A gentle introduction to Deep Learning

ECE 590 Seminar, March 10, 2023

Manel Martínez Ramón
Department of Electrical and Computer Engineering
The University of New Mexico

Pablo Picasso versus ChatGPT

- Me:

  Pablo Picasso said *"Computers are useless. They can only give you answers"*.

- ChatGPT:

  He died in 1973, long before computers became a ubiquitous part of daily life.

  While he may have expressed some thoughts on machines and technology, his comments would likely have been focused on earlier forms of machinery rather than modern computers.

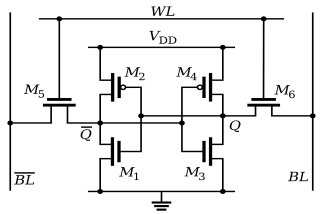
ML: A definition?

Probability

$$p(y|x) = \frac{p(x|y)p(y)}{p(x)}$$

Optimization

$$\nabla_w \mathcal{L}(x, y, w) = 0$$

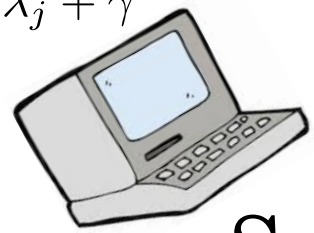


Algebra

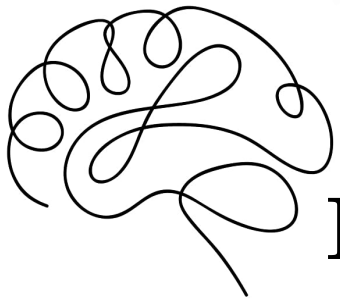
$$(\mathbf{X}^\top \mathbf{X} + \gamma \mathbf{I})^{-1} = \sum_j \frac{1}{\lambda_j + \gamma} \mathbf{v}_j \mathbf{v}_j^\top$$

Computer Engineering

Computer Science

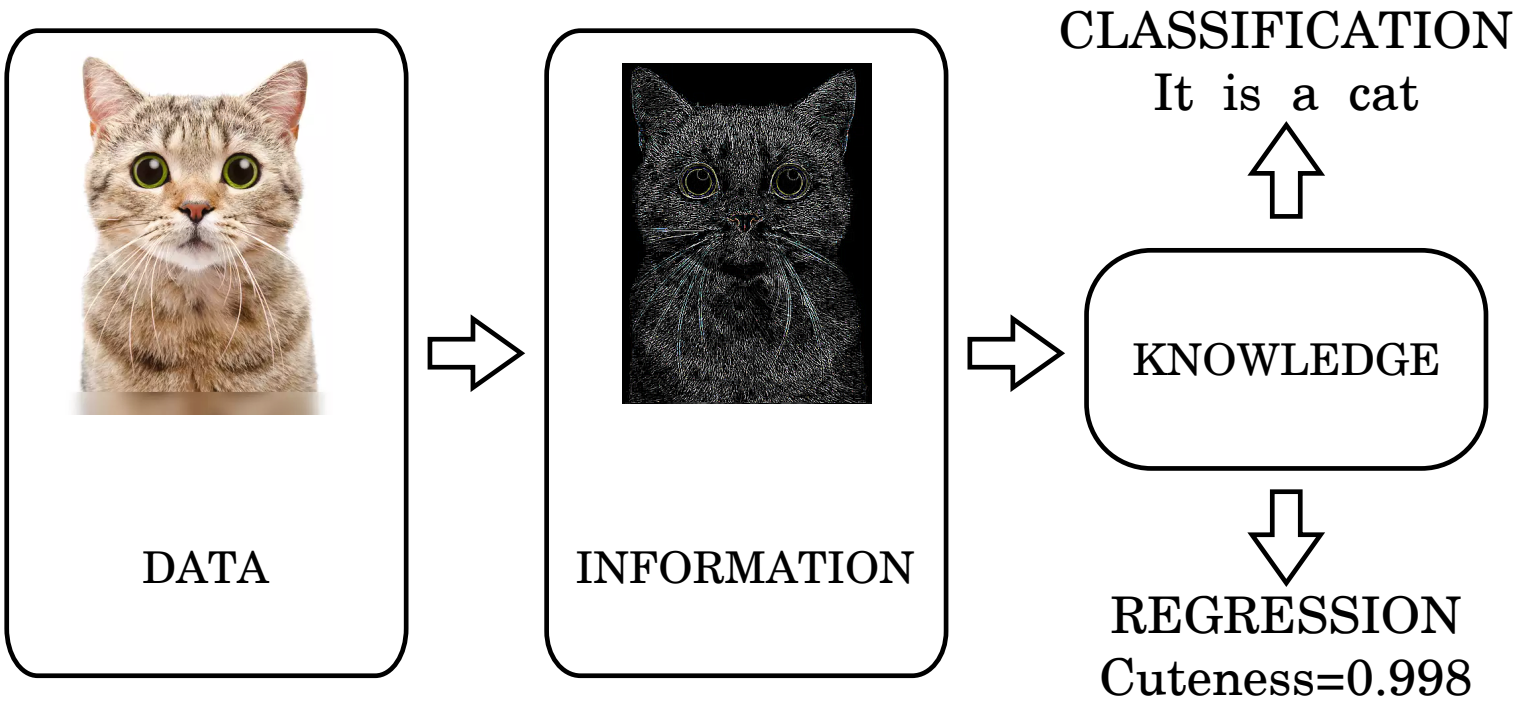


Social Sciences

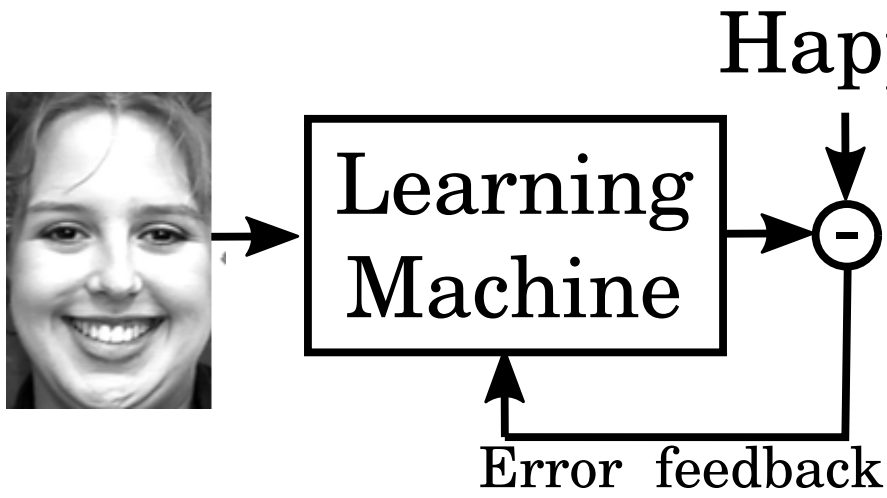


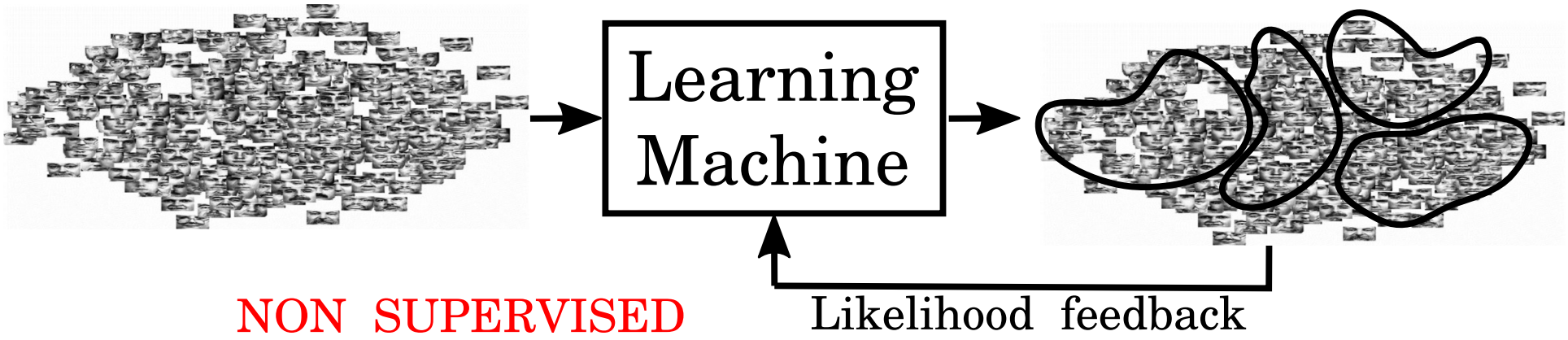
Psychology

ML: A definition?



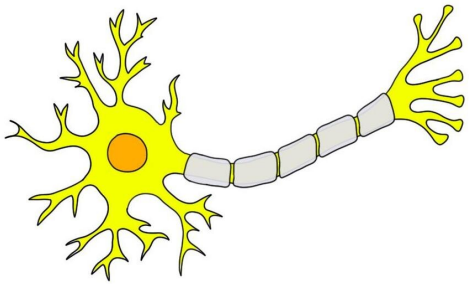
Supervised vs. non supervised

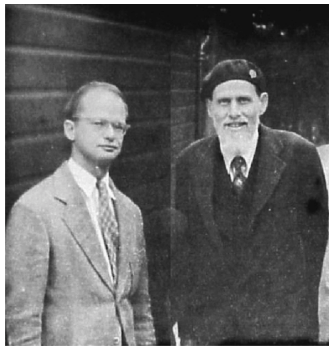




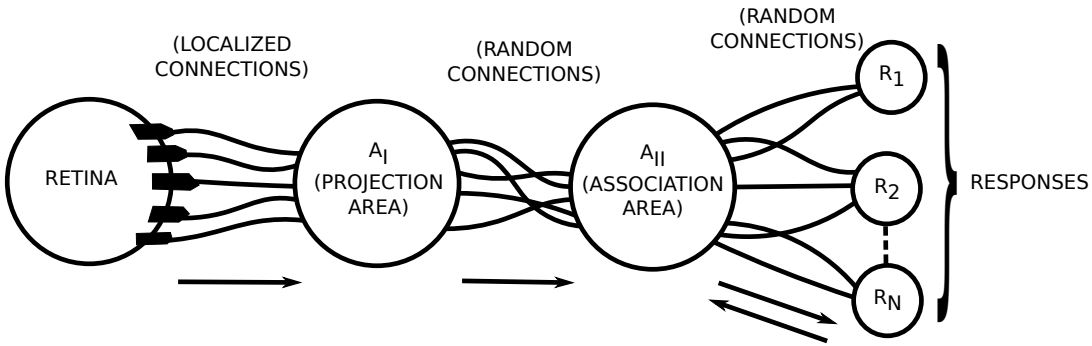
The concept of neuron

- Inspired in the nervous system (McCulloch and Pitts, 1943).
- Described as an element with two possible states (0, 1).



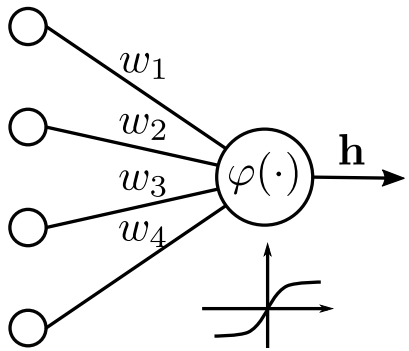


- If a linear combination of the inputs (dendrites) is above a threshold, the output (axon) will be activated.
- Artificial neural network: introduced by Rosemblatt (1958)



Ok, yes, but what about deep learning?

It is based on the concept of neuron. Its structure is different from the previous one.

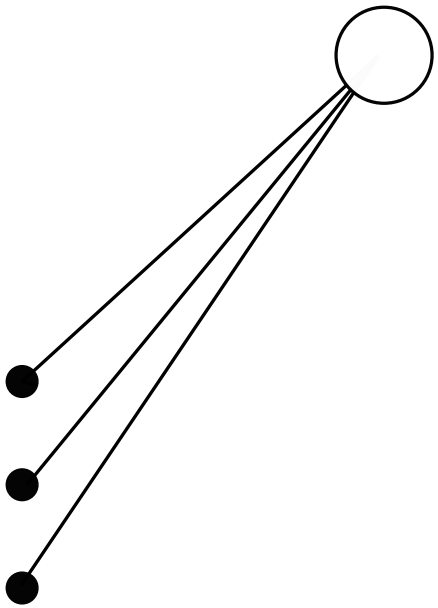


This is a neuron...

The input is connected to the output node through a **vector** of parameters.

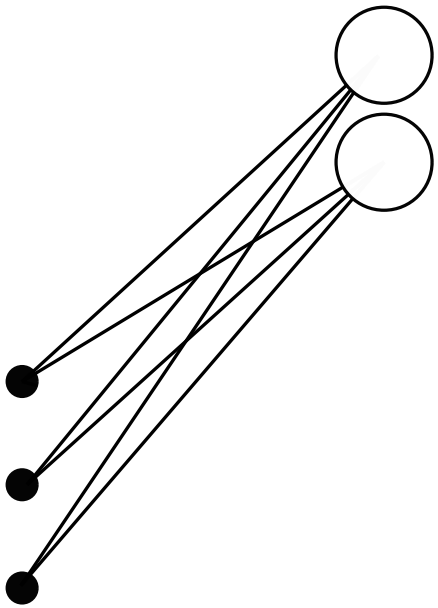
OK, but what about deep learning?

We can construct a complex multilayer structure with this "neuron".



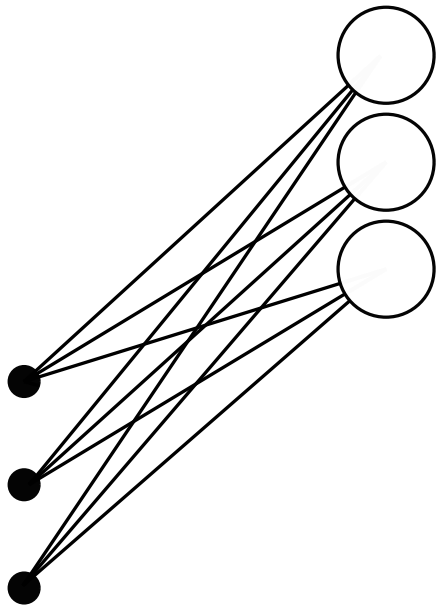
OK, but what about deep learning?

We can construct a complex multilayer structure with this "neuron".



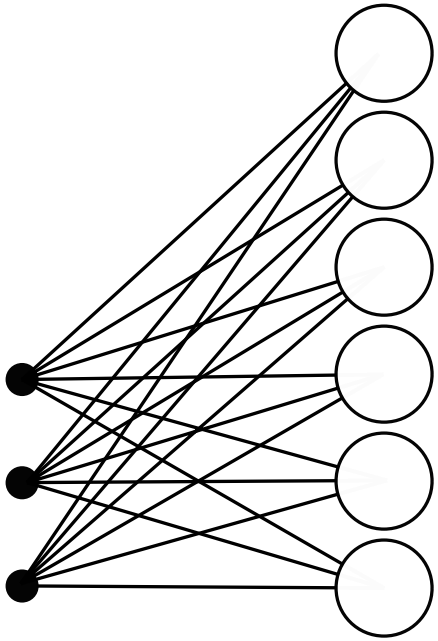
OK, but what about deep learning?

We can construct a complex multilayer structure with this "neuron".



OK, but what about deep learning?

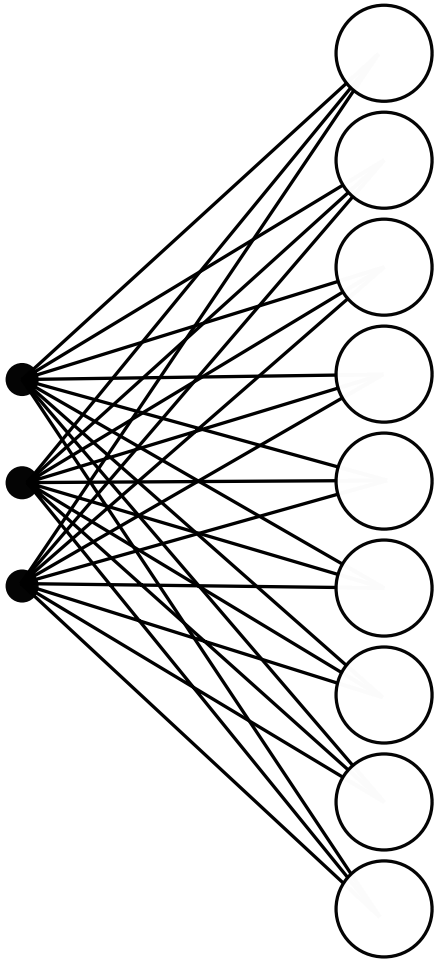
We can construct a complex multilayer structure with this "neuron".



OK, but what about deep learning?

THE UNIVERSITY OF NEW MEXICO

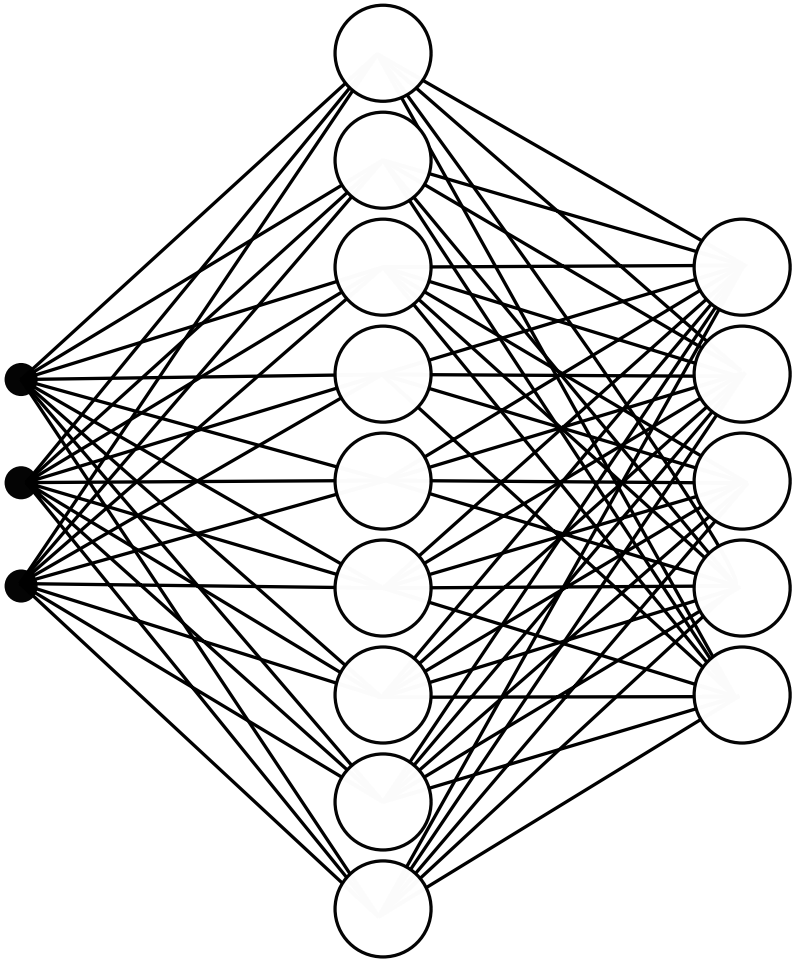
We can construct a complex multilayer structure with this "neuron".



OK, but what about deep learning?

THE UNIVERSITY OF NEW MEXICO

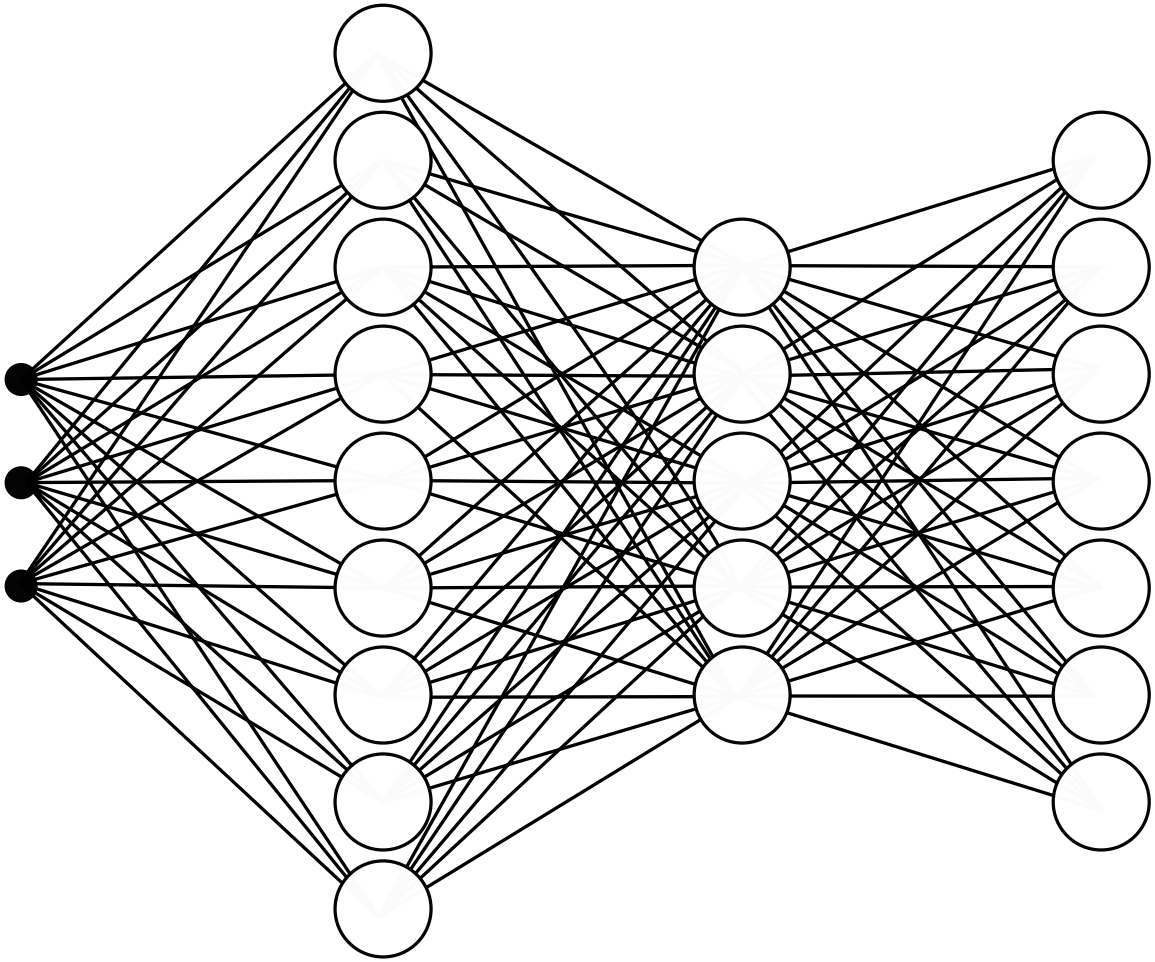
We can construct a complex multilayer structure with this "neuron".



OK, but what about deep learning?

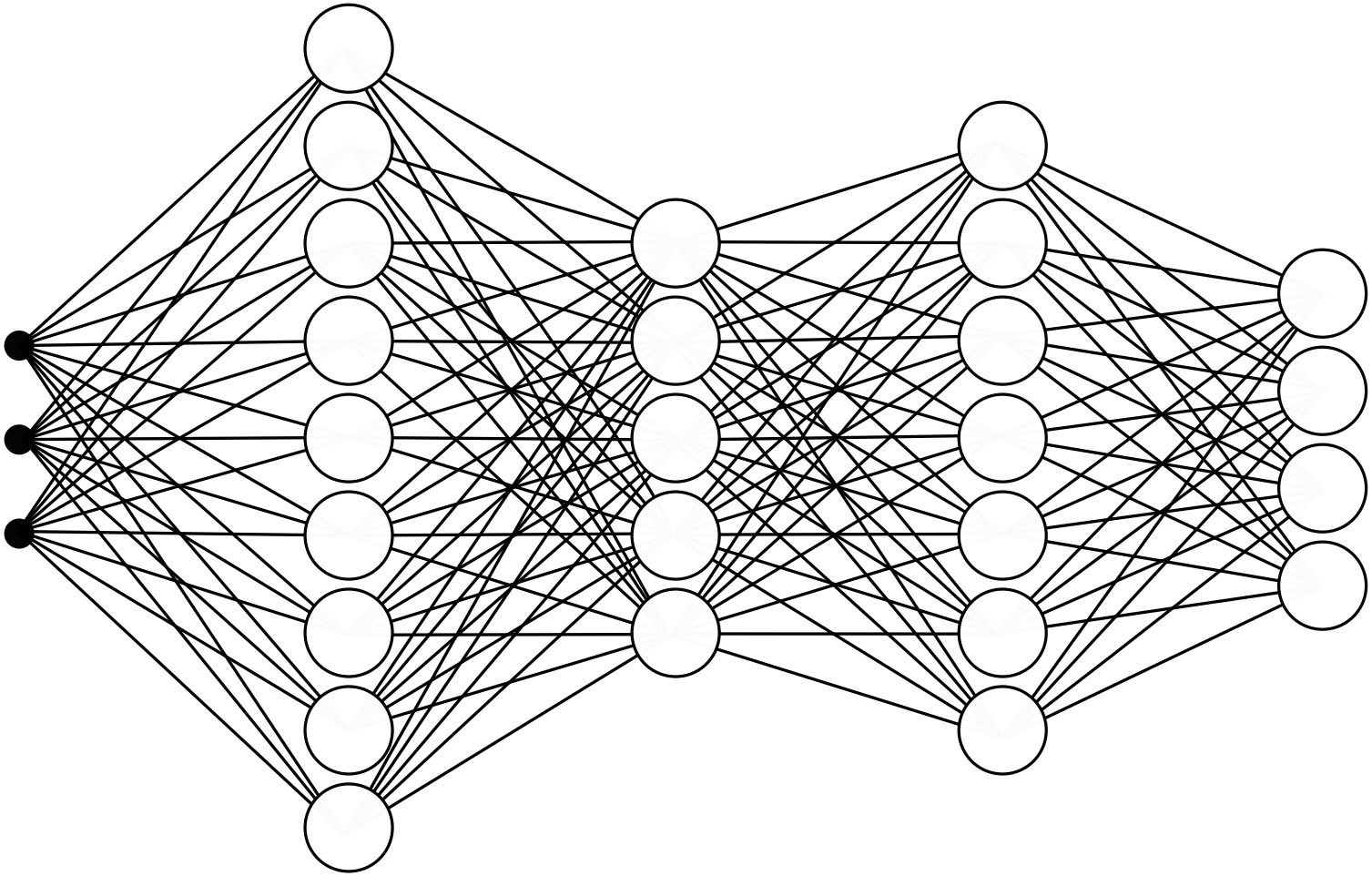
THE UNIVERSITY OF NEW MEXICO

We can construct a complex multilayer structure with this "neuron".



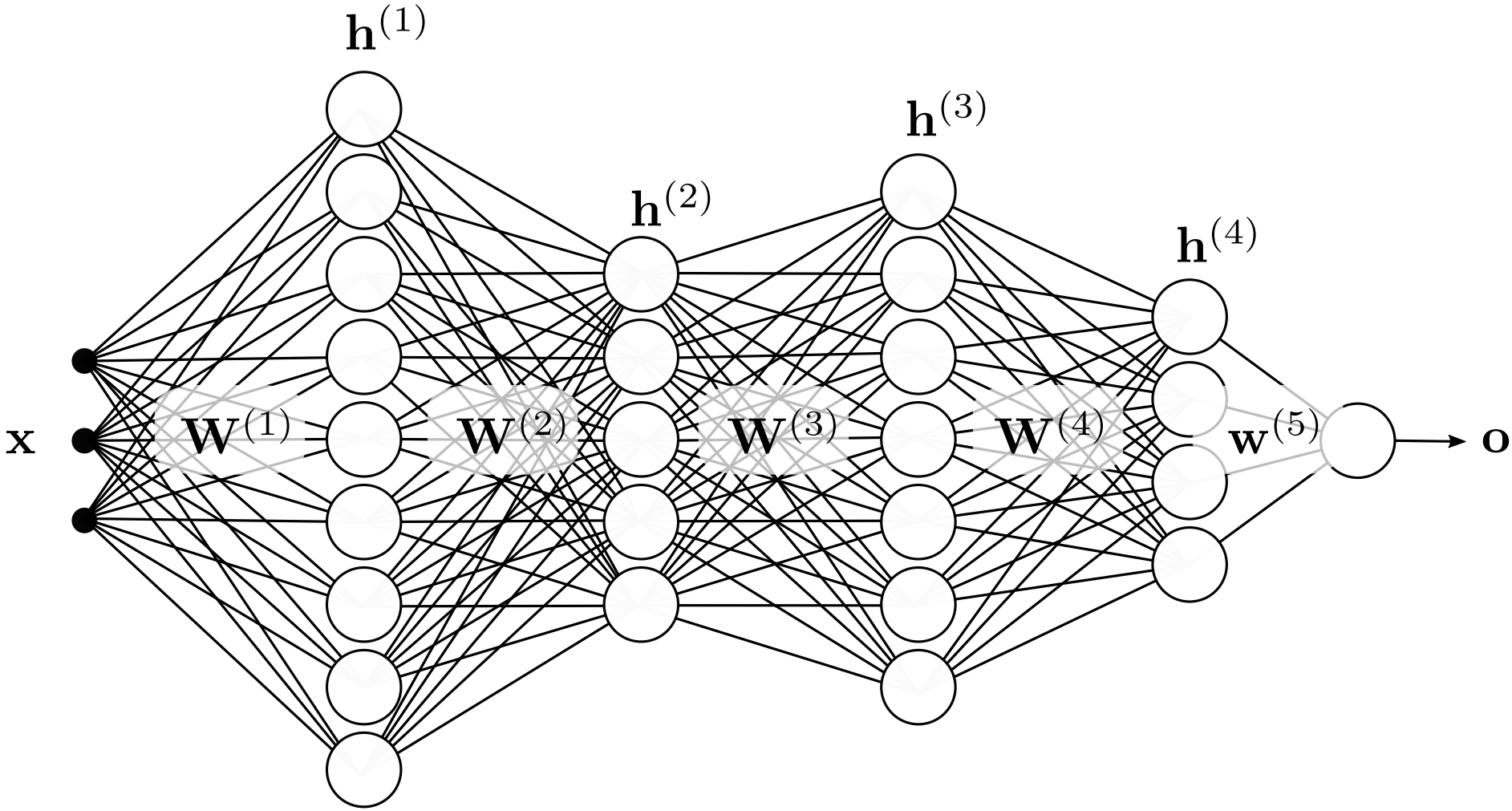
OK, but what about deep learning?

We can construct a complex multilayer structure with this "neuron".

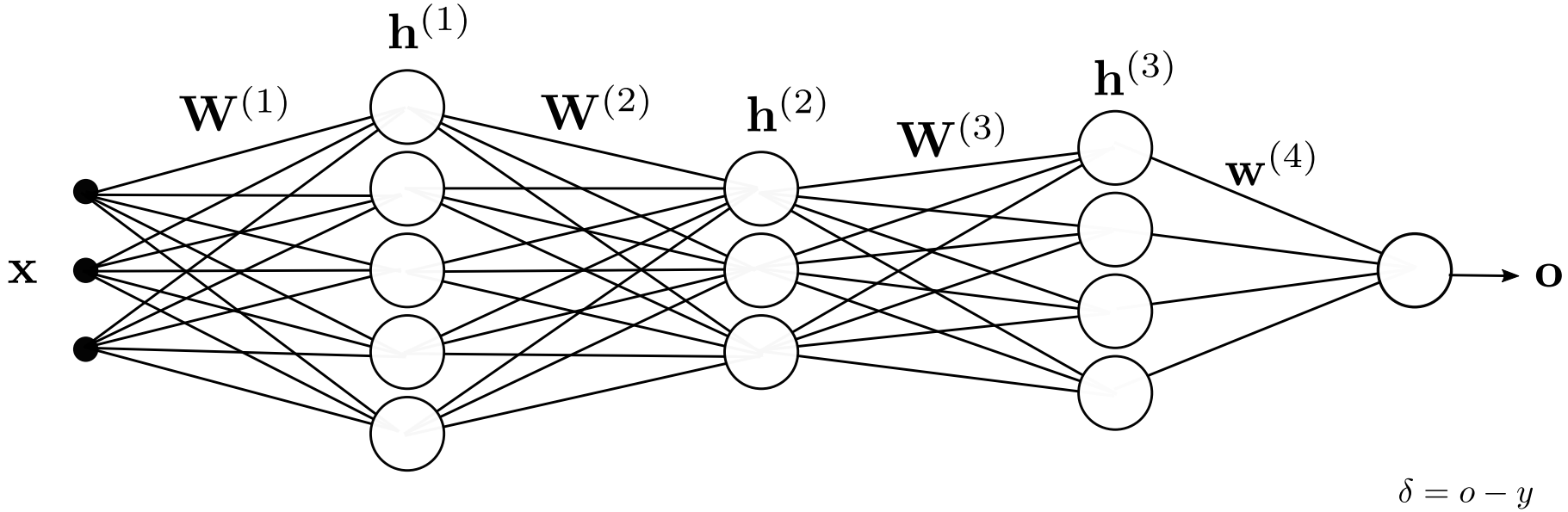


This is one deep learning structure

A neural network can be very complex and have high capabilities.

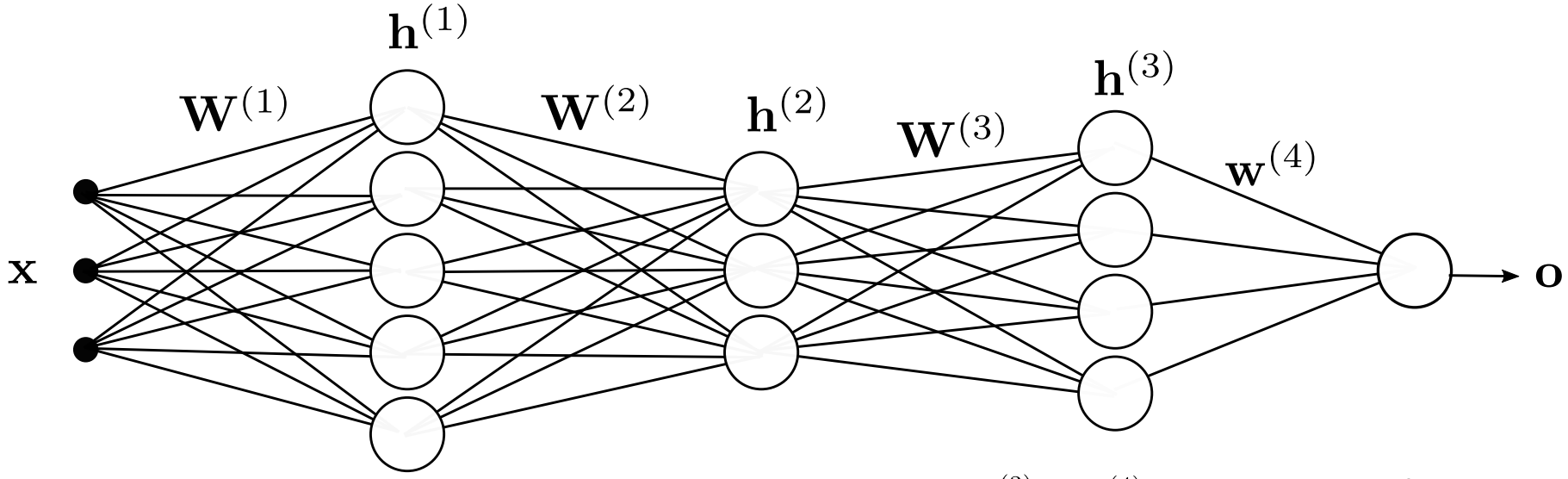


The training algorithm is called backpropagation.



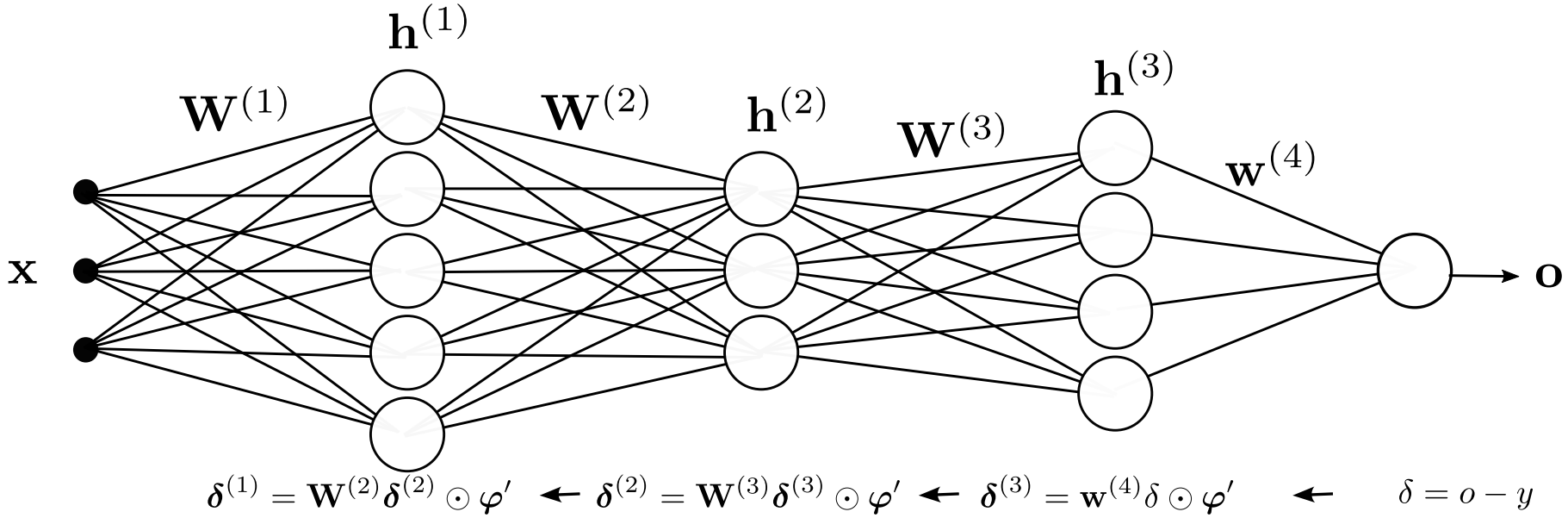
Given an input $\mathbf{x}$ and a desired output y,
we first compute the error δ

The training algorithm is called backpropagation.



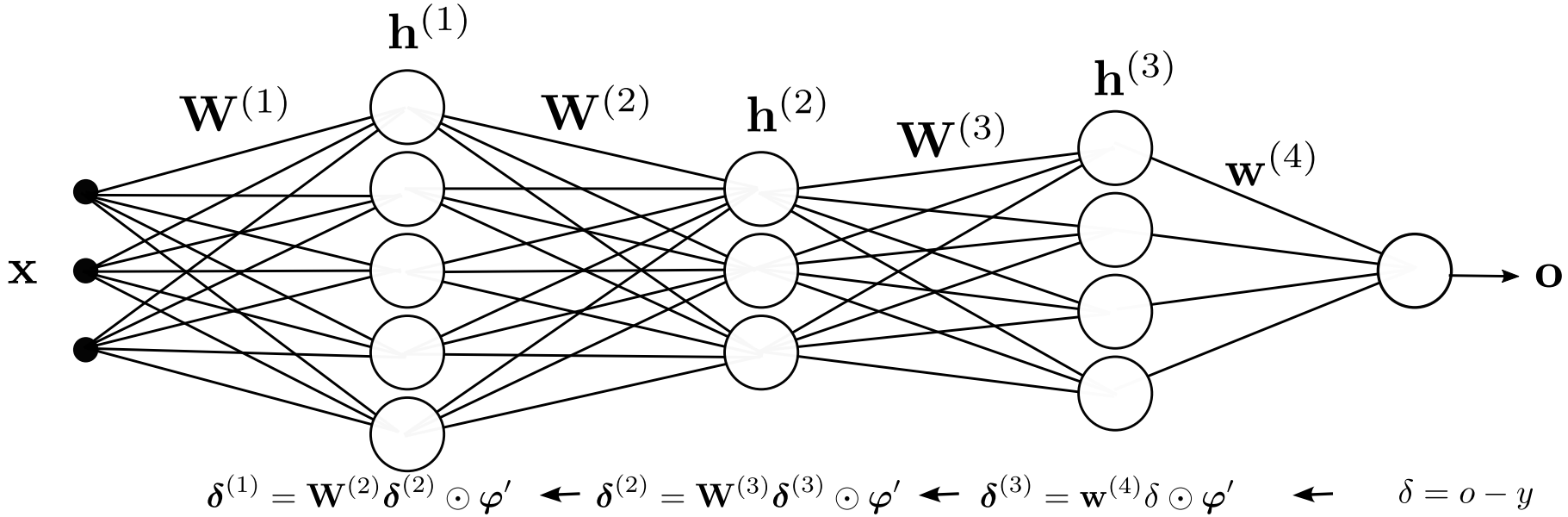
The error is "propagated" back to the previous layer through weight vector $\mathbf{w}^{(4)}$.

The training algorithm is called backpropagation.



The process is iterated to the input with the rest of connection matrices.

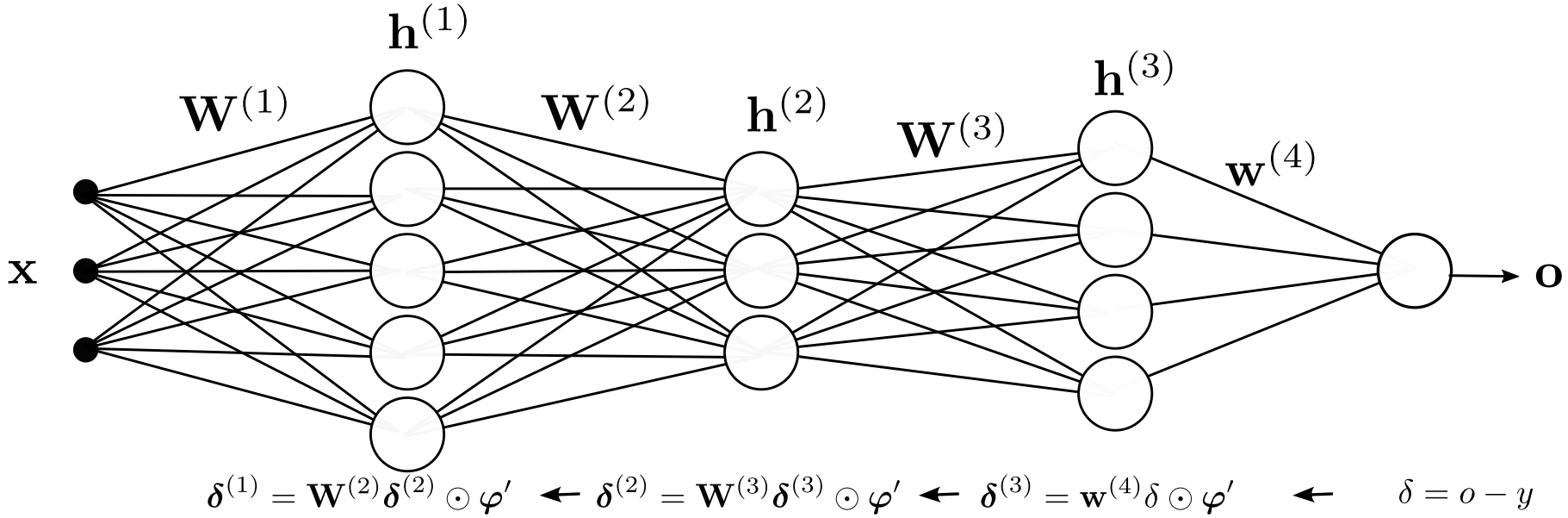
The training algorithm is called backpropagation.



Now we update the weights. The rule is: substract a fraction μ of the input times the error.

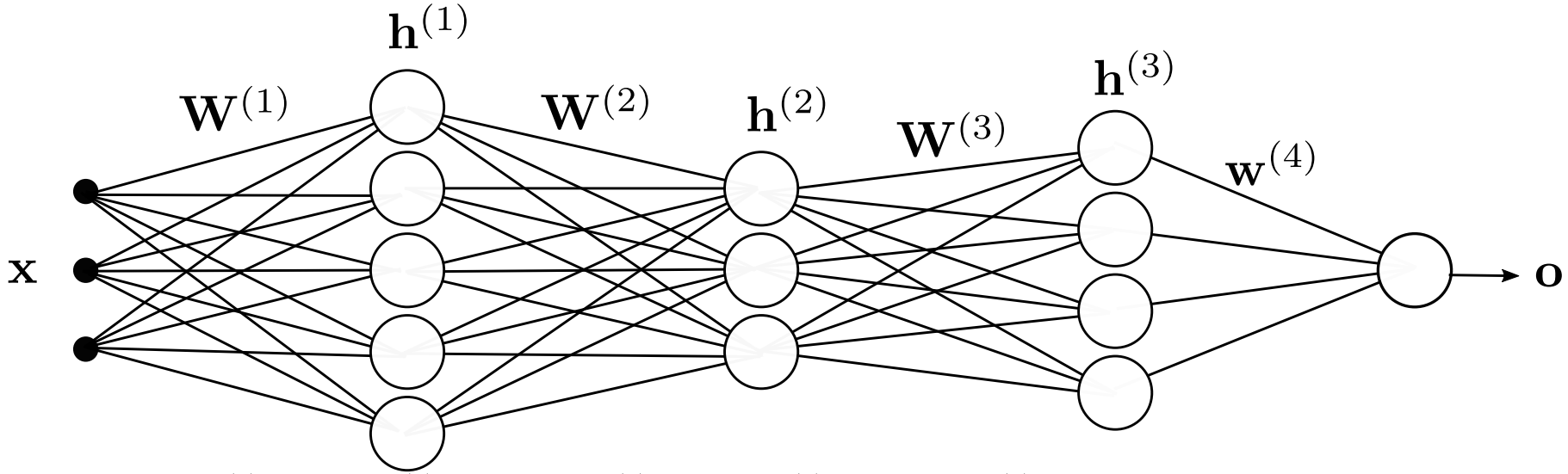
$$\mathbf{w}^{(4)} \leftarrow \mathbf{w}^{(4)} - \mu \mathbf{h}^{(3)} \delta$$

The training algorithm is called backpropagation.



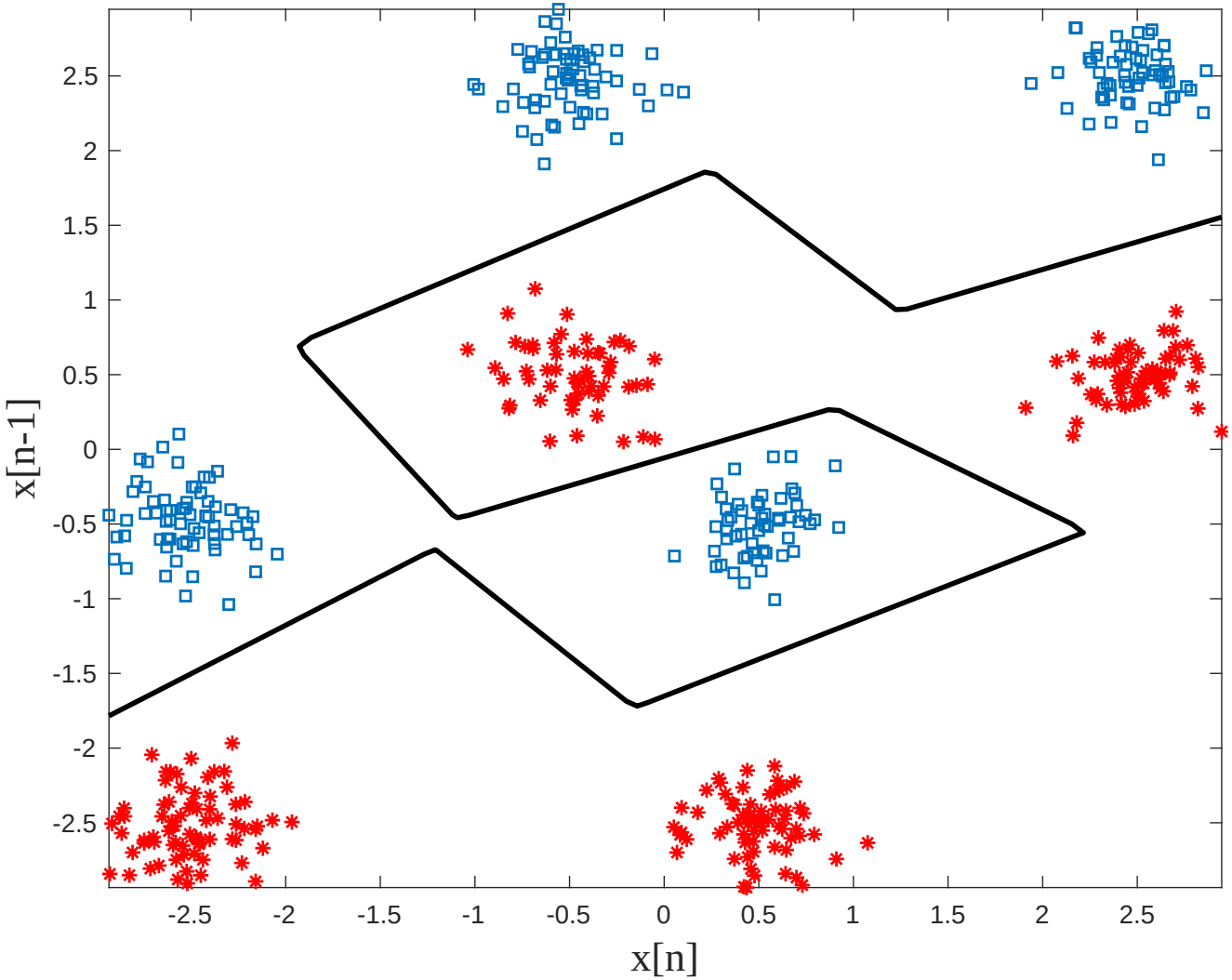
Iterate the process in the previous layer.
Input: $\mathbf{h}^{(2)}$; error: $\boldsymbol{\delta}^{(3)}$.

The training algorithm is called backpropagation.

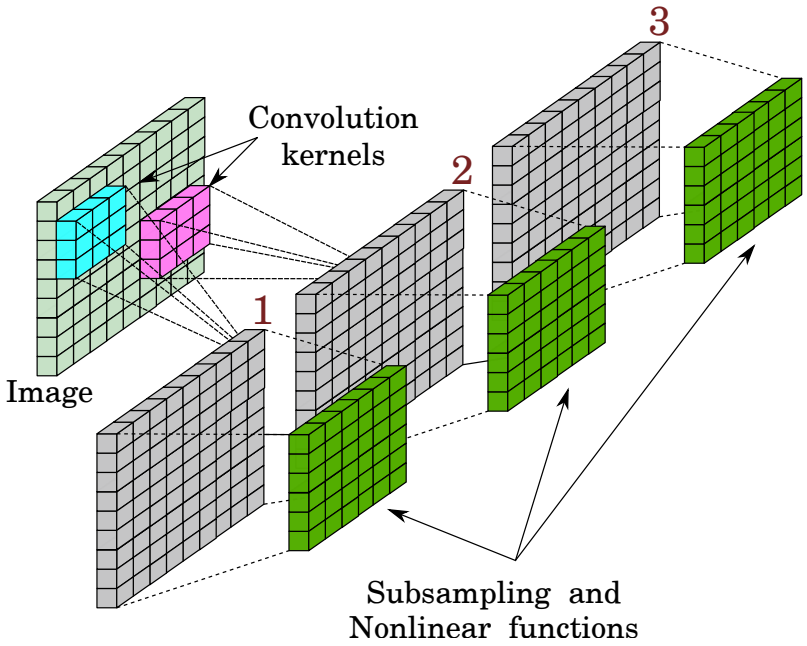


Example of neural network behaviour

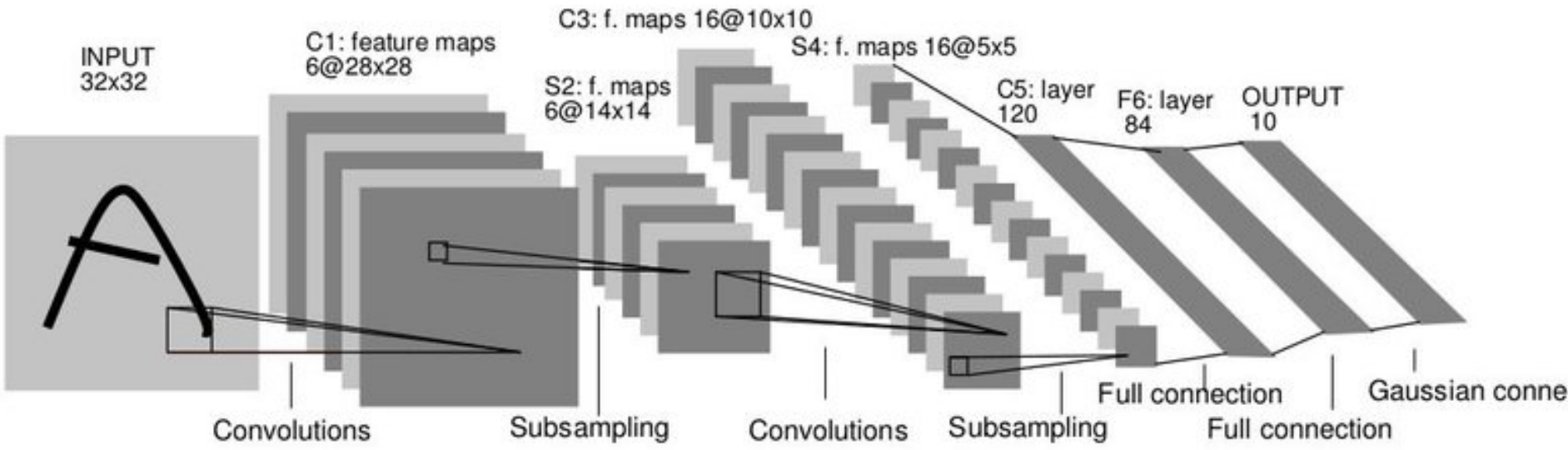
Neural network with 2 hidden layers



- CNNs were introduced in the '80s by LeCun and co-workers.
- It was intended to classify handwritten digits.
- It introduces the concept of local features.



Convolutional neural networks



- 340,000 trainable, 60,000 free parameters.
- Convolution section: 168 parameters; 100,000 connections.
- Error rate: 0.9%

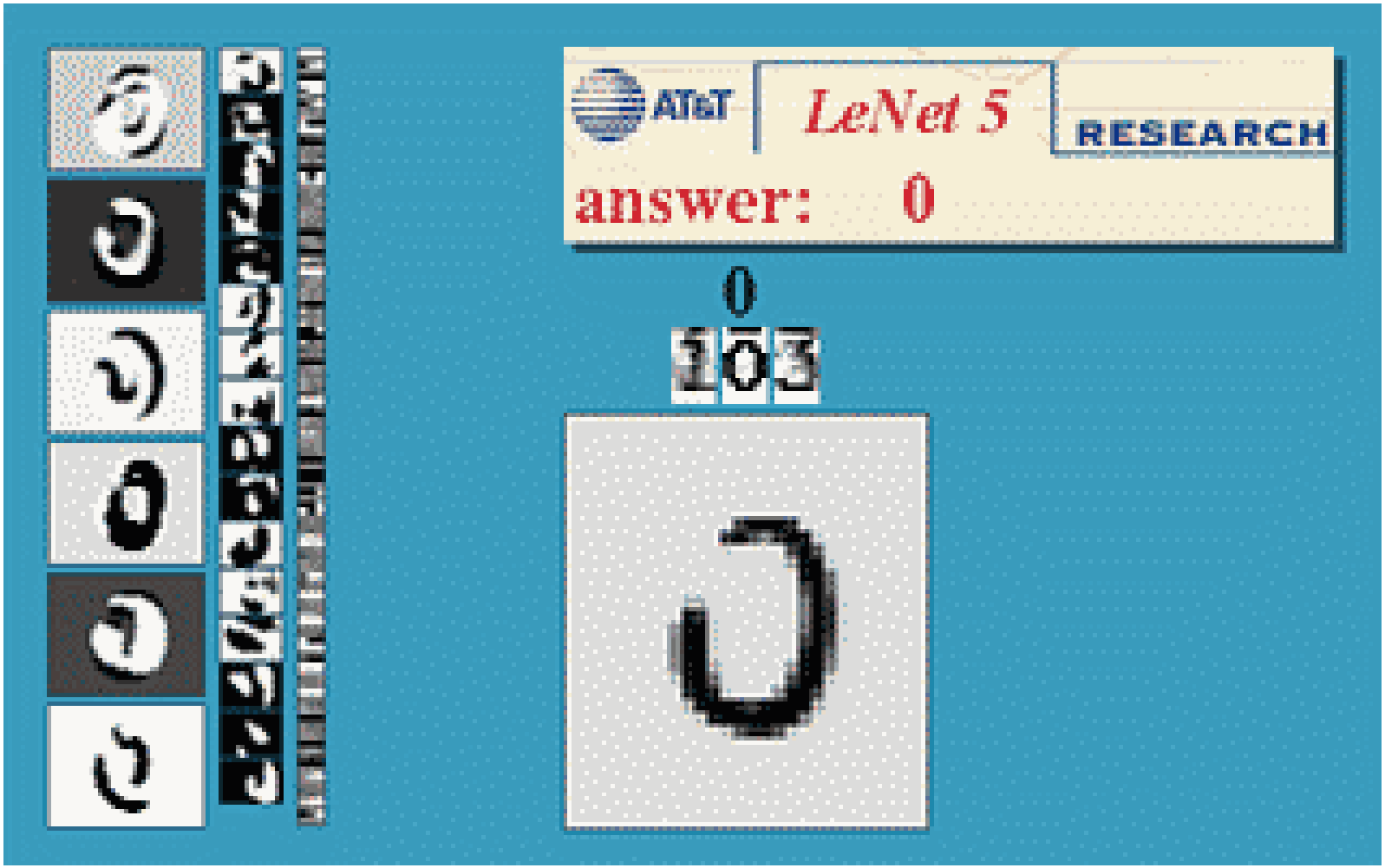
Yann LeCun et al. "Comparison of learning algorithms for handwritten digit recognition." International conference on artificial neural networks. Vol. 60. 1995.

- In a CNN an image with C channels is convolved with a number of convolution kernels to give a new image with a different number of channels

$$\mathbf{H}_k = \varphi\left(\sum_{j=0}^{C_I-1} \mathbf{W}_{j,k} * \mathbf{I}_j + \mathbf{B}_k\right) \tag{1}$$

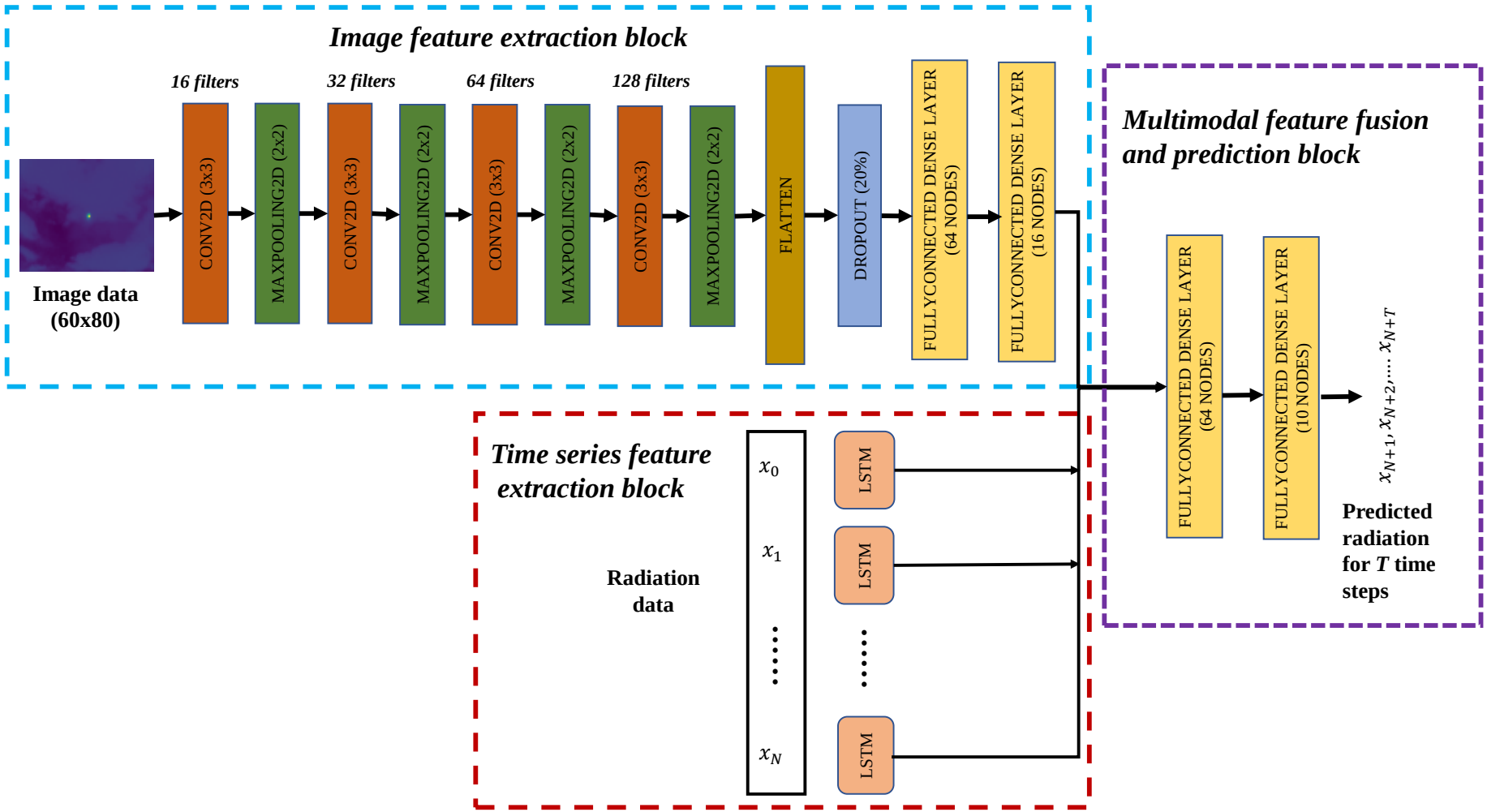
- The output is then passed through a nonlinearity and the process is repeated.
- The training is identic to the training of a NN

$$\boldsymbol{\Delta}^{(l-1)} = \mathbf{W}_{j,k}^{(l)} * \boldsymbol{\Delta}^{(l)} \odot \varphi' \tag{2}$$



http://yann.lecun.com/exdb/lenet/

Application example to smart grid



Single Image Multi-Modal CNN LSTM fusion network

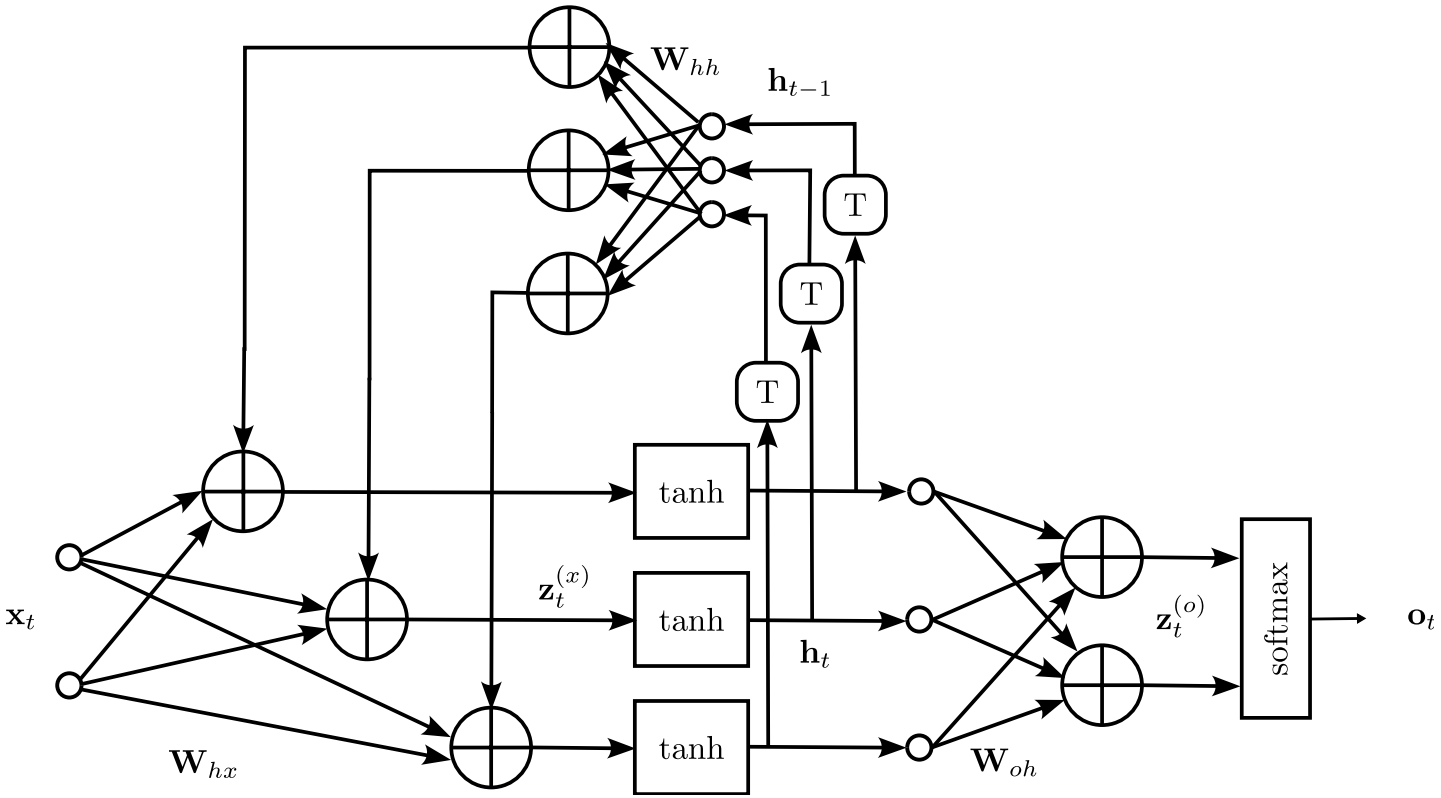
Based on the structure of the CNN, many successful variants have been published:

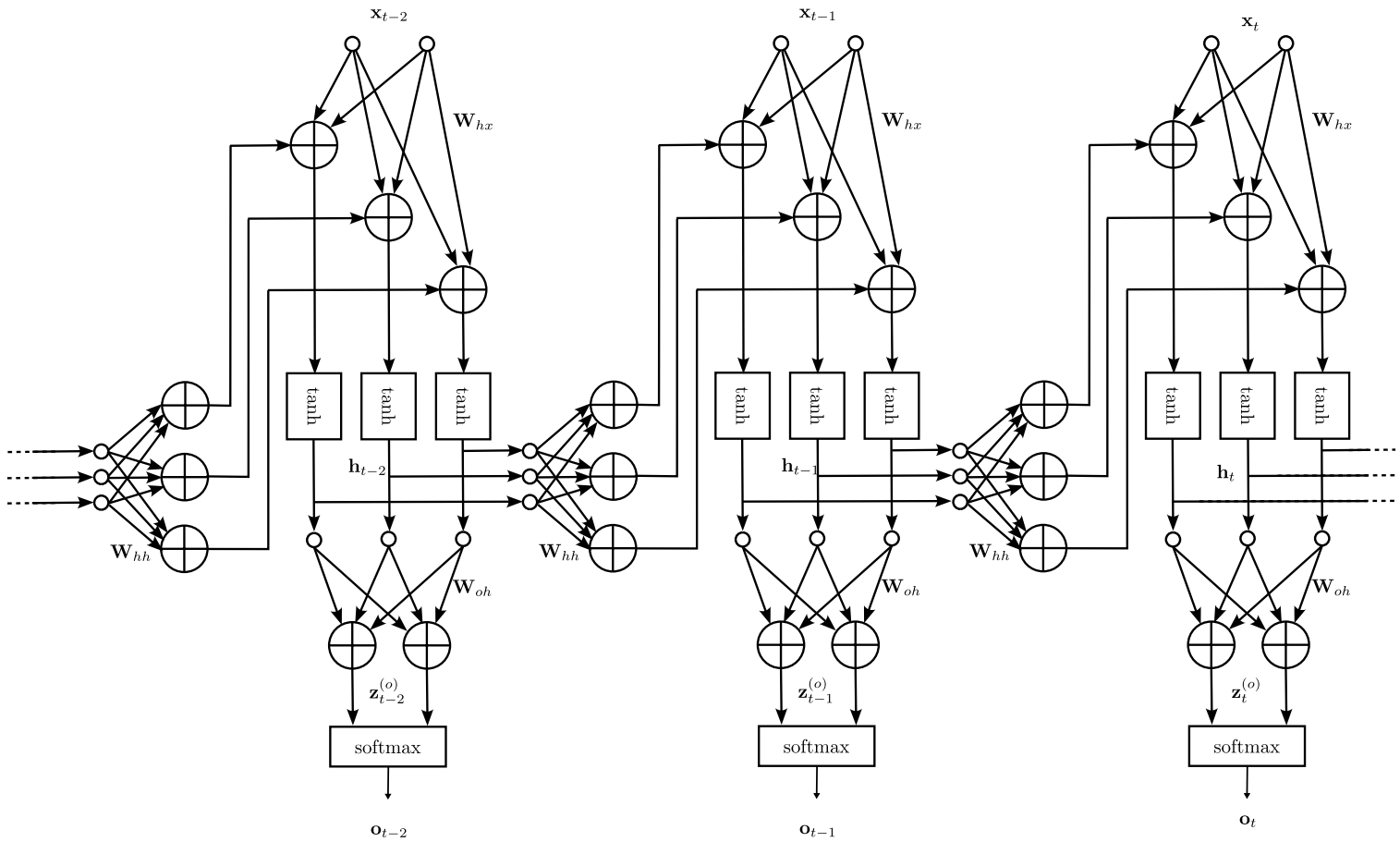
- Alexnet: Image classification (1000 classes) of the ImageNet database.
- Visual Geometric Group (VGG): Won the 2014 Imagenet Large Scale Visual Recognition Challenge competition.
- Inception Network. 5M parameters, also won the ILSVR.
- ResNet: Won the Imagenet localization and detection and the Common Objects in Context (COCO) segmentation and detection contests of ILSVR in 2015.
- Xception: Google, 2017. Inspired in the Inception network.
- Others: Mobilenet, Densenet, EfficientNet...

Neural networks for sequences

- Sequential data requires mechanisms to retain past data information to predict the sequence's future values.
- Recurrent NNs were introduced in the 80s for this purpose.
- Examples of use are:
  - Prediction of future values of a time signal.
  - Predictive text generation.

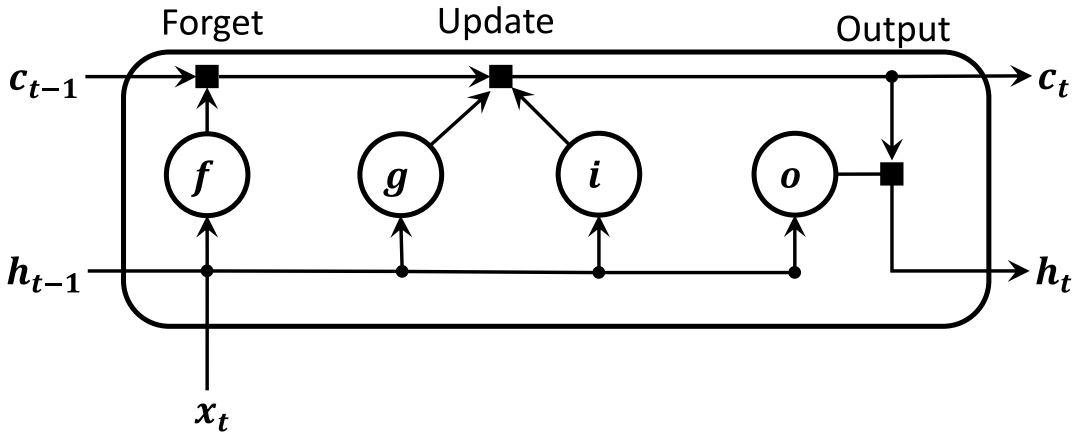
The structure of an RNN has the aspect of an autoregressive model.



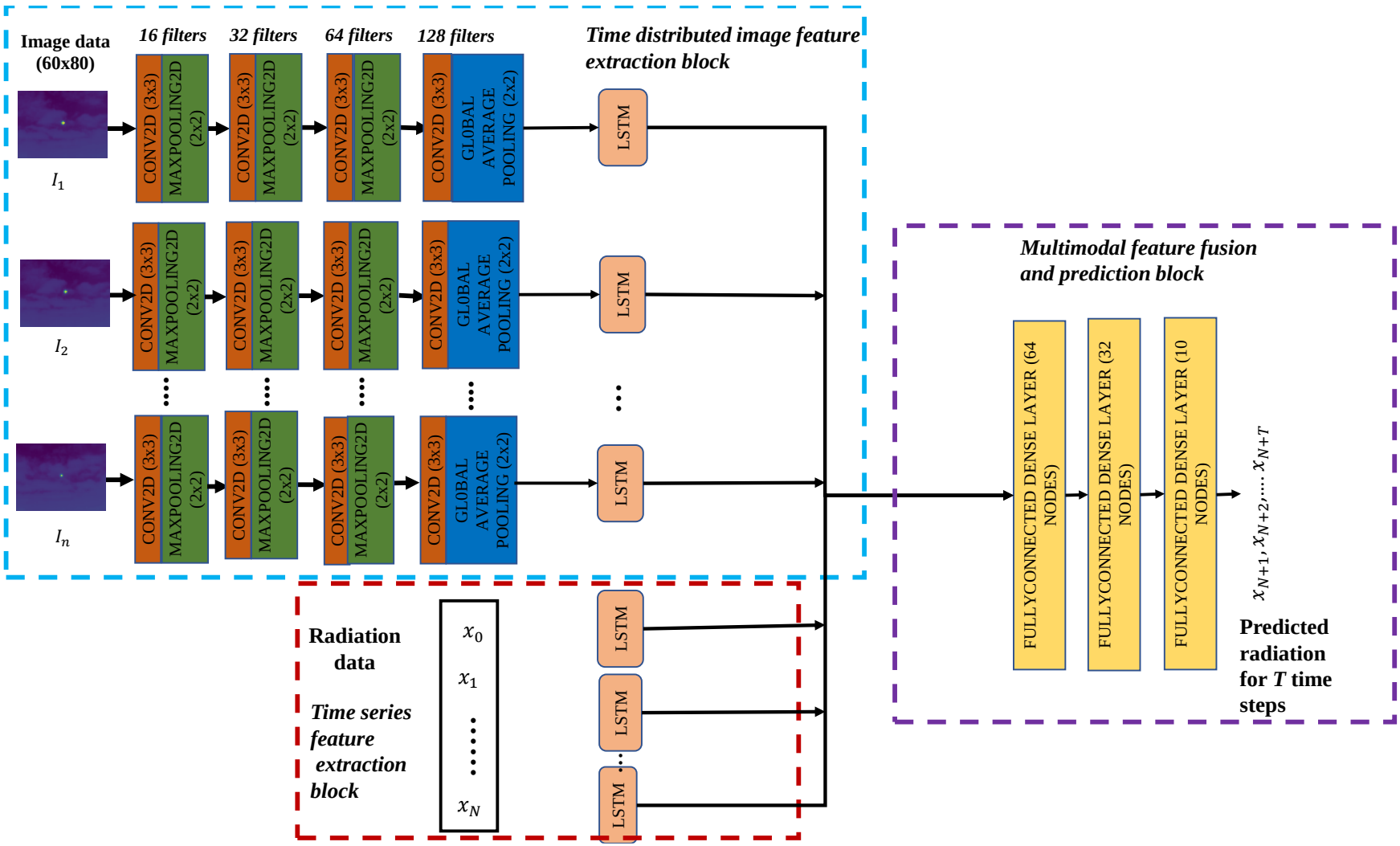


- Input: $\mathbf{x}[n-N], \cdots, \mathbf{x}[n]$. Prediction $\mathbf{o}[n-N], \cdots, \mathbf{o}[n]$.
- The backpropagation is formally equal to the one of a NN.
- Here, it is called *backpropagation in time*.

- In spite of its simplicity, the RNN is difficult to train.
  - It shows the phenomena of vanishing and exploding gradients.
  - It cannot store long term dependencies.
  - Other approaches that are more versatile have been developed with the same idea:
    - Deep RNNs
    - Long Short Term Memory Units
    - Gated Recurrent Units

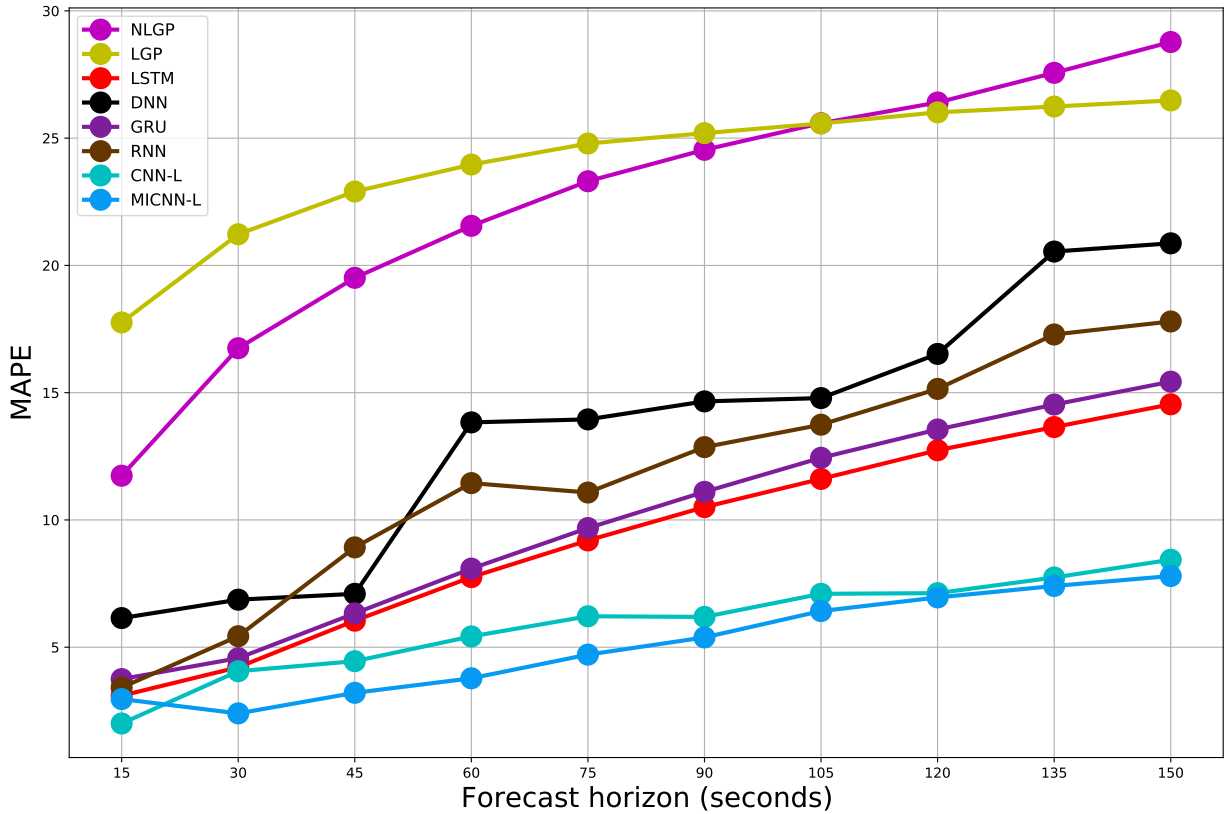


Application example to smart grid



Multiple Image Multi-Modal CNN LSTM fusion network

Application example to smart grid

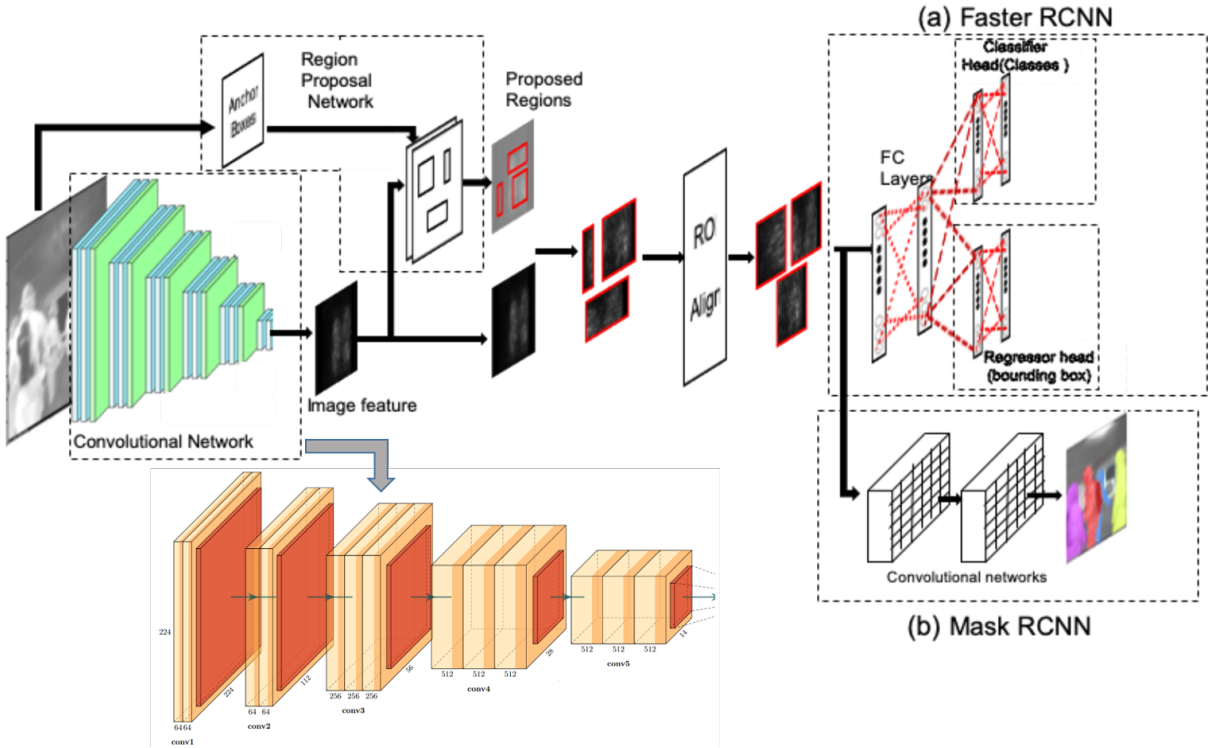


Comparison of MAPE using different forecasting methods. Training and test data with cloudy days.

Meenu Ajith, Manel Martínez-Ramón, Deep learning based solar radiation micro forecast by fusion of infrared cloud image and radiation data, Applied Energy,2022.

Use of ML in real-time with human teams.

- Purpose: reduce the fatality risk of first responders



Objective: Object detection and segmentation.

A Region Proposal Network (RPN) detects the significant regions

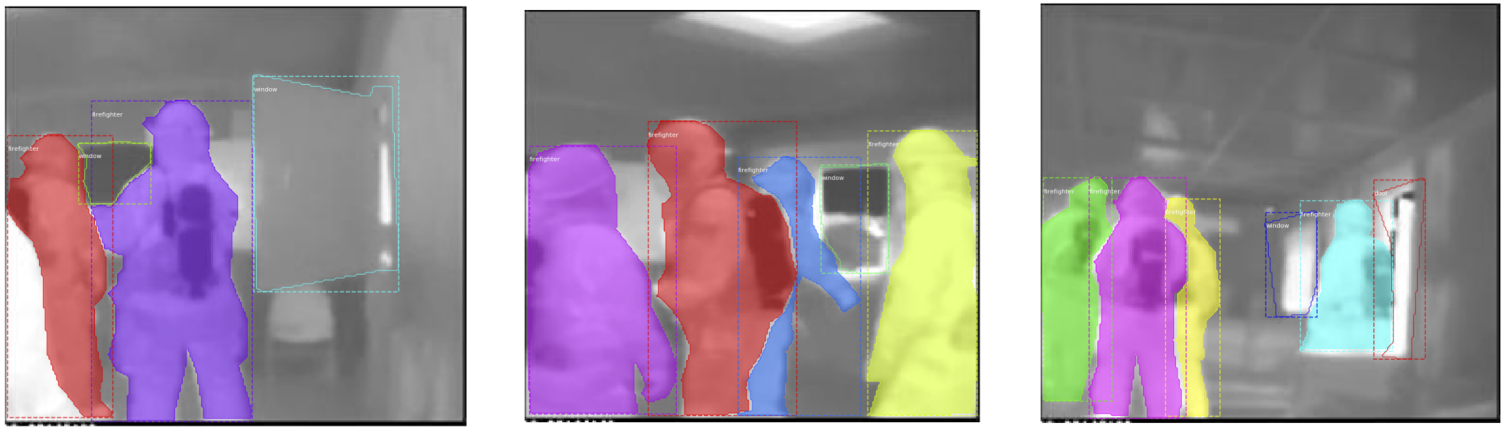
For object detection/tracking,

* a neural network classifies the objects
* a regressor estimates bounding boxes.

For image segmentation, a CNN generates a mask overlayed with the original image.

Bhattarai, Martínez-Ramón. "A Deep Learning Framework for Detection of Targets in Thermal Images to Improve Firefighting." IEEE Access 8 (2020).

Object detection and tracking with Faster RCNN.

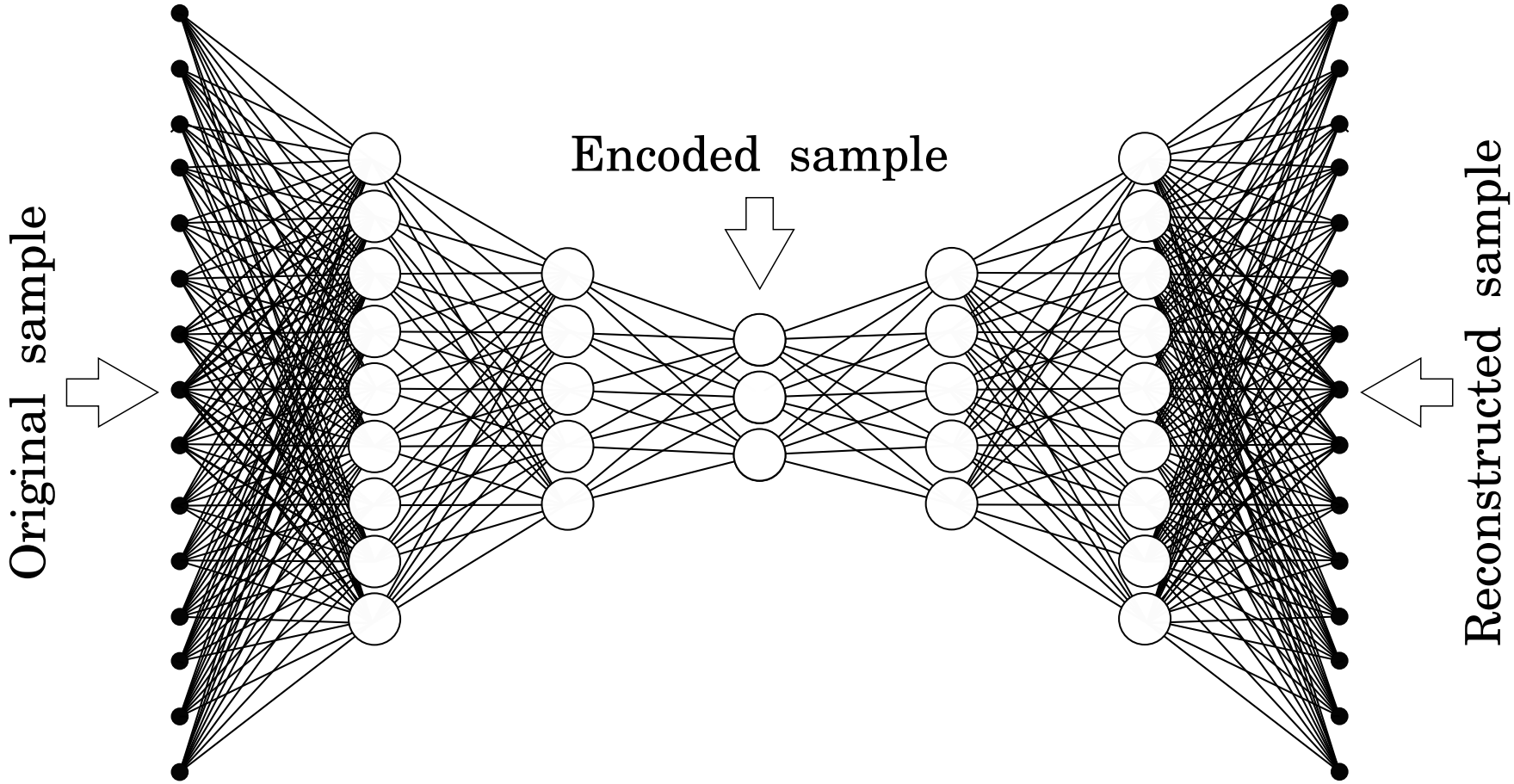


Object instance segmentation with Mask RCNN.

Bhattarai, Martínez-Ramón, "An embedded deep learning system for augmented reality in firefighting applications", in preparation.

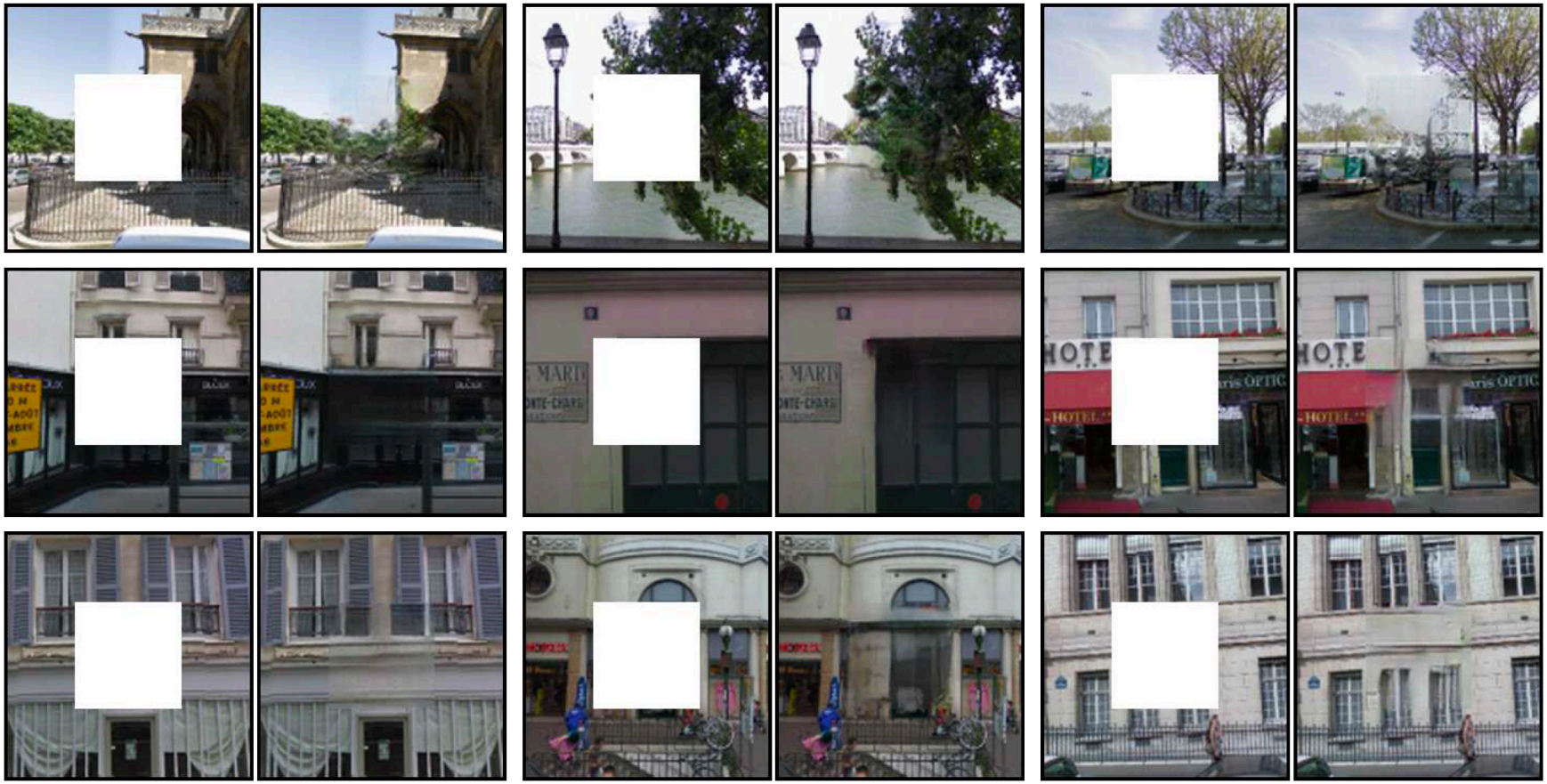
Deep belief networks and autoencoders

Autoencoders are intended to code the information of a pattern in a reduced dimension space.



Hinton, Salakhutdinov. "Reducing the dimensionality of data with neural networks." Science 313.5786 (2006).

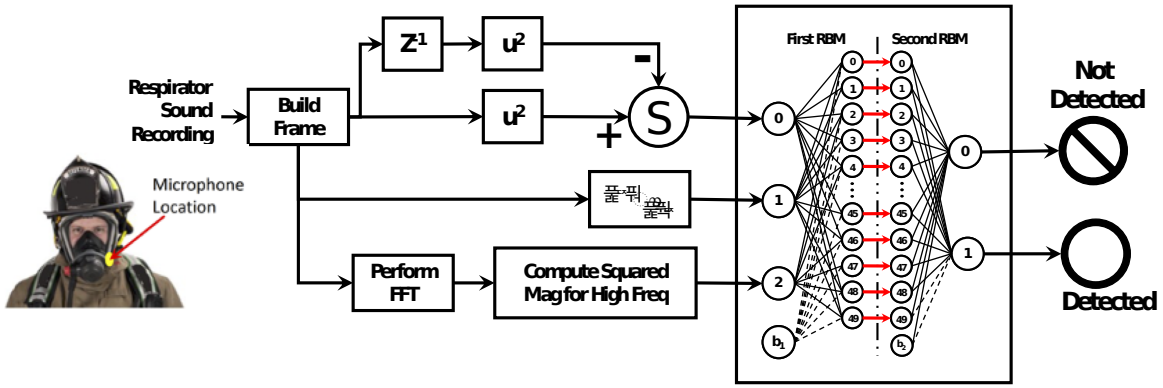
Image processing

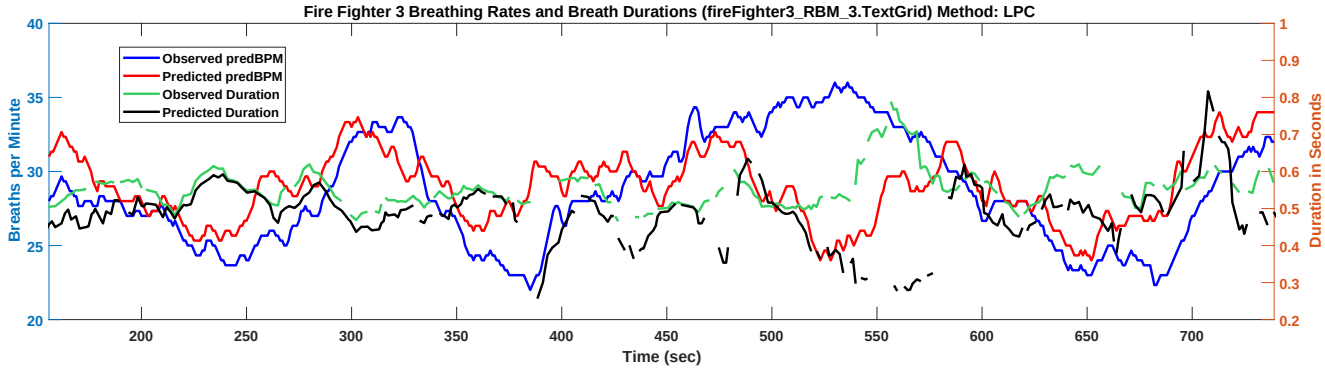


Pathak, Deepak, et al. "Context encoders: Feature learning by inpainting." Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. 2016.

Respiration/voice analysis with deep learning

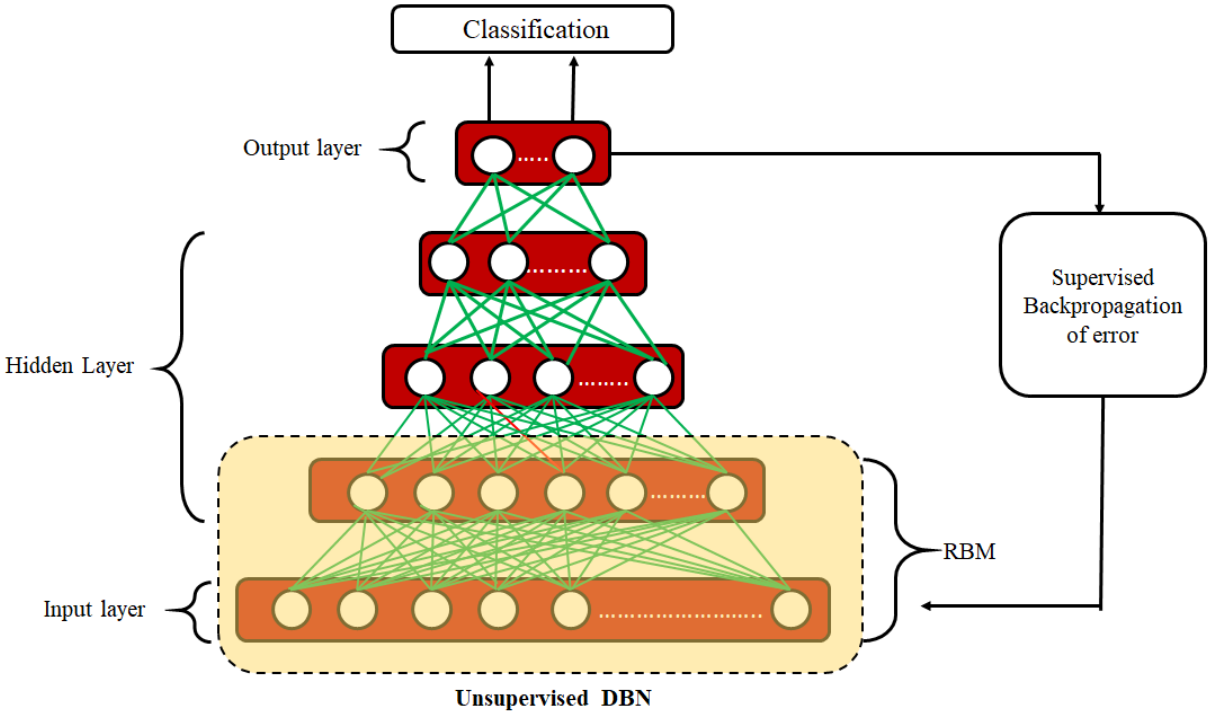
Feature extraction for no invasive analysis of firefighter condition (exhaustion, stress) can reduce accidents, strokes, and others.

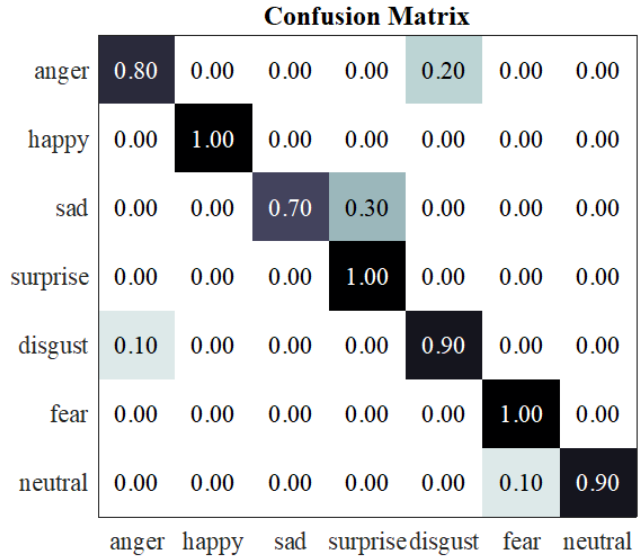




Hamke, Martínez-Ramón, Nafchi, Jordan. "Detecting breathing rates and depth of breath using LPCs and Restricted Boltzmann Machines." Biomedical Signal Processing and Control 48 (2019).

Emotion recognition





Database	Deep learning	SVM
CK+	98.57	28.57
RaFD 135	91.95	84.00
RaFD 90	94.50	88.20
RaFD 45	92.75	86.28

- Training time: 0.06 seconds per image.
- Test time: 0.02 seconds per image.

Kurup, Ajith, Martínez-Ramón. "Semi-supervised facial expression recognition using reduced spatial features and Deep Belief Networks." Neurocomputing 367 (2019).

Conclusion

- Well established, powerful. Allows machines to learn from data.
- Applications in many fields of science and engineering, growing.
- Research is not closed and it continues producing results.
- Future will depend on the available computational resources.
- ML is publicly available research done for humankind to use it.
- AI and ML are a key technology for peace engineering.

Peace Engineering