BUILDING CONSTRUCTIONS, BUILDINGS AND ENGINEERING STRUCTURES

СТРОИТЕЛЬНЫЕ КОНСТРУКЦИИ, ЗДАНИЯ И СООРУЖЕНИЯ

Check for updates

Original Empirical Research

UDC 624.04
https://doi.org/10.23947/2949-1835-2025-4-4-44-52

EDN: BCUPRW

Predicting the Load-Bearing Capacity of Square-Section Pipe-Concrete Columns Using Machine Learning Methods

Tatiana N. Kondratieva ✉, **Anton S. Chepurnenko**
Don State Technical University, Rostov-on-Don, Russian Federation
✉ ktn618@yandex.ru

Abstract
Introduction. In this paper, we consider the problem of predicting the strength of square-section centrally compressed short concrete-filled tubular columns using machine learning methods. Traditional methods, such as the finite element method and the theoretical-experimental approach involving selection of empirical formulas require significant computational resources and time. At the same time, these methods are not always capable of accounting for complex nonlinear dependencies between the parameters. The key objective is to develop a high-precision model capable of predicting the load-bearing capacity of columns using the major parameters.
Materials and Methods. For the current study, a database was generated containing the results of numerical experiments on calculating the load-bearing capacity of square-section concrete-filled tubular columns in a physically nonlinear formulation. As part of the study, models based on machine learning methods were designed and implemented using the Jupyter Notebook interactive computing platform. The main method is the CatBoost mechanism (Gradient Boosting Regressor). The resulting models were trained by means of nonlinear optimization methods.
Results. The article evaluates the degree of impact of each of the input parameters on the final predictions of the model. The results on the degree of impact for the CatBoost and Random Forrest Regressor (RFR) models are obtained. The quality of the resulting models evaluated using the R^2 value was 98% for CatBoost and 94% for RFR.
Discussion and Conclusions. The resulting approach has proved to be highly efficient in predicting the load-bearing capacity of concrete-filled tubular columns, providing a balance between the accuracy of the results and computational complexity.

Keywords: concrete-filled tubular columns, machine learning methods, prediction, load-bearing capacity, artificial intelligence, artificial neural networks

Acknowledgements. The authors appreciate the reviewers, whose critical assessment of the submitted materials and suggestions helped to significantly improve the quality of the project.

For citation: Kondratieva TN, Chepurnenko AS. Predicting the Load-Bearing Capacity of Square-Section Pipe-Concrete Columns Using Machine Learning Methods. *Modern Trends in Construction, Urban and Territorial Planning.* 2025;4(4):44–52. https://doi.org/10.23947/2949-1835-2025-4-4-44-52

Оригинальное эмпирическое исследование

Прогнозирование несущей способности трубобетонных колонн квадратного сечения при помощи методов машинного обучения

Т.Н. Кондратьева ✉, **А.С. Чепурненко**
Донской государственный технический университет, г. Ростов-на-Дону, Российская Федерация
✉ ktn618@yandex.ru

Аннотация
Введение. В данной работе рассматривается задача прогнозирования прочности центрально сжатых коротких трубобетонных колонн квадратного сечения с использованием методов машинного обучения. Традиционные методы,

такие как метод конечных элементов и теоретико-экспериментальный подход с подбором эмпирических формул, требуют значительных вычислительных ресурсов и времени. В то же время эти методы не всегда способны учитывать сложные нелинейные зависимости между параметрами. Основная цель — разработка высокоточной модели, способной предсказывать несущую способность колонн на основе ключевых параметров.
Материалы и методы. Для исследования была сгенерирована база данных, состоящая из результатов численных экспериментов по расчету несущей способности трубобетонных колонн квадратного поперечного сечения в физически нелинейной постановке. В рамках проведенного исследования построены модели на основе методов машинного обучения, реализованные с использованием интерактивной вычислительной платформы Jupyter Notebook. Основным методом является механизм CatBoost (Gradient Boosting Regressor). Обучение построенных моделей произведено с использованием методов нелинейной оптимизации.
Результаты исследования. В статье проведена оценка степени влияния каждого входного параметра на итоговые предсказания модели. Получены результаты по величине степени влияния для моделей CatBoost и Random Forrest Regressor (RFR). Оценка качества построенных моделей по величине R^2 составила 98 % для CatBoost и 94 % — для RFR.
Обсуждение и заключение. Разработанный подход демонстрирует высокую эффективность в задаче прогнозирования несущей способности трубобетонных колонн, обеспечивая баланс между точностью результатов и вычислительной сложностью.

Ключевые слова: трубобетонные колонны, методы машинного обучения, прогнозирование, несущая способность, искусственный интеллект, искусственные нейронные сети.

Благодарности. Авторы выражают благодарность редакции и рецензентам за внимательное отношение к статье и указанные замечания, которые позволили повысить ее качество.

Для цитирования: Кондратьева Т.Н., Чепурненко А.С. Прогнозирование несущей способности трубобетонных колонн квадратного сечения при помощи методов машинного обучения. *Современные тенденции в строительстве, градостроительстве и планировке территорий*. 2025;4(4):44–52. https://doi.org/10.23947/2949-1835-2025-4-4-44-52

Introduction. Evaluating technical condition of monolithic reinforced concrete structures is remaining an urgent and common task, particularly considering there is a need to ensure they are durable and safe. This can be addressed not only by means of analytical and computational methods [1–3], but also using the more up-to-date reputable methods of artificial intelligence (AI) and machine learning (ML) [4–6].

The commonly used finite element method (FEM) enables complex physical processes such as the nonlinear behavior of materials, interaction of steel and concrete [7], as well as the influence of a range of loads to be accounted for [8]. However, the major drawback of the FEM is its high computational complexity and the need for a large number of parameters in order to calibrate the model.

ML methods are a modern data analysis tool allowing one to identify complex nonlinear dependencies between input and output parameters [9–11]. Unlike empirical formulas, machine learning enables patterns to be automatically identified in large amounts of data making it a more versatile and efficient prediction method.

In [12], the authors examine the formation of defects in reinforced concrete structures by means of artificial intelligence algorithms such as random forest (RF), support vector machine (SVM), classification and regression tree (CART), and gradient boosting.

In modern practice, convolutional neural networks (CNN) are being increasingly used in order to predict the strength of reinforced concrete structures [13–15]. E.g., in [13], the authors developed a CNN that is capable of two-dimensional full-scale prediction of crack formation at early stages and description of the entire fracture process. A model capable of predicting both crack initiation and propagation was set forth in [14]. In order to monitor the condition of reinforced concrete structures at complex construction sites, the authors of [15] made use of a fully convolutional neural network (FCN) to segment images and localize cracks on concrete surfaces, accounting for the heterogeneity of concrete properties. The resulting FCN model minimizes false positive and false negative results and is also of high quality, which enables small and complex cracks to be segmented.

In [16], an automated classifier was developed that also functions as a tool for automatic detection and classification of cracks in reinforced concrete columns of different levels of complexity by means of deep CNN (DCNN) methods. The suggested DCNN model analyzes complex textures as well as noises and displays high crack detection accuracy of 96% due to the depth of the model layers and expansion of each layer in a parallel manner.

In order to predict cracks in time, the authors of [17] made a step forward in their research and designed a hybrid model combining DCNN and recurrent neural networks (RNN).

Hence there is no doubt that ML algorithms have a few advantages, such as identifying patterns in large amounts of data, detecting hidden patterns and dependencies accounting for the multidimensional nature of data, automatic analysis of evaluating the condition of reinforced concrete structures using the major parameters, optimization of ML algorithms and parallel computing.

However, these ML algorithms have some drawbacks, such as inaccuracy or weakness, limited generalization ability, and low-speed operation [18, 19]. One of the key ones is the dependence of machine learning models on the quality of training data and its volume.

While training most artificial intelligence models, data from field experiments is used in order to predict the strength of concrete-filled tubular columns [20–22]. Such experiments are typically conducted on samples with relatively small cross-sectional dimensions compared to actual structures. Considering the poor ability of machine learning methods to extrapolate data, serious errors are likely while predicting the load-bearing capacity of actual structures. A solution would be to make use of a combined approach, where training data is generated by means of a finite element calculation of structures with actual dimensions using a method validated based on experimental data.

The objective of the study is to develop machine-learning models for predicting the strength of centrally compressed square-section concrete-filled tubular columns using the data obtained as described above.

Materials and Methods. A database was generated for this study representing the results of numerical experiments in order to calculate the load-bearing capacity of short square-section concrete-filled tubular columns according to the methodology described in [23]. That data was used in order to develop and analyze models combining the traditional methods of structural mechanics and machine learning algorithms.

The input parameters describing the basic geometric as well as physical and mechanical characteristics of the columns were generated in uniform increments in the ranges typical of actual structures, which enabled a wide range of possible combinations to be covered.

The key parameters are as follows: a is the external size of the column cross section, mm; t is the wall thickness of a steel square tube, mm; R_y is the yield strength of steel, MPa; R_b is the compressive strength of concrete, MPa.

The output parameter is the load-bearing capacity of the concrete-filled tubular columns N_{ult}, kN. This indicator was obtained as a result of numerical experiments performed according to the methodology described in [24]. The calculations accounted for the complex interaction of the steel tube and the concrete core, including the joint operation of the materials and their deformation behavior.

The analyzed data array is partially shown in Table 1. The training sample included the total of 22,308 items.

Table 1

Table of the generated data

No.	a, mm	t, mm	R_y, MPa	R_b, MPa	N_{ult}, kN
1	100	3.00	220	10	349.71
2	100	3.45	220	10	385.27
3	100	3.91	220	10	420.72
4	100	4.36	220	10	455.76
5	100	4.82	220	10	490.38
6	100	5.27	220	10	524.59
7	100	5.27	220	10	524.59
8	100	5.73	220	10	558.38
9	100	6.18	220	10	591.76
10	100	6.64	220	10	625.31
11	100	7.09	220	10	657.89
…	…	…	…	…	…
22.299	500	10.55	840	120	44248.28
22.300	500	11.82	840	120	45887.56
22.301	500	13.09	840	120	47511.10
22.302	500	14.36	840	120	49118.97
22.303	500	15.64	840	120	50759.78
22.304	500	16.91	840	120	52338.05
22.305	500	18.18	840	120	53900.79
22.306	500	19.45	840	120	55501.41
22.307	500	20.73	840	120	57089.77
22.308	500	22.00	840	120	58609.37

In order to improve the quality of the models, data preprocessing was performed: normalization, data separation, and cross-validation. The values of each parameter were scaled in the range (0–1) in order to prevent the features with large values from dominating. The generated data was divided into the training (80%) and test (20%) arrays for the model training and evaluation.

The following machine learning algorithms were used to analyze the data and design models for predicting the strength of centrally compressed square-section concrete-filled tubular columns: Linear Regression, Decision Tree, Gradient Boosting, Random Forest Regressor, RFR.

The regularization method was used to normalize the parameters, the Optuna method for optimization, and GridSearchCV and RandomizedSearchCV for hyperparameter selection. The range of parameter values for the CatBoost model was: iterations — 1000–1500; depth — 4–8; learning_rate — 0.1–0.6; l2 reg_lambda — 1.9–4.9. For RFR: n_estimators — 100–250; max_depth — 10–20; min_samples_leaf — 1–4. As the RFR model has no iteration tracking option, model training is possible with a different number of trees and root mean square error (MSE) analysis. With a small number of trees, the RFR model is undertrained and displays a low quality score. As the number of trees increases, the MSE score gets stabilized and the quality score of the model becomes satisfactory.

For the trained models, the significance of the features was also analyzed by evaluating the degree of impact of each input parameter on the final predictions of the model. This approach allowed us to identify an extent to which the prediction results change when do the values of a certain feature.

Research Results. The statistical characteristics of the initial data set are shown in a table (Table 2). The main indicators are as follows: sample size, sample mean, variation dispersion, extremes of the variable values. All of these indicators help to statistically analyze the variables, identify their range in relation to their centre, show the asymmetry of the distribution, and deduce the distribution laws of these variation rows.

Table 2

Table of the statistical characteristics

Parameter	a, mm	t, mm	R_y, MPa	R_b, MPa	N_{ult}, kN
Number	22308	22308	22308	22308	22308
Mean	253.85	9.92	530.00	65.0	10564.50
Standard deviation	128.40	5.06	196.07	34.3	10419.09
min	100.00	3.00	220.00	10.0	349.71
max	500.00	22.00	840.00	120.0	58609.37

Fig. 1 shows the correlation between the model parameters. There is a strong correlation ($0.6 \le |\rho| \le 0.9$) between the parameters: the external size of the column cross-section and the wall thickness of the steel square tube ($\rho_{a/t} = 0.7$); the external size of the column cross-section and the load-bearing capacity of the concrete-filled tubular columns ($\rho_{a/N_{ult}} = 0.88$); the wall thickness of the steel square tube and the load-bearing capacity of concrete-filled tubular columns ($\rho_{t/N_{ult}} = 0.73$).

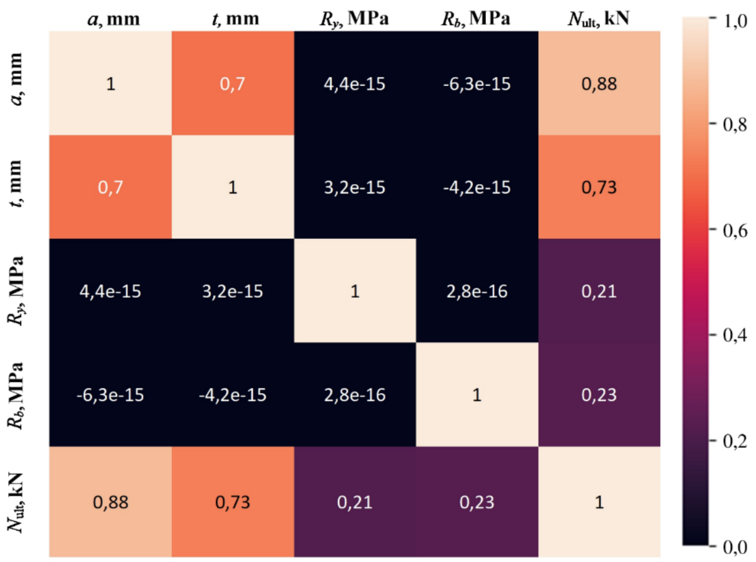

Fig. 1. Correlation matrix

The study focused on the CatBoost gradient boosting algorithm that showed the best results among the tested algorithms (R^2 = 0.98).

The most significant parameter of the CatBoost model is the external cross-sectional dimension of the column, its significance is 96%, the impact of the compressive strength of concrete was 33%, the yield strength of steel was 28%, and the wall thickness of the square steel tube was 20%. The most significant parameters of the RFR model and their degree of importance were distributed as follows: the external cross-sectional dimension of the column was 92%, the compressive strength of concrete was 21%, the yield strength of steel was 17%, and the wall thickness of the square steel tube was 14%. The significance of the influencing factors for both models coincides, and the quantitative assessment of the contribution of each feature is clearly shown in Fig. 2 and 3, respectively.

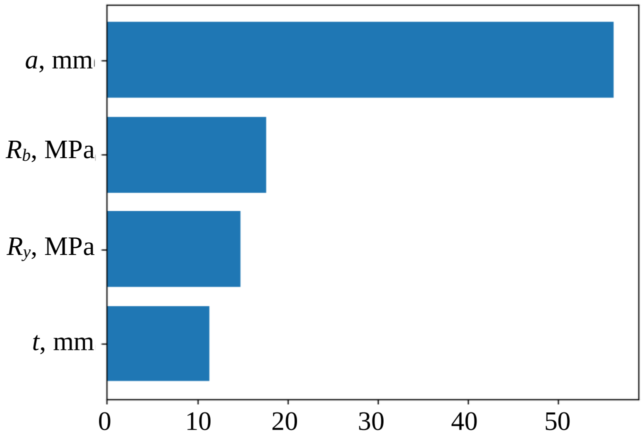

Fig. 2. Evaluation of the signifance of the features for CatBoost

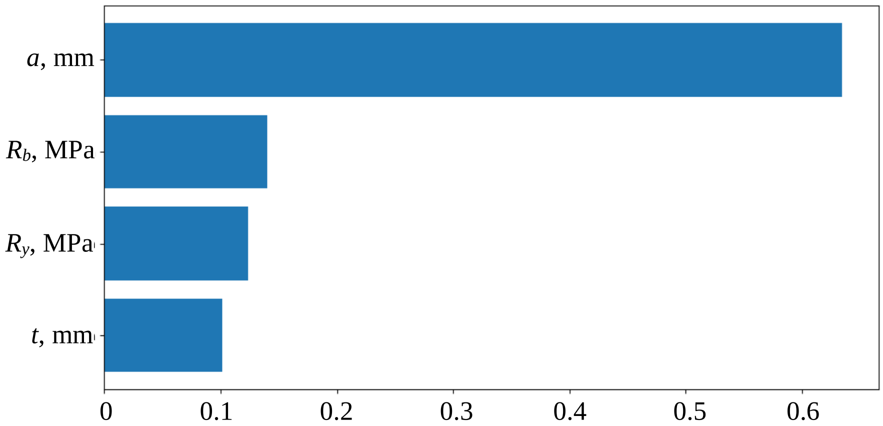

Fig. 3. Evaluation of the signifance of the features for RFR

The optimal parameter values obtained during the model training are shown in Table 3.

Table 3

Optimal values of the model parameters

Model	Parameter	Value
CatBoost	Iterations	1500
	Depth	5
	Learning rate	0.4
	l2 leaf reg	2.8
RFR	N estimators	180
	Max depth	6
	Min samples leaf	1

The evaluation of the quality of the models is shown in Table 4.

Table 4

Model quality metrics

Metrics/Model	CatBoost	RFR
MAE	3.1	7.8
MSE	5.4	4.5
MAPE, %	0.015	0.007
R^2	0.98	0.94

Fig. 4 and 5 show the error histograms: the actual values along the ordinate axis and the predicted ones along the abscissa axis.

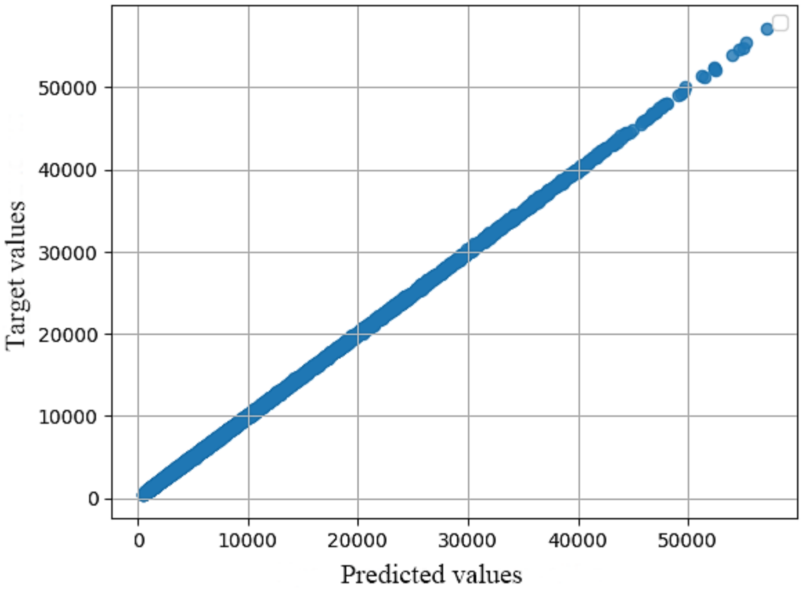

Fig. 4. Error histogram for CatBoost

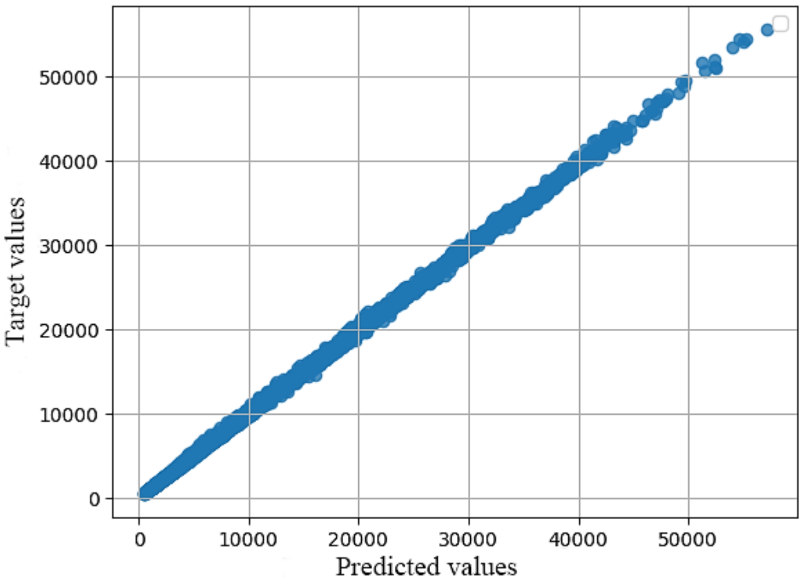

Fig. 5. Error histogram for RFR

Discussion and Conclusion. This study has provided a comprehensive overview of the existing methods for predicting the strength of concrete-filled tubular columns and outlines the advantages of using machine learning in this field.

The use of machine learning methods, particularly CatBoost, has enabled us to identify the precise dependencies between the parameters outperforming the traditional empirical methods. The prediction reliability using the R^2 value for the model based on the CatBoost algorithm was 0.98. The model based on the Random Forest Regressor method displayed a lower accuracy ($R^2 = 0.94$).

According to the analysis of the significance of the features, the external cross-sectional dimension of a concrete-filled tubular column is the major parameter that has the greatest impact on its load-bearing capacity.

In further studies, the range of model parameters based on the current results is going to be expanded considering some additional factors. The additional parameters can include the eccentricity of a longitudinal force, the flexibility of an element, the proportion of prolonged loads in the total loading, etc.

References

1. Abd-El-Nabi E, El-Helloty A, Summra A. Numerical Analysis of Reinforced Concrete Buildings Subjected to Blast Load. *Structural Concrete.* 2023;24(3):3727–3743. https://doi.org/10.1002/suco.202200726

2. Ning J, Meng F, Ma T, Xu X. Failure Analysis of Reinforced Concrete Slab under Impact Loading Using a Novel Numerical Method. *International Journal of Impact Engineering.* 2020;144:103647. https://doi.org/10.1016/j.ijimpeng.2020.103647

3. Kumar V, Kartik KV, Iqbal MA. Experimental and Numerical Investigation of Reinforced Concrete Slabs under Blast Loading. *Engineering Structures*. 2020;(206):110125. https://doi.org/10.1016/j.engstruct.2019.110125

4. Beskopylny A.N., Stel'makh S.A., Shcherban' E.M., Mailyan L.R., Mes-khi B., Razveeva I. et al. Prediction of the Compressive Strength of Vibrocentri-fuged Concrete Using Machine Learning Methods. Buildings. 2024;14(377). https://doi.org/10.3390/buildings14020377

5. Beskopylny A.N., Stel'makh S.A., Shcherban' E.M., Razveeva I., Ko-zhakin A, Kondratieva T.N. et al. Prediction of the Properties of Vibro-Centrifuged Variatropic Concrete in Aggressive Environments Using Machine Learning Methods. Buildings 2024;(14):1198. https://doi.org/10.3390/buildings14051198

6. Kondratieva TN, Chepurnenko AS. Prediction of Rheological Parameters of Polymers by Machine Learning Methods. *Advanced Engineering Research*. 2024;24(1):36–47. https://doi.org/10.23947/2687-1653-2024-24-1-36-47

7. Klemczak B, Żmij A. Insight into Thermal Stress Distribution and Required Reinforcement Reducing Early-Age Cracking in Mass Foundation Slabs. *Materials*. 2021;14(3):477. https://doi.org/10.3390/ma14030477

8. Talaat M, Yehia E, Mazek SA, Genidi M, Sherif A. Finite Element Analysis of RC Buildings Subjected to Blast Loading. *Ain Shams Engineering Journal*. 2022;13(4):101689. https://doi.org/10.1016/j.asej.2021.101689

9. Chepurnenko AS, Kondratieva TN, Deberdeev TR, Akopyan VF, Avakov AA. Prediction of Rheological Parameters of Polymers Using the CatBoost Gradient Boosting Algorithm. *Polymer Science*, *Series D*. 2024;17(1):121–128. https://doi.org/10.1134/S199542122370020X

10. Kondratieva TN, Chepurnenko AS, Poliakova KA, Rodionov KA. CatBoost Algorithms to Predict the Load-Bearing Capacity of Centrally Compressed Short CFST Columns of Circular Cross-Section. *E3S Web Conf.* 2024;583:06009. https://doi.org/10.1051/e3sconf/202458306009

11. Chepurnenko AS, Kondratieva TN, Al-Wali E Processing of Polymers Stress Relaxation Curves Using Machine Learning Methods. *Baghdad Science Journal*. 2023;20(6):2488–2488. https://doi.org/10.21123/bsj.2023.8819

12. Althoey F, Amin MN, Khan K, Usman MM, Khan MA. Machine Learning Based Computational Approach for Crack Width Detection of Self-Healing Concrete. *Case Studies in Construction Materials*. 2022;(17):e01610. https://doi.org/10.1016/j.cscm.2022.e0161

13. Koopas RN, Rezaei S, Rauter N, Ostwald R, Lammering R. A Spatiotemporal Deep Learning Framework for Prediction of Crack Dynamics in Heterogeneous Solids: Efficient Mapping of Concrete Microstructures to its Fracture Properties. *Engineering Fracture Mechanics*. 2025;(314):110675. https://doi.org/10.1016/j.engfracmech.2024.110675

14. Ali R, Chuah JH, Talip M, Mokhtar N, Shoaib MA. Structural Crack Detection Using Deep Convolutional Neural Networks. *Automation in Construction*. 2022;(133):103989. https://doi.org/10.1016/j.autcon.2021.103989

15. Ye X, Jin T, Li ZX, Chen Pengyu, Ma SY, Ding Y, Ou YH. Structural Crack Detection from Benchmark Data Sets Using Pruned Fully Convolutional Networks. *Journal of Structural Engineering*. 2021;147(11):04721008. https://doi.org/10.1061/(ASCE)ST.1943-541X.0003140

16. Wang W, Hu W, Wang W, Xu X, Wang M, Shi Y. Automated Crack Severity Level Detection and Classification for Ballastless Track Slab Using Deep Convolutional Neural Network. *Automation in Construction*. 2021;(124):103484. https://doi.org/10.1016/j.autcon.2020.103484

17. Geetha GK, Yang HJ, Sim SH. Fast Detection of Missing Thin Propagating Cracks during Deep-Learning-Based Concrete Crack/Non-Crack Classification. *Sensors*. 2023;23(3):1419. https://doi.org/10.3390/s23031419

18. Chen Q, Xie Y, Ao Y, Li T, Chen G, Ren S. A Deep Neural Network Inverse Solution to Recover Pre-Crash Impact Data of Car Collisions. *Transportation Research Part C: Emerging Technologies*. 2021;(126):103009. https://doi.org/10.1016/j.trc.2021.103009

19. Chen G, Li T, Chen Q, Ren S, Wang C, Li S. Application of Deep Learning Neural Network to Identify Collision Load Conditions Based on Permanent Plastic Deformation of Shell Structures. *Computational Mechanics*. 2019;(64):435–449. https://doi.org/10.1007/s00466-019-01706-2

20. Tran VL, Thai DC, Nguyen DD. 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*. 2020;(151):106720. https://doi.org/10.1016/j.tws.2020.106720

21. Karatas CA. Artificial Neural Network (ANN) Based Prediction of Ultimate Axial Load Capacity of Concrete-Filled Steel Tube Columns (CFSTCs). *International Journal of Steel Structures*. 2022;22(5):1341–1358. https://doi.org/10.1007/s13296-022-00645-8

22. Bardhan A, Biswas R, Kardani N, Iqbal M, Samui P, Singh MP. A Novel Integrated Approach of Augmented Grey Wolf Optimizer and ANN for Estimating Axial Load Carrying-Capacity of Concrete-Filled Steel Tube Columns. *Construction and Building Materials*. 2022;(337):127454. https://doi.org/10.1016/j.conbuildmat.2022.127454

23. Chepurnenko AS, Yazyev BM, Turina VS, Akopyan VF. Artificial Intelligence Models for Determining the Strength of Centrally Compressed Pipe-Concrete Columns with Square Cross-Section. *Magazine of Civil Engineering*. 2024;17(6):13008. https://doi.org/10.34910/MCE.130.8

24. Chepurnenko A, Yazyev B, Meskhi B, Beskopylny A, Khashkhozhev K, Chepurnenko V. Simplified 2D Finite Element Model for Calculation of the Bearing Capacity of Eccentrically Compressed Concrete-Filled Steel Tubular Columns. *Applied Sciences*. 2021;11(24):11645. https://doi.org/10.3390/app112411645

About the Authors:

Tatiana N. Kondratieva, Cand.Sci. (Eng.), Associate Professor of the Department of Mathematics and Computer Science at the Don State Technical University (1 Gagarin Square, Rostov-on-Don, 344003, Russian Federation), ORCID, ResearcherID, ScopusID, ktn618@yandex.ru

Anton S. Chepurnenko, Dr.Sci. (Eng.), Associate Professor, Professor of the Department of Structural Mechanics and Theory of Structures at the Don State Technical University, (1 Gagarin Square, Rostov-on-Don, 344003, Russian Federation), ORCID, ResearcherID, ScopusID, anton_chepurnenk@mail.ru

Claimed contributorship:

TN Kondratieva: formation of the basic concept, objectives of the study, calculations, analysis of the research results.

AS Chepurnenko: scientific supervision, revision of the manuscript, correction of the conclusions.

Conflict of interest statement: **the authors do not have any conflict of interest.**

All authors have read and approved the final version of manuscript.

Об авторах:

Кондратьева Татьяна Николаевна, кандидат технических наук, доцент, доцент кафедры математики и информатики Донского государственного технического университета (344003, Российская Федерация, г. Ростов-на-Дону, пл. Гагарина, 1), ORCID, ResearcherID, ScopusID, ktn618@yandex.ru

Чепурненко Антон Сергеевич, доктор технических наук, доцент, профессор кафедры строительной механики и теории сооружений Донского государственного технического университета (344003, Российская Федерация, г. Ростов-на-Дону, пл. Гагарина, 1), ORCID, ResearcherID, ScopusID, anton_chepurnenk@mail.ru

Заявленный вклад соавторов:

Т.Н. Кондратьева: формирование основной концепции, цели и задачи исследования, проведение расчетов, анализ результатов исследований.

А.С. Чепурненко: научное руководство доработка текста, корректировка выводов.

Конфликт интересов: **авторы заявляют об отсутствии конфликта интересов.**

Все авторы прочитали и одобрили окончательный вариант рукописи.

Received / Поступила в редакцию 06.08.2025
Revised / Поступила после рецензирования 01.09.2025
Accepted / Принята к публикации 07.09.2025