COMPUTATIONAL MECHANICS 2023
38th conference with international participation
Srní
October 23 - 25, 2023

About the appropriate neural network size for the engineering applications

P. Kovář[a,b], A. Tater[a,b], J. Pařez[b], J. Fürst[a]

[a] *Department of Technical Mathematics, Faculty of Mechanical Engineering, Czech Technical University in Prague, Karlovo náměstí 13, 121 35 Praha 2, Czech Republic*

[b] *Center of Aviation and Space Research, Faculty of Mechanical Engineering, Czech Technical University in Prague, Technická 4, 160 00 Praha 6, Czech Republic*

1. Introduction

Deep learning approaches became very popular in recent years. In terms of computational effectivity and time required for the learning process, number of degrees of freedom in the proposed neural network plays significant role. Thus, an apriori information about the appropriate neural network size for a given problem could be very promising tool in machine learning tasks.

In the contrast to the standard machine learning approaches aimed to deep learning, present contribution deals with shallow higher order neural networks. A comparison of the ability to capture more demanding engineering task using different neural networks is presented and basic idea of the neural network size estimation for the task is discussed.

2. Objective statement

Natural convection phenomenon, governed by system of Navier-Stokes equations, in the annular section representing a part of the aircraft engine where turbines are housed, was selected as the engineering complex training task. The goal is to find a neural operator $\mathcal{N}(\bullet)$ that is able to approximate temperature distribution on the outer tube T_{D_2} with permissible error and the simplest architecture as possible. In order to obtain training data set for neural network and replace experimental measurement, various numerical simulations with different geometrical setups, specifically D_1/D_2 ratios, thickness of the outer tube t_2 and temperature T_{D_1} as boundary condition on the inner tube were performed as described in detail in [5].
There is a sketch of the computational domain in Fig. 1. Based on the theoretical and experimental knowledge [4], it is necessary to assume that temperature distribution on the outer tube is also function of the angle denoted by φ.

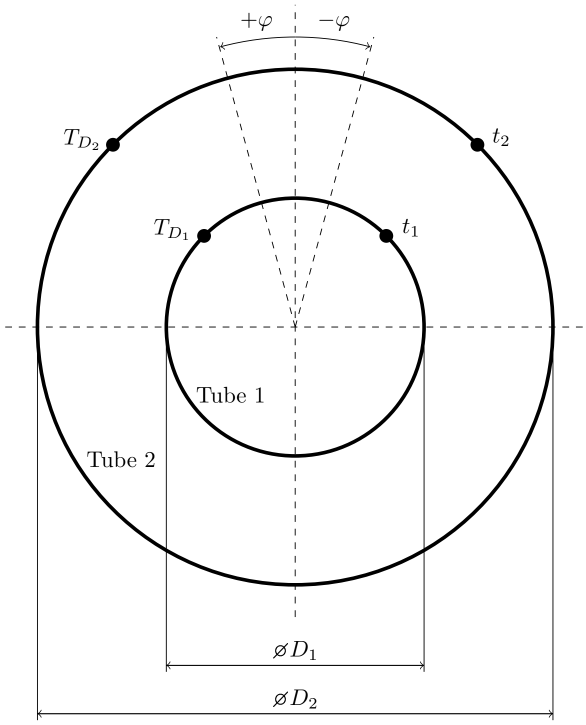

Fig. 1. Sketch of the computational domain

3. Neural networks

Referring to previous considerations, the desired function is assumed in the form

$$T_{D_2} = \mathcal{N}(\bullet) = f(D_1/D_2, t_2, T_{D_1}, \varphi). \tag{1}$$

Neural output of the neural unit is consisted of two different operations as it is shown in Fig. 2 (left). Product of the somatic operation $\widetilde{y}$, in general, can be expressed as [2]

$$\widetilde{y} = \sigma(s). \tag{2}$$

Let us assume N-th order neural unit, then product of synaptic operation can be written as [2]

$$s = w_0 x_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n}\sum_{j=i}^{n} w_{ij} x_i x_j + \cdots + \sum_{i_1=1}^{n} \cdots \sum_{i_N=i_{N-1}}^{n} w_{i_1 i_2 \ldots i_n} x_{i_1} x_{i_2} \ldots x_{i_n}, \tag{3}$$

where $x_0 = 1$ denotes threshold and n stands for length of input feature vector.

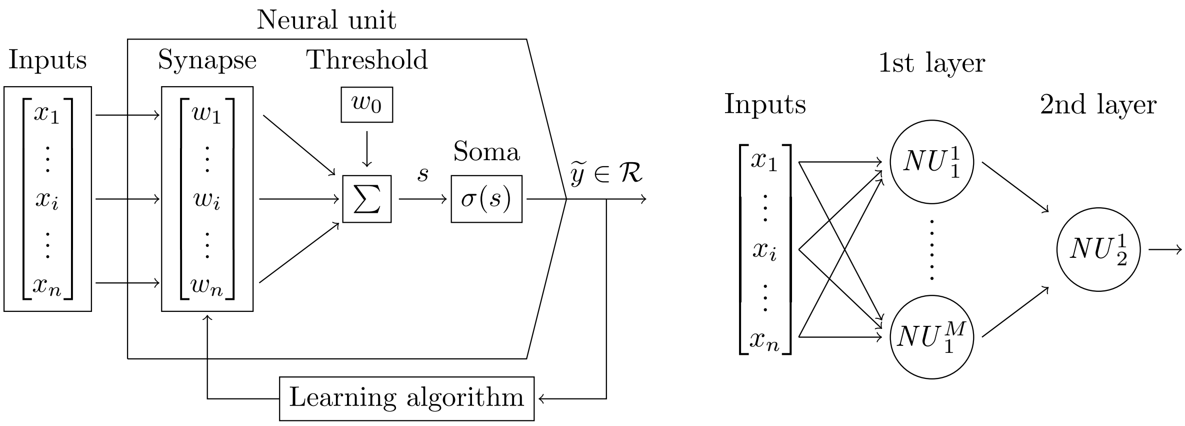

Fig. 2. Neural network: single neural unit (left); shallow neural network (right)

For clarity, let us introduce the labelling of neural networks as it is described in expression (4). It is consisted of network depth L, i.e., number of layers X_L and additional parameters which stand for size of individual layers S, i.e., number of neurons in each layer X_S, Y_S, order of somatic operation and activation function used in individual layers.

$$Name_{ANN} = LX_L\_SX_SY_S\_OX_OY_O \tag{4}$$

It is obvious that number of adaptable weight is dependent on number of neurons in the previous layer but it is also strongly dependent on the order of synaptic operation used in the neurons. Total number of optimizable parameters can be obtained as [1]

$$\sum_{j=0}^{N} \binom{n+j-1}{j} = \sum_{j=0}^{N} \frac{(n+j-1)!}{j!(n-1)!}, \tag{5}$$

where N denotes maximal order of synaptic operation. In Fig. 3, there is a dependency of DoFs based on the designed neural network architecture. Only shallow two layered networks are assumed, as it is indicated in Fig. 2 (right), with maximal thickness in the first layer equals to five neurons. In the output layer, there is a single neuron in all cases. Higher orders up to third only are considered.

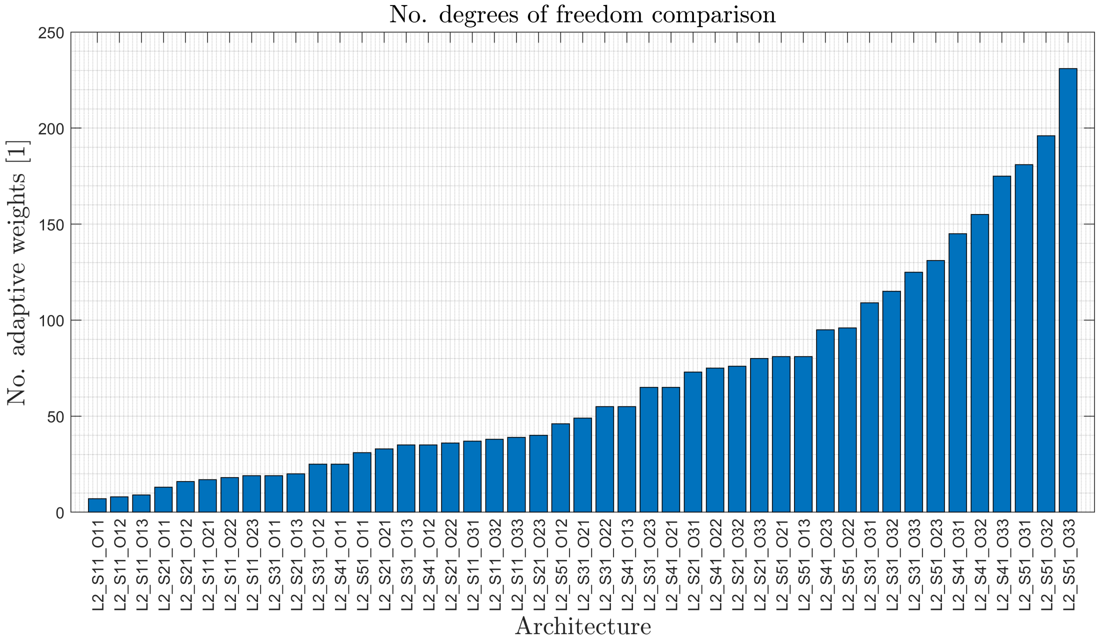

Fig. 3. Proposed neural network architectures: degrees of freedom comparison

4. Results

In Fig. 4, there are results of the learning performed on the training data set with different neural network architectures. Learning rate was set to $\mu = 0.2$ and total number of epochs to $1e3$ in all cases. Pareto front of conflicting criteria, accuracy and number of degrees of freedom, was found and as it can be seen, not all more complex compounding necessarily lead to better performance in sense of testing error value.

Four neural network architectures were found in the Pareto front. In the case of three simplest networks in this set, i.e., architectures *L2_S11_O11*, *L2_S21_O12* and *L2_S31_O13*, expected behaviour is observed. As the number of DoFs is increased, the training error is by half

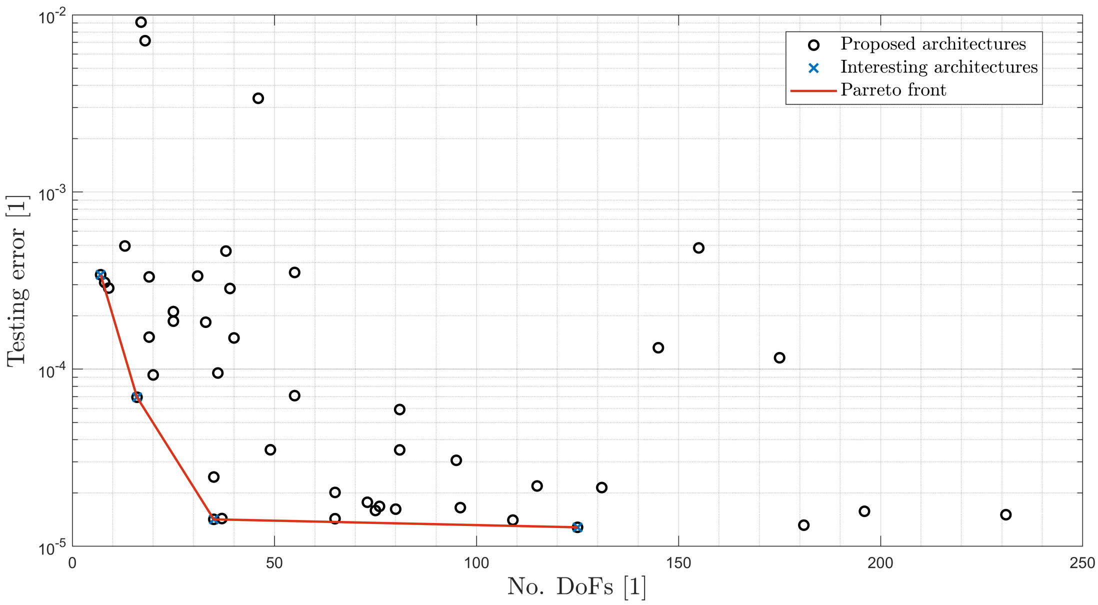

Fig. 4. Proposed neural network architectures: results of the learning

an order of magnitude lower as it is listed in Table 1. Different behaviour can be seen in case of the last architecture on the Pareto front. Architecture *L2_S31_O33* is more than threefold more complex but the improvement of the result accuracy is almost negligible.

Table 1. Networks in the Pareto front

Network	DoFs	Error
L2_S11_O11	7	3.41e-4
L2_S21_O12	16	6.94e-5
L2_S31_O13	35	1.42e-5
L2_S31_O33	125	1.28e-5

5. Conclusions

Learning of the real engineering task using different neural network architectures was presented. It turns out that a more complex neural network does not necessarily better approximate the given problem. Although formula for the apriori estimation of the network complexity was not found, it turned out that more than threefold simpler neural network can approximate the task almost similarly accurate.

It should be noticed that quality of the approximation is strongly dependent on the quality of the training data set. Especially in cases related to CFD simulations where the error can be estimated [6], a bound of permissible error and appropriate network size can be chosen.

More complex formula for the apriori estimation of required neural network complexity based on the complexity of the approximated task should be aim of further research. Thumbling stone of this topic is a method how the complexity of desired pattern can be quantified or if there is a relation between non-linear patterns and approximation using even or odd orders of neurons.

Acknowledgments

This work was supported by the grant agency of the Czech Technical University in Prague, grant No. SGS22/148/OHK2/3T/12.

Authors also acknowledge support from the ESIF, EU Operational Programme Research, Development and Education, and from the Center of Advanced Aerospace Technology (CZ.02.1.01/ 0.0/0.0/16_019/0000826), Faculty of Mechanical Engineering, Czech Technical University in Prague.

References

[1] Gupta, M. M., Bukovsky, I., Homma, N., Solo, A. M., Hou, Z. G., Fundamentals of higher order neural networks for modeling and simulation. In: Artificial Higher Order Neural Networks for Modeling and Simulation, IGI Global, 2013, pp. 103-133.

[2] Gupta, M. M., Liang J., Noriyasu H., Static and dynamic neural networks: From fundamentals to advanced theory, John Wiley & Sons, 2004.

[3] Kovář, P., Fürst, J., Nozzle shape optimization based on machine learning using higher order neural networks, Proceedings of the International Conference Experimental Fluid Mechanics 2022, Safari Park Dvůr Králové nad Labem, 2022, pp. 138-145.

[4] Pařez, J., Tater, A., Kovář, P., Mačák, P., Vampola, T., Influence of geometrical design on the cooling process of double annular turbine section, International Journal of Engine Research 24 (8) (2023) 3707-3719.

[5] Pařez, J., Kovář, P., Tater, A., Modelling of temperature field distribution on single annular using higher order neural network, Proceedings of the 18th Youth Symposium on Experimental Solid Mechanics, Prague, Institute of Theoretical and Applied Mechanics, AS CR, 2023, pp. 1-18.

[6] Roache, P. J., Verification and validation in computational science and engineering, Hermosa Publishers, Albuquerque, 1998.