Journal of Science and Technology in Civil Engineering, NUCE 2020. 14 (3): 1–14

A HYBRID MODEL FOR PREDICTING MISSILE IMPACT DAMAGES BASED ON K-NEAREST NEIGHBORS AND BAYESIAN OPTIMIZATION

Quoc Hoan Doan[a], Duc-Kien Thai[a,b,*], Ngoc Long Tran[b]

[a] *Department of Civil and Environmental Engineering, Sejong University, Gwangjin-gu, Seoul, South Korea*
[b] *Department of Civil Engineering, Vinh University, 82 Le Duan street, Vinh city, Nghe An, Vietnam*

Article history:
Received 11/05/2020, Revised 23/07/2020, Accepted 24/07/2020

Abstract

Due to the increase of missile performance, the safety design requirements of military and industrial reinforced concrete (RC) structures (i.e., bunkers, nuclear power plants, etc.) also increase. Estimating damage levels in the design stage becomes a crucial task and requires more accuracy. Thus, this study proposed a hybrid machine learning model which is based on k-nearest neighbors (KNN) and Bayesian optimization (BO), named as BO-KNN, for predicting the local damages of reinforced concrete (RC) panels under missile impact loading. In the proposed BO-KNN, the hyperparameters of the KNN were optimized by using the BO which is a well-established optimization algorithm. Accordingly, the KNN was trained on an experimental dataset that consists of 254 impact tests to predict four levels (or classes) of damages including *perforation, scabbing, penetration*, and *no damage*. Due to the unbalance of the number of tests in each damage class, an over-sampling technique called BorderlineSMOTE was employed as a balancing solution. The predictability of the proposed model was investigated by comparing with the benchmark models including non-optimized KNN, multilayer perceptron (MLP), and decision tree (DT). Accuracy, F1-score, and area under the receiver operating characteristic (ROC) curve (AUC) were utilized to evaluate the performance of these models. The implementation results showed that the proposed BO-KNN model outperformed the other benchmark models with the average class accuracy of 68.05%, F1-score = 0.641, and AUC = 85.8%. Thus, the proposed model can be introduced as a foundation for developing a tool for predicting the local damage of RC panels under the missile impact in the future.

Keywords: impact damage; k-nearest neighbors; Bayesian optimization; oversampling; imbalanced data; RC panel.

https://doi.org/10.31814/stce.nuce2020-14(3)-01

1. Introduction

In the practical design, a reinforced concrete (RC) structure is often locally damaged when subjects to a missile impact loading. Many levels of damage have been observed in the experiment [1, 2]. Among them *scabbing* and *perforation* damage are often used for the design limit state as required in the American Concrete Institute (ACI 349-01) [3]. Thus, the prediction of damage in the designing stage is a crucial task for a structure resisting missile impact.

In this study, a well-known supervised learning algorithm, namely k-nearest neighbors (KNN) was employed to build a classification model for predicting the local damages of RC panels under

*Corresponding author. *E-mail address:* thaiduckien@gmail.com (Thai, D.-K.)

missile impact loading. Its hyperparameters were optimized by using a Bayesian optimization (BO) method. This forms a hybrid model to predict the missile damage levels, called BO-KNN. Although the KNN classification algorithm has been widely used in the field of computer science or statistics [4–6], their application in the field of structural engineering still has a lot of potentials [7, 8]. Especially, its advantages have not yet been fully explored in the missile impact loading field. This study employed an extensive impact experiment database of RC panels adopted from the work of Thai et al. [9]. This dataset consists of 254 tests collected from the literary works with 17 input features. The dataset was divided into five folds using the cross-validation process which includes one testing set for performance evaluation and the remaining four folds for training and model selection. This process may help to generate more reliable results [10].

Four classes corresponding to four damage levels were classified which including *no damage, penetration, scabbing, perforation*. The number of instances in these classes had an imbalanced distribution. Classifying an imbalanced dataset may result in a biased prediction which mainly reflects the majority classes [11]. It is still a challenging research area [12]. Thus, in this study, a well-known effective oversampling technique called Borderline synthetic minority over-sampling (BorderlineSMOTE) was used to generate more data for the minority classes [13]. The oversample techniques help to balance the instances in the four damage classes. This contributed to improving the performance of the prediction models. The valid of the proposed BO-KNN model was investigated by comparing to the benchmark models including base KNN models (with and without oversampling technique), multilayer perceptron (MLP) model, and decision tree (DT) model. The prediction performances of the investigated models were evaluated by class accuracy, F1-score, Receiver Operating Characteristic (ROC) curve, and Area Under ROC curve (AUC) [14–16]. These evaluation metrics are helpful and needed to fully assess the multiclass imbalanced dataset classification problem [17, 18].

2. Research significance

Aforetime, local impact damages have primarily been measured using an experimental method [19, 20]. This is a basic and important approach to studying the conduct of new materials or systems under impact loads. In this method, the damage levels are explored by estimating penetration or perforation depth through several possible analytical and empirical formulations. Nevertheless, this method can not carry out a detailed parametric analysis due to the high experimental costs and time consumption [21].

To address these limitations, a significant number of computational analysis-based studies have been proposed [22–25] based on the reliable measurement capability of the numerical simulation software. One of the main benefits of this approach is that a more precise prediction form of penetration or depth of perforation can be considered for many other experimental parameters [26]. Nevertheless, if all experimental input parameters are taken into account, this method will face a challenge in terms of computational costs. Moreover, there is therefore still a weak generalization of the penetration depth prediction capacity of the proposed formulas.

To tackle these drawbacks, a data-driven approach which recently, has been successfully applied in the civil engineering field [27–30], was established that benefits from experimental data [9] to develop a prediction model based on machine learning (ML) algorithms. The learned model will identify the damages explicitly and take the effect of all experimental parameters into account. This approach has significantly saved more time than the parametric analysis in the simulation approach. However, the applications of ML in this filed is still inceptive. Significant works of validation and improvement on the effect of this approach are needed. One of the main factors that affect the performance of

ML models is are the model-controlled parameters which are also known as hyperparameters. Many hyperparameters optimization methods have been proposed [31–34]. Among them, the Bayesian optimization (BO) algorithm which has been presented as an effective algorithm in many practical fields [35–37]. However, as the authors' knowledge, the BO algorithm has never been explored in the field of impact damage prediction. Thus, this study contributes a method to improve the KNN model by optimizing its hyperparameters based on the Bayesian optimization algorithm.

3. Missile impact test and data pre-processing

3.1. Missile impact test description

In the experimental approach, many missile impact tests on RC panels/slabs/walls have been conducted to evaluate the local damages. Accordingly, the missiles can be shot into the RC panels from different angles, especially the perpendicular angle, which is a typical angle that was carried out in many works. This study also considered the impact tests based on this type of impact angle. The input features of an impact test are varied depending on the studying purposes. Typically, they include five groups: panel dimension, boundary condition, reinforcement, concrete properties, missile characteristics. By changing the parameter of these input features, we can investigate different behaviors or damage levels of the structure. The detailed features of a missile impact test are demonstrated in Fig. 1.

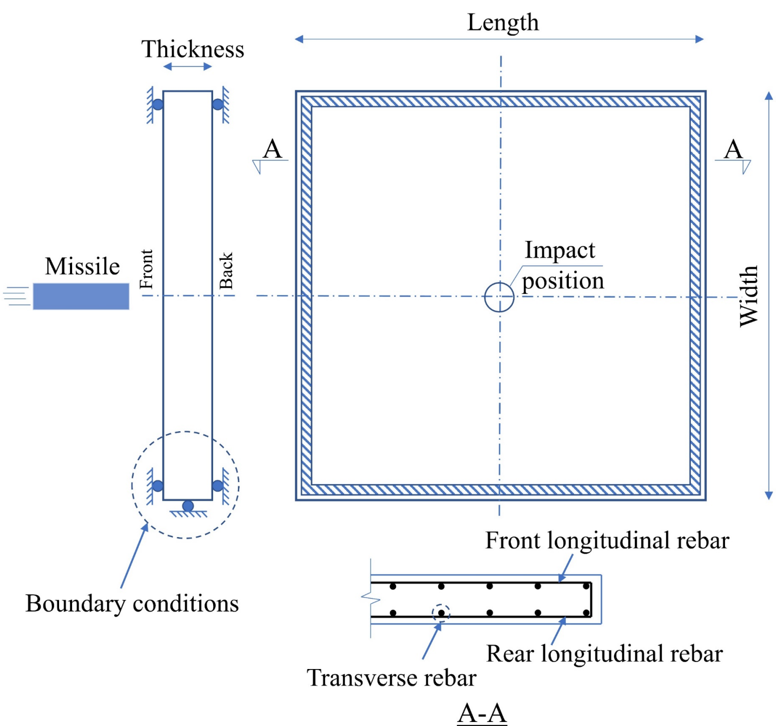
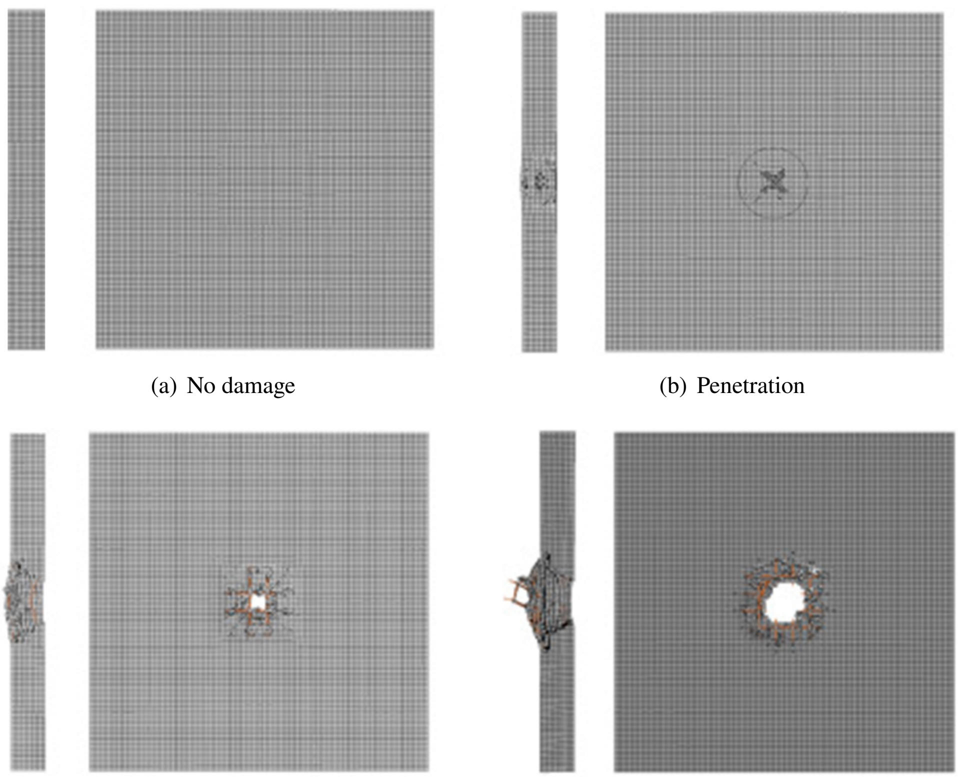


Figure 1. Description of RC panel features

When subjected to a missile impact loading, an RC panel can be damaged locally or globally. With a high striking velocity of the missile onto a large area of the target surface, the local damages are often observed. Thus, many studies focused on investigating the effect of local impact on an RC target [19, 38]. Different levels of damages have been observed and introduced such as *perforation, scabbing, radial cracking, spalling, cone cracking and plugging, penetration*, etc. [2]. Normally, in practical design, only four damage levels are considered as the design limit state, which include *perforation, scabbing, penetration*, and *no damage*. For instance, the American code for designing nuclear-safety concrete structures (ACI 349-01) [3] stated that the design limit state for a structure subjected to the missile impact loading should be *scabbing* or *perforation*. The demonstration of the four damage levels is presented in Fig. 2. Herein, the *perforation* damage is the worst case where the missile went through the RC target. In this study, the four damage levels were predicted by training the proposed BO-KNN model with a dataset of missile impact tests.



(a) No damage

(b) Penetration

(c) Scabbing

(d) Perforation

Figure 2. Missile damage levels

3.2. Data pre-processing

The data of missile impact tests on RC panels were collected from the literature from 1978 to 2017 [39–52]. It consisted of 254 instances classified into four output classes: *perforation*-126 instances, *scabbing*-69 instances, *penetration*-45 instances, *no damage*-14 instances. The input contained 17 features which include both numerical and categorical types. The categorical features were encoded into the digits. Then, all feature values were normalized into [0, 1] range for a proper training. The detail of the input features is presented in Table 1. A brief experimental dataset used is shown in Table 2.

Table 1. Description of the input and output features for the model training

	Description	Notation	Features	Data type*
Input	Length	L	x1	N
	Width	W	x2	N
	Thickness	H	x3	N
	Type of panel: One way (1), Two ways (2)	Ptype	x4	C
	Boundary condition: Connecting 4 corners (0.0), Clamping 4 edges (1.0)	BCtype	x5	C
	Pre-stress	Ptr	x6	N
	Strength of steel	Fs	x7	N
	Front longitudinal rebar ratio	FLr	x8	N
	Rear longitudinal rebar ratio	RLr	x9	N
	Transverse rebar ratio	TRr	x10	N
	Compressive strength	Fck	x11	N
	Tensile strength	Fts	x12	N
	Missile type : Soft missile (0.0), Hard missile (1.0)	Mtype	x13	C
	Missile diameter	Md	x14	N
	Missile mass	Mm	x15	N
	Missile nose type: Flat (0.72), Blunt (0.84), Spherical (1.00), Hollow/flat (1.03), Bi-conic (1.05), Ogival (1.10), Sharp (1.14)	Mntype	x16	C
	Impact velocity	Mv	x17	N
Output	Damage levels: No damage (0.0), Penetration (1.0), Scabbing (2.0), Perforation (3.0)		y1	C

*N: Numerical variable; C: Categorical variable.

Table 2. Brief experimental dataset

Parameters	Features	Unit	No. of specimens			
			1	2	3	4
L	x1	mm	2000	450	750	5400
W	x2	mm	2000	450	750	5400
H	x3	mm	250	60	120	700
Ptype	x4		2	2	1	2
BCtype	x5		1	1	0	1
Ptr	x6	MPa	10	4.09	0	0
Fs	x7	MPa	534	415	472	420
FLr	x8	%	0.35	0.00	0.24	0.39
RLr	x9	%	0.35	1.05	0.24	0.77
TRr	x10	%	1.396	0	0	0.25
Fck	x11	MPa	62.8	48.0	28.7	30.0
Fts	x12	MPa	3.7	3.6	2.5	2.2
Mtype	x13		1	1	1	0
Md	x14	mm	168	19	45	600
Mm	x15	kg	47.000	1.000	0.5	1016.0
Mntype	x16		0.84	1.1	1	0.84
Mv	x17	m/s	155.0	75.0	215.0	172.2
Damage levels	y1		Perforation	Penetration	Scabbing	No damage

4. Methodology

4.1. k-nearest neighbors algorithm (KNN)

The KNN model is known as a non-parametric approach. It calculates the distances of k nearest existing instances to the new instance, then classifies it into a class that most frequently appear among k instances. According to this classifying mechanism, the KNN algorithm can be easily applied for the multiclass problem as presented in this study. The main advantages of the KNN algorithm are useful for nonlinear data and simple to implement or interpret [53]. However, it can be computationally expensive when the number of instances is big. Because the algorithm has to store all the training instances and use them for the testing stage. In this study, the total number of instances is 254, thus the training time was not a significant issue. Another obvious drawback of the KNN algorithm is its sensitivity to a skewed dataset. It tends to predict a new instance according to the voting of the majority class. Thus, the obtained results can be overoptimistic [54]. The performance of the KNN algorithm mainly depends on two hyperparameters including the number of nearest neighbors k and the distance calculating function. Therefore, to find the optimal values of these hyperparameters for the KNN model, the BO method was employed.

4.2. Bayesian optimization (BO)

Bayesian optimization [55] a well-known method in the practical machine learning field, which has been primarily used for tuning the hyperparameters of the machine learning models. BO is known as a sequential model-based approach to solving the problem of finding global extrema of an unknown function $f(x)$ on some bounded domain χ.

$$x^* = \arg\max_{x\in\chi} f(x) \tag{1}$$

BO typically works by constructing a probabilistic surrogate model of $f(x)$ which contains a prior distribution that simulates the behavior of $f(x)$. Then the uncertainty of the potential values of the surrogate model is used to produce an acquisition function $a(x)$. The next examined point x_t is determined by optimizing the $a(x)$ function $x_t = \arg\max_x a(x)$. After that, the performance of the $f(x)$ function is evaluated with the updated hyperparameter x_t. The process is then repeated until obtaining the best hyperparameter.

In this study, the Gaussian process (GP) was selected as the surrogate model due to its powerful prior distribution and flexibility. The GP is defined by the property that any finite set of N points $\{x_i \in \chi\}_{i=1}^N$ induces a multivariate Gaussian distribution on $\mathbb{R}^N$. It is characterized by a mean $\mu(x)$ and a variance $\sigma^2(x)$.

Regarding the acquisition function, in general, it depends on the previous observation and the GP hyperparameters. There are different popular choices of acquisition function such as probability of improvement, expected improvement (EI), upper confidence bounds (UCB), etc. This work focused on the EI function due to its good performance in minimization problems and no requirement of tuning its own parameters. The EI function can be expressed as follows:

$$\begin{aligned} a(x) = EI(x) &= \begin{cases} (\mu(x) - f(\hat{x}))\Phi(Z) + \sigma(x)\phi(Z), & \text{if } \sigma(x) > 0 \\ 0, & \text{if } \sigma(x) = 0 \end{cases} \\ \text{with } Z &= \frac{\mu(x) - f(\hat{x})}{\sigma(x)}, \end{aligned} \tag{2}$$

where $\hat{x}$ is the best hyperparameter observed so far; $\Phi(.)$ and $\phi(.)$ are the cumulative distribution function and probability density function of a standard Gaussian distribution. The EI includes two terms when $\sigma(x) > 0$ that can be interpreted as a tradeoff between exploitation of known optimal areas and exploration of unexplored areas of the objective function.

4.3. BorderlineSMOTE-An oversampling technique

Due to the imbalance of the dataset, a well-established oversampling technique called BorderlineSMOTE was adopted and employed [13]. This technique works as a data generator based on the Synthetic minority over-sampling technique (SMOTE) |cite 56. Since the instances near the borderline (where the instances of a class are close to other class ones) are more prone to be misclassified than the ones far from the borderline. Thus, these instances have higher weight and need to spend more attention. Accordingly, the minority class that is near the borderline is over-sampled based on the data sampling mechanism of SMOTE.

In this work, the dataset was divided into five folds by using the k-fold cross-validation procedure. Among them, one fold was held out for testing and the remaining folds were used for training. To prevent the overoptimistic problem [56], BorderlineSMOTE was employed inside the cross-validation loop. All classes were over-sampled excluding the majority class, here, the perforation class. In particular, the no damage class, penetration class, and scabbing class were oversampled up to 100 instances from 11 instances, 36 instances, and 55 instances, respectively.

4.4. The proposed BO-KNN model

In the present study, the local impact damages were predicted primarily based on the KNN model. Two main hyperparameters including the number of neighbors k and the distance metric functions often have a significant effect on the performance of the KNN model. Thus, Bayesian optimization was employed to determine the best value of these hyperparameters which are then used to construct the final model for missile impact damage prediction, called the BO-KNN model. Three popular distance metric functions including Euclidean, Manhattan, and Minkowski were used to measure the distance between an unknown instance and its k-nearest neighbors. Their mathematical formulation can be expressed as follows:

$$\text{Euclidean distance: } d_{ED}(x,y) = \sqrt{\sum_{i=1}^{m} |x_i - y_i|^2} \tag{3}$$

$$\text{Manhattan distance: } d_{MD}(x,y) = \sum_{i=1}^{m} |x_i - y_i| \tag{4}$$

$$\text{Minkowski distance: } d_{MK}(x,y) = \sqrt[p]{\sum_{i=1}^{m} |x_i - y_i|^p} \tag{5}$$

where m is the number of calculated points; p is a positive value. As can be seen, when $p = 1$, the Minkowski distance becomes Manhattan distance, and when $p = 2$, it becomes Euclidean distance. Thus, p now becomes an alternative hyperparameter that needs to optimize. The procedure of the proposed BO-KNN is accomplished through six steps as shown in Fig. 3.

+ Step 1: Preparing the missile impact dataset

In this step, the dataset was collected and pre-processed according to the method presented in the "Data pre-processing" section.

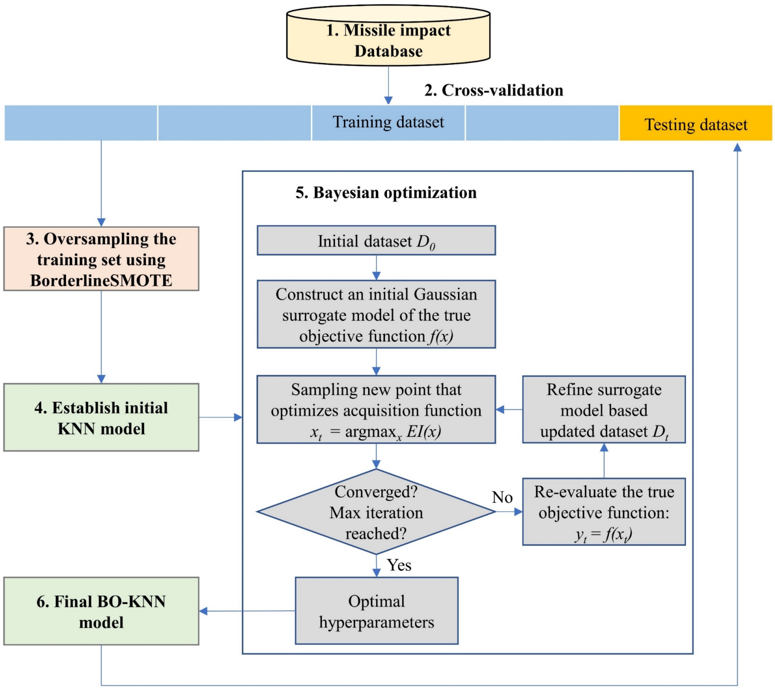


Figure 3. Scheme of the proposed BO-KNN model for predicting missile impact damage

+ Step 2: Splitting the dataset using the k-fold cross-validation method.

The data was divided into five stratified folds which include one testing fold and four training folds. In this cross-validation process, the BO-KNN model was independently trained and tested five times. The testing fold was in-turn replaced by another fold after each iteration. The results will be the mean of the testing results over five times. With an imbalanced dataset in this study, the cross-validation process helped to reduce bias and overfitting problems.

+ Step 3: Oversampling the training folds

In this step, the number of instances in each class was balanced using the BorderlineSMOTE method. New synthetic data points were generated based on the relation between the existing ones.

+ Step 4: Establishing the initial KNN algorithm as a based model.

+ Step 5: Bayesian optimization

This step included the optimization procedure for the two hyperparameters k and p using BO. The search space for k and p were [7, 51] and [1, 11], respectively. These search spaces were selected after implementing some first optimization procedure to investigate the possible range of the hyperparameters. It should be noted that due to the use of the cross-validation process, five optimal hyperparameter sets can be achieved. However, only the dominant one was selected for constructing the final model. Due to the imbalance of the dataset, the objective function was set to the maximization of F1-score instead of minimization of loss which often causes bias toward the majority class.

+ Step 6: Constructing the final BO-KNN model using the obtained optimal hyperparameters. Then the final model was tested on the holdout testing fold.

After that, the procedure was repeated from Step 2 where another train-test set is generated by the cross-validation process. The entire procedure of the proposed model was implemented using Python language.

5. Results and discussion

In the present section, the results of the missile damage prediction models were highlighted. The proposed BO-KNN model was compared to the benchmark models including a non-optimized KNN model or Base KNN model, multilayer perceptron (MLP) model, and decision tree (DT) model. The base KNN model was investigated which includes and not include the oversampling technique. All the hyperparameters selected for the above models were carefully selected to avoid the overfitting problem. In the case of the KNN model, the overfitting problem can occur when using a too-small number of neighbors k. Because the model can over-optimistically classify the damages when considers only a few neighbors at a time. Thus, the searching range of the number of neighbors k for the BO was set in the range of [7, 51]. Besides, the cross-validation process was applied during the optimization to avoid the overfitting problem [57]. In the case of other models, the hyperparameters which were found by a trial-and-error process were selected so that the training and validation errors are closed to each other.

Accordingly, the base KNN model had *k_neighbors* = 11 and $p = 1$. Multilayer perceptron model was configured with *number_of_hidden_layer* = 1, *number_of_neurons* = 100, *l2_regularization* = 0.001, *batch_size* = 16, *learning_rate* = 0.001. In this model, the early-stopping criterion was applied to prevent the overfitting problem. In which, the learning process will be terminated when the validation error starts to increase while the training error is decreasing. This technique helps to constrain the training and validation error to be closed to each other, thus prevent the overfitting problem. In Decision tree model, we set *max_depth* = 3, *criterion* = 'entropy', *min_samples_split* = 0.3. All the hyperparameters were obtained that produced the best performance in each model. For the proposed BO-KNN model, after optimizing using the BO method, the optimal hyperparameters were *k_neighbors* = 9 and $p = 3$ (in Minkowski distance metric).

Table 3. Results of the missile damage prediction models in terms of AUC

Missile damage prediction models	AUC (%)
Base KNN model (without oversampling)	84.7 ± 1.8
Base KNN model (with oversampling)	85.1 ± 2.5
Multilayer perceptron (MLP)	80.4 ± 2.1
Decision Tree (DT)	81.2 ± 2.6
The proposed BO-KNN model	85.8 ± 3.3

The obtained results in terms of AUC were presented in Table 3. The receiver operating characteristic (ROC) curve which is based on the True positive rate (FPR) and False positive rate (FPR) was presented in Fig. 4. The ROC curve indicates the classifying capability of the models between the classes which is a well-known evaluation metric for the multiclass problem. As can be seen, the ROC curve of the proposed BO-KNN model covered most of the other model one. The quantitative evaluation of these curves was presented through the area under the curve, namely AUC. The AUC represents the aggregate measurement of separability and performance of the model across all classification thresholds. The higher AUC introduces a better prediction capacity model. The obtained values of the AUC of each model were shown in Table 3. As can be seen, the proposed BO-KNN model had the highest AUC value with 85.8%. While the base KNN model with and without oversampling had the lower AUC value with 85.1% and 84.7%, respectively. The AUC of the BO-KNN model was also higher than the other benchmark models including MLP model (AUC = 80.4%) and DT model (AUC = 81.2%). These results showed a higher prediction capacity of the BO-KNN model over other investigated models for missile impact damages.

Besides, the mean F1-score results over the five separated folds were demonstrated in Fig. 5. It can be observed from the graph that, the F1-score of the BO-KNN model was higher than other

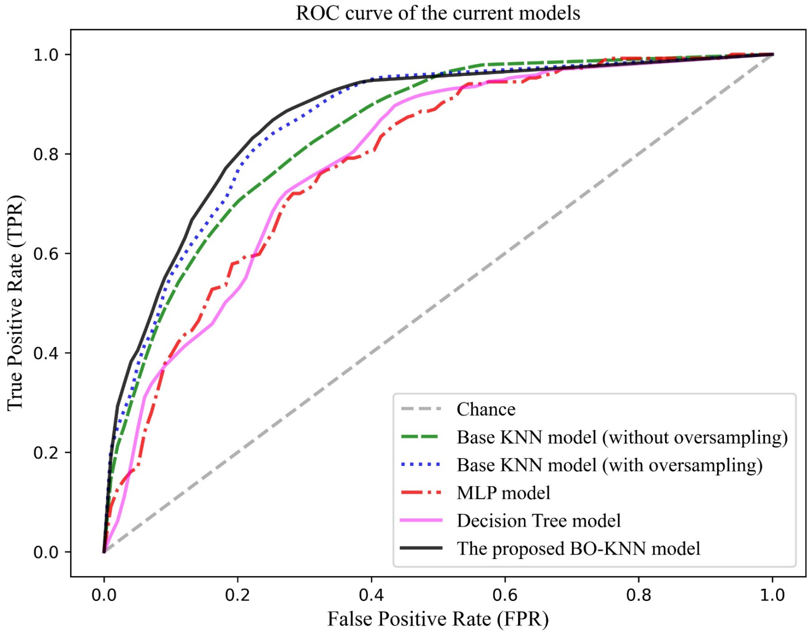


Figure 4. ROC curves of the investigating models

models. In particular, the BO-KNN model obtained F1-score = 0.641, while the base KNN model with and without oversampling, MLP model, and DT model achieved F1-score that equal to 0.586, 0.614, 0.547, and 0.551, respectively. This result enforced the prediction capacity of the BO-KNN model on the missile impact dataset.

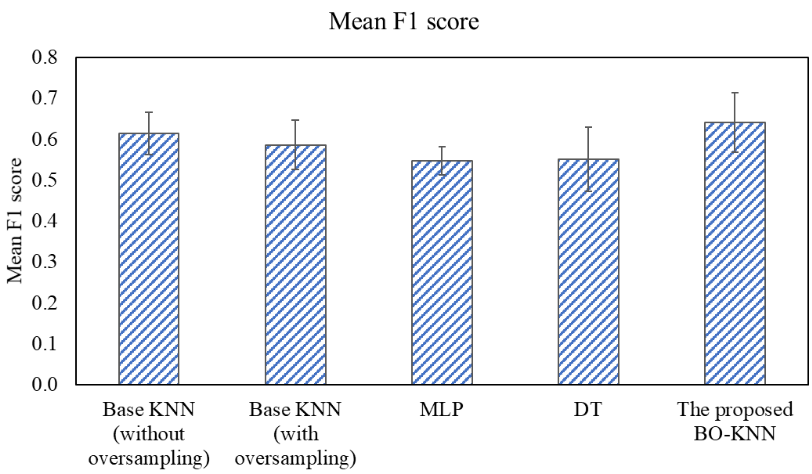


Figure 5. Mean F1-score results of the investigating models

The detailed accuracy corresponding to each damage level was presented in Fig. 6. The average class accuracy of the BO-KNN model was 68.05%. The one obtained from the base KNN model with and without oversampling was 63.4% and 56.45%, respectively. MLP model achieved 56.3% and DT model one was 57.65%. As can been seen, the base KNN model with oversampling biasedly predict the no damage class. This is because the oversampling technique generates many synthetic data samples in only a small region of this class. So in the test phase, when a new sample is in or near this region, the majority number of samples in this class is covered when considering the k near nearest neighbors. Although the accuracy of the perforation and no damage class of the BO-KNN model was smaller than other models, its mean accuracy was still higher. Moreover, in general, the accuracy of the four damage levels in the proposed model was more balanced than the others. It is helpful when predicting the unseen design input features in the future.

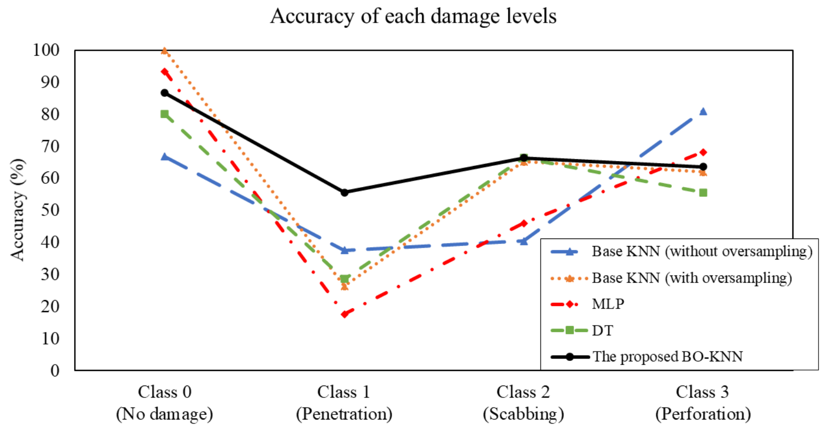


Figure 6. Accuracy of each damage level obtained by the investigating models

6. Conclusions

This study proposed a new hybrid machine learning-based model for predicting the local damages of RC panels under missile impact loading. The proposed model was constructed according to the *k*-nearest neighbor (KNN) and Bayesian optimization (BO), namely BO-KNN, with high prediction performance. The outcomes of this work are as follows:

- The proposed BO-KNN model obtained a high AUC value with 85.8% that outperforms other benchmark models including base KNN model (with and without oversampling), multilayer perceptron (MLP), and decision tree (DT).

- The BO method well contributed to finding the best hyperparameters for the KNN model that achieved a higher damage prediction capacity in terms of F1-score (0.641) and average class accuracy (68.05%).

- The BO-KNN model can be used as a tool in the practical design for predicting the local damage levels, especially in the initial design stage.

The study was limited to consider only two hyperparameters of the KNN model and a few types of distance metrics. The performance of this model can be enhanced if other hyperparameters or more robust oversampling techniques are considered.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2018R1C1B5086385).

References

[1] Hashimoto, J., Takiguchi, K., Nishimura, K., Matsuzawa, K., Tsutsui, M., Ohashi, Y., Kojima, I., Torita, H. (2005). *Experimental study on behavior of RC panels covered with steel plates subjected to missile impact*.

[2] Li, Q. M., Reid, S. R., Wen, H. M., Telford, A. R. (2005). Local impact effects of hard missiles on concrete targets. *International Journal of Impact Engineering*, 32(1-4):224–284.

[3] American Concrete Institute (2001). *Code Requirements of Nuclear Safety Related Concrete Structures (ACI 349-01)*.

[4] Goldberg, D. E., Holland, J. H. (1988). Genetic algorithms and machine learning. *Machine Learning*, 3 (2/3):95–99.

[5] Friedman, J., Hastie, T., Tibshirani, R. (2001). *The elements of statistical learning*. Springer series in statistics New York.

[6] Biship, C. M. (2006). *Pattern recognition and machine learning (information science and statistics)*. Springer-Verlag, Berlin, Heidelberg.

[7] Wang, X., Guan, S., Hua, L., Wang, B., He, X. (2019). Classification of spot-welded joint strength using ultrasonic signal time-frequency features and PSO-SVM method. *Ultrasonics*, 91:161–169.

[8] Akpinar, P., Khashman, A. (2017). Intelligent classification system for concrete compressive strength. *Procedia Computer Science*, 120:712–718.

[9] Thai, D.-K., Tu, T. M., Bui, T. Q., Bui, T.-T. (2019). Gradient tree boosting machine learning on predicting the failure modes of the RC panels under impact loads. *Engineering with Computers*, 1–12.

[10] Yang, T., Vojislav, K., Longbing, C., Chengqi, Z. (2010). Testing adaptive local hyperplane for multi-class classification by double cross-validation. In *The 2010 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 1–5.

[11] Kotsiantis, S., Kanellopoulos, D., Pintelas, P. et al. (2006). Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering*, 30(1):25–36.

[12] Ramyachitra, D., Manikandan, P. (2014). Imbalanced dataset classification and solutions: a review. *International Journal of Computing and Business Research (IJCBR)*, 5(4).

[13] Han, H., Wang, W.-Y., Mao, B.-H. (2005). Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. In *International Conference on Intelligent Computing*, Springer, 878–887.

[14] Ghamrawi, N., McCallum, A. (2005). Collective multi-label classification. In *Proceedings of the 14th ACM international conference on Information and knowledge management*, 195–200.

[15] Huang, J., Ling, C. X. (2005). Using AUC and accuracy in evaluating learning algorithms. *Ieee Transactions on Knowledge and Data Engineering*, 17(3):299–310.

[16] Fan, J., Upadhye, S., Worster, A. (2006). Understanding receiver operating characteristic (ROC) curves. *Canadian Journal of Emergency Medicine*, 8(1):19–20.

[17] Menardi, G., Torelli, N. (2014). Training and assessing classification rules with imbalanced data. *Data Mining and Knowledge Discovery*, 28(1):92–122.

[18] Hossin, M., Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, 5(2):1.

[19] Kojima, I. (1991). An experimental study on local behavior of reinforced concrete slabs to missile impact. *Nuclear Engineering and Design*, 130(2):121–132.

[20] Haldar, A., Miller, F. J. (1982). Penetration depth in concrete for nondeformable missiles. *Nuclear Engineering and design*, 71(1):79–88.

[21] Riera, J. D. (1989). Penetration, scabbing and perforation of concrete structures hit by solid missiles. *Nuclear engineering and design*, 115(1):121–131.

[22] Thai, D.-K., Kim, S.-E. (2015). Numerical simulation of reinforced concrete slabs under missile impact. *Structural Engineering and Mechanics*, 53(3):455–479.

[23] Thai, D.-K., Kim, S.-E. (2014). Failure analysis of reinforced concrete walls under impact loading using the finite element approach. *Engineering Failure Analysis*, 45:252–277.

[24] Kong, X., Fang, Q., Wu, H., Peng, Y. (2016). Numerical predictions of cratering and scabbing in con-

crete slabs subjected to projectile impact using a modified version of HJC material model. *International Journal of Impact Engineering*, 95:61–71.

[25] Liu, Y., Ma, A., Huang, F. (2009). Numerical simulations of oblique-angle penetration by deformable projectiles into concrete targets. *International Journal of Impact Engineering*, 36(3):438–446.

[26] Thai, D.-K., Kim, S.-E. (2016). Prediction of UHPFRC panels thickness subjected to aircraft engine impact. *Case Studies in Structural Engineering*, 5:38–53.

[27] Tung, P. T., Hung, P. T. (2020). Predicting fire resistance ratings of timber structures using artificial neural networks. *Journal of Science and Technology in Civil Engineering (STCE)-NUCE*, 14(2):28–39.

[28] Nguyen, T. T., Dinh, K. (2019). Prediction of bridge deck condition rating based on artificial neural networks. *Journal of Science and Technology in Civil Engineering (STCE)-NUCE*, 13(3):15–25.

[29] Hung, T. V., Viet, V. Q., Thuat, D. V. (2019). A deep learning-based procedure for estimation of ultimate load carrying of steel trusses using advanced analysis. *Journal of Science and Technology in Civil Engineering (STCE)-NUCE*, 13(3):113–123.

[30] Hung, D. V., Hung, H. M., Anh, P. H., Thang, N. T. (2020). Structural damage detection using hybrid deep learning algorithm. *Journal of Science and Technology in Civil Engineering (STCE)-NUCE*, 14(2): 53–64.

[31] Palar, P. S., Zuhal, L. R., Shimoyama, K. (2019). On the use of metaheuristics in hyperparameters optimization of gaussian processes. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 263–264.

[32] Pedregosa, F. (2016). Hyperparameter optimization with approximate gradient. *arXiv preprint arXiv:1602.02355*.

[33] Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., Talwalkar, A. (2017). Hyperband: A novel bandit-based approach to hyperparameter optimization. *The Journal of Machine Learning Research*, 18(1): 6765–6816.

[34] Yu, T., Zhu, H. (2020). Hyper-parameter optimization: A review of algorithms and applications. *arXiv preprint arXiv:2003.05689*.

[35] Cheng, H., Ding, X., Zhou, W., Ding, R. (2019). A hybrid electricity price forecasting model with Bayesian optimization for German energy exchange. *International Journal of Electrical Power & Energy Systems*, 110:653–666.

[36] Wang, Y., Kandeal, A., Swidan, A., Sharshir, S. W., Abdelaziz, G. B., Halim, M., Kabeel, A., Yang, N. (2020). Prediction of tubular solar still performance by machine learning integrated with Bayesian optimization algorithm. *arXiv preprint arXiv:2002.03886*.

[37] Sun, D., Wen, H., Wang, D., Xu, J. (2020). A random forest model of landslide susceptibility mapping based on hyperparameter optimization using Bayes algorithm. *Geomorphology*, 362:107201.

[38] Kosteski, L. E., Riera, J. D., Iturrioz, I., Singh, R. K., Kant, T. (2015). Assessment of empirical formulas for prediction of the effects of projectile impact on concrete structures. *Fatigue & Fracture of Engineering Materials & Structures*, 38(8):948–959.

[39] Orbovic, N., Elgohary, M., Lee, N., Blahoianu, A. (2009). Tests on reinforced concrete slabs with pre-stressing and with transverse reinforcement under impact loading. In *20th International Conference on Structural Mechanics in Reactor Technology (SMiRT 20)*, IASMiRT.

[40] Rajput, A., Iqbal, M. A. (2017). Ballistic performance of plain, reinforced and pre-stressed concrete slabs under normal impact by an ogival-nosed projectile. *International Journal of Impact Engineering*, 110: 57–71.

[41] Vepsä, A., Saarenheimo, A., Tarallo, F., Rambach, J., Orbovic, N. (2011). IRIS_2010–Part II: Experimental data. *Transactions, SMiRT*, 21:6–11.

[42] Almusallam, T. H., Siddiqui, N. A., Iqbal, R. A., Abbas, H. (2013). Response of hybrid-fiber reinforced concrete slabs to hard projectile impact. *International Journal of Impact Engineering*, 58:17–30.

[43] Abdel-Kader, M., Fouda, A. (2014). Effect of reinforcement on the response of concrete panels to impact of hard projectiles. *International Journal of Impact Engineering*, 63:1–17.

[44] Orbovic, N., Sagals, G., Blahoianu, A. (2015). Influence of transverse reinforcement on perforation resistance of reinforced concrete slabs under hard missile impact. *Nuclear Engineering and Design*, 295:

716–729.

[45] Stephenson, A. E., Sliter, G. E., Burdette, E. G. (1978). Full-scale tornado-missile impact tests. *Nuclear Engineering and Design*, 46(1):123–143.

[46] Nachtsheim, W., Stangenberg, F. (1983). Interpretation of results of Meppen slab tests—comparison with parametric investigations. *Nuclear Engineering and Design*, 75(2):283–290.

[47] Hanchak, S. J., Forrestal, M. J., Young, E. R., Ehrgott, J. Q. (1992). Perforation of concrete slabs with 48 MPa (7 ksi) and 140 MPa (20 ksi) unconfined compressive strengths. *International Journal of Impact Engineering*, 12(1):1–7.

[48] Sugano, T., Tsubota, H., Kasai, Y., Koshika, N., Itoh, C., Shirai, K., Von Riesemann, W. A., Bickel, D. C., Parks, M. B. (1993). Local damage to reinforced concrete structures caused by impact of aircraft engine missiles Part 2. Evaluation of test results. *Nuclear Engineering and Design*, 140(3):407–423.

[49] Zhang, M. H., Shim, V. P. W., Lu, G., Chew, C. W. (2005). Resistance of high-strength concrete to projectile impact. *International Journal of Impact Engineering*, 31(7):825–841.

[50] Dancygier, A. N., Yankelevsky, D. Z., Jaegermann, C. (2007). Response of high performance concrete plates to impact of non-deforming projectiles. *International Journal of Impact Engineering*, 34(11): 1768–1779.

[51] Orbovic, N., Blahoianu, A. (2011). Tests on concrete slabs under hard missile impact to evaluate the influence of transverse reinforcement and pre-stressing on perforation velocity. *SMiRT_21 Proceedings, New Delhi, India, SMiRT-21*.

[52] Pires, J. A., Ali, S. A., Candra, H. (2011). Finite element simulation of hard missile impacts on reinforced concrete slabs. *Transactions of the 21st SMiRT, paper ID*, 777.

[53] Cunningham, P., Delany, S. J. (2020). k-Nearest neighbour classifiers. *arXiv preprint arXiv:2004.04523*.

[54] Saini, I., Singh, D., Khosla, A. (2013). QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases. *Journal of Advanced Research*, 4(4):331–344.

[55] Pelikan, M., Goldberg, D. E., Cantú-Paz, E. et al. (1999). BOA: The Bayesian optimization algorithm. In *Proceedings of the genetic and evolutionary computation conference GECCO-99*, volume 1, Citeseer, 525–532.

[56] Santos, M. S., Soares, J. P., Abreu, P. H., Araujo, H., Santos, J. (2018). Cross-validation for imbalanced datasets: Avoiding overoptimistic and overfitting approaches [research frontier]. *IEEE Computational Intelligence Magazine*, 13(4):59–76.

[57] Moore, A. W. (2001). Cross-validation for detecting and preventing overfitting. *School of Computer Science Carneigie Mellon University*.