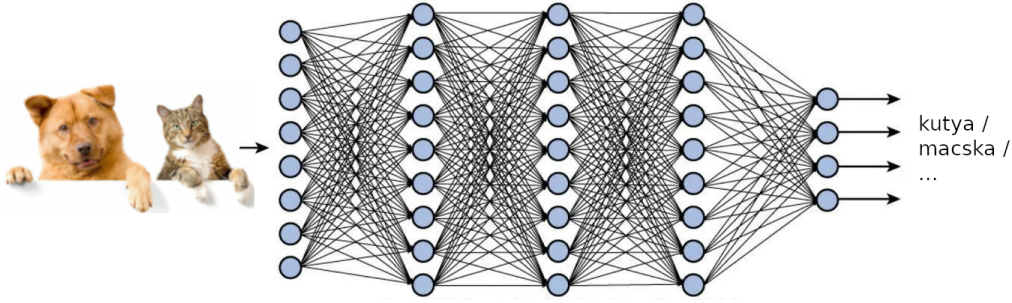
$$|w_i|$$

- for complex / *black-box* models (e.g. neural networks) is not so straightforward

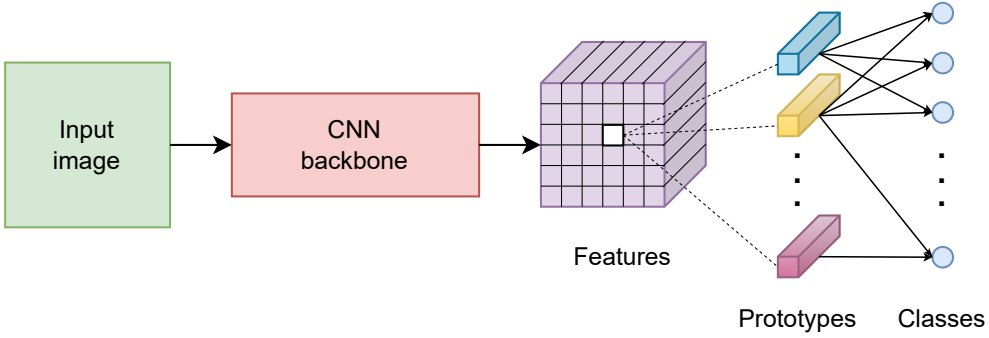

Interpretability
?
Model complexity

Why interpretability is important?

- increases user confidence in the system
- transparency is good for finding what causes an error/misprediction
- improving on the performance of the model can become easier



$\Rightarrow$

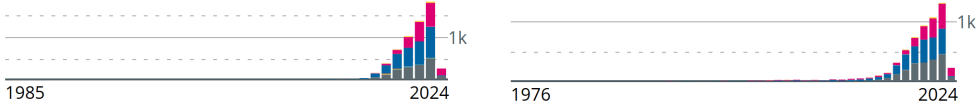


Explanations in prototype-based self-interpretable neural networks.

- AI Act of the EU – 2021 / 2024 – it will come into force in May
  3 categories: banned practices, high-risk systems, others
  requirements in the case of *high-risk* systems:
  - high-quality data
  - documentation
  - traceability
  - **transparency**
  - human oversight
  - accuracy
  - robustness

  available in 24 languages:
  https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

- there is a significant increase in explainable/interpretable AI/ML publications in the last years (https://dblp.org/)



(a) “explainable”, (b) “interpretable”

- some of the most successful / best-performing models:
  - 2016: LIME
  - 2017: Grad-CAM
  - 2019: ProtoPNet
  - 2019: BagNet
  - 2023: PIP-Net

Thank you for your attention!