1 2 9 0

UNIVERSIDADE Ð
COIMBRA

David Alexandre Mendes Carreira

EXAMPLE OF A TITLE WRITTEN IN VERSALETE BOLD 22PT

SUBTITLE VERSALETE 20PT

Dissertação no âmbito do <nome completo do mestrado> orientada pelo/a Professor/a Doutor/a <nome completo do orientador> e apresentada <à nome da unidade orgânica>/<ao nome do departamento, se aplicável da/do nome da unidade orgânica>. AH

Month 20XX

Contents

Acknowledgments

Abstract

Resumo

Chapter 1

Introduction

Chapter 2

Background Section

2.1 History of AI

The following sections present a broad overview of the history of Artificial Intelligence (AI) without specifying or detailing too much on particular topics of this theme. The main objective is to present some context by presenting important articles in order for the reader to be able to have a notion of the progress that has been made over the past decades, the hardships encountered and how important AI is in our lives.

2.1.1 Philosophy

On October 1950, in his article *Computing Machinery and Intelligence*, Alan Turing questioned: ”Can machines think?” [39]. At the time, the question was too meaningless to answer since not only the theory but also the technology available weren't developed enough. Nonetheless, Turing still predicted that in the future there would be computers that could, effectively, display human-like intelligence and discernment under the conditions proposed on the aforementioned article.

2.1.2 Relevant events to the birth of AI

The breakthroughs of AI are predominant, and its importance in our everyday life is undeniable, but the theory behind it has several early roots. The interest in the area grew immensely with, for example, all the Turing's theoretical research, the proposal of the first mathematical Artificial Neuron model in 1943 by Warren McCulloch and Walter Pitts (based of binary inputs and output) [25] and in 1949 Donald Hebb revolutionized the way the artificial neurons were treated by proposing what is known as the Hebb's rule[1]. Taking into consideration the latter two, but specially Hebb's proposals, Belmont Farley and Westley Clark implemented in 1954 one of the first successful Artificial Neural Networks (ANN), also called Perceptron, composed of two layers of 128 artificial neurons with weighted inputs [10]. Over the span of approximately ten years, multiple researches were performed attempting to computerize the human brain. However, only in 1956, during the *Dartmouth Summer Research Project on Artificial Intelligence* [24], was the term "Artificial Intelligence" firstly proposed by John McCarthy *et al.*, beginning what is now considered to be the birth of AI [47].

2.1.3 The fading of general interest

The succeeding two decades following the Dartmouth conference were filled with important developments, with special emphasis in the works published in 1958 by Frank Rosenblatt (generalized the Farley and Clark training to multilayer networks rather than only two) [31], the 1959 General Problem Solver implemented by Allen Newel *et al.* (a program intended to work as a universal problem

[1] "When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased." [13], meaning that when two neurons fire together their relation is strengthened.

solver that was capable of solving exercises such as the Towers of Hanoi[2]) [27] and the ELIZA a natural language processing tool program developed by Joseph Weizenbaum between 1964 and 1966 [41]. Unfortunately, part of the interest and development around AI met an unforeseen fade after criticisms about the exaggerated public funding [12] and the Marvin Minksy and Seymour Papert 1969 book *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain* [26] that reported on the problems of the Perceptron network. The overall sentiments regarding this topic of research was of doubt and fear of no progress, mainly due to the spending and two issues raised by Minsky and Papert: the ANN couldn't solve linear inseparable problems[3] and there were limitations due to a lack of sufficient computing power to handle the processing of multi-layer large networks.

2.1.4 A better approach

Minksy and Papert raised important questions, but it shouldn't have discouraged other researchers from further trying, since they failed to acknowledge alternative approaches that had already solved those exact problems. As previously stated, the model proposed by McCulloch and Pitts, later improved by the Farley-Clark implementation and, finally, Rosenblatt, couldn't handle linearly inseparable classes. A possible solution for cases like this started being studied in the 1960s [17, 32] and, although it didn't produce relevant results, in 1965 Alexey Ivakhnenko and Valentin Lapa [15] were, indeed, successful in implementing what is nowadays considered to be the first deep learning network of its kind [34]. In 1971 Ivakhnenko also published an article describing a deep learning network with 8 layers that was already able to create hierarchical internal representations [16].

[2] *The Towers of Hanoi* is a game with 3 stacks of increasingly smaller disks. The goal is to stack them one at a time, so that they are arranged in a decreasing radius manner.

[3] That is, if two sets X and Y in $\mathbb{R}^d$ can't be divided by a hyperplane such that the elements of X and Y stay on opposing sides, then we're dealing with linear inseparable classes [9]

The years progressed, in 1979 Kunihiko Fukushima introduced the first Convolutional Neural Network (CNN) in a structural sense, due to its similarity to the architecture of modern ones of this category. Ten years later, Yann LeCun *et al.* applied for the first time a revolutionizing training algorithm called Backpropagation to a CNN [19], creating what is now a pillar for most of the modern competition winning networks in computer vision [34] and employing the term ”convolution” for the first known time [22]. He also introduced the MNIST (**M**odified **N**ational **I**nstitute of **S**tandards and **T**echnology) dataset, a collection of handwritten digits [21], that to this day is still one of the most famous benchmarks in Machine Learning. Backpropagation can be traced back many decades, but the modern version was first described by Seppo Linnainmaa (1970) [23], implemented for the first time by Stuart Dreyfus (1973) [8] and, finally in 1986, David Rummelhart *et al.* popularized it in the Neural Network's (NN) domain by demonstrating the growing usefulness of internal representations [33].

2.1.5 The importance of Convolutional Neural Networks

The study on Neural Networks continued and there were improvements on all types of architectures [14, 42] with special highlight to pioneering Neural Networks processed by GPUs[4] (standard NN in 2004 by [28] and CNN in 2006 by [3]). But there's a well deserved particular attention related to the developments of CNNs due to their great performance in image related tasks when compared to others networks, as proven by LeCun in his 1998 paper [21]. Some relevant examples: in 2003 the MNIST record was broken by Patrice Simard *et al.* [35], achieving an error rate of 0.4% (whereas a non-convolutional neural network by the same authors took the second place with 0.7%); three years later, the same benchmark had a new set low of 0.39% by Marc'Aurelio Ranzato *et al.* [30]; in 2009 a CNN by Yang *et al.* was the first network of this type to win an official international competition (TRECVid) [45]; a GPU implementation of a CNN [5] achieved superhuman vision

[4] Graphics Processing Unit

performance in a competition (IJCNN 2011) in a *German Traffic Sign Recognition Benchmark* with a 0.56% error rate (0.78% for the best human performance, 1.69% for the second-best neural network contestant and 3.86% for the best non-neural method [36]). This last example conjoined with non-convolutional methods [29, 7] and the previously cited [3, 28], reinforces how fundamental GPUs were to further develop neural networks. To supplement even more the importance of CNNs and GPUs, only a year later, Alex Krizhevsky *et al.* proposed a Deep CNN trained by GPUs that was the first one to win the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), achieving an error rate of 15.3% while the second place obtained 26.2% [18].

The year of 2012 was very important for Deep Learning, CNNs and Computer Vision, due to all the attention brought to many researches on this topic after several systems of this kind won image analysis competitions ([4, 6] and the very important previously mentioned [18]), beginning what's considered to be the start of the new wave, we're currently in, of interest in Artificial Intelligence, specially in the aforesaid topics [22].

2.2 Fundamentals

There are several types of Neural Networks architectures, but Convolutional Neural Networks (CNNs or Convnets) are probably the most widely implemented model overall [44, 22] with successful applications in the domains of Computer Vision [18, 37, 38, 48] or Natural Language Processing[1, 40, 43]. In the CNN category itself there are different variants, but they all abide the fundamental structure of a feedforward hierarchical multi-layer network. Feedforward because the information only flows in a singular direction without cycling [46], hierarchical because the higher complexity internal representations are learned from lower ones [20, 49] and multi-layer because it is composed of a series of stages, blocks or layers: the raw data is fed to an input layer, forwarded to a sequence of intercalating convolutional and pooling layers, proceeded to a stage of one or more fully-connected layers [20, 44, 11, 2].

Using CNNs for Computer Vision tasks is not an arbitrary choice, but due to the fact that the network design can benefit from the intrinsic characteristics of the input data [20] they perform really well in image related applications. In the first place, images have an array-like structure with numerous elements, namely, each pixel has an assigned value organized in a grid like manner [44], matching the type of input for these networks [20]. In the second place, there's an inherent correlation between local groups of values, which creates distinguishable motifs. Finally, the local values of images are invariant to location, that is, a certain composition should have the same value independently of the spatial location in the picture.

This model architecture is based on the visual cortex ventral pathway, therefore, it is capable of automatically extracting spatial feature hierarchies, from a lower to higher complexity (in contrast to conventional machine learning) [20, 44, 11, 2].

Key features such as local connections/receptive fields, shared weights, subsampling and the use of many layers allows this network to be invariant to shift, scale and distortions.

Chapter 3

Related work

Chapter 4

Methodology

Chapter 5

Results

Chapter 6

Conclusion

Bibliography

[1] Ossama Abdel-Hamid et al. “Convolutional Neural Networks for Speech Recognition”. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22.10 (Oct. 2014), pp. 1533–1545. ISSN: 2329-9304. DOI: 10.1109/TASLP.2014.2339736 (cit. on p. 7).

[2] Laith Alzubaidi et al. “Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions”. In: *Journal of Big Data* 8.1 (Mar. 2021), p. 53. ISSN: 2196-1115. DOI: 10.1186/s40537-021-00444-8 (cit. on p. 7).

[3] Kumar Chellapilla, Sidd Puri, and Patrice Simard. “High Performance Convolutional Neural Networks for Document Processing”. In: () (cit. on pp. 5, 6).

[4] D.C. Ciresan et al. “Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images”. In: *NIPS* 25 (2012). Export Date: 26 January 2023; Cited By: 92, pp. 2852–2860 (cit. on p. 6).

[5] Dan Cireşan et al. “A Committee of Neural Networks for Traffic Sign Classification”. In: *The 2011 International Joint Conference on Neural Networks*. July 2011, pp. 1918–1921. DOI: 10.1109/IJCNN.2011.6033458 (cit. on p. 5).

[6] Dan C. Cireşan et al. “Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks”. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*. Ed. by Kensaku Mori et al. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2013, pp. 411–418.

ISBN: 978-3-642-40763-5. DOI: 10.1007/978-3-642-40763-5_51 (cit. on p. 6).

[7] Dan Claudiu Cireşan et al. "Deep, Big, Simple Neural Nets for Handwritten Digit Recognition". In: *Neural Computation* 22.12 (Dec. 2010), pp. 3207–3220. ISSN: 0899-7667. DOI: 10.1162/NECO_a_00052 (cit. on p. 6).

[8] Stuart E. Dreyfus. "The Computational Solution of Optimal Control Problems with Time Lag". In: *IEEE Transactions on Automatic Control* 18.4 (1973). Cited by: 32, pp. 383–385. DOI: 10.1109/TAC.1973.1100330 (cit. on p. 5).

[9] D. Elizondo. "The Linear Separability Problem: Some Testing Methods". In: *IEEE Transactions on Neural Networks* 17.2 (Mar. 2006), pp. 330–344. ISSN: 1045-9227. DOI: 10.1109/TNN.2005.860871 (cit. on p. 4).

[10] B. Farley and W. Clark. "Simulation of Self-Organizing Systems by Digital Computer". In: *Transactions of the IRE Professional Group on Information Theory* 4.4 (1954), pp. 76–84. DOI: 10.1109/TIT.1954.1057468 (cit. on p. 3).

[11] Jiuxiang Gu et al. "Recent Advances in Convolutional Neural Networks". In: *Pattern Recognition* 77 (May 2018), pp. 354–377. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2017.10.013 (cit. on p. 7).

[12] Michael Haenlein and Andreas Kaplan. "A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence". In: *California Management Review* 61 (July 2019), p. 000812561986492. DOI: 10.1177/0008125619864925 (cit. on p. 4).

[13] Donald Olding Hebb. *The Organization of Behavior: A Neuropsychological Theory*. Wiley, 1949. ISBN: 978-0-471-36727-7 (cit. on p. 3).

[14] Sepp Hochreiter and Jürgen Schmidhuber. "Long Short-Term Memory". In: *Neural Computation* 9.8 (Nov. 1997), pp. 1735–1780. ISSN: 0899-7667. DOI: 10.1162/neco.1997.9.8.1735 (cit. on p. 5).

[15] A G Ivakhnenko and V G Lapa. "Cybernetic Predicting Devices". In: () (cit. on p. 4).

[16] A. G. Ivakhnenko. "Polynomial Theory of Complex Systems". In: *IEEE Transactions on Systems, Man, and Cybernetics* SMC-1.4 (1971), pp. 364–378. DOI: 10.1109/TSMC.1971.4308320 (cit. on p. 4).

[17] Roger David Joseph. *Contributions to Perceptron Theory.* Cornell Aeronautical Laboratory, 1960 (cit. on p. 4).

[18] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems.* Vol. 25. Curran Associates, Inc., 2012 (cit. on pp. 6, 7).

[19] Y. LeCun et al. "Backpropagation Applied to Handwritten Zip Code Recognition". In: *Neural Computation* 1.4 (1989), pp. 541–551. DOI: 10.1162/neco.1989.1.4.541 (cit. on p. 5).

[20] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep Learning". In: *Nature* 521.7553 (May 2015), pp. 436–444. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature14539 (cit. on p. 7).

[21] Yann LeCun et al. "Gradient-Based Learning Applied to Document Recognition". In: (1998) (cit. on p. 5).

[22] Zewen Li et al. "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects". In: *IEEE Transactions on Neural Networks and Learning Systems* 33.12 (Dec. 2022), pp. 6999–7019. ISSN: 2162-2388. DOI: 10.1109/TNNLS.2021.3084827 (cit. on pp. 5–7).

[23] Seppo Linnainmaa. “The Representation of the Cumulative Rounding Error of an Algorithm as a Taylor Expansion of the Local Rounding Errors”. PhD thesis. Master’s Thesis (in Finnish), Univ. Helsinki, 1970 (cit. on p. 5).

[24] J McCarthy et al. “A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE”. In: () (cit. on p. 3).

[25] Warren S Mcculloch and Walter Pitts. “A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY”. In: () (cit. on p. 3).

[26] Marvin Minsky and Seymour Papert. *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA, USA: MIT Press, 1969 (cit. on p. 4).

[27] Allen Newell, John C Shaw, and Herbert A Simon. “Report on a General Problem Solving Program”. In: *IFIP Congress*. Vol. 256. Pittsburgh, PA. 1959, p. 64 (cit. on p. 4).

[28] Kyoung-Su Oh and Keechul Jung. “GPU Implementation of Neural Networks”. In: *Pattern Recognition* 37.6 (June 2004), pp. 1311–1314. ISSN: 0031-3203. DOI: 10.1016/j.patcog.2004.01.013 (cit. on pp. 5, 6).

[29] Rajat Raina, Anand Madhavan, and Andrew Y. Ng. “Large-Scale Deep Unsupervised Learning Using Graphics Processors”. In: *Proceedings of the 26th Annual International Conference on Machine Learning*. ICML ’09. New York, NY, USA: Association for Computing Machinery, June 2009, pp. 873–880. ISBN: 978-1-60558-516-1. DOI: 10.1145/1553374.1553486 (cit. on p. 6).

[30] Marc’ aurelio Ranzato et al. “Efficient Learning of Sparse Representations with an Energy-Based Model”. In: *Advances in Neural Information Processing Systems*. Vol. 19. MIT Press, 2006 (cit. on p. 5).

[31] F. Rosenblatt. “The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain.” In: *Psychological Review* 65 (1958),

pp. 386–408. ISSN: 1939-1471(Electronic),0033-295X(Print). DOI: 10.1037/h0042519 (cit. on p. 3).

[32] Frank Rosenblatt. *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms.* Spartan Books, 1962 (cit. on p. 4).

[33] DE Rumelhart, GE Hinton, and RJ Williams. *Learning Internal Representations by Error Propagation", in\Parallel Distributed Processing", DE Rumelhart, JL McClelland Eds.* 1986 (cit. on p. 5).

[34] Jürgen Schmidhuber. "Deep Learning in Neural Networks: An Overview". In: *Neural Networks* 61 (Jan. 2015), pp. 85–117. ISSN: 08936080. DOI: 10.1016/j.neunet.2014.09.003 (cit. on pp. 4, 5).

[35] P.Y. Simard, D. Steinkraus, and J.C. Platt. "Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis". In: *Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings.* Vol. 1. Edinburgh, UK: IEEE Comput. Soc, 2003, pp. 958–963. ISBN: 978-0-7695-1960-9. DOI: 10.1109/ICDAR.2003.1227801 (cit. on p. 5).

[36] J. Stallkamp et al. "Man vs. Computer: Benchmarking Machine Learning Algorithms for Traffic Sign Recognition". In: *Neural Networks.* Selected Papers from IJCNN 2011 32 (Aug. 2012), pp. 323–332. ISSN: 0893-6080. DOI: 10.1016/j.neunet.2012.02.016 (cit. on p. 6).

[37] Yaniv Taigman et al. "DeepFace: Closing the Gap to Human-Level Performance in Face Verification". In: *2014 IEEE Conference on Computer Vision and Pattern Recognition.* Columbus, OH, USA: IEEE, June 2014, pp. 1701–1708. ISBN: 978-1-4799-5118-5. DOI: 10.1109/CVPR.2014.220 (cit. on p. 7).

[38] Jonathan Tompson et al. *Efficient Object Localization Using Convolutional Networks.* Comment: 8 pages with 1 page of citations. June 2015. arXiv: 1411.4280 [cs] (cit. on p. 7).

[39] A. M. Turing. "I.—COMPUTING MACHINERY AND INTELLIGENCE". In: *Mind* LIX.236 (Oct. 1950), pp. 433–460. ISSN: 1460-2113, 0026-4423. DOI: 10.1093/mind/LIX.236.433 (cit. on p. 2).

[40] Mingxuan Wang et al. "genCNN: A Convolutional Architecture for Word Sequence Prediction". In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Beijing, China: Association for Computational Linguistics, July 2015, pp. 1567–1576. DOI: 10.3115/v1/P15-1151 (cit. on p. 7).

[41] Joseph Weizenbaum. "ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine". In: *Communications of the ACM* 9.1 (Jan. 1966), pp. 36–45. ISSN: 0001-0782. DOI: 10.1145/365153.365168 (cit. on p. 4).

[42] J. Weng, N. Ahuja, and T.S. Huang. "Cresceptron: A Self-Organizing Neural Network Which Grows Adaptively". In: *[Proceedings 1992] IJCNN International Joint Conference on Neural Networks*. Vol. 1. June 1992, 576–581 vol.1. DOI: 10.1109/IJCNN.1992.287150 (cit. on p. 5).

[43] Lingyun Xiang et al. "A Convolutional Neural Network-Based Linguistic Steganalysis for Synonym Substitution Steganography". In: *Mathematical Biosciences and Engineering* 17.2 (2020), pp. 1041–1058. ISSN: 1551-0018. DOI: 10.3934/mbe.2020055 (cit. on p. 7).

[44] Rikiya Yamashita et al. "Convolutional Neural Networks: An Overview and Application in Radiology". In: *Insights into Imaging* 9.4 (Aug. 2018), pp. 611–629. ISSN: 1869-4101. DOI: 10.1007/s13244-018-0639-9 (cit. on p. 7).

[45] Ming Yang et al. "Detecting Human Actions in Surveillance Videos". In: 2009 TREC Video Retrieval Evaluation Notebook Papers. Cited by: 26. 2009 (cit. on p. 5).

[46] Andreas Zell. "Simulation Neuronaler Netze". In: 1994 (cit. on p. 7).

[47] Caiming Zhang and Yang Lu. "Study on Artificial Intelligence: The State of the Art and Future Prospects". In: *Journal of Industrial Information Integration* 23 (Sept. 2021), p. 100224. ISSN: 2452414X. DOI: `10.1016/j.jii.2021.100224` (cit. on p. 3).

[48] Yu-Dong Zhang et al. "Improved Breast Cancer Classification Through Combining Graph Convolutional Network and Convolutional Neural Network". In: *Information Processing & Management* 58.2 (Mar. 2021), p. 102439. ISSN: 0306-4573. DOI: `10.1016/j.ipm.2020.102439` (cit. on p. 7).

[49] Xinqi Zhu and Michael Bain. *B-CNN: Branch Convolutional Neural Network for Hierarchical Classification*. Comment: 9 pages, 8 figures. Oct. 2017. DOI: `10.48550/arXiv.1709.09890`. arXiv: `1709.09890 [cs]` (cit. on p. 7).