Operations Research and Engineering Letters
Vol. 2, No. 1, 2023, pp. 1-9
ISSN: 2812-9318

A Novel Ann Technique for Fast Prediction of Structural Behavior

Filip Đorđević [1]
[1] Faculty of Civil Engineering, University of Belgrade, Belgrade Serbia,
e-mail: fdjordjevic@grf.bg.ac.rs

Article Info

Article history:

Received November 17, 2022
Revised January 11, 2023
Accepted January 31, 2023

Keywords:

Machine learning,
Transfer learning,
Artificial neural networks,
CFST columns,
Eurocode 4.

ABSTRACT

In recent decades, different concepts of machine learning (ML) have found applications in solving many engineering problems. Less time consumption in performing analyses, better optimization of the quality-price ratio and maintaining high model accuracy are just some ML advantages compared to traditional modeling procedures. There are currently a significant number of pre-trained machine learning models based on classification or regression tasks. However, there is a tendency to improve them through the implementation of the transfer learning (TL) approach. This article proposes an upgrade of the existing, pre-trained artificial neural network (ANN) model for the evaluation of the ultimate compressive strength of square concrete-filled steel tubular (CFST) columns. The aim of the improved TL model is to adapt to the problem of predicting the axial capacity of rectangular CFST columns in a more optimal way. The attractiveness of the TL is reflected through the possibility of overcoming certain shortcomings of classical models. Quick adaptation to the problem with small modifications of the existing surrogate model, better overcoming of potential overfitting even with a small dataset, and improved convergence towards the required solutions are some of the advanced TL strategies. The robustness of the proposed model was demonstrated through verification with experimental solutions and validation with the Eurocode 4 (EC4) design code. The application of such innovative paradigms can also be ensured for other research fields in a similar manner.

Corresponding Author:

Filip Đorđević,
PhD Candidate, Teaching assistant - Faculty of Civil Engineering, University of Belgrade.
Email: fdjordjevic@grf.bg.ac.rs

1. Introduction

Hands-on implementation of ML algorithms currently reaching its pinnacle in many areas. Along with the rapid development of computers, the possibilities offered by this artificial intelligence (AI) technique are becoming enormous. One of the newest techniques that capture the researcher’s attention is based on taking knowledge of existing models to solve approximate problems, named as transfer learning.

In the past few years, many authors have emphasized the development of non-destructive methods for predicting the axial capacity of CFST members, using a number of ML algorithms. (Đorđević & Kostić, 2022c) proposed Decision Tree (DT) and Random Forest (RF) algorithms for the evaluation of the ultimate compressive strength of circular CFST columns, but on a small amount of data (236 stub columns and 272 slender columns). The lack of samples is visible through the given depth of the architecture and the number of elements in each node of both algorithms, which leads to doubts about the objectivity of the obtained results. Certain improvements are visible during the application of artificial neural networks suggested by

(Đorđević & Kostić, 2022a, 2022b). The investigation in those cases was based on the application of the advanced second-order Levenberg-Marquardt (LM) algorithm for the estimation of the axial capacity of square and circular CFST columns. The model that was developed for square columns with 1022 samples was just used as a pre-trained model, which is improved through the application of TL in the present article. The efficient application of the same LM algorithm was derived by (Zarringol et al., 2020), simultaneously on a square and rectangular CFST columns. (Tran et al., 2020, 2021; Tran & Kim, 2020) proposed ML solutions for the similar problem of CFST columns with ultra-high-performance concrete (UHPC), double-skin steel tube and columns with elliptical sections, respectively. Some studies have focused on the application of hybrid and alternative ML techniques. The combinations of ANNs and genetic algorithm (GA) or particle swarm optimization (PSO) were successfully applied by (Nguyen & Kim, 2021; Nikoo et al., 2015).

Some applications of ML with TL (Weiss et al., 2016) were developed mainly in the areas of Natural Language Processing (NLP) (Prettenhofer & Stein, 2010; Zhou et al., 2014) or Image Recognition (IR) (Duan et al., 2012; Zhu et al., 2008). The application of TL is also recognized in material science for overcoming the problem of a limited amount of data as described by (Yamada et al., 2019). A successful attempt of optimization of torsion design for CFST columns using a two-stage TrAdaBoost, transfer learning-based algorithm was conducted by (H. Huang et al., 2022). The TrAdaBoost algorithm gave better results and outperformed the basic extreme gradient boosting (XGBoost) model.

The advantages of the model derived in this paper is the possibility of its quick implementation, without the traditionally long exploration procedure, less sensitivity on overfitting and the opportunity to obtain reliable results even with a small number of experimental results. The updated existing model using TL shows superiority over the same TensorFlow (TF) model created from scratch using the interaction between Python and Matlab software, as well as over the EC4 solutions. The goal of this study is to discover a novel ML strategy and to open new possible perspectives of its application in the civil engineering practice.

2. Pre-trained ANN model

The single ANN model previously developed by (Đorđević & Kostić, 2022a) for the purpose of predicting the axial capacity of square CFST columns, was used as a starting point for training the target task of rectangular members. The initial model formed from scratch through the TensorFlow paradigm was created by careful selection of network parameters using the trial and error method. The most stable pretrained model has shown a distinct power of predicting the ultimate load capacity of square CFST columns with R^2 values of 0.984, 0.980, 0.976 and 0.982 on training, validation, test set and all data, respectively. The original model was designed on the formulation of supervised learning using the back-propagation (BP) rule with a train/validation/test split of 70/15/15%. Based on the fast Levenberg-Marquardt algorithm, the model has shown a better fit to the regression line than the results generated by EC4 (R^2=0.953) over the entire range. Practically applicable empirical equations were also derived from the mentioned study. Obtained weights and biases from the best ANN model, were used for the application of TL in this study. Transfer learning has particular importance in the case of a deficit in the number of samples, as in this case.

2.1. Datasets

The dataset (1022 samples) of a pre-trained model was established on five input parameters (B, t, L, f_y, f_c') – section width, thickness of the steel tube, length of column, steel yield stress, concrete compressive strength and one output parameter (N_{exp}), i.e. the ultimate compressive strength. An additional database (418 samples) of rectangular columns has the same input and output dimensions but with a small modification of the section width i.e. $B_{eq} = \sqrt{0.5 \cdot (B^2+H^2)}$ as recommended by (Zarringol et al., 2020). The database of rectangular columns was created according to collected experimental results by (Denavit, 2005; Goode, 2008; Thai et al., 2019). Table 1 shows the basic information on distributions of pre-trained and target data. The largest differences in mean values and standard deviations between databases were observed for length of column and for ultimate compressive strength of CFST columns.

Table 1. Distribution of pre-trained / target (TL) model parameters

Parameter	Unit	Mean	St.Dev.	Min.	Max.
$B(B_{eq})$	mm	157.7 / 143.1	70.3 / 47.5	60 / 70.2	750 / 446.5
t	mm	4.5 / 4.2	2.3 / 1.6	0.7 / 0.7	16 / 10
L	mm	936.8 / 741.9	859.5 / 592.4	180 / 200	4500 / 3050
f_y	MPa	388.2 / 371.1	162.1 / 106.4	115 / 145	835 / 550
f_c'	MPa	52.1 / 52.2	31 / 18.4	7 / 7	164.1 / 108.6
N_{exp}	kN	2318.1 / 1860.2	2302.6 / 946.5	105.4 / 182	24294 / 7091

Figure 1 demonstrates the heatmap of the coefficients of correlation between variables. It can be seen that the most influential parameters on the output results are section width B_{eq} (0.68), the thickness of the steel tube t (0.46) and steel yield stress f_y (0.35), similarly to the case in the pre-trained task with the values of 0.79, 0.61 and 0.35 respectively. Such values open up the possibility of acquiring certain knowledge and its employment in the new, improved model.

However, there are certain deviations of mean values and standard deviations of individual parameters between databases, as delivered in Table 1, which indicates the need for additional adaptation of the existing pre-trained model. Further confirmation of this conclusion can be seen if we look at the quantile-quantile (Q-Q) probability plots and frequency plots presented in Figure 2 and 3. One of the prerequisites for applying TL is that the input and output parameters of both databases follow the same/approximately the same distribution, which is fulfilled in this case (see Figure 2 and 3).

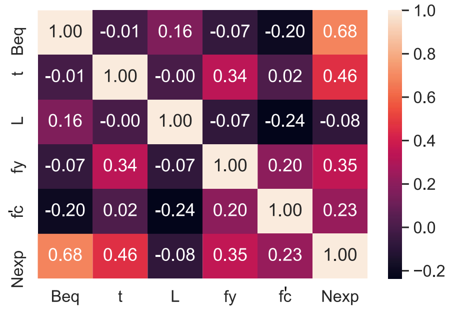


Figure 1. Heatmap of the correlation coefficients

The closest distributions were observed for the thickness of the steel tube, as well as between the output variable. In general, larger or smaller differences between the distributions will represent an equivalent effort that the target ANN will have to overcome. This small drawback is generated in the process of taking knowledge from the existing pre-trained model, but it can be easily controlled by adjusting the network hyperparameters.

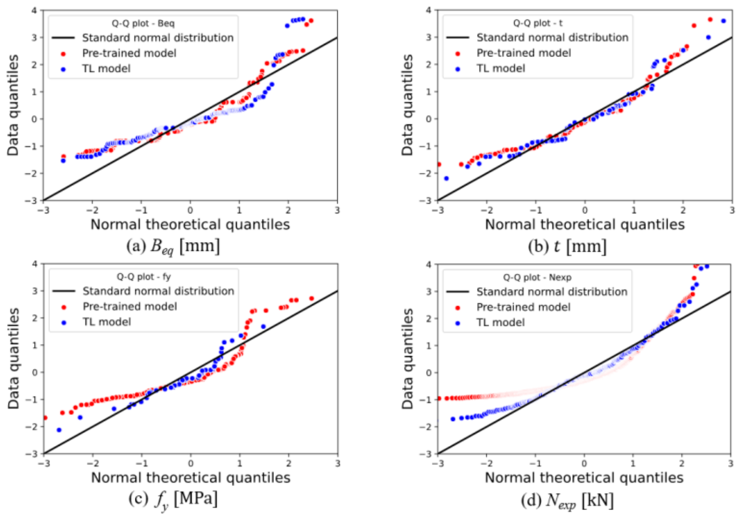
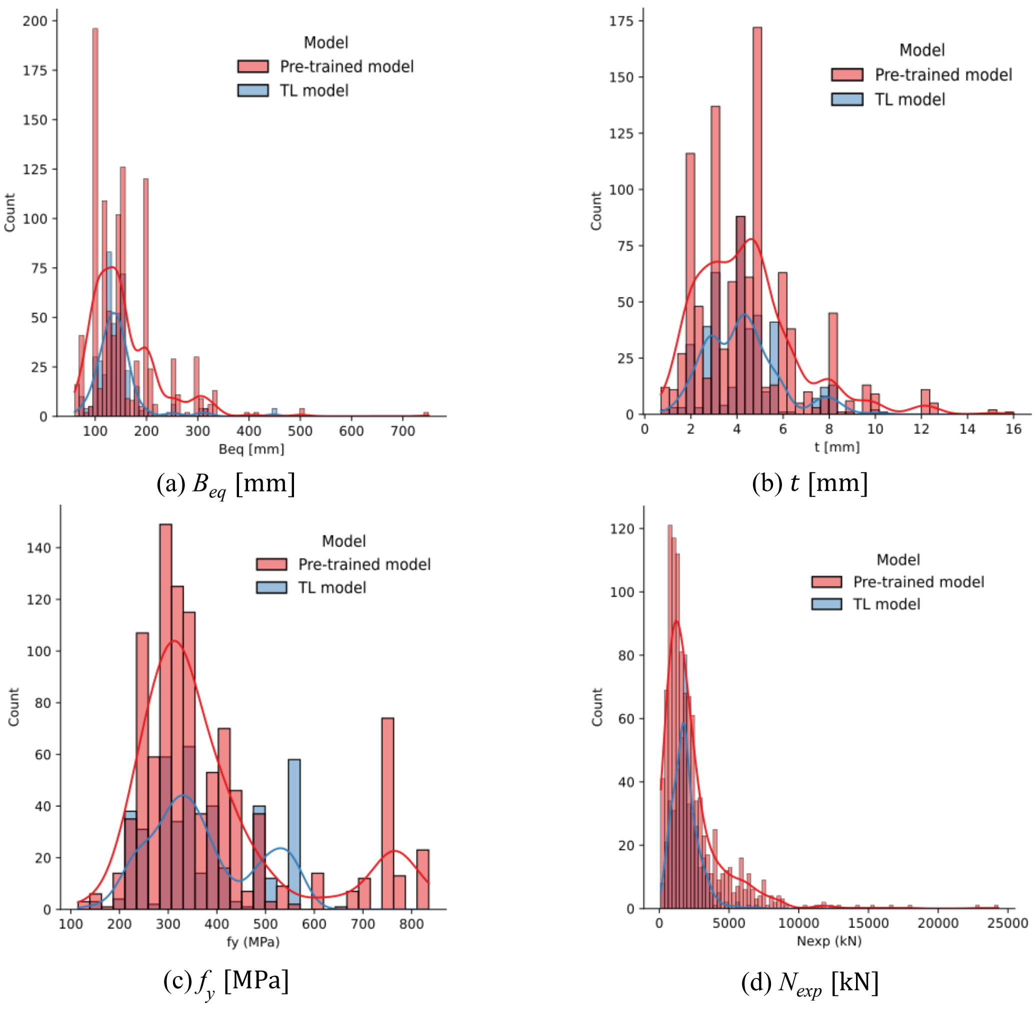


Figure 2. Quantile-quantile plots

Q-Q plot is a graphical method for comparing two distributions by plotting their quantiles. They compare the shapes between the current distribution of data and the equivalent theoretical standard normal distribution. Figure 2 presents multiple Q-Q plots, for the most important input and output parameters, in order to compare variables of initial and new datasets, as well as to compare their distributions simultaneously with the equivalent Gaussian distribution (K.-W. Huang et al., 2019).

The potential normal distribution of the data is observed by matching the points with the identity line x=y, as presented in Figure 2. Each input and output parameter in this paper tends more or less to follow exponential distribution (see Figure 3). By matching the variables of the old and new dataset, in the preprocessing phase, another confirmation was made for the possible use of transfer learning in this analysis. Also, there is an agreement between the Q-Q plots and the frequency plots.

Based on matching the distribution of parameters, a good adaptation of the existing model to the new task is expected. Fine-tuning of the newly created model through slight modification of the hyperparameters is described in subsection 3.1.



(a) B_{eq} [mm] (b) t [mm]

(c) f_y [MPa] (d) N_{exp} [kN]

Figure 3. Frequency plots

3. Transfer learning

Improving the ANN algorithm in order to adapt it to a certain task can be done by transferring information from a related domain. The previous sections showed us distinct matches between related databases of square and rectangular columns, which opened the way for the adaptation of the original model to a new task. Existing TL method will tend to further minimize the differences between the distributions. The benefit of the TL model compared to the one developed from scratch is the possibility of training such models even when there is a lack of sufficient data for the training of the target task, as recognized by (Choi et al., 2017; Yamada et al., 2019). In this article, the mentioned advantage of TL is reflected through the practical implementation of both approaches. By directly training the original model on a small data set, overfitting is inevitable due to the interaction of a reduced data set and a strong model, without previous experience, based on random initialization of network parameters.

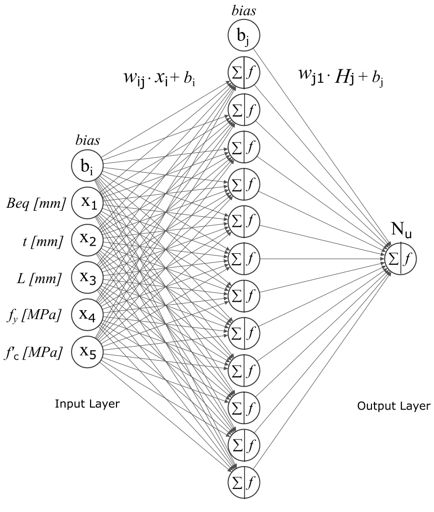
There are two common approaches for ANN transfer learning, freezing and fine-tuning (Guo et al., 2019; Vrbančič & Podgorelec, 2020). The second procedure is applied in this paper, and refers to the possibilities when all or only certain network parameters can be re-trainable. In this case the complete, already learned set of parameters was retrained since the goal was to fully adapt only to the new task. A common property for the pre-trained and TL models is to maintain the same order of inputs as well as the same dimensionality of the problem. It is noticeable that mastering the tasks is much easier, that it is performed in a better way, with significant time savings and increased work speed brought by the TL model.

Train/Validation/Test split is the same in this case (70/15/15%), but considering the amount of data, almost three times less number of samples (293 versus 715) will now be exposed to training. This number is closely related to the strength of the pre-trained model and more relevant target solutions could be expected with an effective and powerful source model. Paradigms based on such recommendations are very often encountered in NLP and IR problems.

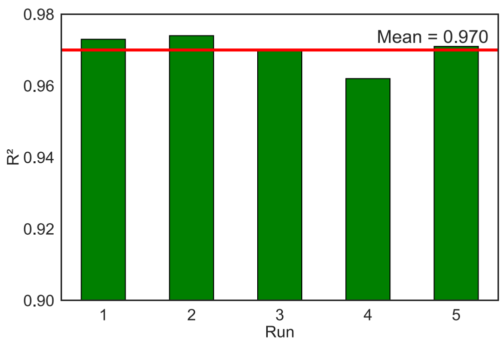
3.1. Hyperparameters fine-tuning

The fine-tuning strategy for the TL task involves replicating a large part of the original model, taking into account network architecture, data scaling, activation functions, supervised learning strategy with LM algorithm following the back-propagation rule, etc. The most optimal network architecture of the pre-trained model had one hidden layer with twelve neurons i.e. 5-12-1 as presented in Figure 4(a). Features scaling was done by mapping the ranges of the training set extracted from the database for square columns to a range -1 to 1. Tangent-hyperbolic and pure linear activation functions were proposed for the hidden and output layer, respectively. Additionally, LM algorithm has also three important hyperparameters, a damping factor μ and related factors μ_{dec} and μ_{inc} with values of 0.1, 0.01 and 10 for square columns. According to the recommendations derived by (Yamada et al., 2019), it is necessary to reduce the learning factor in the TL procedure, which was done here, more precisely, the updated values of the main hyperparameters are 1e-0.5, 1e-0.5 and 7, respectively.

The relevance of the improved model was checked by the multiple runs procedure and the results of coefficients of determination are graphically illustrated in Figure 4(b). It can be seen that the histogram bins show small relative deviations and a high mean value for the validation dataset. The results are distributed for the test set similarly.



(a) ANN architecture



(b) Validation set – multiple runs

Figure 4. ANN architecture and TL multiple runs results (Validation set)

3.2. Eurocode 4

The design value of ultimate compressive strength (N_u^{EC4}) of rectangular and square CFST columns, according to the EC4 propositions, calculates as follows (see Eq.1):

$$N_u^{EC4} = \chi \cdot N_{pl,Rd} = \chi \cdot (A_s \cdot f_y + A_c \cdot f_c') \quad (1)$$

where χ is reduction factor, A_s and A_c are the area of the structural steel section and cross-sectional area of concrete and $N_{pl,Rd}$ named as plastic resistance to compression

The reduction factor for the relevant (flexural) buckling mode depends on the relative slenderness and calculates as presented by (Đorđević & Kostić, 2022a). EC4 also prescribes certain limitations of geometric and material characteristics in accordance with the shape of the cross-section of the column, which partially limits its predictive power against ANN. Traditionally Eurocode 4 is conservative on average over the entire range of results, as will be demonstrated in the following section.

4. Results

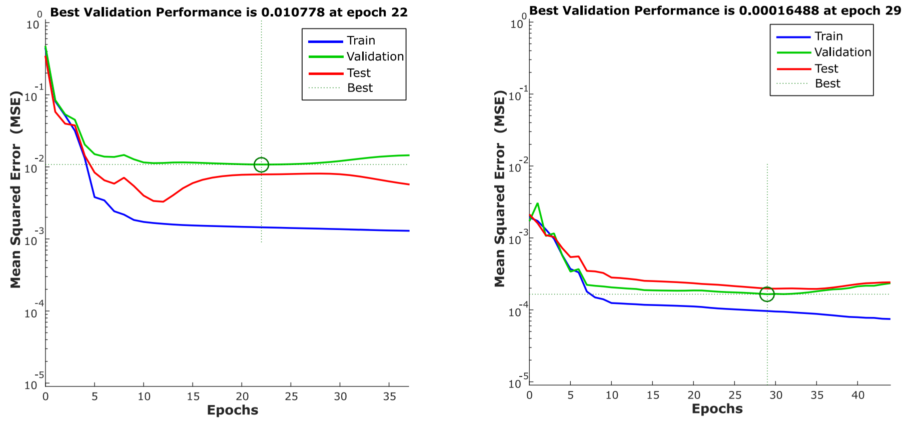
Confirmation of the advantage that TL-ANN has against traditional pure ANN models, especially in environments with small datasets can be seen from the results in Table 2. Figure 5 illustrates the relation between MSE and the number of epochs for TF and TL algorithms for all three subsets of data (training, validation and test set). TensorFlow model from scratch expresses overfitting due to its robustness against dataset size. TL model is superior in the whole range of data, without significant deviations from the regression line, as illustrated in Figure 6. In addition to the (R^2), comparison of the results was made by considering other indicators, mean squared error (MSE) and root mean squared error (RMSE). Higher values of coefficient of determination and lower error values indicate to better model performance and vice versa.

TL model gave significantly better R2 results than the TF model from scratch, with values of 0.984, 0.970, 0.977 and 0.980 for training, validation, test set and all data, respectively. In contrast, the best TF model was generated R2 values of 0.985, 0.864, 0.884 and 0.958 for the same data. It is obvious that the overfitting occurred in the second case, as expected. The same conclusion applies to the values of MSE and RMSE errors. It is evident that the TL model kept the stability that the pre-trained model had on square samples with corresponding R2 values (0.984, 0.980, 0.976 and 0.982). At the same time, EC4 gave a worse prediction for rectangular columns (R2 = 0.927) than for square columns (R2 = 0.953) for all data.

Table 2. Performances of TF model from scratch, TL model and EC4

Data	Set	R^2			MSE ($\cdot 10^{-4}$)		RMSE ($\cdot 10^{-2}$)	
		TF	TL	EC4	TL	EC4	TL	EC4
Rectangular columns	Train	0.985	0.984	-	0.963	-	0.982	-
	Valid	0.864	0.970	-	1.976	-	1.406	-
	Test	0.884	0.977	-	1.649	-	1.284	-
	All	0.958	0.980	0.927	1.219	93.763	1.104	9.683

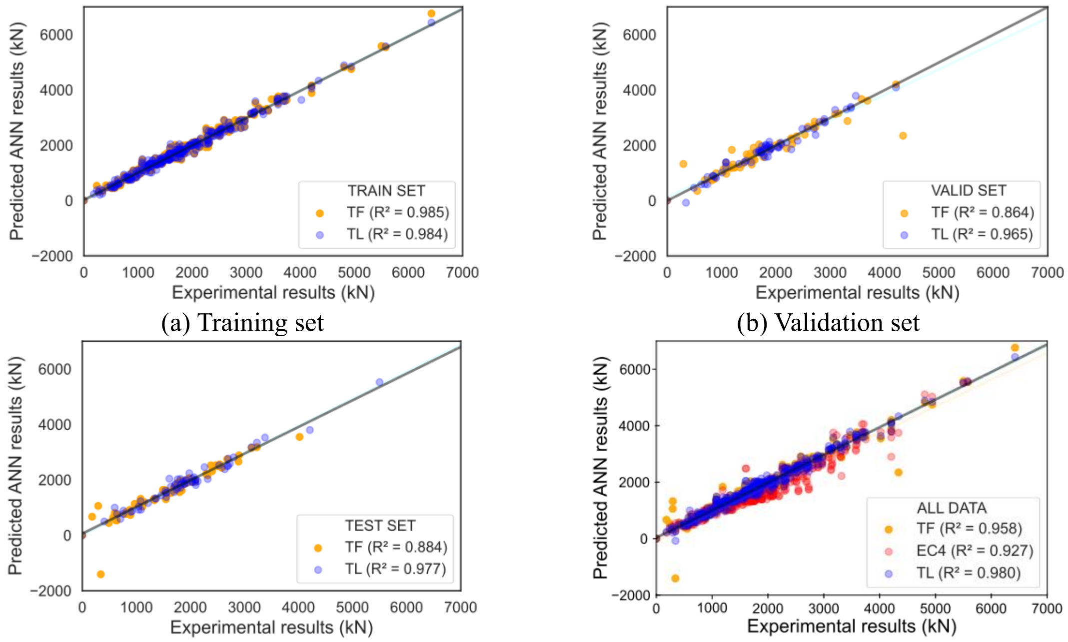
The TF model built from scratch for rectangular columns is the same as the pre-trained model for square columns without any additional changes. This model showed vulnerability in prediction, especially on the test set, which indicates the need for additional time spend in setting up the network for a new task. Regardless, such a model would certainly not give a good generalization of the problem of predicting the axial capacity of rectangular columns.



(a) TensorFlow model (b) TL model

Figure 5. Mean squared error curves of the TensorFlow model from scratch and TL model

The pronounced overfitting of the TF model is so great that the error values on the training and test set differ by an order of magnitude (see Figure 5(a)). TF results show a significantly higher scatter of points around the identity line, while the TL procedure showed an exceptional match between the experimental and predictive results. Considering the obtained results, it is suggested to use the TL approach, considering the much higher degree of reliability of the outputs it possesses. This trend can be especially noticed in Figure 6(b)-(d). EC4 confirmed the conservatives of the results on the entire data set.



(a) Training set (b) Validation set

(c) Test set (d) All data

Figure 5. Comparison of the results of the TensorFlow model from scratch, EC4 and TL procedure

In the process of adapting the pre-trained model to the new task, the difference in weights and biases of the network is clearly visible, i.e. certain connections between neurons became more important for the target solutions.

5. Conclusion

This article once again demonstrated the power of applying machine learning techniques in construction industry with a focus on artificial neural networks and transfer learning procedures. Apart from the higher speed and accuracy it possesses, such a novel approach also shows great practical applicability even compared to conventional commercial techniques based on finite element method (FEM), but also in relation to current design codes such as Eurocodes (EC4). Applying advanced softwares based on FEM even for such simpler models, would require significantly more consumption of time for conducting analyzes than the ANN model, and especially than the improved TL model. The improved TL procedure breaks the barrier of lack of experimental data and design code limitations, and takes an additional step towards the potential development of future fast and efficient software solutions. Future development of computers will enable even more precise capture of the linear and non-linear behavior of complex structures in the civil engineering.

In further research, it is definitely necessary to pay more attention to the connection between robust single models and new programming procedures with experimental data, but also to their availability to the wider community.

References

Choi, K., Fazekas, G., Sandler, M., & Cho, K. (2017). Transfer learning for music classification and regression tasks. Proceedings of the 18th International Society for Music Information Retrieval Conference, ISMIR, 141–149.

Denavit, M. D. (2005). “Steel-concrete composite column database.”

Đorđević, F., & Kostić, S. M. (2022a). Axial Strength Prediction of Square CFST Columns Based on The ANN Model. First Serbian International Conference on Applied Artificial Intelligence.

Đorđević, F., & Kostić, S. M. (2022b). Estimation of Ultimate Strength of Slender CCFST Columns Using Artificial Neural Networks. 16th Congress of Association of Structural Engineers of Serbia, Arandjelovac, Serbia.

Đorđević, F., & Kostić, S. M. (2022c). Prediction of ultimate compressive strength of CCFT columns using machine learning algorithms. The 8th International Conference “Civil Engineering – Science and Practice”.

Duan, L., Xu, D., & Tsang, I. W. (2012). Learning with augmented features for heterogeneous domain adaptation. Proceedings of the 29th International Conference on Machine Learning, ICML 2012, 1, 711–718.

[dataset] Goode, C. D. (2008). 1819 tests on concrete-filled steel tube columns compared with Eurocode 4. Struct Eng, 8(33), 86.

Guo, Y., Shi, H., Kumar, A., Grauman, K., Rosing, T., & Feris, R. (2019). Spottune: Transfer learning through adaptive fine-tuning. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 4800–4809. https://doi.org/10.1109/CVPR.2019.00494

Huang, H., Xue, C., Zhang, W., & Guo, M. (2022). Torsion design of CFRP-CFST columns using a data-driven optimization approach. Engineering Structures, 251. https://doi.org/https://doi.org/10.1016/j.engstruct.2021.113479

Huang, K.-W., Qiao, M., Liu, X., Liu, S., & Dai, M. (2019). Computer Vision and Metrics Learning for Hypothesis Testing: An Application of Q-Q Plot for Normality Test. 1–14. http://arxiv.org/abs/1901.07851

Nguyen, M. S. T., & Kim, S. E. (2021). A hybrid machine learning approach in prediction and uncertainty quantification of ultimate compressive strength of RCFST columns. Construction and Building Materials, 302(February). https://doi.org/10.1016/j.conbuildmat.2021.124208

Nikoo, M., Torabian Moghadam, F., & Sadowski, Ł. (2015). Prediction of Concrete Compressive Strength by Evolutionary Artificial Neural Networks. Advances in Materials Science and Engineering. https://doi.org/10.1155/2015/849126

Prettenhofer, P., & Stein, B. (2010). Cross-Language Text Classification Using Structural Correspondence Learning. Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, 1118–1127. https://aclanthology.org/P10-1114

Thai, S., Thai, H. T., Uy, B., & Ngo, T. (2019). Concrete-filled steel tubular columns: Test database, design and calibration. Journal of Constructional Steel Research, 157, 161–181. https://doi.org/10.1016/j.jcsr.2019.02.024

Tran, V. L., Jang, Y., & Kim, S. E. (2021). Improving the axial compression capacity prediction of elliptical CFST columns using a hybrid ANN-IP model. Steel and Composite Structures, 39(3), 319–335. https://doi.org/https://doi.org/10.12989/scs.2021.39.3.319

Tran, V. L., & Kim, S. E. (2020). Efficiency of three advanced data-driven models for predicting axial compression capacity of CFDST columns. Thin-Walled Structures, 152(April). https://doi.org/10.1016/j.tws.2020.106744

Tran, V. L., Thai, D. K., & Nguyen, D. D. (2020). Practical artificial neural network tool for predicting the axial compression capacity of circular concrete-filled steel tube columns with ultra-high-strength concrete. Thin-Walled Structures, 151(January). https://doi.org/10.1016/j.tws.2020.106720

Vrbančič, G., & Podgorelec, V. (2020). Transfer learning with adaptive fine-tuning. IEEE Access, 8(December), 196197–196211. https://doi.org/10.1109/ACCESS.2020.3034343

Weiss, K., Khoshgoftaar, T. M., & Wang, D. D. (2016). A survey of transfer learning. In Journal of Big Data, 3(1). Springer International Publishing. https://doi.org/10.1186/s40537-016-0043-6

Yamada, H., Liu, C., Wu, S., Koyama, Y., Ju, S., Shiomi, J., Morikawa, J., & Yoshida, R. (2019). Predicting Materials Properties with Little Data Using Shotgun Transfer Learning. ACS Central Science, 5(10), 1717–1730. https://doi.org/10.1021/acscentsci.9b00804

Zarringol, M., Thai, H. T., Thai, S., & Patel, V. (2020). Application of ANN to the design of CFST columns. Structures, 28(August), 2203–2220. https://doi.org/10.1016/j.istruc.2020.10.048

Zhou, J. T., W.Tsang, I., Pan, S. J., & Tan, M. (2014). Heterogeneous Domain Adaptation for Multiple Classes. In S. Kaski & J. Corander (Eds.), Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics, 33, pp. 1095–1103). PMLR. https://proceedings.mlr.press/v33/zhou14.html

Zhu, Y., Chen, Y., & Lu, Z. (2008). Heterogeneous Transfer Learning for Image Classification. 1304–1309.