Deep learning for image analysis quick introduction

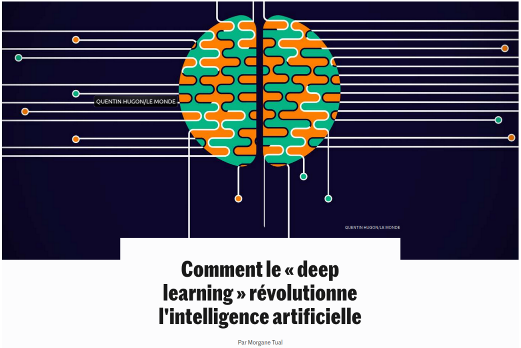
E. Decencière

MINES ParisTech
PSL Research University
Center for Mathematical Morphology

Contents

Contents

The rise of deep learning



Le Monde, juillet 2015

The rise of deep learning

Nature, 2016

The rise of deep learning

Le prix Turing récompense trois pionniers de l'intelligence artificielle (IA)

L'association américaine ACM a remis son prestigieux prix aux chercheurs français, canadien et britannique : Yann LeCun, Yoshua Bengio et Geoffrey Hinton.

Par David Larousserie · Publié le 27 mars 2019 à 11h01 - Mis à jour le 29 mars 2019 à 12h11

Le Monde, mars 2019

The rise of deep learning

Pour Elon Musk, l'intelligence artificielle pourrait menacer la civilisation

L'entrepreneur américain, qui a fondé Tesla, a alerté les politiques américains sur la nécessité de réguler l'intelligence artificielle.

Par **Le Figaro**
Publié le 18/07/2017 à 06:00, mis à jour le 18/07/2017 à 11:25

Le Figaro, juillet 2017

Contents

Artificial neural networks and deep learning history

- 1958: Rosenblatt's perceptron [Rosenblatt, 1958]

Artificial neural networks and deep learning history

- 1958: Rosenblatt's perceptron [Rosenblatt, 1958]
- 1979: Neocognitron (convolutional neural network architecture) [Fukushima, 1979, Fukushima, 1980]

Artificial neural networks and deep learning history

- 1958: Rosenblatt's perceptron [Rosenblatt, 1958]
- 1979: Neocognitron (convolutional neural network architecture) [Fukushima, 1979, Fukushima, 1980]
- 1980's: the backpropagation algorithm ([Werbos, 1982]; see also the work of LeCun [LeCun, 1985])

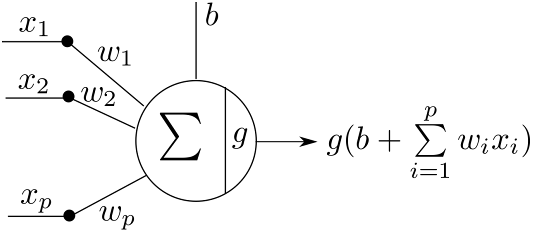
Artificial neural networks and deep learning history

- 1958: Rosenblatt's perceptron [Rosenblatt, 1958]
- 1979: Neocognitron (convolutional neural network architecture) [Fukushima, 1979, Fukushima, 1980]
- 1980's: the backpropagation algorithm ([Werbos, 1982]; see also the work of LeCun [LeCun, 1985])
- 2006-: CNN implementations using Graphical Processing Units (GPU): up to a 50 speed-up factor.

Artificial neural networks and deep learning history

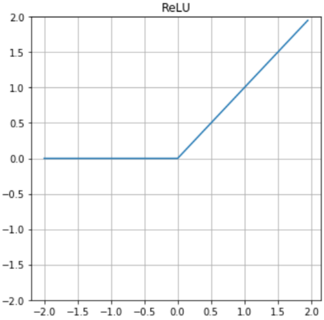
- 1958: Rosenblatt's perceptron [Rosenblatt, 1958]
- 1979: Neocognitron (convolutional neural network architecture) [Fukushima, 1979, Fukushima, 1980]
- 1980's: the backpropagation algorithm ([Werbos, 1982]; see also the work of LeCun [LeCun, 1985])
- 2006-: CNN implementations using Graphical Processing Units (GPU): up to a 50 speed-up factor.
- 2012: Imagenet image classification won by a CNN with AlexNet [Krizhevsky et al., 2012].

Artificial neuron



Activation: rectified linear unit (ReLU)

$$g(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$$

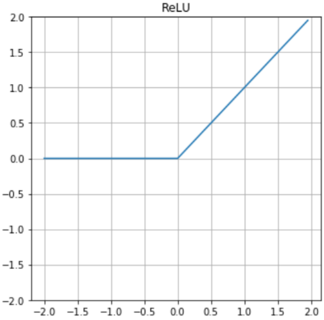


Remarks

+ Usable gradient when activated
+ Fast to compute
+ High abstraction

Activation: rectified linear unit (ReLU)

$$\mathrm{g}(x) = \begin{cases} x, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}$$

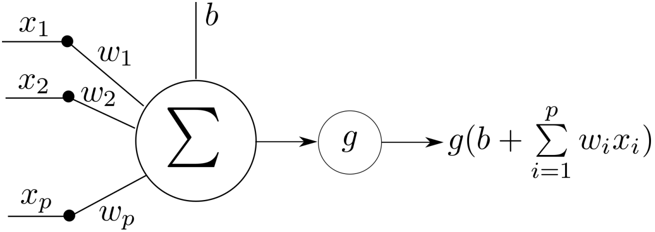


Remarks

- + Usable gradient when activated
- + Fast to compute
- + High abstraction

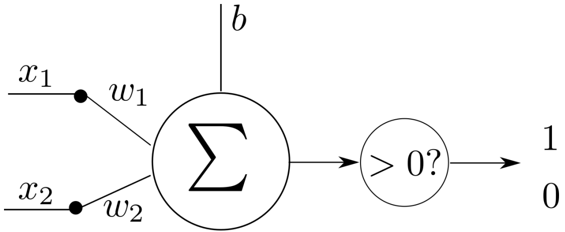
ReLU is the most commonly used activation function.

What can an artificial neuron compute?



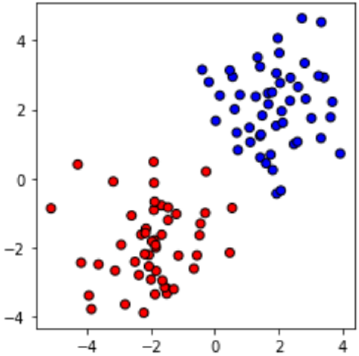
In $\mathbb{R}^p$, $b + \sum_{i=1}^{p} w_i x_i = 0$ corresponds to a hyperplane H. For a given point $\mathbf{x} = \{x_1, \ldots, x_p\}$, decisions are made according to the side of the hyperplane it belongs to.

Example of what we can do with a neuron

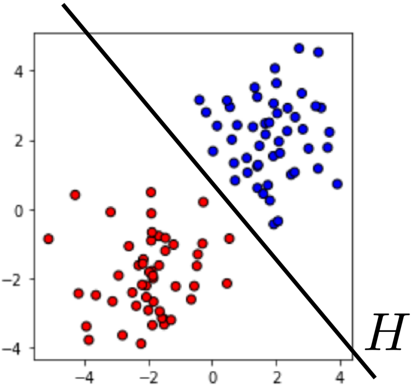


- $p = 2$: 2-dimensional inputs (can be represented on a screen!)
- Activation: binary
- Classification problem

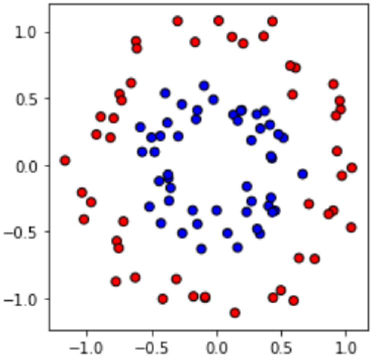
Gaussian clouds



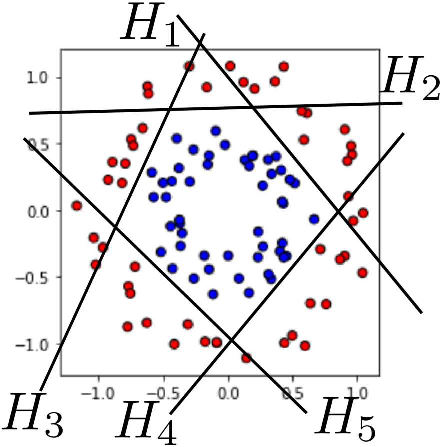
Gaussian clouds



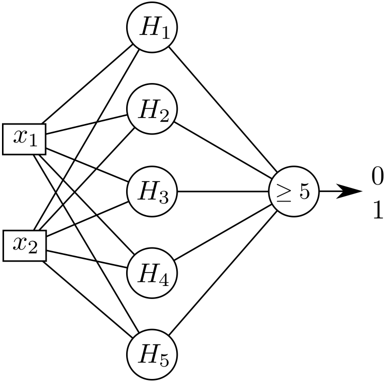
Circles

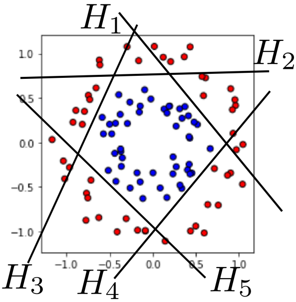


Circles



Solution



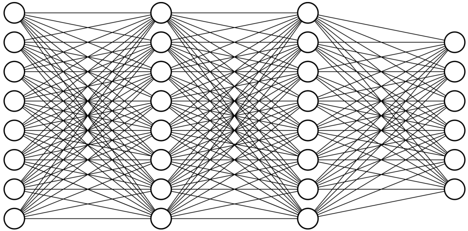


Intuition

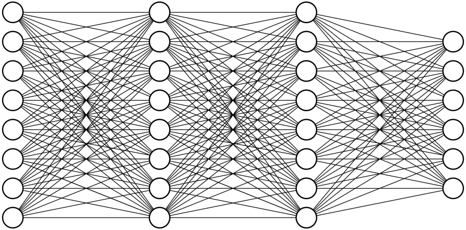
Combining several neurons one can build complex classifiers.

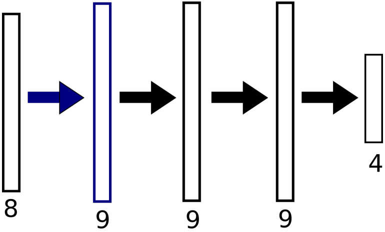
Fully-connected layers

- A layer is said to be fully-connected if each of its neurons is connected to all the neurons of the previous layer



Graphical representation of NNs





- Data is organized into arrays, linked with operators
- A layer corresponds to an operator between arrays.

Learning with artificial neural networks

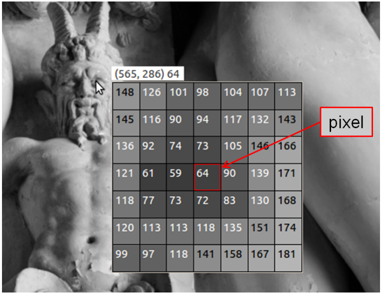
- Constitute your learning, validation and test sets
- Define a loss function
- Use gradient descent (backpropagation) to minimize the loss

Contents

A picture is worth a thousand words

Definition

- Classically, an image is a matrix of values belonging to $[0, \ldots, 255]$ (grey level images) or to $[0, \ldots, 255]^3$ (color images).
- More generally, an image is a q-dimensional array of values belonging to R^d.



Grey level values around the left eye of the faun

The role of annotated image databases

Image databases including *annotations* (typically some kind of high level information) are essential to the development of *supervised* machine learning methods for image analysis.

Annotations

- Image class
- Measure(s) obtained from the image
- Position of objects within the image
- Segmentation

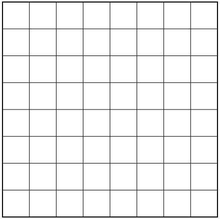
MNIST database [Lecun et al., 1998]

- The Modified National Institute of Standards and Technology (MNIST) database contains $60\,000$ training images of hand-written digits, and $10,000$ test images.
- Image size: 28×28

Credits: Images from MNIST assembled by Josef Stepan (licensed under CC BY-SA 4.0)

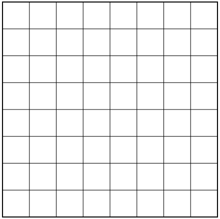
Layers representation

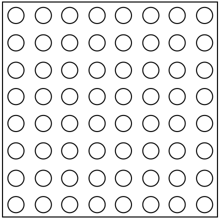
For illustration purposes, in the following slides images and filters will be displayed as rows of neurons – these can be seen as 1D arrays or as sections of 2D arrays.
We represent some connections between neurons. Each such connection is associated to a weight. The bias are not represented, to avoid clutter, but must not be forgotten.



Layers representation

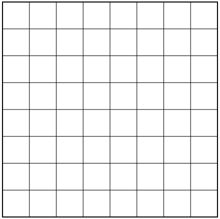
For illustration purposes, in the following slides images and filters will be displayed as rows of neurons – these can be seen as 1D arrays or as sections of 2D arrays.
We represent some connections between neurons. Each such connection is associated to a weight. The bias are not represented, to avoid clutter, but must not be forgotten.

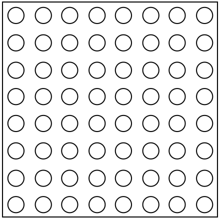




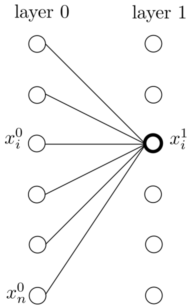
Layers representation

For illustration purposes, in the following slides images and filters will be displayed as rows of neurons – these can be seen as 1D arrays or as sections of 2D arrays.
We represent some connections between neurons. Each such connection is associated to a weight. The bias are not represented, to avoid clutter, but must not be forgotten.



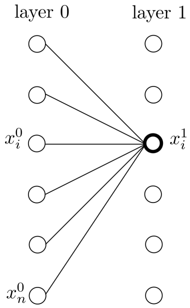


Towards convolutional layers

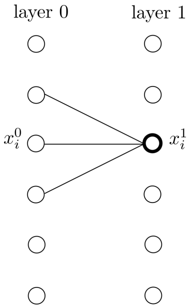


Fully connected layer:
$n(n+1)$ weights

Towards convolutional layers

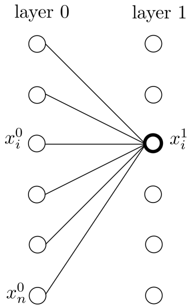


Fully connected layer:
$n(n+1)$ weights

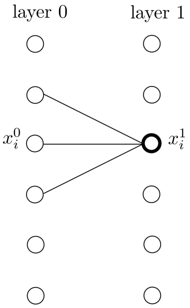


Locally conn. layer:
$n(s+1)$ weights

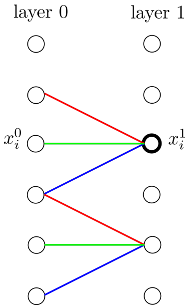
Towards convolutional layers



Fully connected layer:
$n(n+1)$ weights



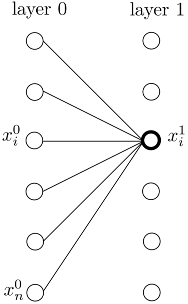
Locally conn. layer:
$n(s+1)$ weights



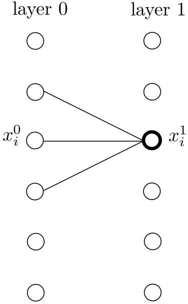
Weight replication: $s+1$ weights.
Convolutional layer.

Towards convolutional layers: some figures

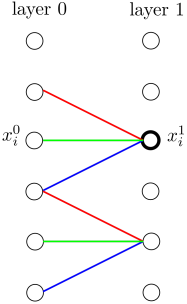
- 3×3 convolutions: $s = 9$
- Toy image: $n = 28 \times 28 = 784$
- Typical image: $n = 1000 \times 1000 = 10^6$



Fully connected layer:
$n(n+1)$ weights
$\approx 6.10^5$
$\approx 10^{12}$



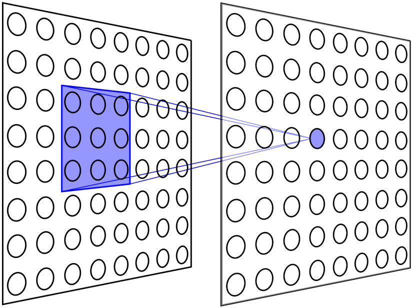
Locally conn. layer:
$n(s+1)$ weights
7840
10^7



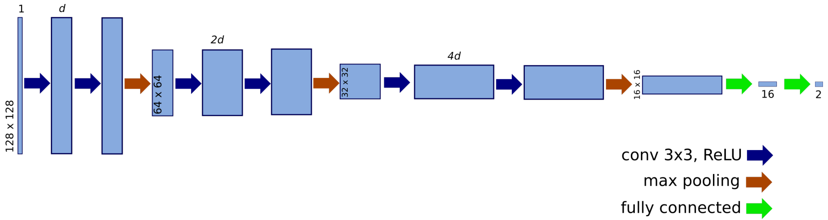
Weight replication: $s+1$ weights.
10
10

Convolutional layer illustration in 2D

- Illustration of a convolution of size 3×3

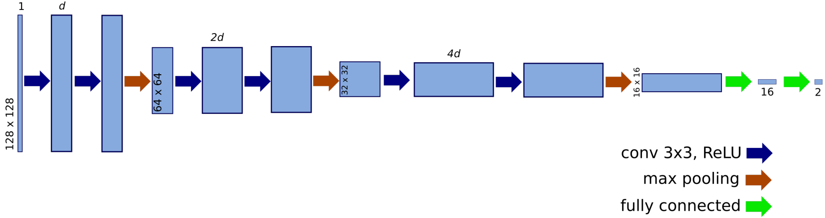


1D representations



Credits: NN is work of Robin Alais et al.
Fundus image by Mikael Häggström, used
with permission (CC0).

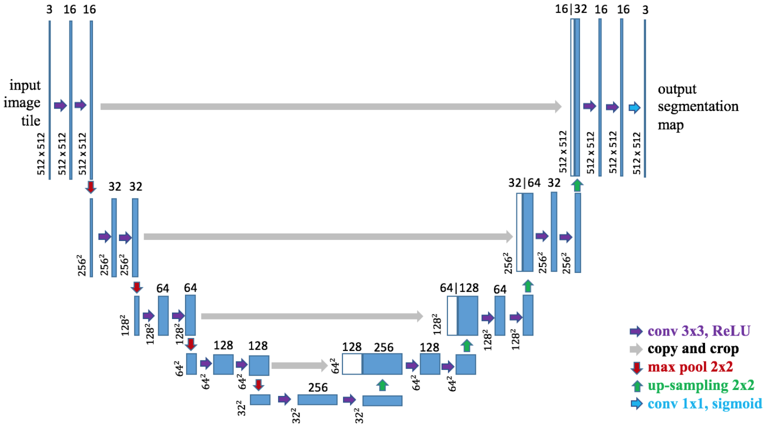
1D representations



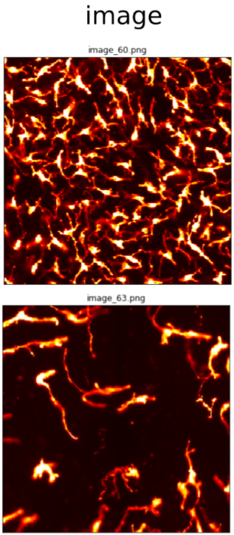
This NN was used to estimate the position of the center of the macula on fundus images.

Credits: NN is work of Robin Alais et al.
Fundus image by Mikael Häggström, used with permission (CC0).

U-Net architecture [Ronneberger et al., 2015]

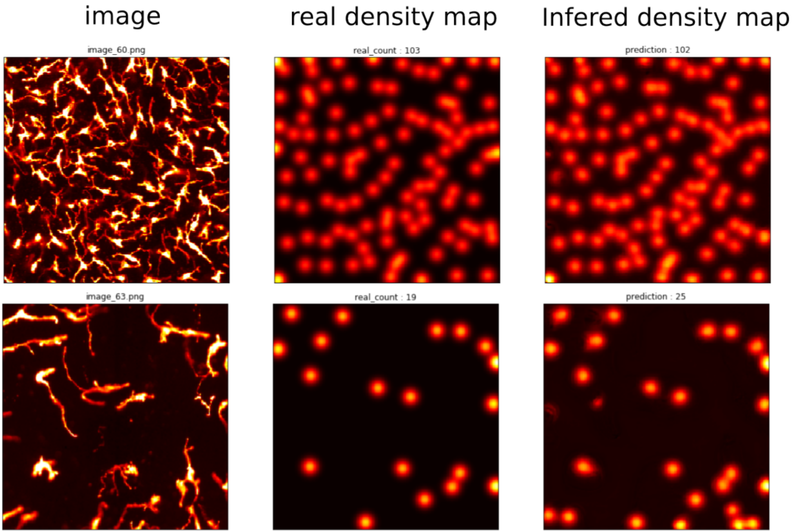


Example: counting cells



Credits: Tristan Lazard, master thesis. In collaboration with L'Oréal.

Counting cells

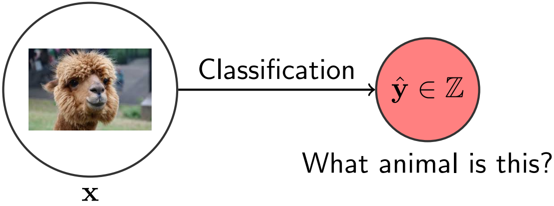


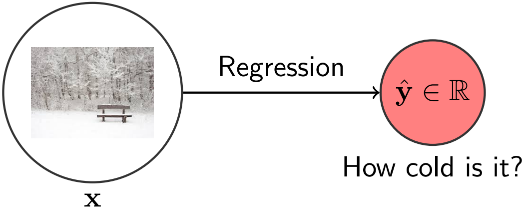
Credits: Tristan Lazard, master thesis. In collaboration with L'Oréal.

Contents

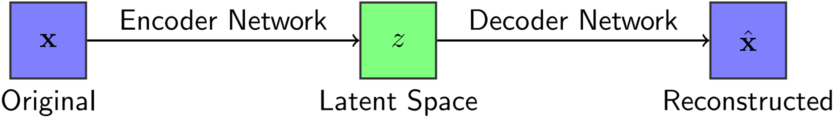
Supervised Learning

Given a labeled dataset $(\mathbf{X}, \mathbf{Y})$, we would like to learn a mapping from data space to label space.



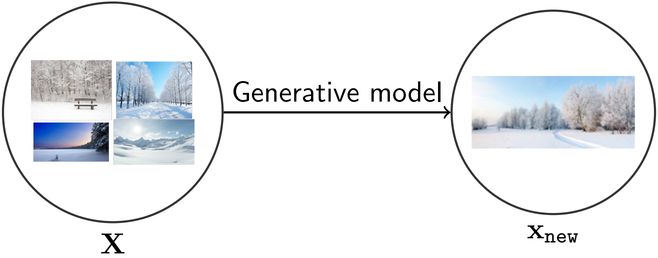


Autoencoders

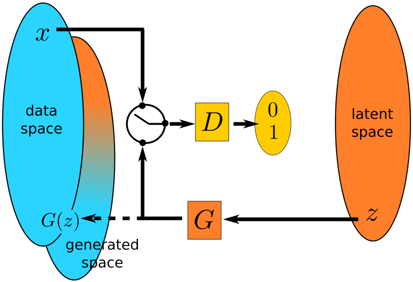


Unsupervised learning: Generative Models

Given an unlabeled dataset $(\mathbf{X})$, we would like to learn: How to generate a new observation from the same distribution (unknown) of dataset?



Generative adversarial networks [Goodfellow et al., 2014]



- The discriminator D is optimized so that it correctly classifies images as real (1) or fake (0)
- The decoder or generator G is optimized so that the produced images are classified as real by the discriminator

Which face is real?

Contents

Conclusion

- Deep learning allows to learn complex transformations between tensors, thanks to:
  - Smart methods and algorithms
  - Lots of data
  - Specialized hardware
- General artificial intelligence is still far away

References I

[Bishop, 2006] Bishop, C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg.

[Fukushima, 1979] Fukushima, K. (1979). Neural Network Model for a Mechanism of Pattern Recognition Unaffected by Shift in Position- Neocognitron. *ELECTRON. & COMMUN. JAPAN*, 62(10):11–18.

[Fukushima, 1980] Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4):193–202.

[Goodfellow et al., 2014] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative Adversarial Nets. In *Advances in Neural Information Processing Systems 27*, pages 2672–2680. Curran Associates, Inc.

[Hastie et al., 2009] Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning: data mining, inference and prediction*. Springer, 2 edition.

[Krizhevsky et al., 2012] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. In Pereira, F., Burges, C. J. C., Bottou, L., and Weinberger, K. Q., editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc.

References II

[LeCun, 1985] LeCun, Y. (1985). Une procedure d'apprentissage pour reseau a seuil asymmetrique (A learning scheme for asymmetric threshold networks). In *proceedings of Cognitiva 85*.

[Lecun et al., 1998] Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

[Ronneberger et al., 2015] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In Navab, N., Hornegger, J., Wells, W. M., and Frangi, A. F., editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, number 9351 in Lecture Notes in Computer Science, pages 234–241. Springer International Publishing.

[Rosenblatt, 1958] Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408.

[Simard et al., 1993] Simard, P., LeCun, Y., and Denker, J. S. (1993). Efficient pattern recognition using a new transformation distance. In *Advances in neural information processing systems*, pages 50–58.

References III

[Werbos, 1982] Werbos, P. J. (1982). Applications of advances in nonlinear sensitivity analysis. In Drenick, R. F. and Kozin, F., editors, *System Modeling and Optimization*, Lecture Notes in Control and Information Sciences, pages 762–770. Springer Berlin Heidelberg.