https://doi.org/10.52326/ic-ecco.2022/EL.07

Prospects Overview of the Superconducting Neural Networks

Maria Lupu [1], ORCID: 0000-0001-6768-9357
[1] D.GHITU Institute of Electronic Engineering and Nanotechnologies, Technical University of Moldova¸ Chisinau, Moldova, e-mail: simplex.prof.nanotech@gmail.com

Abstract—**The long-term efforts of many research groups have led to the fact that by now a large number of different "learning rules" and architectures of neural networks, their hardware implementations and methods of using neural networks to solve applied problems have been accumulated. These intellectual inventions exist in the form of a "technopark" of neural networks. Each network from the technopark has its own architecture, training rules and solves a certain set of tasks. Moreover, specialized high-speed devices can be created on its basis. There are several levels of alienation of a neural network from a universal computer: from network learning on a universal device and the use of rich possibilities in manipulating a task book, learning algorithms and modifying architecture, to complete alienation without learning and modification capabilities, only the functioning of a trained network.**

Keywords— superconducting neural networks; dynamic processes; physics-based models; deep neural networks

I. INTRODUCTION

For progress in the field of high–performance computing and artificial intelligence, it is necessary to improve the energy efficiency and density of integration of existing circuits, which can be realized only with the use of a new element base - superconducting neurons and synapses. The proposed study is relevant due to the possibility of developing new energy-efficient computers with non-von Neumann architecture based on elements of superconducting spintronics.

Indeed, the best modern systems on specialized semiconductor microprocessors simulate the work of about 1 million neurons and a quarter of a billion synapses. However, the largest and most ambitious projects state the goals of 10^{10} neurons and 10^{14} synapses. The key problem on the way to such goals is the reduction of energy release in all active elements of a neuromorphic computing system.

For this reason, the use of superconducting materials seems to be the most promising direction that meets these tasks. Traditionally, in superconducting logic and memory, information is associated with a quantum of magnetic flux, which, firstly, limits the degree of integration (a cell must contain one quantum of flux), and secondly, determines the localization of information, which complicates the physical implementation of information processing parallelization algorithms. These limitations lead to a low functional density of existing superconducting circuits and make it difficult to develop circuits based on non-classical principles of information processing, such as deep neural networks, which are key components in the creation of artificial intelligence.

II. PHYSICAL REALIZATION OF NEURONS AND CONNECTIONS IN THE ANNS

Neural network (also artificial neural network, INS) — a mathematical model, as well as its software or hardware implementation, built on the principle of organization and functioning of biological neural networks — networks of nerve cells of a living organism. This concept arose when studying the processes occurring in the brain, and when trying to simulate these processes. The first such attempt was the neural networks of U. McCulloch and W. Pitts.

After the development of learning algorithms, the resulting models began to be used for practical purposes: in forecasting tasks, for pattern recognition, in control tasks, etc.

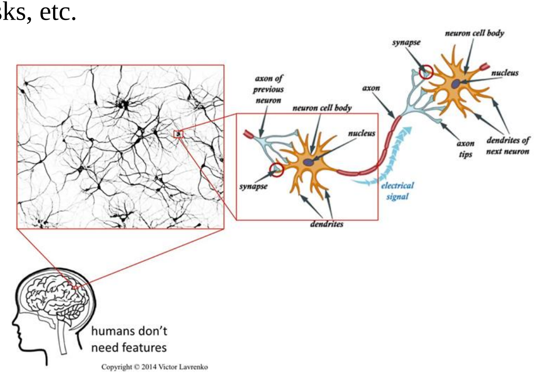

Figure 1. Biological and artificial neural networks [2].

From the point of view of machine learning, a neural network is a special case of pattern recognition methods, clustering methods, etc. From the point of view of mathematics, training neural networks is a multiparametric problem of nonlinear optimization. From the point of view of cybernetics, a neural network is used in adaptive control tasks and as algorithms for robotics. From the point of view of the development of computing and programming, a neural network is a way to solve the problem of effective parallelism. From the point of view of artificial intelligence, ANN is the main direction in the structural approach to study the possibility of constructing (modeling) natural intelligence using computer algorithms [9].

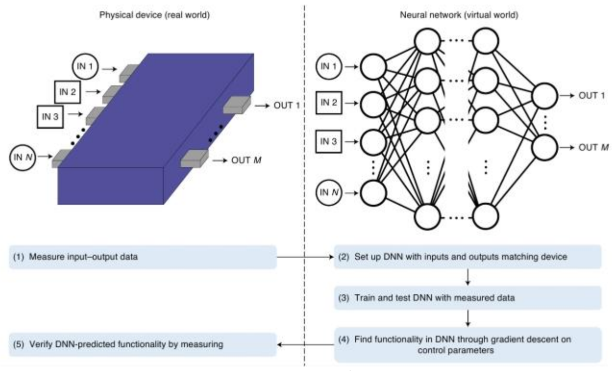

Figure 2. Examples of ANNs [4].

Neural networks are not programmed in the usual sense of the word, they are trained. The possibility of learning is one of the main advantages of neural networks over traditional algorithms.

Technically, learning consists in finding the coefficients of connections between neurons. During the learning process, the neural network is able to identify complex dependencies between input and output data, as well as perform generalization. This means that in case of successful training, the network will be able to return the correct result based on data that was missing in the training sample, as well as incomplete and/or "noisy", partially distorted data.

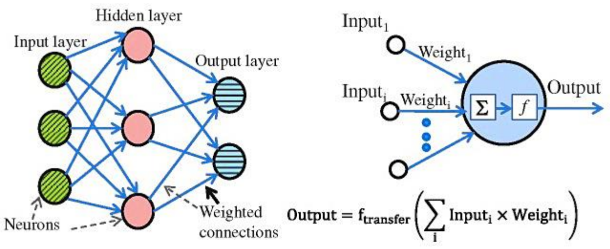

Figure 3. Technical realization of ANNs [6].

Neurons interact through a series of pulses lasting several milliseconds, each pulse is a frequency signal with a frequency of several units to hundreds of hertz. This is unimaginably slow compared to modern computers, but at the same time, the human brain can process analog information much faster than a machine, such as: recognize images, taste, recognize sounds, read someone else's handwriting, operate with qualitative parameters. All this is realized through a network of neurons connected by synapses. In other words, the brain is a system of parallel processors that works much more efficiently than the now popular sequential computing.

A neural network is a collection of neurons connected to each other in a certain way. Each neuron is an element that calculates the output signal (according to a certain rule) from a set of input signals. That is, the main sequence of actions of one neuron is as follows:

· Receiving signals from previous network elements;
· Combination of input signals
· Output signal calculation
· Transmission of the output signal by the next element of the neural network

Neurons can be connected to each other in absolutely different ways, this is determined by the structure of a particular network. An output signal (or several output signals) is generated based on the totality of the signals coming to the network input.

A neuron is characterized by its state and, by analogy with a real neuron, can be either excited or inhibited.
The neuron has a group of synapses – unidirectional input connections connected to the outputs of other neurons, and also has an axon – the output connection of this neuron, from which the signal (excitation or inhibition) enters the synapses of the following neurons. Each synapse is characterized by the magnitude of the synaptic connection or its weights.

In physical terms, the weight of a synaptic connection is the electrical conductivity of a given synapse.

The current state of a neuron is defined as the weighted sum of its inputs. The value at the input of the synapse is multiplied by the weight of this synapse, then all these values are summed up and we get the current state of the neuron (Fig. 1, 2)

In artificial neural networks, the activation function of a neuron determines the output signal, which is determined by an input signal or a set of input signals.

A standard computer chip can be considered as a digital network of activation functions, which can take the values "ON" (1) or "OFF" (0) depending on the input. This is similar to the behavior of a linear perceptron in neural networks. However, only nonlinear activation functions allow such networks to solve non-trivial problems using a small number of nodes. In artificial neural networks, this function is also called a transfer function.

Each function has its own distinctive properties, advantages and disadvantages. None of the functions is universal, it is impossible to say unequivocally in which case a linear rectifier, a sigmoid or a hyperbolic tangent should be used.

Knowing some characteristics of the function to be approximated, you should choose an activation function that approximates the desired function as accurately as possible and will lead to rapid learning [8].

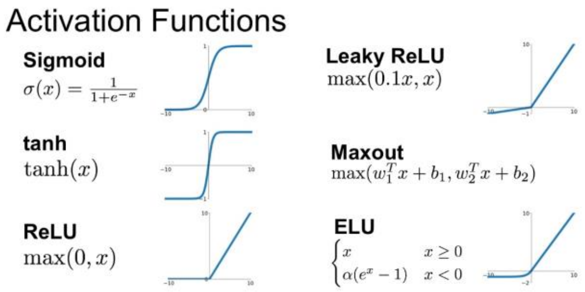

Figure 4,5. Examples of the activation functions [7].

Energy efficient memory has been the main detractor for multiple superconducting digital projects in the past. Recently, fundamental physics research in superconductor-ferromagnet thin-film tunnel structures created a new opportunity to solve this long-standing problem. Superconductivity and ferromagnetism, two deeply antagonistic electronic properties, can co-exist in form of Magnetic Josephson Junctions (MJJs). The superconducting-ferromagnetic MJJs are electrically compatible with traditional superconductor-insulator-superconductor (SIS) Josephson junctions (JJs) used for digital energy-efficient single flux quantum circuits. Both MJJ and JJ circuits have similar fabrication process and can be integrated on a single chip. As a result, a combination of MJJs and JJs can be used to form addressable memory cells, energy-efficient memory periphery circuits and programmable logic elements [9].

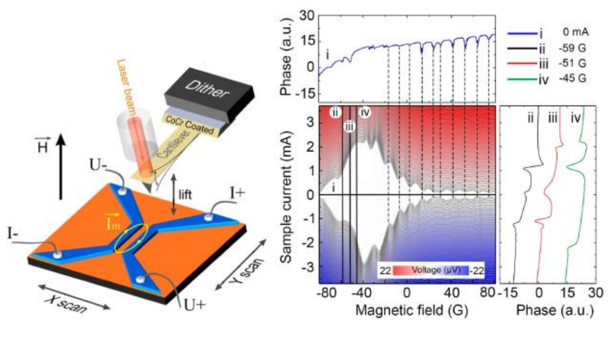

Figure 6. Reducing the size of the base elements [2].

III. APPLICATION OF ANNS

Classification of images. The task is to indicate whether the input image (for example, a speech signal or a handwritten symbol) represented by a feature vector belongs to one or more predefined classes. Well-known applications include letter recognition, speech recognition, electrocardiogram signal classification, and blood cell classification.

Clustering/categorization. When solving the clustering problem, there is no training sample with class labels. The clustering algorithm is based on the similarity of images and places close images in one cluster. There are known cases of clustering for knowledge extraction, data compression and data properties research.

Approximation of functions. Suppose there is a training sample $((x_1,y_1), (x_2,y_2)..., (x_n,y_n))$ (input-output data pairs), which is generated by an unknown function (x) distorted by noise. The task of approximation is to find an estimate of an unknown function (x). Approximation of functions is necessary when solving numerous engineering and scientific modeling problems.

Prediction/forecast. Let n discrete samples $\{f(t_1), f(t_2)..., f(t_n)\}$ be given at consecutive time points $t_1, t_2,...,$ tn . The task is to predict the value of $f(t_n+1)$ at some future time t_n+1. Prediction/forecast has a significant impact on decision-making in business, science and technology. Stock exchange price prediction and weather forecast are typical applications of prediction/forecasting techniques.

Optimization. Numerous problems in mathematics, statistics, engineering, science, medicine and economics can be considered as optimization problems. The task of the optimization algorithm is to find a solution that satisfies the system of constraints and maximizes or minimizes the objective function. The traveling salesman problem is a classic example of an optimization problem.

Memory addressable by content. In the von Neumann model of computation, memory access is available only through an address that does not depend on the memory content. Moreover, if an error is made in calculating the address, completely different information can be found. Associative memory, or memory addressable by content, is available at the direction of the specified content. The contents of memory can be called even by partial input or distorted content. Associative memory is extremely desirable when creating multimedia information databases.

Management. Consider a dynamic system given by the set $\{f_1(t), f_2(t)\}$, where $f_1(t)$ is the input control action, and $f_2(t)$ is the output of the system at time t. In control systems with a reference model, the purpose of control is to calculate such an input impact $f_1(t)$, in which the system follows the desired trajectory dictated by the reference model. An example is optimal engine management [1,3,5].

IV. Practical Implementations: Reducing the Size of the Basic Elements, Scientific Novelty

The developed cells of adiabatic Josephson transmission lines as part of neural network signal processing units allow for four orders of magnitude (up to attojoule scales) to reduce the release of energy during the functioning of neurons.

The technique of analyzing macroscopic quantum effects in multi-contact and multi-circuit superconducting quantum interferometers was used for the first time to study the possibilities of integrating artificial neural networks into digitized signal processing systems with built-in magnetic Josephson memory (Fig. 5)

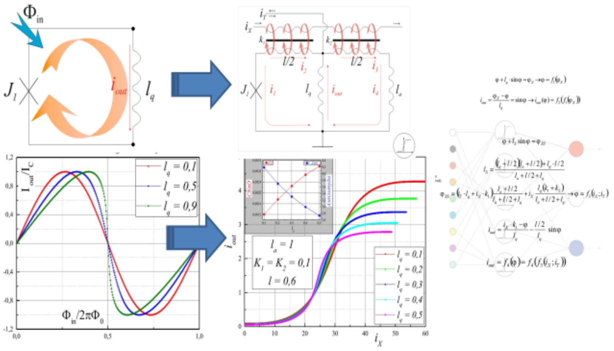

Figure7. Expression for the activation function [2].

The developed compact neurons allow for one operation (on subnanosecond time scales) "calculate" the activation function of a neuron. The developed compact synapses allow for four orders of magnitude (up to attojoule scales) reduce the energy release during the passage of a single pulse.

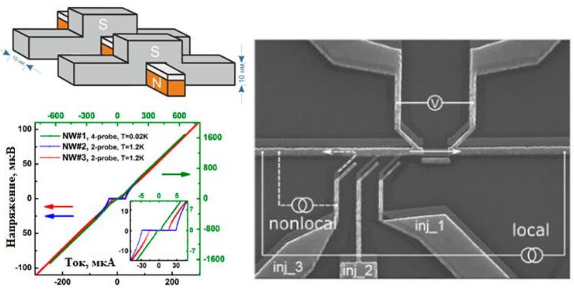

Figure 8. Development of technological solutions [2].

Methods have been developed and tested for the analysis of charge transfer processes in compact Josephson cells and phase batteries (taking into account the peculiarities of the influence of topology during the transition to nanoscale structures), which are part of both the SHP ADC and the signal processor, neural network and quantum signal processing unit (Fig. 6) [1,3,5].

Acknowledgment

This review article was supported by D.GHITU Institute of Electronic Engineering and Nanotechnologies, Technical University of Moldova¸ M.V. Lomonosov Moscow State University, Skobeltsyn Institute of Nuclear Physics, D.GHITU Institute of Electronic Engineering and Nanotechnologies, Lobachevsky State University of Nizhni Novgorod, Laboratory of Functional Nanostructures, Orel State University named after I.S. Turgenev, Moscow Institute of Physics and Technology, Moscow Technical University of Communication and Informatics (MTUCI), Dukhov All-Russia Research Institute of Automatics, Institute of Solid State Physics RAS, Solid State Physics Department, Lukin Scientific Research Institute of Physical Problems, Zelenograd, Moscow.

The study was financially supported by the Russian Science Foundation Grant (RSF) No. 20-79-10018. and partially by the Moldova State Program Project «Nanostructuri şi nanomateriale funcţionale pentru industrie şi agricultură» no. 20.80009.5007.11.

References

[1] S. Bakurskiy, M. Kupriyanov, N. Klenov, I. Soloviev, A. Schegolev, R. Morari, Yu. Khaydukov, A. Sidorenko, „Controlling the proximity effect in a Co/Nb multilayer: the properties of electronic transport”, *Beilstein J. Nanotechnol.* 2020, 11, pp.1336–1345. https://doi.org/10.3762/bjnano.11.118

[2] S. Bakurskiy, M. Kupriyanov, N. Klenov, I. Soloviev, A. Schegolev, M. Kupriyanov, presentation “The element base of neuromorphic signal processors based on macroscopic quantum effects in superconductors”.

[3] T. E. Golikova, F. Hubler, D. Beckmann, N. V. Klenov, S. V. Bakurskiy, M. Yu. Kupriyanov, I. E. Batov, V. V. Ryazanov, “Critical current in planar SNS Josephson junctions” // JETP Letters. — 2013. — Vol. 96, no. 10. — P. 668–673.

[4] H. R. Euler, M. N. Boon, J. T. Wildeboer, B. Ven, T. Chen, Hajo Broersma, P. A. Bobbert and W. G. van der Wiel, “A deep-learning approach to realizing functionality in nanoelectronic devices”, *J. Nature Nanotechnology*. https://doi.org/10.1038/s41565-020-00779-y

[5] N. Klenov, Y. Khaydukov, S. Bakurskiy, R. Morari, I. Soloviev, V. Boian, T. Keller, M. Kupriyanov, A. Sidorenko, B. Keimer, “Periodic Co/Nb pseudo spin valve for cryogenic memory”, *Beilstein J. Nanotechnology*, 10 (2019), pp. 833–839. https://doi.org/10.3762/bjnano.10.83

[6] https://ebrary.net/98123/economics/artificial_neural_networks

[7] https://www.datasciencepreparation.com/blog/articles/what-is-an-activation-function-what-are-commonly-used-activation-functions/

[8] https://www.sciencedirect.com/topics/earth-and-planetary-sciences/artificial-neural-network

[9] https://www.sciencedirect.com/science/article/pii/S1875389212020639