Advances in Engineering Software 184 (2023) 103502

Contents lists available at ScienceDirect

Advances in Engineering Software

journal homepage: www.elsevier.com/locate/advengsoft

Research paper

Speed estimation of a car at impact with a W-beam guardrail using numerical simulations and machine learning

Dawid Bruski [a,*], Lukasz Pachocki [a], Adam Sciegaj [a,b], Wojciech Witkowski [a]

[a] Department of Mechanics of Materials and Structures, Faculty of Civil and Environmental Engineering, Gdańsk University of Technology, Gdańsk, Poland
[b] EkoTech Center, Gdańsk University of Technology, Gdańsk, Poland

ARTICLE INFO

Keywords:
Road traffic safety
Numerical modeling
Crash tests
Accident
Machine learning
Intelligent transportation systems

ABSTRACT

This paper aimed at developing a new method of estimating the impact speed of a passenger car at the moment of a crash into a W-beam road safety barrier. The determination of such a speed based on the accident outcomes is demanding, because often there is no access to full accident data. However, accurate determination of the impact speed is one of the key elements in the reconstruction of road accidents. A machine learning algorithm was used to create the speed estimation model. The model was based on regression trees algorithms, with base regressors forming a final voting ensemble. The model was trained, validated, and tested using a database containing results from full-scale crash tests and numerical simulations. The developed machine learning model had a mean absolute error of 6.76 km/h with a standard deviation of 1.01 km/h on the cross-validation set, and a coefficient of determination, R^2, of 0.85. This model was used to estimate the impact speed of the vehicle in three real road accidents with the W-beam barrier, and then the determined speeds were used in additional simulations to verify the results. A good quantitative and qualitative agreement between the simulation and accident outcomes was achieved, and this confirmed that the proposed method and the developed ML models combined with numerical simulations and full-scale crash tests can be effective tools for estimating the speed of the vehicle at impact with a roadside barrier.

1. Introduction

Road traffic safety is one of the most important aspects of social life. In Poland, 2245 deaths and 26 415 severe injury cases stemmed from 22 816 road accidents in 2021 [1]. Even though the death rate on roads in Poland follows a downward trend, it is a challenge to ensure its further drop.

One way to reduce the likelihood and mitigate the effects of severe accidents is the use of road safety barriers (RSBs) [2]. There are many types of barriers, the most popular of which are W-beam guardrails [3,4], cable barriers [5,6], and concrete barriers [7,8] . Nevertheless, accidents usually involve damage to the barrier and the vehicle. The extent of the damage may depend on many factors, among which the impact energy is one of the most important [9,10]. This energy depends on the mass of the vehicle and its ballast, impact angle, and vehicle speed. The mass and the impact angle can usually be easily assessed after the crash. On the other hand, a reliable estimation of the impact speed still poses a challenge in the scientific and expert community [11]. The knowledge of the impact speed is crucial to determine the energy of the impacting car. This can improve the design and maintenance of roadside barriers, and allows for the proper choice of a barrier type and design depending on local road and traffic conditions [12]. It enables also verification of the impact speed requirements set by the EN1317 standard for design and testing of barriers. Determining the impact speed is also an important element in road accident reconstruction. Lastly, the knowledge of the impact speed might help the committees investigating road accidents in the determination of the causes of the accidents and the corresponding liabilities.

Numerical analyses using the Finite Element Method (FEM) are widely used to improve road safety [13–16]. To ensure the credibility of simulation results, it is necessary to link the results to the experiments and validate the model. The examples of the validation process of a crash test with a roadside barriers can be found in [17–21]. Numerical simulations can also be used to predict the speed of the vehicle impact. For instance, the impact speed estimation method based on crack patterns on the windshield due to the pedestrian head impact is presented in [22].

Machine Learning (ML) is another method that attracts interests of both researchers and practitioners. ML algorithms have already been successfully used in materials science [23] and intelligent transportation systems [24]. For instance, Ji et al. [25] used the Decision Tree

* Corresponding author.
E-mail address: dawbrusk@pg.edu.pl (D. Bruski).

https://doi.org/10.1016/j.advengsoft.2023.103502
Received 27 March 2023; Received in revised form 9 May 2023; Accepted 2 June 2023
Available online 14 June 2023

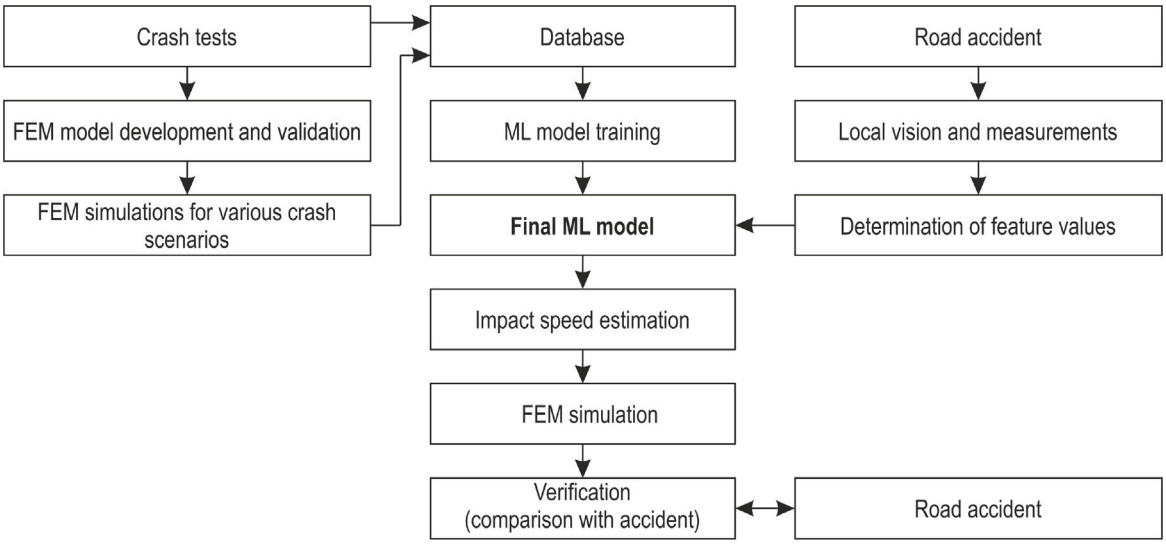

Fig. 1. Flowchart of the study.

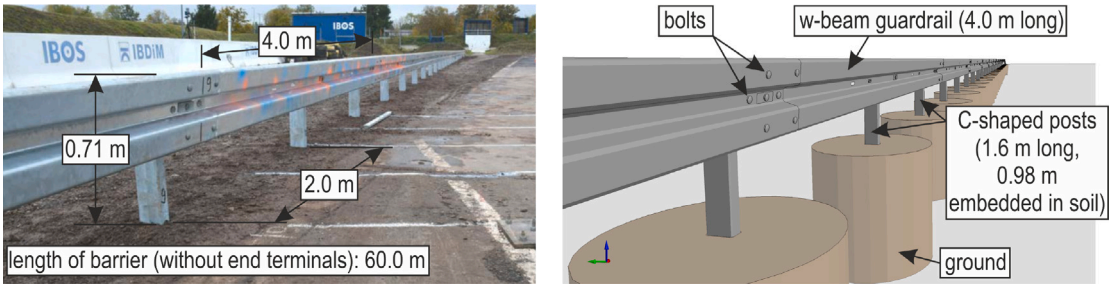

Fig. 2. Comparison between the N2/W4/A road safety barrier and its corresponding numerical model.

algorithm to predict the speeding of the vehicle based on the data acquired from 903 traffic accidents. However, successful performance of those methods is contingent on being able to provide large amounts of data for the algorithms to be trained on. As access to traffic accident data is limited, the use of numerical simulations is additionally beneficial.

A literature review on vehicle-road safety equipment accident reconstruction indicates that the number of developed methods for impact speed determination is still insufficient. There are several methods for road accident reconstruction, including numerical simulation, video recordings, and crash data retrieval [26]. Coon et al. [9] presented a technique for determining the impact speed for accidents with roadside barriers. The impact energy and impact speed were estimated by adding the energy dissipated during an impact to the kinetic energy of the vehicle at the departure from the barrier. The method was validated against a full-scale crash test. Similarly, a procedure for determining the initial speed of a vehicle crashing a guardrail end terminal was proposed [11]. This method is based on the conservation of momentum and energy. The developed procedure was verified in comparison with full-scale crash tests. On another note, a reconstruction procedure for estimation of the absorbed energy and impact vehicle velocities in crashes involving cable barriers was studied in [10]. The method provides charts determined on the basis of crash tests and component tests and was verified in three full-scale crash tests. A two-dimensional viscoelastic model for reconstruction of a vehicle crash with a roadside barrier was developed in [27]. The results obtained from the model were compared with results from a full-scale crash test. The authors also proposed suggestions on how to further develop this model to make it more accurate and general. However, those require nontrivial computations, e.g., related to the dissipated energy during an accident. Hence, the development of a technique that would quickly give information about the impact speed and impact energy of the vehicle with a minimum amount of data from a crash could be helpful in the process of reconstructing road accidents.

This study presents a new method for the estimation of the speed of a car at the moment of an impact with W-beam guardrail, which is the key to the crash reconstruction process. The proposed method utilizes machine learning and non-linear FEM simulations, allowing for a quick and immediate estimation of the impact speed. Further advantage of the developed model is that it needs only a few simple input features, which can be easily determined from the accident scene.

The flowchart of this study is presented in Fig. 1. First, a full-scale crash test was conducted and the corresponding numerical model was developed. Afterwards, the numerical model was validated against the results from the crash test. Subsequently, a series of simulations with various initial conditions were carried out using the validated model, and the results from the crash tests and simulations were used to create a database that was used for training, validation, and testing of ML models. After the model was trained, it was used to estimate the impact speed for 3 real-world road accidents with a W-beam guardrail. Finally, the predictions were verified in additional simulations, where the results and views of the damaged barrier from the simulation were compared with the results and views from the real accidents.

The paper is structured as follows: Section 2 describes the W-beam guardrail, its FEM model, and the numerical model of the car. The validation of the FEM model against the crash test is also presented. The result database and input features that were used to train the ML models are described in Section 3. Section 4 gives an overview of the developed ML models for impact speed estimation and presents selected performance metrics of the models. The ML model is used to predict

the speed in three different real-world accidents in Section 5, and the results are subsequently verified with pertinent numerical simulations. The paper ends with Section 6, which contains conclusions and final remarks.

2. Numerical model

2.1. Numerical model description

2.1.1. Road safety barrier

The road safety barrier in the current study was certified according to EN1317 [28] standard for the classes N2/W4/A and H1/W4/A. The system comprised steel W-beam guardrails, posts, spacers, and bolts. Each segment of the barrier was 4.0 m long, and the spacing between the posts was 2.0 m. Steel posts were 1.6 m long and they were fixed 0.98 m into the ground. The total height of the system was 0.71 m. The total length of the straight/horizontal section of the system was 60.0 m. The corresponding FEM model geometry was recreated based on the available technical drawings. W-beams, posts, and spacers were made of S235JR steel and were modeled using 4-node fully-integrated shell elements with piecewise linear plasticity material law and strain rate effects. The rate effects were modeled using Cowper–Symonds rate term [29,30] with $C = 40.4$ and $p = 5.0$. The bolts in the system were of grades 4.6 and 8.8 [31]. They were pretensioned to the force of 9.9 kN and modeled using 8-node constant-stress solid elements with piecewise linear plasticity material law. The material properties of steel were obtained from quasi-static tensile tests conducted on specimens cut from a section of the considered system. The ground around the posts was modeled as cylinders using 8-node constant-stress solid elements with the material model and properties according to [6,32,33]. The outer surface of the cylinders was fixed in all directions. The approximate element sizes in the model were as follows: W-beams, posts, and spacers — from 10 to 15 mm, bolts — from 6 to 10 mm, and soil cylinders — from 10 to 60 mm. The complete model of the system comprised 449 700 nodes and 419 157 finite elements. The view of the safety system and its corresponding numerical model is presented in Fig. 2.

2.1.2. Vehicle

A 1500 kg BMW E34 520 was used in this study. The selected vehicle fulfilled the requirements of the TB32 crash test [28]. The numerical model of the vehicle was initially developed by Transpolis (France) and then modified for the current application. The modifications mostly considered re-meshing of parts that were in direct contact with the barrier. A parameter was also introduced to the model that scaled the mass of the seats and concentrated masses so that the vehicle could be uniformly adjusted to the masses of 1300, 1500, and 1800 kg. Most parts of the vehicle were modeled using fully-integrated shell elements with piece-wise linear plasticity material law and strain rate effects. The modified model consisted of 29 172 nodes and 28 011 finite elements. The dimensions of the model also fulfilled the requirements for TB32 crash test. The view of the vehicle and its corresponding numerical model is presented in Fig. 3.

2.2. Numerical model validation

The road safety barrier model was validated against two modified TB32 full-scale crash tests [28]. The modification of the first test was the impact angle which was modified from 20 to 7 degrees, the other impact conditions remained unchanged, i.e., the impact speed was 110 km/h, the car mass was 1500 kg. The modification in the second crash test concerned only the configuration of the installation of the barrier system, i.e., the barrier was installed on the horizontal convex curve of the road with 400 m radius, the impact conditions corresponded to the TB32 test (110 km/h, 20 degrees, 1500 kg). Both experiments were conducted by the Road and Bridge Research Institute (IBDiM, www.ibdim.edu.pl) in the Research Institute for Protective Systems (IBOS, www.ibos.com.pl). The contact between elements in the corresponding numerical simulations was modeled using the penalty-based approach with Coulomb kinematic friction. The coefficients of static and kinematic friction between steel elements were 0.2 and 0.1, respectively. Similarly, coefficients of static and kinematic friction steel posts and the surrounding soil were 0.4 and 0.24, respectively. Regarding friction between tires and ground, the coefficients of static and kinematic friction were 0.6 and 0.42, respectively. The decay coefficient in all contacts equaled 0.001. The simulations were carried out using R10.1 LS-Dyna explicit solver with double precision on four 12-core Intel Xeon E5 v3 @ 2.3 GHz processors. The experimental and simulation results are compared in Table 1. The indices used for the comparison, such as Acceleration Severity Index (ASI), Theoretical Head Impact Velocity (THIV), working width, dynamic deflection, and length of contact are all specified in EN1317 [28]. The differences between real and virtual crash tests were calculated and compared in accordance with EN16303 [34]. All conditions were met except one, where in the first virtual test the working width was 0.05 m too small. However, the measurement uncertainty of the working width was 0.1 m and the considered test was not a standard test of the EN16303 standard [34], but its modification. Hence, the acquired validation result was assumed sufficient for the current application.

Table 1
Comparison of the results from two full-scale crash tests and their corresponding simulations.

	Crash test 1[a]	Simulation 1[a]	Crash test 2[b]	Simulation 2[b]
ASI	0.6	0.5	0.7	0.8
THIV, km/h	19	19	23	26
Working width, m	0.7 ± 0.1	0.5	1.0 ± 0.1	1.1
Dynamic deflection, m	0.4 ± 0.1	0.3	1.0 ± 0.1	0.8
Length of contact, m	7.86 ± 0.02	6.3	10.05 ± 0.02	8.6

[a]Impact conditions for crash test 1 and simulation 1: 110 km/h, 7 degrees, 1500 kg, crash into straight section of barrier.
[b]Impact conditions for crash test 2 and simulation 2: 110 km/h, 20 degrees, 1500 kg, crash into curved section of barrier, radius 400 m, impact side: convex.

3. Database

The estimation of the impact speed between a car and a W-beam guardrail should be possible based on several measurements and/or estimations at accident site. Measurement and/or estimation of such features should not require expertise and could in principle be carried out by anyone. Furthermore, it should be possible for the chosen features to be efficiently estimated by automated computer vision systems or to be acquired manually from photographs of the impact site taken by the authorities. Thus, a set of 5 input features is proposed for the problem, i.e., (1) total mass of the impacting vehicle, (2) impact angle, (3) permanent displacement of the W-beam guardrail, (4) number of the damaged posts, and (5) number of the damaged guardrail segments. The features are described in Section 3.1. To determine the feature values for various impact speeds, FEM simulations and crash tests were employed, which are presented in Section 3.2.

3.1. Input features

3.1.1. Mass of the vehicle

The first input feature was the total mass of the vehicle. The mass is easy to estimate once the model of the car is known, as this information is usually readily available. The mass of the occupants should be added to the mass of the vehicle. This could be estimated by knowing the number, age, and gender of the occupants. The mass of other objects inside the vehicle at the time of the crash, if they are known, should be also included.

Fig. 3. Comparison between the vehicle and its corresponding numerical model.

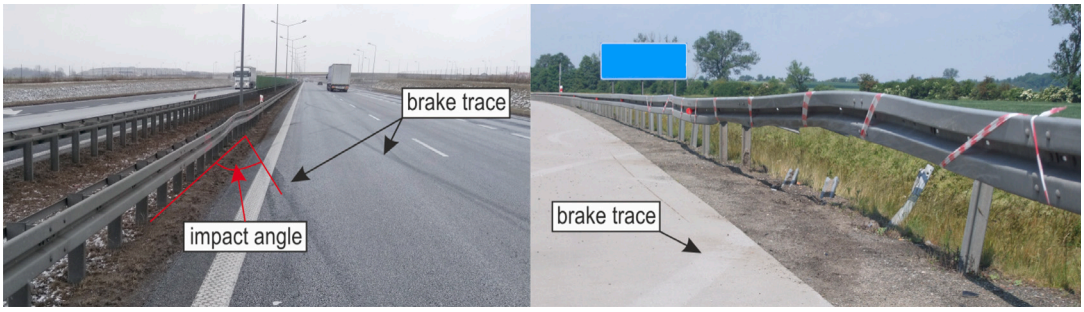

Fig. 4. Examples of measurement of the impact angle from the crash scene, source of photos: GDDKiA, Poland.

Fig. 5. Examples of measurement of the displacement from the crash scene, source of photos: GDDKiA, Poland.

3.1.2. Impact angle

The second input feature was the impact angle. This is the angle at which the vehicle hit the roadside barrier, relative to the barrier traffic face. The angle can usually be determined if there is a clear brake trace on the road or in the ground by the barrier. An example picture from which the impact angle could be estimated is shown in Fig. 4.

3.1.3. Permanent displacement of the W-beam guardrail

The third input feature was the maximum lateral deflection of the W-beam guardrail traffic face. This value is defined as the maximum lateral distance between the face of the intact guardrail and the face of the guardrail after the accident. This feature usually needs to be measured on-site or estimated from the available photographs. An example picture with the definition of lateral displacement is shown in Fig. 5.

3.1.4. Number of damaged posts and guardrail segments

The last two input features are the number of damaged posts and segments of the guardrail. These can easily be counted, and the convention was to use a greedy approach, i.e., if any part of a segment is damaged, the segment is counted as damaged. An example picture is shown Fig. 6. In this specific accident, 3 segments of the guardrail and 8 posts were considered as damaged. Note that the post does not need to be ripped off the ground to be counted as damaged, it is enough that an excessive deformation or torsion is observed.

3.2. Database description

A database containing the results from various accidents was created. The values of the described input features were obtained with the help of FEM simulations and full-scale crash tests. The number of simulations was assumed to be approximately 300. As the computational time of a typical simulation was approximately 6 h and the computations were run on 48 CPU cores, which yielded around 86 400 core hours (2.5 core months) for the whole simulation set. Additionally, the following ranges while creating the simulation dataset:

- impact speed: from 70 to 160 km/h,
- impact angle: from 4 to 30 degrees,
- mass of the car: 900, 1300, 1500, and 1800 kg.

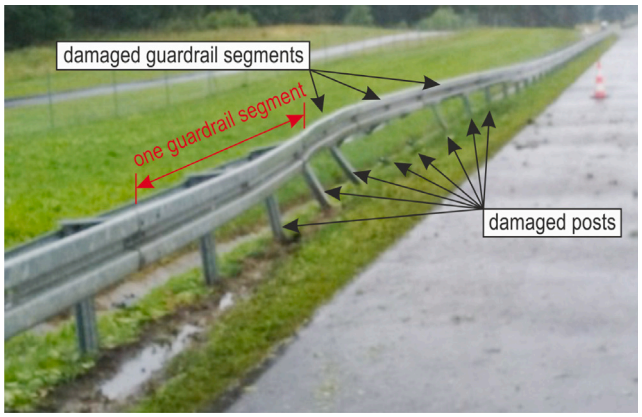

Fig. 6. Examples of the determination of the number of damaged posts and guardrails, source of photo: GDDKiA, Poland.

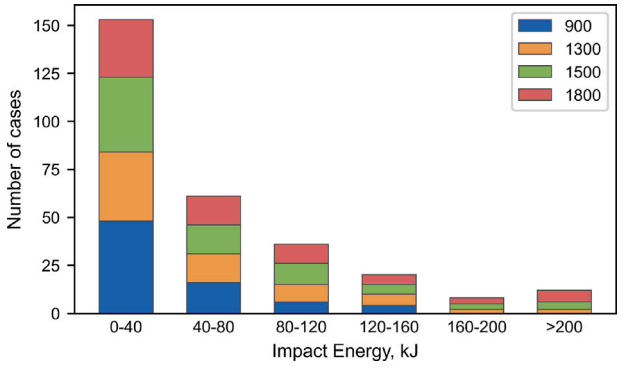

Fig. 7. The structure of the data from simulations.

The created array of simulations contained 290 cases in total. In addition, the results from two crash tests of the considered barrier were added to the database. A histogram of the impact energy distribution for the considered crash test scenarios is shown in Fig. 7. The impact energy EI (also known as impact severity, [35]) was calculated as

$$EI = \frac{M \cdot (V \cdot \sin\alpha)^2}{2}, \tag{1}$$

where M is the total mass of the vehicle, V is the impact speed, and α is the impact angle.

Additionally, the cases in which the vehicle got stuck under the barrier, overrode the barrier, or broke the barrier were excluded from the database. Therefore, the final database contained a total of 253 crash cases, including real and numerical tests. For all crash cases, the input features (i.e., the impact angle, car mass, permanent guardrail displacement, number of damaged posts, and number of damaged guardrail sections) were extracted from the tests and the vehicle impact speed values were recorded. Out of the total 253 samples, 64 samples (25%) were set aside as a test set, to evaluate the developed models. The remaining 189 samples were used for training the models.

4. Machine learning model

For the modeling, ensemble methods based on regression trees were chosen. In particular, the model considered three different base regressors: an extremely randomized forest [36], gradient boosted regression trees [37] and an adaptive boosting (AdaBoost) regressor [38,39].

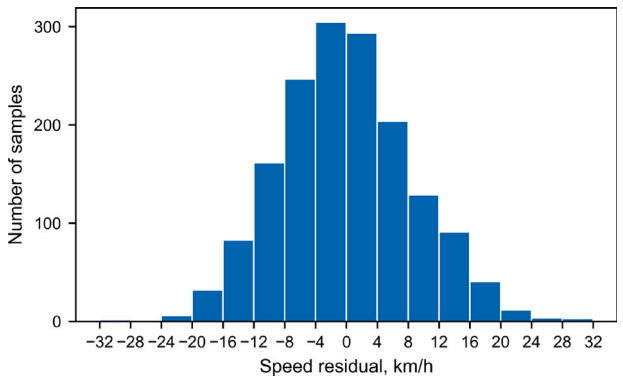

Fig. 8. Histogram of speed residuals (difference between true and predicted value) in the test set for the developed model.

4.1. Extremely randomized forest

The extremely randomized forest model is based on the classic random forest model [40]. In a random forest, an ensemble of regression trees is trained with bootstrap aggregating, i.e., every predictor is trained on a random subset of the training set, sampled with replacement. Each regression tree is then grown and the training set is split into subsets, so that the mean squared error between the average within each subset and the corresponding targets is minimized. When splitting a node during the growth of the trees, classic random forest searches for the best feature among a random subset of features (e.g., one that reduces the variance the most). After all the trees have been trained, the forest can make predictions by averaging the predictions from all trees.

Unlike regular decision trees, which search for the best possible threshold for each feature while splitting the node, an extremely randomized forest considers a random threshold for each feature [36]. This, along with using a random subset of features at node splitting results in even greater regressor diversity, trading higher bias for lower variance. Furthermore, random thresholds also provide a speedup in training.

An optimal set of model hyperparameters (for which the mean squared error on the validation was smallest) was found using randomized grid search with 5-fold cross-validation on the full training set including 189 training examples (using 20% of the training samples as validation set). In the course of cross-validation, the values of the hyperparameters used to train the model were repeatedly sampled from a predefined distribution. In the end, the optimal set was chosen as the one yielding the lowest mean squared error on the cross-validation set.

4.2. Gradient boosted regression trees

The general idea of boosting methods is to sequentially train predictors, where every following predictor tries to correct its predecessor [37]. In this aspect, the gradient-boosting regressor fits a new regression tree to the residual error made by previous predictors. In this setting, a regression tree is used as the base regressor, and in each training stage, a new regression tree is fit on the negative gradient of the loss function (mean squared error). In the end, the ensemble prediction can be obtained as the sum of the prediction of all regressors.

An optimal set of model hyperparameters was found using randomized grid search with 5-fold cross-validation on the full training set including 189 training examples (using 20% of the training samples as validation set). In the course of cross-validation, the values of the hyperparameters used to train the model were repeatedly sampled from a predefined distribution. In the end, the optimal set was chosen as the one yielding the lowest mean squared error on the cross-validation set.

Fig. 9. General view of the accident no 1, source of photo: GDDKiA, Poland.

4.3. Adaptive boosting of regression trees

Similar to the previous model, the adaptive boosting (AdaBoost) model trains regression trees sequentially so that each following predictor tries to correct its predecessor. In contrast to gradient boosting, AdaBoost adjusts the weights of the training examples according to the error of the current predictions. As a result, subsequent predictors focus more on examples that are difficult to predict. The final prediction from the regressor is then obtained through a weighted sum of the predictions made by individual regressors.

An optimal set of model hyperparameters was found using randomized grid search with 5-fold cross-validation on the full training set including 189 training examples (using 20% of the training samples as validation set). In the course of cross-validation, the values of the hyperparameters used to train the model were repeatedly sampled from a predefined distribution. In the end, the optimal set was chosen as the one yielding the lowest mean squared error on the cross-validation set.

4.4. Model summary

After the optimal sets of hyperparameters have been found for all base regressors, the final model can be built. In particular, we consider an ensemble estimator that uses several base regressors and averages their prediction to form a final prediction, i.e., a voting ensemble. The final prediction is then obtained as a weighted average of the predictions of the extremely randomized forest (15% weight), gradient boosted regression trees (75% weight), and adaptive boosting (10% weight).

To evaluate model performance, the original training set (189 training examples), was repeatedly split into training and validation sets (157 and 32 examples, respectively). For each random split, all base estimators were trained on the smaller training set, and the voting ensemble of them was formed. After that, the models made predictions on the validation set, and suitable metrics were computed. For the metrics, the mean absolute error (MAE) between the target and predicted values was considered. It can be computed as

$$\mathrm{MAE} = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n}, \tag{2}$$

where y_i denotes the target value, and $\hat{y}_i$ is the value predicted by the model. The MAE gives a rough estimate of the mean error on the predicted value given by the model. We also define the coefficient of determination, R^2, which can be computed as

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}, \tag{3}$$

where $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$ is the average target value. This score provides a measure of how well observed outcomes are replicated by the model, based on the proportion of total variation of outcomes explained by the model. The best possible R^2 score is 1.0. For each random split of the full training set, the MAE and R^2 on the validation sets were recorded. Afterward, the mean values (and standard deviations) over all validation sets can be used to estimate the value of MAE and R^2 on unseen data. Model evaluation using random splitting of the full training set was done for 50 random splits.

Subsequently, the base regressors (and the voting ensemble of them) were retrained on the full training set. As a final check, the MAE and R^2 were also computed on the test set, which was initially created and set aside for this purpose (64 samples). The values of the MAE and R^2 for the training, validation, and test sets are summarized in Table 2. For the validation set, the values represent the mean and standard deviation over all random splits.

As can be seen in Table 2, the MAE (in km/h) in the validation set is similar to the one on the test set, meaning this metric (and its standard deviation) can be used to estimate the performance of the model on unseen data. The histogram of the residuals (differences between the true and predicted values) on the validation sets is presented in

Table 2
Mean absolute error (MAE), standard deviation (σ) of MAE and coefficient of determination (R^2) for the ensemble model (ERF — extremely ransomized forest, GBRT — gradient boosted regression trees, AdaBoost — adaptive boosting of regression trees, TRE — tree ensemble (voting ensemble).

Metric	Set	ERF (15%)	GBRT (75%)	AdaBoost (10%)	**TRE**
MAE	Train	3.05 km/h	1.62 km/h	7.40 km/h	**2.15 km/h**
	Val	8.78 km/h	6.39 km/h	9.52 km/h	**6.76 km/h**
	Test	8.61 km/h	6.14 km/h	9.03 km/h	**6.42 km/h**
σ	Val	1.18 km/h	0.99 km/h	1.49 km/h	**1.01 km/h**
R^2	Train	0.97	0.99	0.83	**0.99**
	Val	0.75	0.86	0.69	**0.85**
	Test	0.72	0.86	0.67	**0.84**

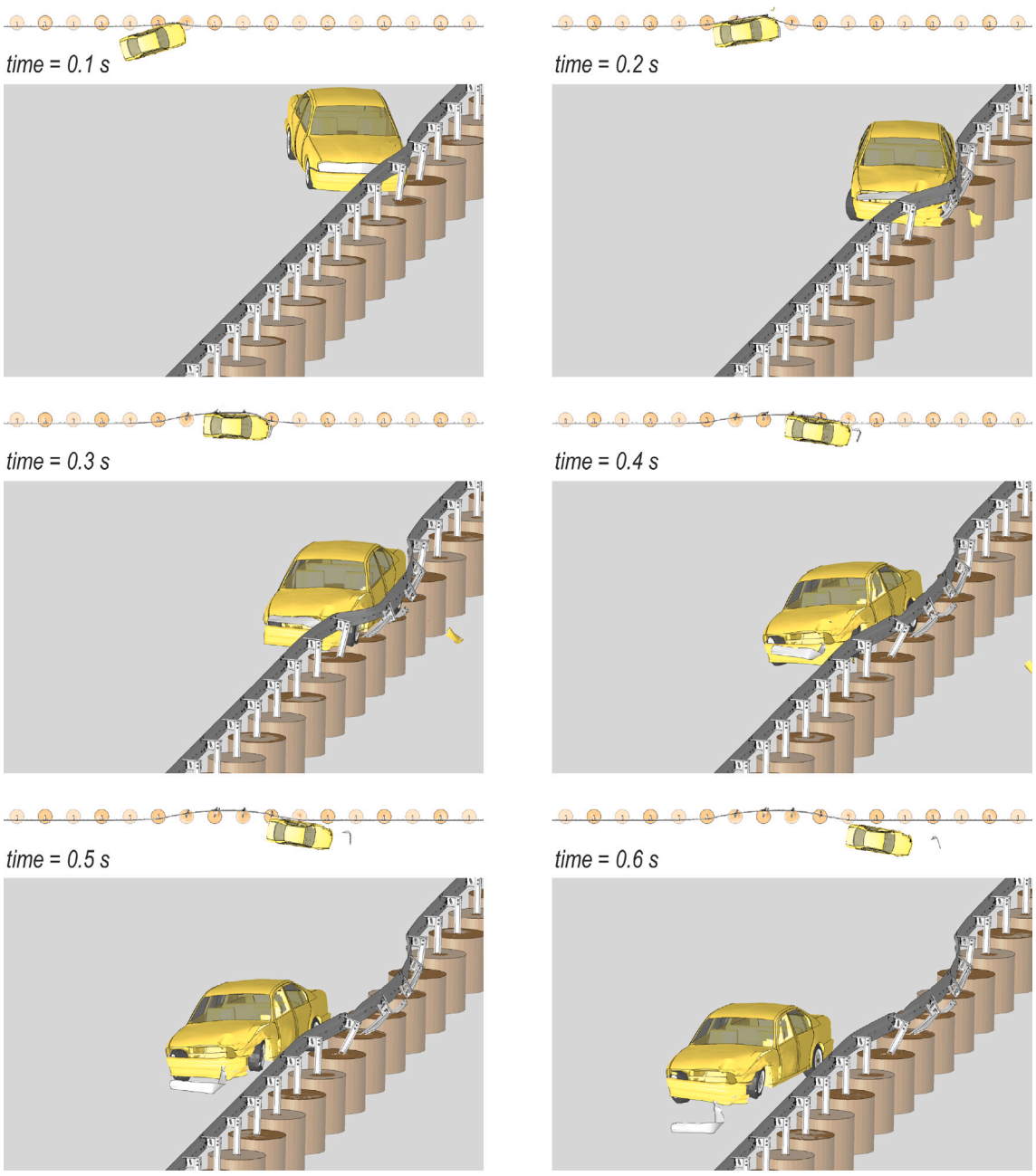

Fig. 10. Crash course of accident no. 1.

Fig. 8. Note that these are all residuals recorded in the course of cross-validation using 50 different random splits of the full training sets (a total of 1600 predictions). It can be seen that the speed residual has an approximately Gaussian distribution. The complete Tree Ensemble model has been released for public use, and can be accessed at [41], where the Tree Ensemble model, at the time of writing, coincides with the one described in this work. The authors, however, plan for future updates and improvements of the models.

5. Examples of application of the developed ML model

5.1. General assumptions

The developed ML models were used to estimate the impact speed of a car into a barrier in three real-world road accidents. Data from actual accidents was obtained from sections of expressways and highways. Unfortunately, these accidents involved a barrier that was not the same barrier used to train the ML model. The barrier involved in the accidents was of class H1/W5/A (according to EN1317 [28]), whereas the barrier used to train the models was of class H1/W4/A. Nevertheless, both road safety systems were steel W-beam barriers and had the same distance between the posts and the length of the guardrail segments. Also, due to scarcity of accident data, all information on the accident, such as the angle of impact and the number of damaged elements, was inferred only from photos of the accident site. The mass of the vehicle was determined in a similar way, where the brand and model of the vehicle were determined based on the photos. As the available data did not contain information about the number of occupants, it was assumed that there were 2 occupants in a car, each weighing 80 kg. It is noteworthy, that the geometry of the vehicles involved in the actual accidents may have differed from those used for learning of the model.

Considering the above assumptions, the impact speed was estimated using the ML model described in Section 4, based on the values of 5

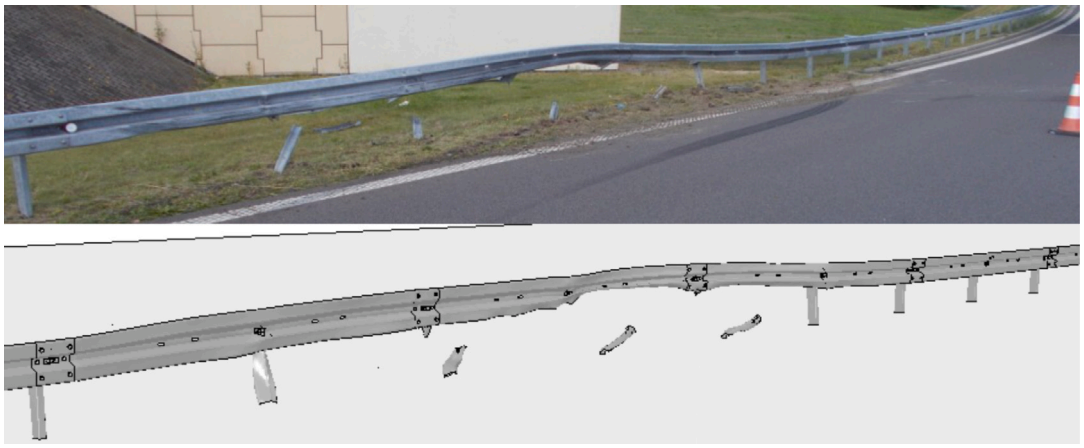

Fig. 11. Comparison of the damaged W-beam guardrails from accident no. 1, source of top photo: GDDKiA, Poland.

Fig. 12. General view of the accident no 2, source of photo: GDDKiA, Poland.

features (see Section 3.1) acquired from accident photos. The estimated impact speeds were rounded to the nearest integer value.

5.2. Accident no. 1

This accident occurred on an expressway. The total mass of the vehicle (with occupants) was 1410 kg, the impact angle was 17 degrees, the displacement of the guardrail was 50 cm. The number of posts and guardrails which were considered as damaged were 7 and 5, respectively. The photo showing the general view from the accident is presented in Fig. 9.

The ML model estimated the impact speed to be 131 km/h, resulting in an impact energy of 91.1 kJ. This estimated speed was verified using the corresponding FE simulation. In the simulation, the collision parameters, such as the total vehicle mass, the impact angle, and the impact speed were set to be the same as the inputs to the ML model. Then, the values of remaining features were extracted. The following results were obtained from the simulation: the permanent displacement of the guardrail was 67 cm, the number of damaged posts was 7, and the number of damaged guardrails was 4. The crash course is shown in Fig. 10, and a comparison of the views of the damaged system from the simulation and real-world accident is presented in Fig. 11.

5.3. Accident no. 2

This accident occurred on an expressway. The total car mass (with occupants) was 1460 kg, and the impact angle was 14 degrees. The permanent displacement of the guardrail was estimated to 90 cm. The crash resulted in 8 posts and 5 guardrails damaged. The general view from the accident is shown in Fig. 12.

The ML model estimated the impact speed to be 153 km/h. resulting in an impact energy of 78.2 kJ. This estimation was verified by a FE simulation, where the final guardrail displacement was 85 cm, the number of damaged posts was 8, and the number of damaged guardrails was 5. The crash course is shown in Fig. 13, and a comparison of the views of the damaged system from the simulation and real-world accident is presented in Fig. 14.

5.4. Accident no. 3

Accident no. 3 occurred on a highway. As the accident documentation did not record the car brand, it was assumed that the impacting vehicle weighed 1400 kg. The total mass, including two 80 kg occupants, was taken as 1560 kg. The impact angle was 8 degrees, and the displacement of the guardrail was 65 cm. In this case, 5 posts and

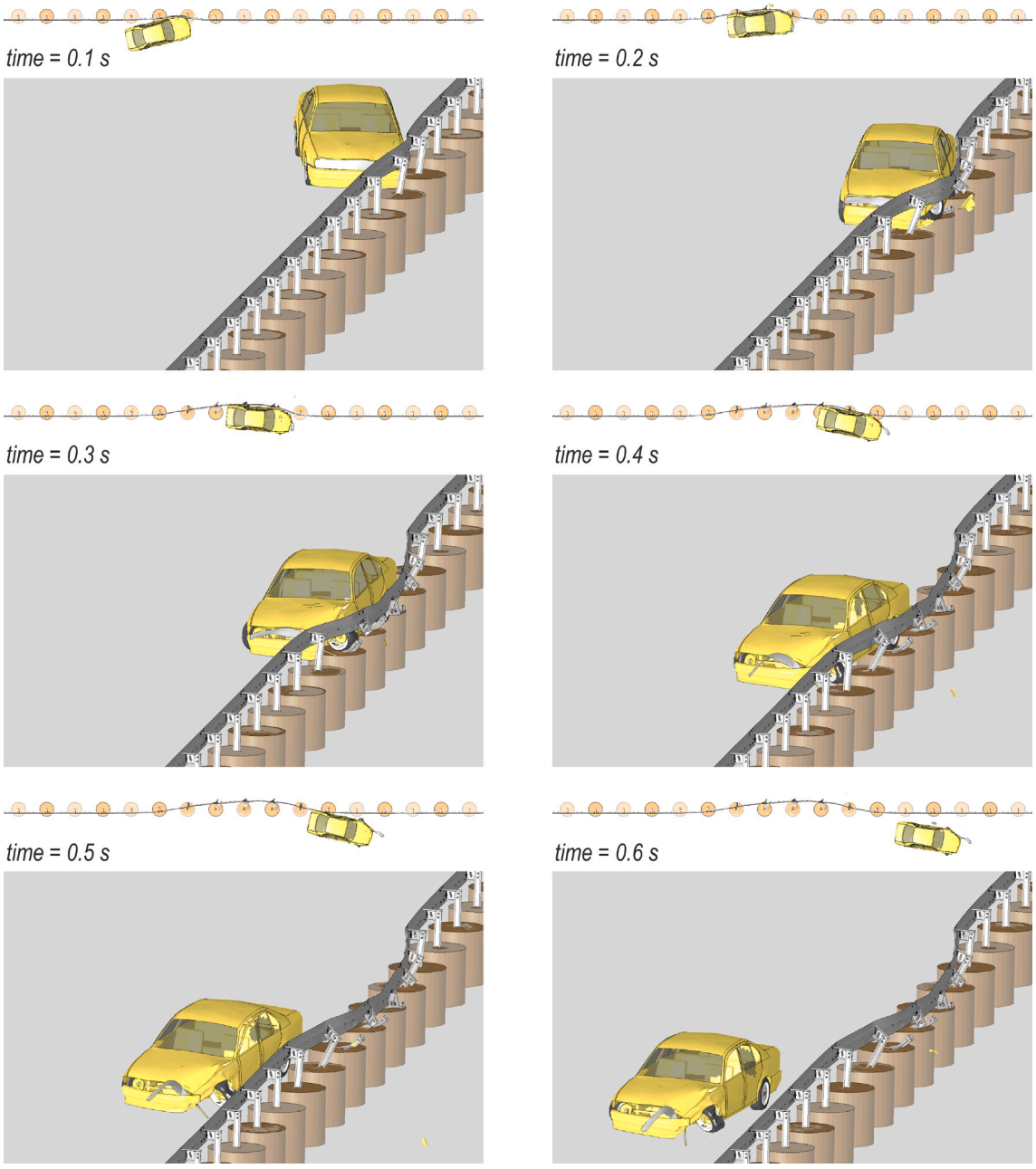

Fig. 13. Crash course of accident no. 2.

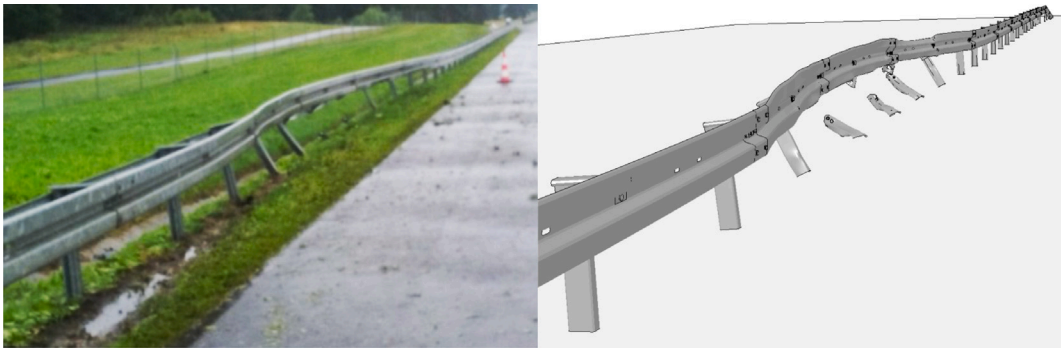

Fig. 14. Comparison of the barrier after accident no. 2, source of left photo: GDDKiA, Poland.

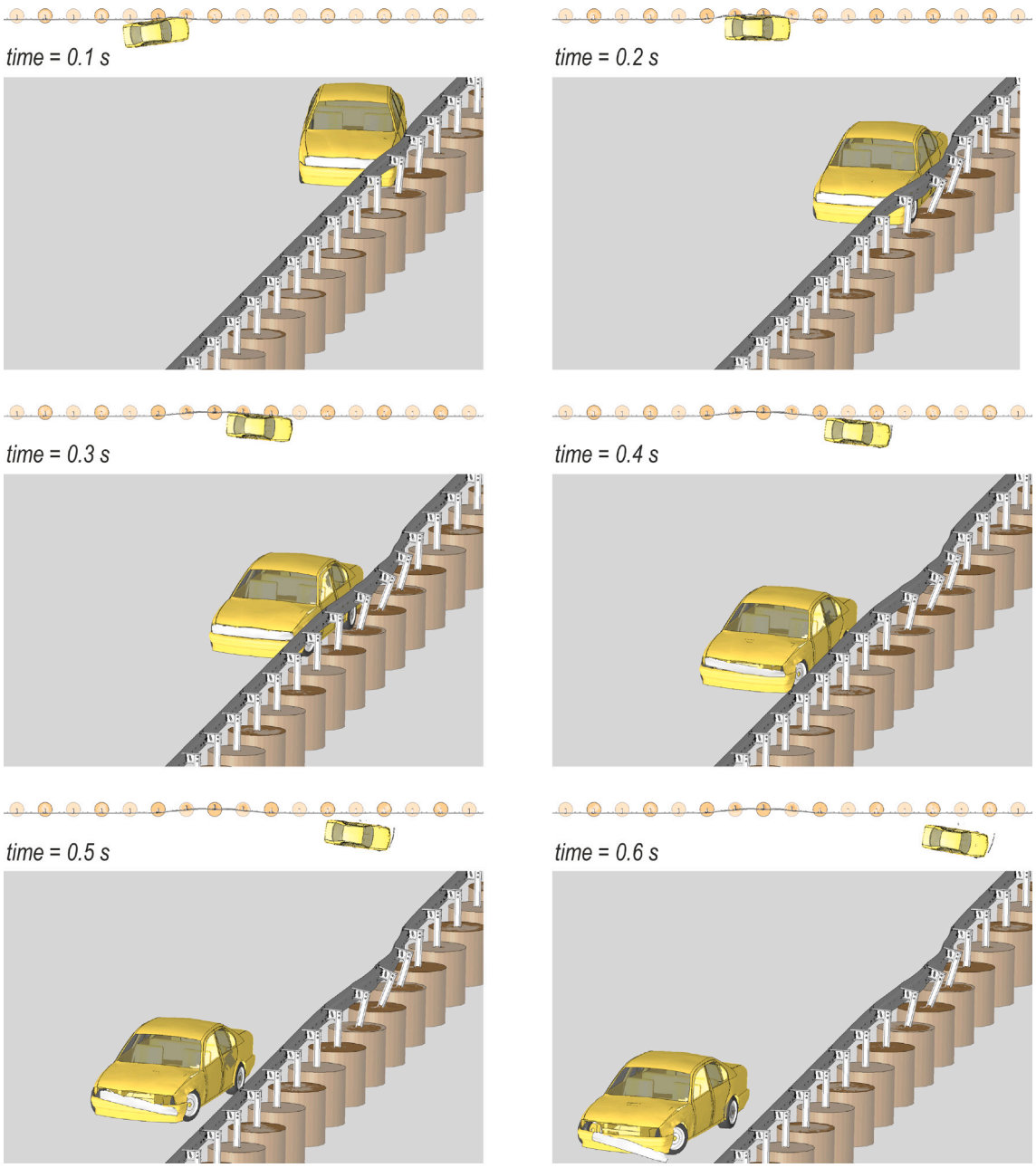

Fig. 15. Crash course of accident no. 3.

4 guardrails were found. The ML model estimated the impact speed to be 145 km/h, thus, resulting in an impact energy of 26.2 kJ. For verification, a FE simulation was conducted, where the final guardrail displacement was 31 cm, the number of damaged posts was 5, and the number of damaged guardrails was 4. The course of crash is shown in Fig. 15, and a comparison of the damage to the barrier system from the simulation and real-world accident is shown in Fig. 16.

5.5. Summary of the estimated impact velocities

The results of studied accidents are summarized in Table 3. The first column contain the ML model features of the accidents that were extracted from site photos. Subsequently, the impact speed estimated by the ML model is given in the second column. The last column contains the results of corresponding FEM simulations. Accidents 1 and 2 happened on the expressways, whereas Accident 3 happened on a highway. In Poland, the speed limit on expressways and highways is 120 km/h and 140 km/h, respectively. Therefore, the estimated speed from the considered accidents exceeded those limits. The results (Table 3) show good agreement with the real accident results. Hence, it could be concluded that the speed estimation model provided a reasonable estimate of the vehicle impact speed. Discrepancies may result from differences between the barrier that was involved in the accident and the one used to develop the ML model. The working width class of the barrier in the accident was W5, whereas the barrier used for the ML model training was W4. Hence, the accident barrier was more flexible compared to the barrier used for model learning. In the case of Accident 3, a significant difference in the guardrail displacement was obtained from the simulation. This can be due to the fact that the mass of the car involved in the crash was not known. Most probably, the assumed mass of the vehicle should be greater than the assumed 1560 kg.

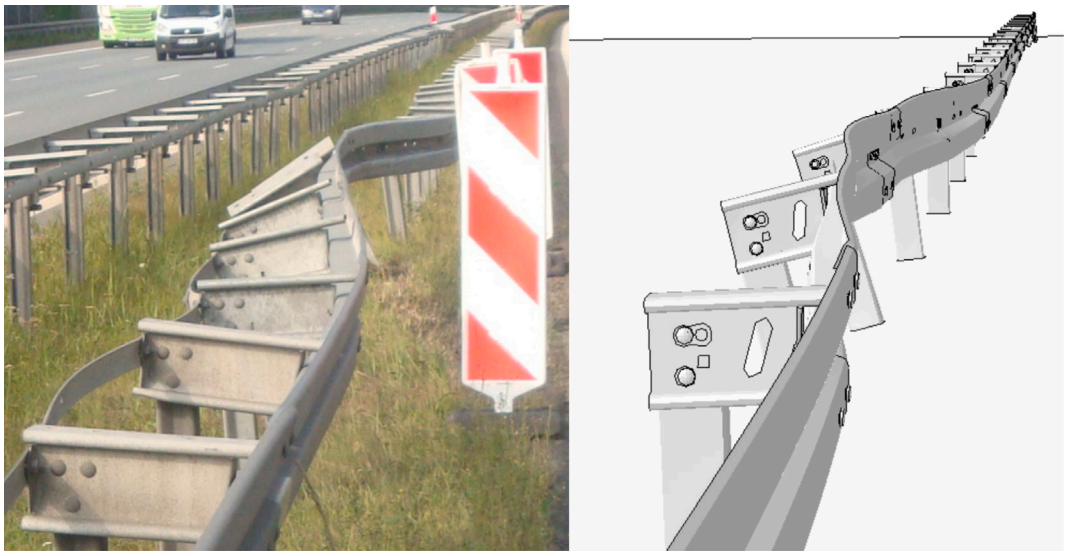

Fig. 16. Comparison of the barrier after accident no. 3, source of left photo: GDDKiA, Poland.

Table 3
Summary of road accident analysis.

Accident						ML	Simulation		
No.	Mass, kg	Angle, °	Disp., cm	Damaged posts	Damaged guardrails	Speed, km/h	Disp., cm	Damaged posts	Damaged guardrails
1	1410	17	50	7	5	131	67	7	4
2	1460	14	90	8	5	153	85	8	5
3	1560	8	65	5	4	145	31	5	4

6. Conclusions

In this study, a new method of developing models to estimate the speed of a car at the time of an impact with a W-beam barrier was presented. The proposed ML model comprised multiple machine learning algorithms and was based on the results from FE simulations and full-scale crash tests. A database of approximately 300 crash tests was created and used to train and validate the ML model. The model was an ensemble of an extremely randomized forest, gradient boosted regression trees, and an adaptive boosting regressor. The ensemble was established to improve the accuracy of the estimation and limit the overfitting of each model, so that the model might perform better on new, unseen data.

The newly developed model showed that FEM and ML methods can be used to successfully determine the impact speed of a vehicle with a barrier. Although the proposed methodology is laborious, the final model allows for quick estimation of the impact speed and energy, which is determined on the basis of simple features. Those features can be easily determined by the authorities present at the site of the accident, or even from the photos. For simulation-based training of the algorithm, only models validated by real full-scale experiments should be used. To further increase the reliability of the model, the database of accidents should be extended by additional cases. Other types of road barriers could also be added in the future. It is also worth checking how the quality of the results would be affected by the data obtained from in-situ measurements, which are certainly a more reliable source of information than the photos themselves. The advantage of photos, however, is that in the future it might be possible to automatically extract input features. With this automatic extraction, it will also be possible to increase the number of features at virtually no cost. The factor influencing the results obtained from the ML model is the mass of the vehicle, which depends on the number of passengers in the car. Knowledge of this data would further increase the credibility of the results. However, access to this type of data is limited, which makes the use of numerical simulations a convenient way to circumvent this.

It should be highlighted, that a similar methodology based on crash tests, numerical simulations, and ML algorithms can also be used to estimate the other initial impact conditions (e.g., impact angle, total impacting mass) as well as the accident outcomes (e.g., length of barrier damage, number of damaged posts and guardrails, severity level for the vehicle occupants, damage extend to the car, the amount of absorbed impact energy, etc.). From a broader perspective, it could be also possible to estimate the injuries of passengers depending on their age, gender, weight, and other factors.

CRediT authorship contribution statement

Dawid Bruski: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing, Supervision. **Lukasz Pachocki:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Adam Sciegaj:** Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Wojciech Witkowski:** Writing – critical revision, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

Numerical simulations were carried out at the Academic Computer Centre in Gdańsk (TASK), Gdańsk University of Technology, Poland.

The BWM vehicle model was developed by Transpolis (formerly LIER), France.

Funding

The research was financially supported by Gdańsk University of Technology, Poland under grant DEC-36/2021/IDUB/I.3.3 within the Argentum Trigerring Research Grants programme, project title: Methods for Speed Estimation of a Vehicle at Impact with a Road Safety. The work regarding development of machine learning models was financially supported by Gdańsk University of Technology, Poland under grant DEC-2/2020/IDUB/I.1 within the Nobelium Joining GUT Research Community programme.

References

[1] Symon E. Wypadki drogowe w polsce w 2021 roku. Tech. rep., Warszawa: Komenda Główna Policji, Biuro Ruchu Drogowego; 2022.
[2] Mehrara Molan A, Rezapour M, Ksaibati K. Investigating the relationship between crash severity, traffic barrier type, and vehicle type in crashes involving traffic barrier. J Traff Transp Eng 2020;7(1):125–36. http://dx.doi.org/10.1016/j.jtte.2019.03.004.
[3] Li Z, Fang H, Fatoki J, Gutowski M, Wang Q. A numerical study of strong-post double-faced W-beam and thrie-beam guardrails under impacts of vehicles of multiple size classes. Accid Anal Prev 2021;159(June):106286. http://dx.doi.org/10.1016/j.aap.2021.106286.
[4] Ferdous MR, Abu-Odeh A, Bligh RP, Jones HL, Sheikh NM. Performance limit analysis for common roadside and median barriers using LS-DYNA. Int J Crashworthiness 2011;16(6):691–706. http://dx.doi.org/10.1080/13588265.2011.623023.
[5] Fang H, Wang Q, Weggel DC. Crash analysis and evaluation of cable median barriers on sloped medians using an efficient finite element model. Adv Eng Softw 2015;82:1–13.
[6] Bruski D, Burzyński S, Chróścielewski J, Jamroz K, Pachocki Ł, Witkowski W, Wilde K. Experimental and numerical analysis of the modified TB32 crash tests of the cable barrier system. Eng Fail Anal 2019;104:227–46. http://dx.doi.org/10.1016/j.engfailanal.2019.05.023.
[7] Dinnella N, Chiappone S, Guerrieri M. The innovative "NDBA" concrete safety barrier able to withstand two subsequent TB81 crash tests. Eng Fail Anal 2020;115(January):104660. http://dx.doi.org/10.1016/j.engfailanal.2020.104660.
[8] Borkowski W, Hryciów Z, Rybak P, Wysocki J. Numerical simulation of the standard TB11 and TB32 tests for a concrete safety barrier. KONES Powetrain Transp 2010;17(4):63–71.
[9] Coon BA, Reid JD. Crash reconstruction technique for longitudinal barriers. J Transp Eng 2005;131(1):54–62. http://dx.doi.org/10.1061/(ASCE)0733-947X(2005)131:1(54).
[10] Asadollahi Pajouh M, Schmidt JD, Meyer CL, Lechtenberg KA, Faller RK. Crash reconstruction technique for cable barrier systems. J Transp Saf Secur 2019;11(3):243–60. http://dx.doi.org/10.1080/19439962.2017.1386251.
[11] Coon BA, Reid JD. Reconstruction techniques for energy-absorbing guardrail end terminals. Accid Anal Prev 2006;38(1):1–13. http://dx.doi.org/10.1016/j.aap.2005.06.016.
[12] Budzyński M, Gobis A, Jamroz K, Jeliński Ł, Ostrowski K. Road restraint systems as a basis for roadside safety improvement. IOP Conf Ser: Mater Sci Eng 2019;471(062029):1–10. http://dx.doi.org/10.1088/1757-899X/471/6/062029.
[13] Li N, Fang H, Zhang C, Gutowski M, Palta E, Wang Q. A numerical study of occupant responses and injuries in vehicular crashes into roadside barriers based on finite element simulations. Adv Eng Softw 2015;90:22–40. http://dx.doi.org/10.1016/j.advengsoft.2015.06.004.
[14] Sybilski K, Małachowski J. Sensitivity study on seat belt system key factors in terms of disabled driver behavior during frontal crash. Acta Bioeng Biomech 2019;21(4). http://dx.doi.org/10.5277/ABB-01421-2019-02.
[15] John J, Klug C, Kranjec M, Svenning E, Iraeus J. Hello, world! VIVA+: A human body model lineup to evaluate sex-differences in crash protection. Front Bioeng Biotechnol 2022;10. http://dx.doi.org/10.3389/fbioe.2022.918904.
[16] Bruski D, Burzyński S, Witkowski W. Analysis of passenger car crash with a cable barrier installed with anti-glare screens on a horizontal convex road curve with 400 m radius. Int J Impact Eng 2023;173(April 2022). http://dx.doi.org/10.1016/j.ijimpeng.2022.104486.
[17] Mohan P, Marzougui D, Meczkowski L, Bedewi N. Finite element modeling and validation of a 3-strand cable guardrail system. Int J Crashworthiness 2005;10(3):267–73. http://dx.doi.org/10.1533/ijcr.2005.0345.
[18] Teng T-l, Liang C-l, Tran T-t. Development and validation of a finite element model for road safety barrier impact tests. Simul: Trans Soc Model Simul Int 2016;92(6):565–78.
[19] Soltani M, Topa A, Karim MR, Sulong NH. Crashworthiness of G4(2W) guardrail system: a finite element parametric study. Int J Crashworthiness 2017;22(2):169–89. http://dx.doi.org/10.1080/13588265.2016.1243636.
[20] Klasztorny M, Zielonka K, Nycz DB, Posuniak P, Romanowski RK. Experimental validation of simulated TB32 crash tests for SP-05/2 barrier on horizontal concave arc without and with composite overlay. Arch Civ Mech Eng 2018;18:339–55.
[21] Pachocki Ł, Bruski D. Modeling, simulation, and validation of a TB41 crash test of the H2/W5/B concrete vehicle restraint system. Arch Civ Mech Eng 2020;20(62). http://dx.doi.org/10.1007/s43452-020-00065-7.
[22] Shahriari M, Saeidi Googarchin H. Prediction of vehicle impact speed based on the post-cracking behavior of automotive PVB laminated glass: Analytical modeling and numerical cohesive zone modeling. Eng Fract Mech 2020;240(May):107352. http://dx.doi.org/10.1016/j.engfracmech.2020.107352.
[23] Bock FE, Aydin RC, Cyron CJ, Huber N, Kalidindi SR, Klusemann B. A review of the application of machine learning and data mining approaches in continuum materials mechanics. Front Mater 2019;6(May). http://dx.doi.org/10.3389/fmats.2019.00110.
[24] Choi JG, Kong CW, Kim G, Lim S. Car crash detection using ensemble deep learning and multimodal data from dashboard cameras. Expert Syst Appl 2021;183:115400. http://dx.doi.org/10.1016/j.eswa.2021.115400.
[25] Ji W, Yang T, Yuan Q, Cheng G, Yu S. Prediction model of accident vehicle speed based on artificial intelligence decision free algorithm. Lect Notes Electr Eng 2023;941 LNEE:317–24. http://dx.doi.org/10.1007/978-981-19-4786-5_44.
[26] Duma I, Burnete N, Todorut A. A review of road traffic accidents reconstruction methods and their limitations with respect to the national legal frameworks. IOP Conf Ser: Mater Sci Eng 2022;1220(1):012055. http://dx.doi.org/10.1088/1757-899x/1220/1/012055.
[27] Pawlus W, Karimi HR, Robbersmyr KG. Reconstruction and simulation of the vehicle to road safety barrier oblique collision based on the levenberg-marquardt algorithm. Int J Crashworthiness 2012;17(6):676–92. http://dx.doi.org/10.1080/13588265.2012.714300.
[28] EN 1317 - European road restraint systems.
[29] Volume I, volume II. In: LS-Dyna R10.0 keyword user's manual. LSTC; 2017.
[30] Hallquist JO. LS-Dyna theory manual. Livermore: Livermore Software Technology Corporation (LSTC); 2006.
[31] ISO 4032:2012 - hexagon regular nuts (style 1) — Product grades A and B.
[32] Wilde K, Bruski D, Burzyński S, Chróścielewski J, Pachocki Ł, Witkowski W. On analysis of double-impact test of 1500-kg vehicle into w-beam guardrail system. Arch Civ Eng 2021;LXVII(2):101–15. http://dx.doi.org/10.24425/ace.2021.137157.
[33] Wolny R, Bruski D, Budzyński M, Pachocki L, Wilde K. Influence of a lighting column in the working width of a W-beam barrier on TB51 crash test. Materials 2022;15(14). http://dx.doi.org/10.3390/ma15144926.
[34] EN 16303 - road restraint systems - Validation and verification process for the use of virtual testing in crash testing against vehicle restraint system.
[35] AASHTO MASH-2, manual for assessing safety hardware (MASH). 2nd ed.. American Association of State Highway Transportation Officials; 2016, p. 259.
[36] Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. Mach Learn 2006;63(1):3–42. http://dx.doi.org/10.1007/s10994-006-6226-1.
[37] Friedman JH. Greedy function approximation: A gradient boosting machine. Ann Statist 2001;29(5):1189–232. http://dx.doi.org/10.1214/aos/1013203451.
[38] Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. J Comput System Sci 1997;55(1):119–39. http://dx.doi.org/10.1006/jcss.1997.1504.
[39] Drucker H. Improving regressors using boosting techniques. In: International conference on machine learning. 1997.
[40] Ho TK. Random decision forests. In: Proceedings of 3rd international conference on document analysis and recognition, Vol. 1. 1995, p. 278–82. http://dx.doi.org/10.1109/ICDAR.1995.598994, vol.1.
[41] Sciegaj A. Speedest - online app for estimating vehicle impact speed. 2023, [Online; accessed 9-May-2023]. http://www.speedest.pl/.