

Structural integrity of aging steel bridges by 3D laser scanning and convolutional neural networks

Georgios Tzortzinis^{ab*}, Angelos Filippatos^c, Jan Wittig^d, Maik Gude^a
Aidan Provost^e, Chengbo Ai^e, Simos Gerasimidis^e

^aInstitute of Lightweight Engineering and Polymer Technology, Technische Universität Dresden, Dresden, 01307
Germany

^bDresden Center for Intelligent Materials, Technische Universität Dresden, Dresden, 01307, Germany

^cDepartment of Mechanical Engineering and Aeronautics, University of Patras, Rio, 26504, Greece

^dDepartment of Engineering Science, University of Oxford, Oxford, OX1 3PJ, United Kingdom

^eDepartment of Civil and Environmental Engineering, University of Massachusetts Amherst, Amherst, MA, 01003,
USA

*Corresponding author: Georgios Tzortzinis; georgios.tzortzinis@tu-dresden.de

Abstract

Evaluating the structural condition of aging assets is a key challenge for resilient infrastructure. For steel bridges, corrosion has historically triggered bridge failures resulting to fatalities and injuries and to enhance public safety and prevent such incidents, authorities mandate in-situ evaluation and reporting of corroded members. Intense labour, traffic delays and poor capacity predictions characterise the current inspection and evaluation protocol. This work combines full scale experimental testing of a decommissioned girder, 3D laser scanning and convolutional neural networks (CNN) to introduce a continuous inspection and evaluation framework. By training both classification and regression CNNs on a databank of 1421 naturally inspired corrosion scenarios, low errors of up to 2.0% and 3.3%, respectively, are achieved. The methodology is validated on eight real corroded ends and implemented for the evaluation of an in-service bridge. This framework promises significant advancements in assessing aging bridge infrastructure with higher accuracy and efficiency compared to analytical or semi-analytical approaches.

Keywords: Aging Infrastructure, Machine Learning, Evaluation, 3D laser scanning, Corroded bridges

1. Introduction

The structural condition of transportation infrastructure is a critical determinant of economic vitality as has been demonstrated lately by the ASCE [1], the FHWA [2] and other organizations. In the US today, according to the FHWA [2], there are about 43,000 bridges in poor condition and 305,000 bridges in fair condition; the number of bridges in fair condition has tripled in the last 30 years. The ongoing challenge lies in maintaining these bridges as they are subjected to a constant deterioration, gradual loss of structural integrity and therefore reduced operational functionality. Deterioration of transportation infrastructure occurs due to a multitude of factors, including weather-related wear and tear, heavy vehicular loads, and insufficient maintenance funding with far-reaching consequences including the compromise of safety, increased travel times, higher fuel consumption, and additional maintenance costs for vehicles. Accurately and efficiently assessing the structural condition of bridges is at the heart of the pursuit for resilient infrastructure. To

address this challenge, national inspection standards [3] have been established, dictating mandatory bridge inspections. While this approach remains crucial, it comes with significant challenges and drawbacks. The current state-of-practice for bridge inspections largely relies on visual observations or low precision equipment [4] which unavoidably brings subjective engineering judgement in the methodology of documentation of deterioration in standardized reports [5]. The reliability and accuracy of data collected by current inspection tools necessitates continuous calibration, maintenance, and quality control. Additionally, the sheer volume of data generated by these technologies can overwhelm inspection teams, making it imperative to develop efficient data management and analysis systems.

As an alternative, 3D laser scanning technology offers rapid three-dimensional spatial data acquisition and has the potential to overcome the limitations of conventional inspection techniques. Based on their mode of operation, laser scanners are classified to terrestrial (TLS), photogrammetric, portable or mounted on UAV devices. Regardless of the type and the technical specifications, the obtained point-cloud data enables the post-processing and the reconstruction of actual surfaces and objects. Consequently, laser scanning technology has enhanced the capture of geometric information on different kinds of bridges [6, 7, 8, 9, 10] (e.g., masonry, timber, cable) and the creation of as-built models of existing bridges, albeit with little focus on the structural condition of individual components [11, 12, 13, 14].

For bridge inspections, post-processed point clouds facilitate the detection of surface damage for concrete and steel components. For concrete, a variety of methodologies based on defect indices [15], unsupervised classification algorithms [16] or Gaussian filtering and parabolic fitting [17] have been proposed for spalling detection and quantification. In a work by Bolourian and Hammad [18], a laser scanner mounted on an unmanned aerial vehicle (UAV) was used to detect locations with surface defects along a four-span concrete bridge. 3D laser scanning technology has been combined with image processing for cracks detection and mapping [19, 20, 21], as well as with virtual reality (VR) to reconstruct a masonry railway arch bridge, enabling visual inspection with consistent and repeatable results [22]. For steel bridges, an early work [23] proposed a framework of bridge inspection using TLS data. The methodology was demonstrated on three metal truss bridges in Ireland and Germany, where deformations and section loss were successfully detected. Corrosion induced deterioration along steel members can be also detected in combination with photogrammetry [24], or exclusively with color image processing and Fourier transform [25].

For bridges with steel girders, corrosion is the major factor of deterioration. Expansion joints located above the beam bearings allow water leakage and deicing chemicals to contaminate the steel load bearing surfaces, resulting in section loss and posting recommendations. For assessing the structural integrity of steel bridges, a significant challenge lies on current analytical tools which are generally poorly performing [26]. Limited efforts [27, 28, 29, 30] have focused on quantifying the effect of deterioration to the structural integrity of corroded steel bridge assets. Wang et al. [27] scanned and conducted four-point bending tests on three H-beams extracted from a naturally corroded truss. A beam with artificial section loss was tested after being scanned with structured-light 3D scanner. Decommissioned stiffened and un-stiffened bridge girders with end deterioration have been scanned with TLS technology and tested by the authors of the current paper [29, 30]. In the same work, a semi-automated method for section loss integration in computational models was proposed. Zhang and Zaghi [31] have also utilized data from a high resolution structured light 3D-scanner for the generation of a computationally efficient finite element model with shell elements. The accuracy of the proposed methodology has been demonstrated with experimental testing of a stiffened girder. Although accurate, a common characteristic of the previously mentioned research efforts is the necessity for development of beam specific high-fidelity computational models which requires both a deep understanding of the finite element method and a significant amount of work. These aspects arise barriers to the wide adoption of the laser scanning technology and the associated benefits in accuracy, robustness, documentation and digitization of the inspection process.

Machine learning (ML) algorithms have been previously used for the development of capacity prediction methods for structural components [32, 33, 34, 35]. Degtyarev and Naser [33] explored five ML algorithms as tools for predicting the elastic shear buckling loads and the ultimate shear strength of cold formed steel channels with staggered web perforations. Both metrics compared excellently with the simulation results. In [34], the extreme gradient boosting algorithm is recommended to evaluate the load-carrying capacity of semi-rigid steel structures. Mojtabei et al. [35] demonstrated the use of Artificial Neural Networks (ANNs) for the calculation of the buckling load and mode (i.e. local, distortional and global) of thin-walled structural elements. In the field of bridge engineering, Mangalathu et al. [36] proposed a methodology to facilitate the post-earthquake operations through the rapid damage state classification of two-span box-girder bridges.

Through the integration of new technologies such as LiDAR scanning and machine learning, modern bridge inspection methodologies can improve the accuracy of condition assessments and enable efficient predictive maintenance strategies. In this study, we present a new framework for the continuous inspection and assessment of corroded steel bridges, which combines 3D laser scanning technology with machine learning algorithms. The presented research relies on the attainability of real corroded girders from the region of New England and the availability of hard-to-find experimental specimens. Point clouds are used to capture the condition of deteriorated girders, while contour maps are generated to visualise the remaining thickness in a format which could be easily processed by inspectors and engineers. Following the full scale testing of a corroded girder from the state of MA, convolutional neural networks (CNNs) are trained on naturally inspired artificial corrosion data. Both regression and classifications models are developed to quantify the deterioration severity of beams with corroded ends, with a minimum error of 2% and 3.3%, respectively. Eight decommissioned girders of varying beam sections from the states of MA and NH are 3D scanned, and the developed models capture the residual capacities with errors in the range of 0 and 13% of the nominal capacities. Finally, the proposed inspection and evaluation methodology is demonstrated on an in-service bridge in the state of MA.

2. Results

2.1. Laser scanning

To demonstrate the high effectiveness of 3D laser scanning for corroded bridge girders, we scanned a girder from a decommissioned bridge in the state of MA (Fig. 1a). According to the inspection reports which were accessible to the authors, the deck of the three span viaduct underside had widespread leakages, which led to widespread surface rust with areas of heavy corrosion damage in the girders. The extent of the phenomenon resulted to the posting of the bridge and its eventual replacement with a new one. To imitate challenges that inspectors face in the field, stiffeners were welded to the girder after demolition.

The method to assess the structural capacity of corroded steel bridges begins with the measuring and documentation of deterioration. The current state-of-practice of bridge inspections typically includes a rough sketch of the beam end in the relevant inspection report, as shown in Fig. 1b. To follow commonly used practice, we measured the girder plate thicknesses of the stiffened girder end using an ultrasonic thickness gauge (PocketMIKE by GE). In most cases, inspectors have to overcome accessibility challenges and use single point instrument readings to describe varying thicknesses over a planar surface, using mainly thickness gauges and calipers. The obtained measurements are documented in the form of written text, while in some cases sketches or photos are also included in the inspection report. However, even when sketches are present, no more than one or two single point measurements are reported. Given that section loss is not uniform in the area of deterioration, the representative points of measurements currently depend on the judgment of the individual inspector.

In contrast, thousands of points can be captured instantly with laser scanning technology. After post-processing the obtained point clouds, contour maps illustrating the remaining thickness profile of the first two web panels is generated (Fig. 1c). During the in-service period, the beam had a steel diaphragm riveted on the upper part of the web and upon its removal, perforations were revealed, as illustrated for y between 300 mm and 400 mm. Additionally, the scan also reveals that the girder's corrosion pattern is characterized by a diagonal area of extensive section loss, which, for x greater than 200 mm, propagates in parallel with the bottom flange (blue color). The minimum remaining thickness is 4 mm, located approximately at a distance of 300 mm from the web's edge.

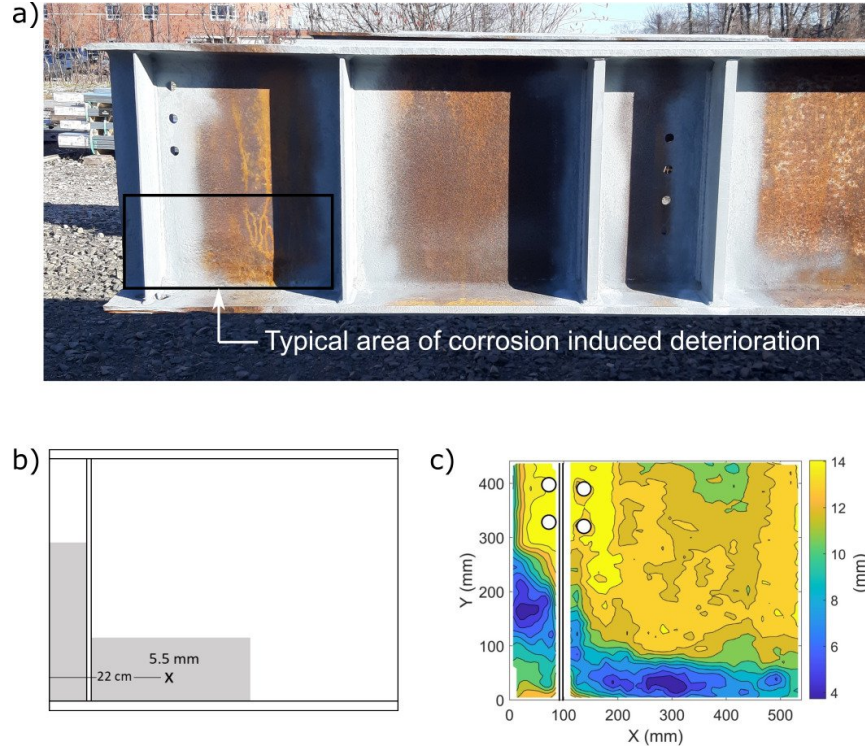


Figure 1: Conventional and proposed data-acquisition inspections for a corroded girder end. a) A naturally corroded girder from a bridge in the state of MA is inspected in laboratory conditions, and its deterioration is documented based on the (b) current and (c) proposed framework. b) Usually a unique thickness measurement is taken and reported in inspection reports today. In some US States documentation is based only on visual observation. c) By post processing the obtained point clouds, detailed thickness contours are developed for the residual thickness profile visualisation including thousands of thickness data points.

2.2. Structural capacity and mode of failure of deteriorated beam

Understanding the mechanical behavior of deteriorated bridge girders is key to developing tools for assessing their capacity. In that direction, we test a naturally corroded girder obtained from another bridge in MA using the testing set up of Fig. 2a. The thickness contour map developed by scanning the girder is shown in Fig. 2b; the specimen is a 62 cm deep I beam (24CB120) found in the superstructure of a simple span steel bridge. For the mechanical testing (Fig. 2a), an upward vertical force is applied by two 890 KN hydraulic jacks at the bearing of the corroded end, simulating the reaction force of an in-service girder. The

specimen is held down by a cross beam anchored to the strong floor and to achieve a shear dominated failure of the web, the cross beam is placed 1.2 m from the loaded end. The specimen is instrumented with two 890 KN load cells located at the lower end of each rod, and a 440 KN one beneath the far end of the girder, to record the applied load.

The beam fails by web buckling at an applied peak load of 478 KN (Fig. 2c and 2e). The failure mode is characterised by the sudden formation of a buckling wave at the lower part of the web that extends in parallel with the bottom flange, a location consistent with the extensive section loss area, illustrated with dark blue shades in Fig. 2b. The obtained failure mode provides strong evidence that the web condition at the lower part of the web governs the residual capacity of the steel girder, and subsequently of the whole bridge's superstructure. This finding is consistent with the outcome of previous research by the authors [37], and determines the area of focus for the developed CNN models.

The developed contour map of Fig. 2b is used to build a finite element model of the beam ([38, 29]) mirroring the condition of the girder (Fig. 2d). This model is then used to simulate the experimental testing and after analysis, Fig. 2e presents comparison between the experimental and computational results. The failure load is captured with an error of just 4.1%, providing evidence that the use of contour maps, and their integration to a computational model enables the accurate residual capacity prediction of corroded girders. This validated model is used later to compute and assign capacities to the artificial corrosion scenarios implemented as the data set for the training of the CNN models. The artificial corrosion scenarios are generated based on real corrosion data, implementing a methodology presented in the Methods section.

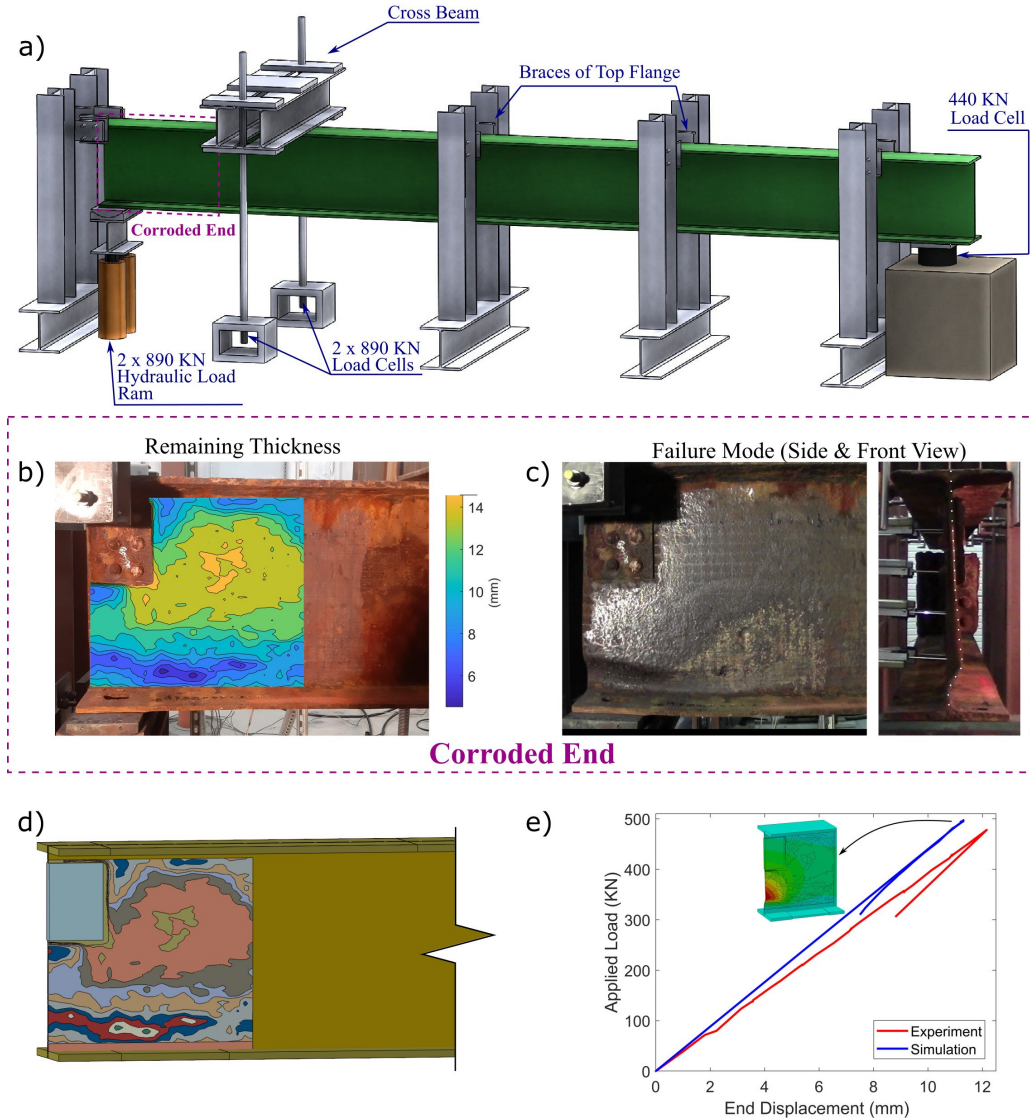


Figure 2: For the generation of the CNN dataset, a finite element model is validated based on the full scale testing of a decommissioned bridge girder. a) The experimental configuration was designed to imitate the loading of in-service girders. **b)** Obtained contour map via 3D laser scanning and projected on the examined end before testing, and **c)** the obtained buckling failure mode. The developed finite element model **d)** encapsulates the exact beam condition by integrating the thickness contour maps and **e)** captures the experimentally obtained capacity with an error of 4.1%.

2.3. CNN modeling

CNNs consist of a class of artificial neural networks (ANNs) able to recognize pattern within images, and in this work, the image input is constituted by the thickness contours located at the lower 60% of the girder web depth (Fig. 3a). We used two separate tools for evaluating the capacity of corroded bridge girders, utilizing CNN classification algorithms and CNN regression algorithms respectively. CNN classification algorithms classify contour maps in groups of structural capacity with a resolution of up to 10% relatively to the structural capacity of an intact girder (no section loss). For the particular problem, we used three classification models with different resolutions: a two-class model (0-60% and 60%-85%), a three-class model (0-55%, 55%-70% and 70%-85%) and a five-class model (0-45%, 45%-55%, 55%-65%, 65%-75% and 75%-85%). On the other hand, CNN regression algorithms directly predict the maximum load that the deteriorated girder can bear before it fails.

For both tools, the CNN models are trained with a dataset of artificial corrosion scenarios which are associated with their respective girder structural capacities found using the finite element model developed previously. The artificial data set includes corrosion contours on a 33WF130 beam, a beam section commonly found on the superstructure of aging bridges in the New England region. The data set of 1421 scenarios is shuffled and split into 64% training, 16% validation and 20% testing data. Fig. 3 presents the results for the two types of evaluation tools developed for predicting the capacity of corroded bridge girders.

The performance of the classification models is assessed by using the unseen test data set for the CNN and visualized with confusion matrices shown in Fig. 3b-d. Using as input just the thickness contour maps, the CNN is able to predict the remaining capacity of deteriorated girders with a remarkable balanced accuracy of 0.91 for five classes, 0.94 for three classes and 0.98 for two classes. Furthermore, the sample imbalance (Fig. 7a-c) in the different remaining capacity classes does not influence the classification performance, which is represented by an equal weighted *f1-score* and balanced accuracy score, as shown in Table 1. It is worth noting that for the five classes classification, which is the most demanding one, the false predictions are distributed similarly over all the remaining capacity classes, except the remaining class 45% – 55%, which shows the lowest *f1-score* of 0.88.

The Root Mean Square Error (RMSE) serves as the metric for assessing the performance of the regression-trained model on unseen external test data. Impressively, the model exhibits an error of just 3.3% for predicting the remaining capacity. Fig. 3e includes the comparison between CNN predicted remaining capacities (depicted in black) and capacities as computed by the finite element model (perfect prediction depicted by the diagonal solid line). The majority of CNN prediction points are very close to the perfect prediction and they also reside in the upper and most conservative half of the plot, showcasing the safety-focused nature of the developed evaluation tool.

To further assess the performance of the CNN models, we are comparing their predictions with the analytical solutions that are currently being used in the state of MA for the rating of deteriorated beam ends [39]. These analytical solutions have been developed in the last years and they have been based on an extensive experimental program on real deteriorated girders [38]. In brief, this analytical solution includes parameters such as the web deviation from straightness (which in this case is set to 50% of the intact web thickness) and the Nd^{-1} ratio, where N denotes the bearing length and d the beam depth. The remaining structural bearing capacity of the deteriorated girder is then calculated by Eq. (1):

$$R_n = [0.32 \sqrt{EF_y t_f} t_{ave}^{1.5} + 0.5 \left(\frac{0.33d}{N} \right) \left(\frac{4N}{d} - 0.2 \right) \frac{\sqrt{EF_y t_f}}{t_f^{1.5}} t_{ave}^3] \left(\frac{CL}{N + 0.1d} \right)^{0.4} \quad (1)$$

where, E is the Young's Modulus, F_y is the steel yield stress, and CL is the length of the corroded region

Residual capacity (%)	Precision	Sensitivity	Specificity	F1-score	Accuracy
Two classes					
0 – 60	0.98	0.97	0.99	0.98	0.98
60 – 85	0.98	0.99	0.98	0.98	0.98
Weighted overall	0.98	0.98	0.98	0.98	0.98
Three classes					
0 – 55	0.97	0.95	0.98	0.96	0.97
55 – 70	0.88	0.91	0.96	0.90	0.95
70 – 85	0.98	0.98	0.98	0.98	0.98
Weighted overall	0.95	0.95	0.98	0.95	0.94
Five classes					
0 – 45	0.91	0.91	0.99	0.91	0.98
45 – 55	0.91	0.85	0.98	0.88	0.95
55 – 65	0.88	0.94	0.97	0.91	0.96
65 – 75	0.89	0.92	0.97	0.91	0.96
75 – 85	0.96	0.94	0.98	0.95	0.97
Weighted overall	0.92	0.92	0.98	0.92	0.91

Table 1: Classification performance with a sample weighted overall score. The equal weighted *f1-score* and the balanced accuracy score denote an equally good performance for all classes. The lowest *f1-score* is captured for the remaining capacity class 45%-55%.

of the bottom part of the web, bound by a rectangle with a base equal to $N + 0.1d$ and a 10 cm height. Due to the absence of holes in the area of interest, t_{ave} describes the remaining web thickness.

Applying the analytical solution, estimations can be made for the same data set. We have considered both the minimum (blue in Fig. 3e) and average (red Fig. 3e) thickness within the area of interest as remaining thickness values. When the minimum thickness is used, the predicted capacities form six clusters of points that run parallel to the y-axis, reflecting the six thickness levels considered in the artificial data. Conversely, the red points are scattered, covering a broad range of loads. The predictions using the minimum thickness as corrosion input in the analytical tool result in overly conservative predictions, and a RMSE of 261.6 KN. On the other hand, when assuming the average thickness, the estimated capacities are dominated by over-estimations, and a RMSE equal to 85.3 KN.

Despite the experimentally demonstrated superiority of the implemented analytical tool over various other equations, as shown by [26], the provisions currently incorporated in the MassDOT LRFD Bridge Manual [39], show a larger scatter in the predictions of the capacities of deteriorated girders when compared to the CNN models. This is rather expected for two main reasons. First, the analytical prediction tools are based on envelope solutions developed by a brute-force finite element analysis of thousands of corrosion scenarios, therefore, the accuracy of the predictions is limited. Secondly, the corrosion input in the analytical solutions is limited to one single value (minimum or average in this study), while the CNN models take advantage of the wealth of information coming from laser scanning.

The developed CNN models demonstrated accurate performance with the artificial test data. However, it could be argued that the proposed tool is beam type specific (33WF130), as well as that both the training and test data appear simplified compared to the random nature of the corrosion on a beam end. For data handling and computational purposes, when creating the implemented artificial scenarios, each contour level represents a change in thickness of 15% of the nominal web thickness; in total this results to six contour levels (Explained in Methods). On the other hand, processing of real point clouds can potentially lead to

visualizations with increased thickness resolution.

To examine whether these aspects can compromise the performance of the proposed evaluation tools when handling real data, corroded girders of different cross sections that have not been implemented in any way on the generation of the artificial scenarios are now used. Five decommissioned beams (Beams I to V - 33WF130) obtained from the state of Maine and three beams (Beams VI to VIII - 16WF40) from the neighbor State of New Hampshire (Fig. 4a) are 3D scanned under laboratory conditions using the Reigl VZ-2000 laser scanner and the Artec Leo 3D scanner. The acquired point clouds were processed to generate two groups of two-dimensional contours for each girder. The first group includes contours with increased resolution in terms of corrosion values (Fig. 4b), and shape complexity, similar to the representation that was implemented to accurately capture the peak load of the experimentally tested beam previously. For the second group of contours (Fig. 4c) a simplified version colormap consistent with the trained model is preferred to serve as the input to the CNN models.

To computationally evaluate the residual capacity of the eight girders, and set the benchmark for the performance evaluation of the developed tools, finite element analysis of all the beams is performed using laser scanning data (Fig. 4b). For the CNN input, the models developed previously used a region covering 60% by 100% of the web height along the height and longitudinal direction, respectively. The obtained point clouds covered a region slightly less than the CNN input domain and to comply with the required input for the trained models, minimum section loss is assumed along the missing parts, resulting to the domain with the uniform orange colour at the right side of the contours at Fig. 4c. In both FEA and CNN cases, 250 MPa bilinear steel material law is assumed due to its frequent use in the New England region during the period that the bridges were constructed.

The performance of regression and classification models on real corrosion scenarios is assessed through a direct comparison between the computationally obtained results and the CNN predicted values (Table 2). The obtained capacities are normalized with respect to the capacity of a girder without section loss for ease of comparison with the predicted values. The regression model shows promising results for Beams I to V, yielding a Root Mean Square Error (RMSE) of only 4.8%. The estimations for Beams II and IV demonstrated exceptional precision, with errors in the range of 0% to 4% of the nominal section capacity (952 KN). Even in the case of Beam I, which displayed the weakest performance, the deviation was limited to 10% of the nominal capacity, highlighting the model's effective ability to predict structural capacities. Equally impressive was the performance of the classification models. In the most demanding model, featuring five classes, four beam capacities were predicted within the respective ranges or at the boundary values of each class, demonstrating an accurate capacity estimation. Beam I was overestimated in both the three and five-class models. Finally, for the two-class model, the failure loads of all five beam ends were predicted correctly.

For Beams VI to VIII, which describe a shallow web section with extensive section loss (16WF40), the developed classification models correctly predicted all three cases. Regarding regression, the capacities of Beams VII and VIII are estimated with an error of 2% and 3% relative to the capacity (402 KN). The weakest performance is observed for Beam VI, which is overestimated by 13%. It is important to note that the analytical equations (Eq. 1) are different for specific ranges of bearing length (N) to web depth ratios (d). In the scenarios we examined, we are exploring a wide range of Nd^{-1} ratios between 0.30 and 0.63 (a bearing length of 25 cm both for 33WF130 and 16WF40) and the developed tools perform remarkably well.

Capacity	Beam I	Beam II	Beam III	Beam IV	Beam V
Type	33WF130	33WF130	33WF130	33WF130	33WF130
Actual	47%	35%	42%	54%	72%
Regression	57%	35%	38%	57%	70%
Two classes	less than 60%	less than 60%	less than 60%	less than 60%	60% – 85%
Three classes	55% – 70%	less than 55%	less than 55%	55% – 70%	70% – 85%
Five classes	55% – 65%	less than 45%	less than 45%	55% – 65%	75% – 85%
Capacity	Beam VI	Beam VII	Beam VIII		
Type	16WF40	16WF40	16WF40		
Actual	14%	40%	33%		
Regression	27%	38%	36%		
Two classes	less than 60%	less than 60%	less than 60%		
Three classes	less than 55%	less than 55%	less than 55%		
Five classes	less than 45%	less than 45%	less than 45%		

Table 2: Capacity evaluation results from validation in laboratory conditions The developed evaluation tools are validated on real corrosion scenarios, shown in Fig. 4. The computationally obtained capacities are labeled as "Actual" and serve as the benchmark against which the results from the trained classification and regression models are compared.

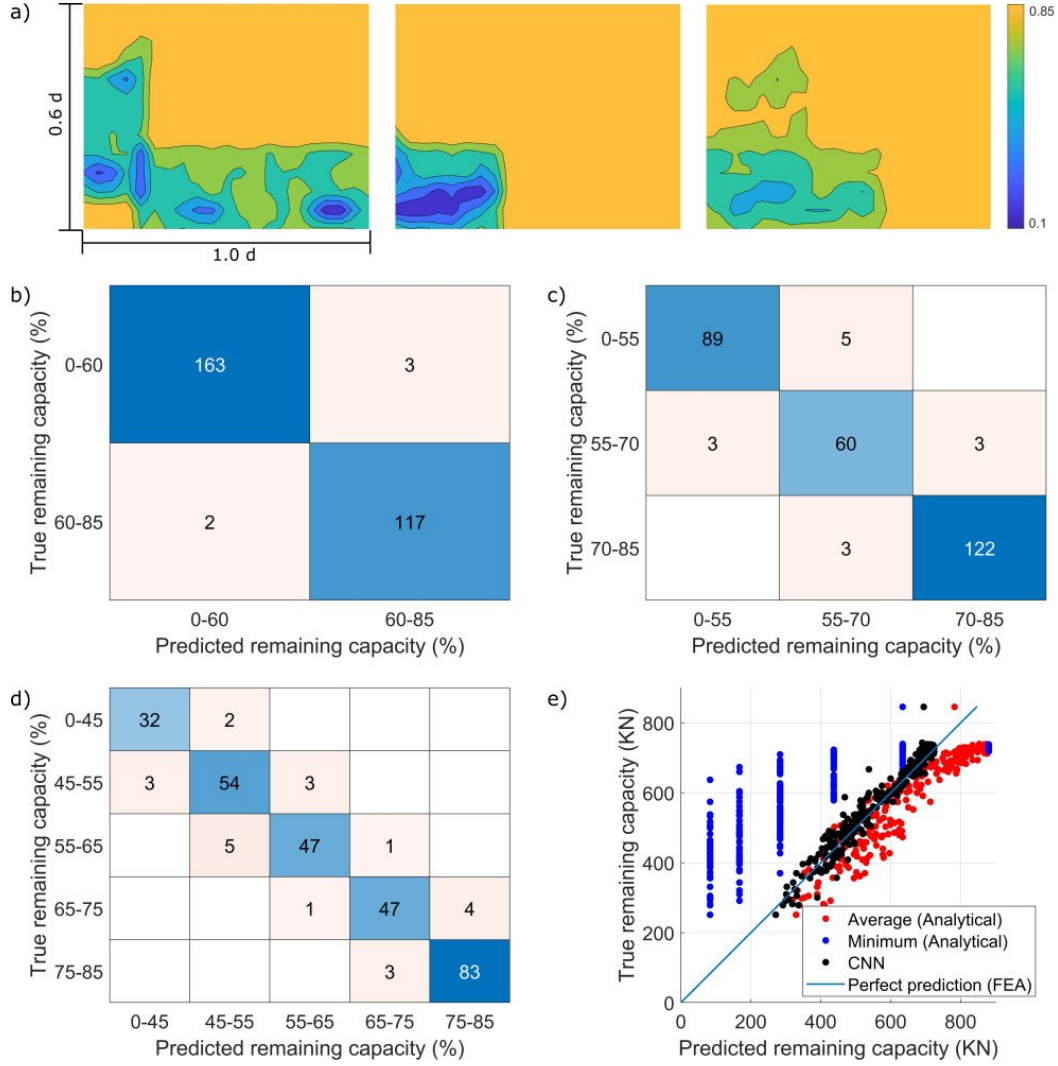


Figure 3: Performance of CNN models in predicting the structural capacity of deteriorated girders. a) The models evaluation is conducted based on artificial corrosion data developed with a methodology presented in the methods section. For the classification models, confusion matrices for **b)** two, **c)** three, and **d)** five classes. Diagonal elements represent the number of points for which the predicted remaining capacity range is equal to the true range. **e)** For the regression model, the blue solid line represents the perfect prediction, while estimations lying above this line underestimate the actual capacities. Black color points illustrate estimations with the proposed CNN Model, points in red and blue are predictions using the average and minimum thickness as corrosion input to an analytical tool, respectively. Nowadays, structural evaluation of corroded girders is conducted with the use of similar equations. Capacities for subfigures **b**, **c** and **d** are normalized with respect to the capacity of 33WF130 girder without section loss.

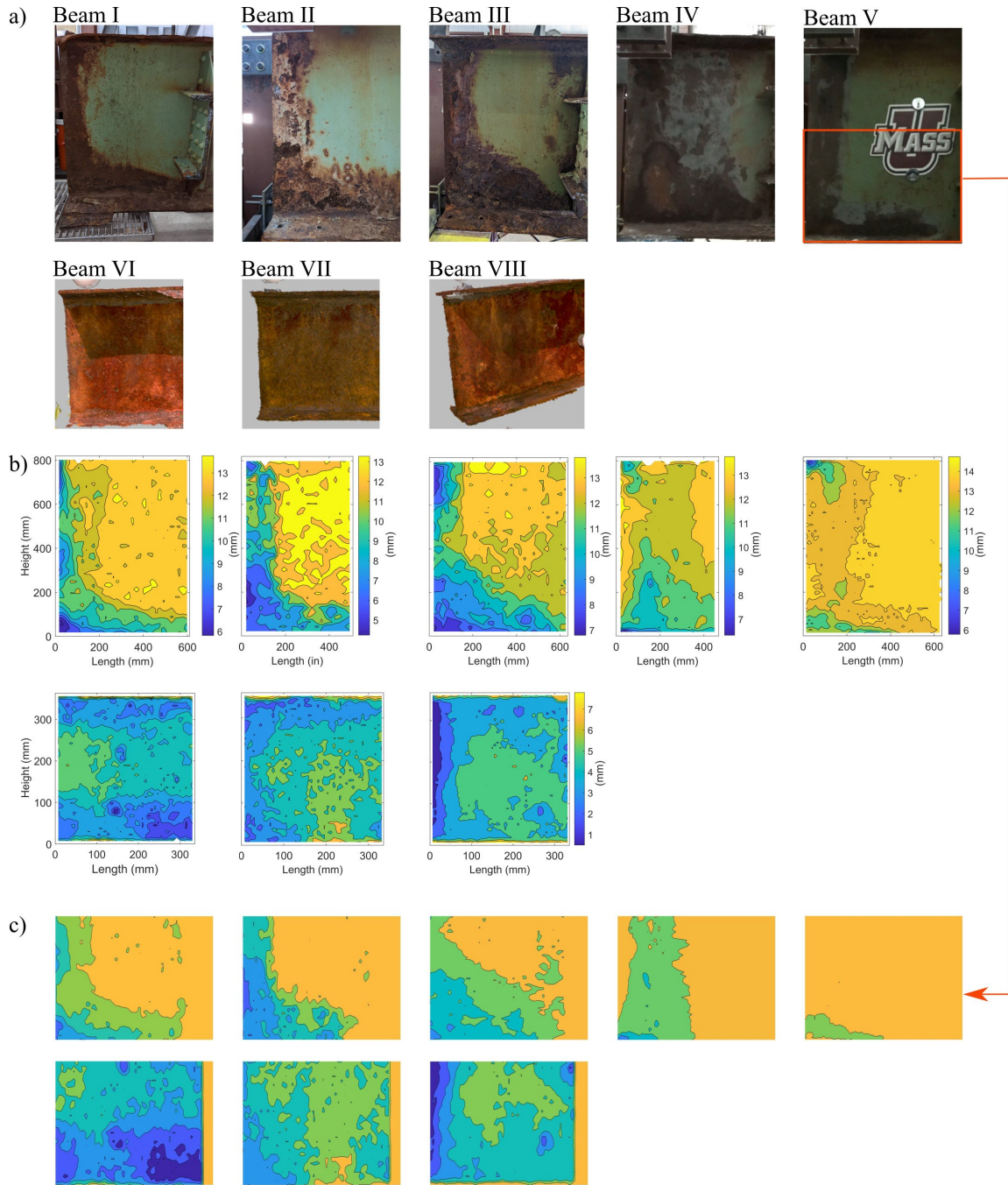


Figure 4: Validation of the developed framework on real corrosion scenarios. a) Eight decommissioned bridge girders are scanned. **b)** Based on the captured point clouds, detailed contour maps are generated for computational calculation (FEM) of the residual bearing capacity. **c)** Contours maps in compliance with the input format of the trained CNN evaluation tool in terms of colourmap, size and number of contour layers are also generated. Results obtained both for the classification and regression models are presented in Table 2.

Capacity	In-Service Beam I	In-Service Beam II	In-Service Beam III	In-Service Beam IV
Type	33Bx11½	33Bx11½	33Bx11½	33Bx11½
Regression	862 KN	934 KN	851 KN	832 KN

Table 3: Capacity evaluation results for corroded girders of an in-service bridge. Capacities are obtained by applying the CNN regression model to the thickness contours presented in Fig. 5c.

2.4. Application on an in-service bridge

Although the developed CNN- and laser scanning-based framework has demonstrated exceptional results in the controlled laboratory environment, its true potential can only be fully assessed when deployed in real-life conditions for typical steel bridges. We investigated its applicability, feasibility and performance concerning equipment mobility, data acquisition and required labor by scanning and evaluating an in-service highway bridge located in Massachusetts, (Fig. 5a). The selected viaduct constitutes the ideal testing ground due to its geometric characteristics. The bridge design includes beam ends partially stiffened with a portion of the end encased by a concrete diaphragm (Fig. 5b), an aspect that hinders data acquisition.

In contrast to the terrestrial laser scanner preferred during the development phase, on site scans were taken with an Artec Leo mobile 3D scanner. Accessing the beam ends was accomplished using a bucket truck and a ladder (Fig. 5a), which are typical methods employed during bridge inspections. The selected scanner is light, wireless and compact enough that it can be carried by an inspector up to a ladder or in the bucket truck. Additionally, the Artec Leo 3D scanner encourages constant scanning with lots of movement. These technical characteristics, have allowed us to capture on site data with limited error from vibrations and other accidental movements during the inspection process. Each beam end of interest was recorded from approximately 0.75 m distance and due to accessibility constraints, each side of the beam ends had to be captured separately and aligned using as reference reflective targets and in-house made algorithms. The corrosion of the scanned beams was also extensively measured by an ultrasonic thickness gauge and a slide caliper as a way to cross check the intact thickness of the observed point cloud and contour map.

Following the field inspection, the acquired point clouds undergo post-processing to create informative representations, which describe the web condition (Fig. 5c). At all four beam ends, the bright yellow colour represents regions of high values due to the presence of partial stiffeners and concrete diaphragms. All the beams share a similar deterioration condition, governed by the design configuration. Distinct areas with section loss are depicted between the outer web edge and the diaphragm, as well as at the lower part of the web above the bottom flange. For capacity prediction purposes, images compatible with the trained CNN models are generated and presented in Fig. 5d. Similarly to conventional inspection and evaluation techniques, the impact of partial stiffeners and diaphragms on the residual bearing capacity is not considered and, therefore, neglected as CNN input.

The main benefit of the developed CNN-based model is to overcome the need for detailed finite element modeling/analysis of the corroded girders. The CNN regression model is able to provide capacity predictions using just the input presented in Fig. 5d. The final step to reach a numeric prediction would be to denormalize the prediction using the capacity of the intact girder. For that reason, we have created a library of intact girder capacities for common beam types (with a maximum web deviation amplitude of 0.5 times the nominal web thickness) and we have used these capacities in conjunction with the CNN prediction to reach the final result. For the girders of the bridge under study the results are given in Table 3. Using these values, a load rating engineer can directly assess the maximum traffic load based on the bridge geometric characteristics.

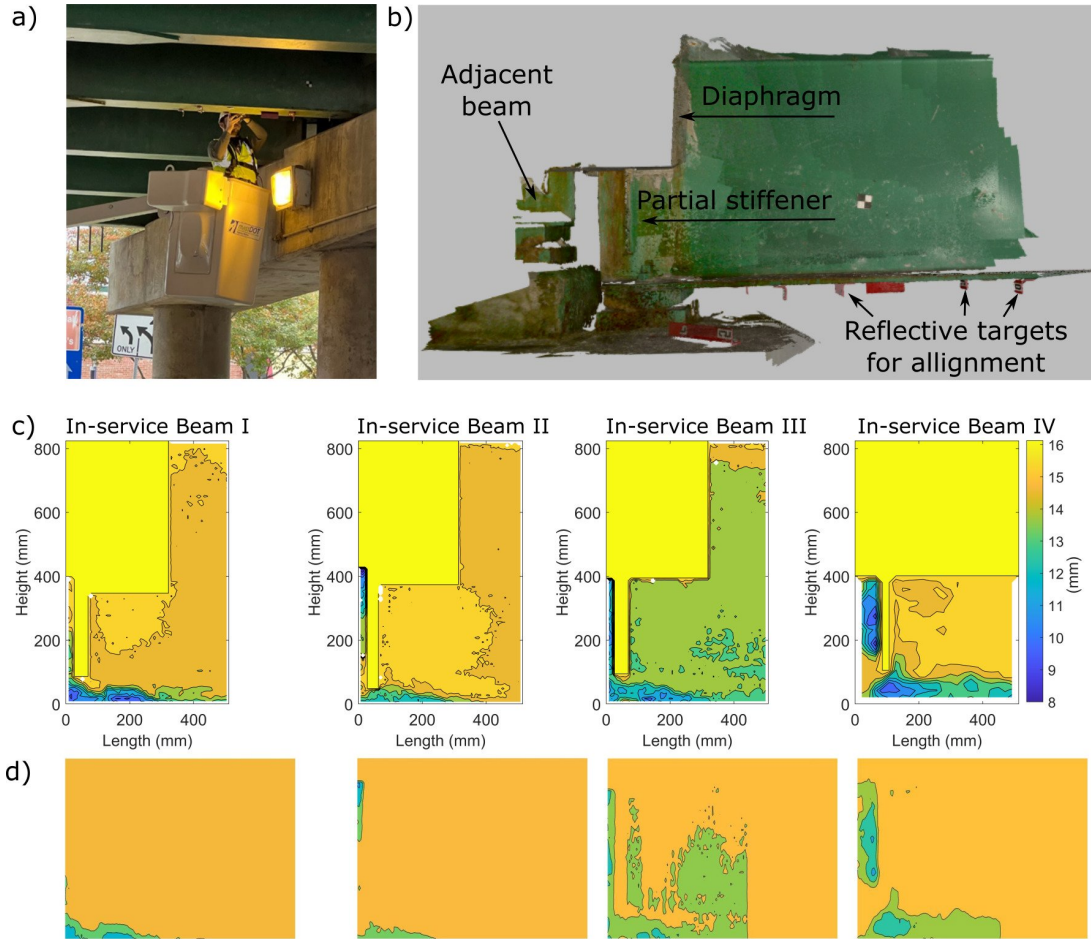


Figure 5: Inspection and structural integrity evaluation of an in-service bridge. a) Field 3D scanning with portable equipment. b) Rendering of a scanned beam end, the four examined beams are partially stiffened with a concrete diaphragm above the bearing. c) Contour maps for documentation of the corrosion phenomenon in the inspection reports. d) Input for the CNN models. The existence of diaphragms and the partial stiffeners is not taken into account similarly to the conventional evaluation techniques.

3. Discussion

This work has developed and validated a complete bridge inspection/rating workflow based on laser scanning and machine learning, which can be applied in real-world conditions. First, we have used 3D laser scanning technology to accurately capture the deterioration of corroded steel girders. Second, a post processing framework generates 2D contours depicting the remaining thickness profile along the web of the girders. Finally, the contours are fed to a CNN-based tool to estimate the remaining bearing capacity of aging girders regardless of the cross section type. To develop this framework we have used naturally corroded girders from the region of New England to obtain point-clouds of the corroded parts, we have experimentally tested the corroded girders to determine their failure mode and their peak capacity and based on these results, we have developed a generalized finite element model to compute the peak loads of deteriorated beam ends. Combining these components with two different CNN models we have put in place the new framework for a quick and accurate load rating methodology of deteriorated steel bridges.

We chose contours as a two-dimensional representation of the remaining thickness profile along the deteriorated area. This approach is easily processed from inspectors and engineers, it provides an overall description of the examined area, and it can be integrated in the inspection reports to upgrade the corrosion mapping. In terms of inspection, laser scanning outperforms the conventional techniques by collecting thousands of points along the corroded domain of the girder. Currently, inspectors are equipped with an ultrasonic thickness gauge and are solely based on their judgement to collect a unique thickness measurement that shall be representative of the whole beam end. As demonstrated in [37], buckling of deteriorated beams typically initiates at the web bottom, particularly at the location with the most severe section loss. However, to pick the location of a representative thickness measurement, the inspector is currently relying solely on visual observation of each individual web face, which may present fundamentally different deterioration profiles. In contrast, the proposed visualization approach considers both web faces and precisely identifies locations that require inspection in future assessments, even if inspectors have basic equipment, such as an ultrasonic thickness gauge. This comprehensive visualization method offers a more reliable and thorough inspection process.

The proposed methodology has limitations in accurately capturing the geometry of delaminated cross-section. The delamination of the steel hinders the data acquisition or leads to overly optimistic estimates, regardless of the inspection technique. With conventional thickness gauges, inspectors must use a hammer in the field to expose the steel before applying a coupling gel and taking single-point measurements. Depending on the severity of the delamination, the proposed methodology may also require removal of the steel fractures if they are within the area of interest. Cleaning the surface of the beams prior to scanning may increase the cost of the inspection activities, but this task is required even with today's techniques. On-going research efforts by the authors examines the trade-off between accurate capacity predictions and presence of delamination to determine the level of delamination that can be tolerated.

The training of the CNN model required the creation of a data set consisting of artificial corrosion scenarios in 2D contours format and the associated capacities. More than 1400 artificial corrosion scenarios were generated by parameterizing real scan data from 3 naturally corroded girders. The main advantage of the proposed CNN-based capacity assessment framework is that it directly uses the post-processed laser scanner output and does not require any further calculations or post-processing. Although the creation of 1 to 1 finite element models that encapsulate the exact web condition can be materialized with the help of the semi-automated scripts-based procedure developed, implementing a finite element model increases the complexity and the time required to build the model and demands challenging theoretical background knowledge. In contrast, the CNN-based model presented in this work is easy to use and to be embedded in existing inspection workflow. When compared to the evaluation approaches using analytical expressions,

the developed machine learning tool considers not only a single thickness value, but a wealth of information coming from the point-clouds for the area of interest.

A finite element model was required to calculate the capacity of the artificial corrosion scenarios. The implemented model was validated based on experimental results obtained from the full scale testing of a naturally corroded decommissioned girder from the state of MA. Comparison of experimentally and numerically obtained results provided credibility to the developed finite element model with the failure load being captured with 4.1% error.

Both classification and regression models are trained for capacity evaluation. For classification two, three and five classes are considered. Each class represents a range of load magnitude to describe the residual bearing capacity of a corroded beam end. For five classes, a maximum resolution of 10% of the nominal bearing capacity is achieved. The developed CNN models are tested with artificial data and achieved accuracy of 0.98, 0.94 and 0.91 for two, three and five classes, respectively. For the regression model an RMSE error of just 0.033 is obtained.

The developed evaluation tools performed exceptionally for different types of cross sections with depth between 84 cm and 41 cm. In addition to the artificial data, the performance of the developed models is tested with contours from eight decommissioned girders that had not been previously used for the generation of artificial scenarios. For each beam a finite element model is built and the obtained computational capacities serve as the benchmark. For two classes, all predictions are accurate, while for three and five classes the capacity of one particular beam (Beam I) is overestimated. The maximum and minimum error of the regression model equals to 0% and 12%, respectively.

Our methodology provides a game changing tool for dealing with the multiple uncertainties that stem from data acquisition, biased engineering judgement and analytical evaluations, which currently result in unreliable estimates of remaining capacity of corroded steel bridge girders. The tools developed will help with both failure prognosis and decision making for retrofitting and upgrading steel bridges, preventing major impacts to the transportation network and the increased emissions associated with them. Closed lanes, detours, and unnecessary repairs not only impact local communities, but also prevent optimal management of funds, especially considering that the Federal Highway Administration estimates a \$125 billion repair backlog on existing bridges. Finally, the proposed method is not only beneficial on purely financial criteria, but also promotes a more sustainable and resilient construction industry. Optimized decisions about bridge closures, repairs or demolitions are an important contribution in this direction.

4. Methods

4.1. Point clouds post processing

In the development phase of the methodology, a decommissioned beam is scanned under laboratory conditions using a Riegl VZ-2000 TLS. Both web surfaces of the same specimen are scanned and the point clouds obtained are aligned in a unique coordinate system, making use of the free source software Cloud Compare. Following, noise removal, a mesh is built from each face based on Delaunay triangulation. Subsequently, one side acts as the reference and the other as the compared mesh. The distances between the vertices of the compared mesh with respect to the reference mesh are computed and the obtained values, which correspond to the thickness of the section, are extracted as a scalar field associated to the reference mesh's points. The least squares algorithm is used to select the best fitting plane to the reference side, which is subsequently rotated with scripts in Matlab to scale down the problem from 4 (*x coordinate*, *y coordinate*, *z coordinate*, *thickness*) to 3 (*x coordinate*, *y coordinate*, *thickness*) parameters. Following an approach previously proposed by the authors [29], the resulting *x*, *y*, and *thickness* data are employed to generate thickness contours, which are presented in Fig. 1c.

4.2. Finite Element Validation for Data Generation

The commercial finite element software ABAQUS [40] is utilized for the simulation of the conducted experiment. The geometry is described as a three dimensional body discretized with S4R shell elements. The corroded end and the bearing areas are meshed with element size of 1.3 cm. For computational efficiency, the rest of the beam is discretized with elements sized up to 6.5 cm. Thickness is assigned as property to a surface in the middle of the section. The exact deterioration condition is integrated into the model by partitioning the web of the beam to reproduce the exact contour lines illustrated in Fig. 2b. Subsequently, for each web region enclosed by contour lines, the level of the surrounding contour with the minimum value is assigned as uniform remaining thickness. Dogbone coupons are extracted and tested for material properties definition.

Two-step analysis is carried out. During the first step, quasi-static analysis is performed to account for the self-weight of the beam, applied as a uniform pressure along the top flange. During the second step, Riks analysis is preferred to capture both the pre and post failure behaviour of the specimen. The operation of the hydraulic rams is idealized as a point load assigned upwards to a hinge at the location of the bearing supporting the specimen. The hinge is tied to the bottom flange along length equal to the length of the bearing. Since no web deviation from straightness was identified during the post-processing of the point clouds, the first eigenmode of the specimen is scaled to ten percent of the intact web thickness (1.4 mm) and imported as imperfection for the two-step analysis.

4.3. Data Generation

Three naturally corroded girders from two decommissioned steel bridges, built in the first half of the 20th century in the state of Massachusetts, form the basis for the design of the parametric space for training and testing the performance of the CNNs, Fig. 6a. Fig. 6b depicts the contours describing the three naturally corroded girder ends. While in service, all specimens had diaphragms attached to the upper half of the web above bearing. At Specimen 1, the diaphragm was still on the beam, at the time of scanning, while for Specimens 2 and 3, diaphragms had been removed revealing perforations.

Previous numerical work by authors [38] has mapped the relationship between corrosion topology and residual capacity. By simplifying the corrosion induced damage as a rectangular region with uniform section loss, an extensive parametric analysis had revealed that the harmful effect of deterioration height is mostly limited to the bottom 30% of the web. Consequently, describing damage along the entire web height would dramatically increase the complexity of the artificially generated data and provide CNN information with minimal impact on capacity estimation. Thus, the parametric space for the artificial data is limited to 60%, and 100% of the web depth along the vertical and horizontal directions, respectively.

To initiate corrosion scenarios generation, the contours maps of the three actual girders are described as a synthesis of multiple single points in the two dimensional space (x, y) by removing the connecting lines. For each beam, several multivariable Gaussian distributions are fitted to the remaining joint points Fig. 6c. This approach takes into account the density and the location of the points, and not the magnitude of the remaining thickness, nevertheless it captures the shape of the corrosion pattern of each beam, allowing the researchers to control the location of maximum section loss by parametrizing the covariance values (mean vector and covariance matrix). The maximum section loss is scaled between 30% and 90% of the intact web thickness. Ultimately, to create a more realistic representation of the topologically non-uniform corrosion phenomenon, we deliberately incorporate noise into the scenarios. The magnitude of this noise is carefully controlled for each case to prevent the generation of negative thickness values. This control involves random reductions in the remaining thickness at the web bottom within a range of 0.14 to 0.8 times the maximum section loss. Fig. 6c illustrates indicative scenarios generated with the use of in-house developed scripts for each of the actual corroded beams.

Observation of numerous inspection reports as well as in-field visits at aging bridges along the six states of New England (Maine, Vermont, New Hampshire, Massachusetts, Connecticut and Rhode Island) has revealed that beams with web deviation from straightness of varying maximum amplitude are still in service. Web deviation is mainly interpreted as a sign of buckling or in some cases as an indication of the early twentieth century manufacturing limitations. Previous experimental and numerical research efforts by the authors [37, 38] have addressed the deleterious effect of web deviation from straightness to bearing capacity, developing analytical expressions for maximum web deviation amplitude of 0.1, 0.5 and 1.0 times the intact web thickness (t_w). In this work, we assume imperfection amplitude of $0.5 t_w$, which is typically combined with section loss along large part of the web above bearing. Consequently, section loss of minimum 15% (illustrated with the orange color in 6c) is considered for the artificially generated data. For the numerical simulations the out of straightness of the web is taken into account by introducing and scaling imperfections into the model based on the eigenmodes shapes. This is consistent with plate theory as imperfection mode shapes are considered among the most deleterious of all imperfection shapes.

Employing the previously validated finite element model, each of the approximately 1400 corrosion scenarios is associated with a computationally determined residual load capacity by projecting and analyzing the studied contours on a 33WF130 beam, a beam section commonly found on the superstructure of aging bridges in the New England states.

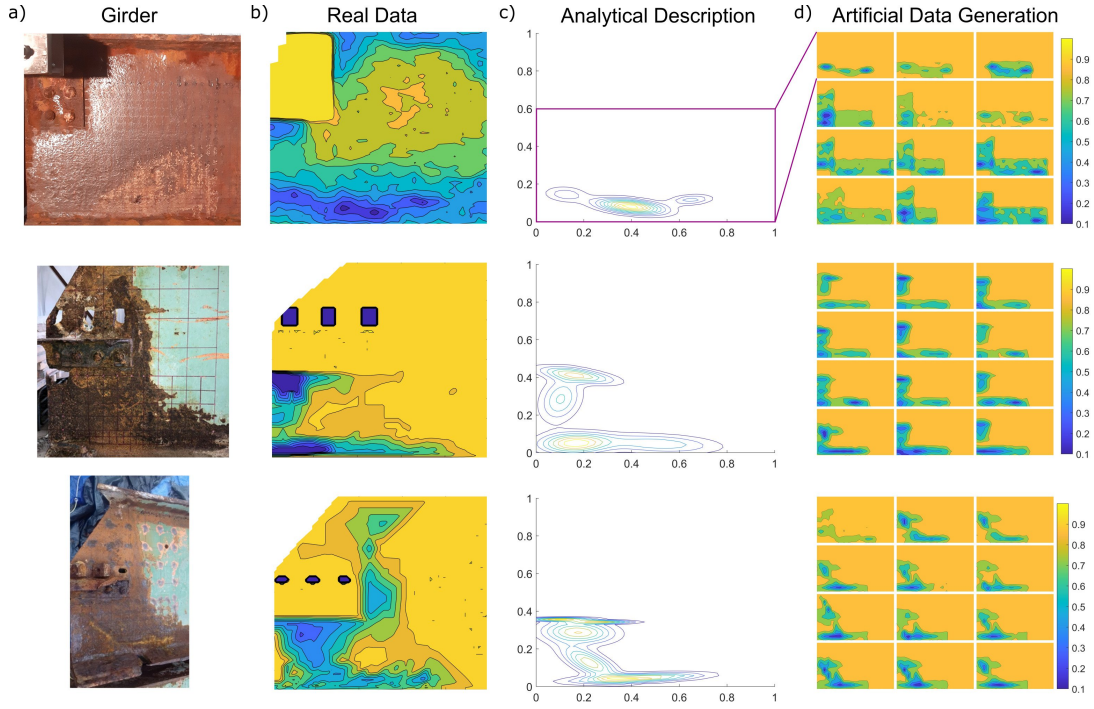


Figure 6: Naturally inspired data set for CNN models training. **a)** Data generation is based on real data from three decommissioned girders. **b)** Thickness contour maps are generated for each beam end, and **c)** several multivariable Gaussian distributions are fitted to analytically describe the section loss profiles. **d)** 1421 naturally inspired corrosion scenarios are generated by parametrizing the distributions locations and their covariance values. Both lengths and thickness are normalized, so no units are illustrated.

4.4. Convolutional Neural Network Training

CNNs are comprised of a high number of interconnected computational nodes (neurons), of which work entwined in a distributed fashion to collectively learn from the input in order to optimise its final output [41]. A CNN is trained through pre-labeled inputs, which act as targets. A data set of deterioration scenarios associated to a unique capacity value each, is created parametrizing real corrosion data and calculating the corresponding capacity making use of a finite element method. The finite element model is validated with full scale experimental testing of a naturally corroded girder obtained from a bridge replacement project in the state of Massachusetts.

The proposed CNNs are trained on images with a size of 539×683 . For ease of reference, each of the 1421 computationally obtained capacities, related to the examined corrosion scenarios, is normalized with respect to the bearing failure load of a 33WF130 girder without section loss. Each scenario is labeled automatically with a tag describing its residual capacity (e.g., between 55% and 65%, or 59% of the as built strength, for classification and regression algorithms, respectively). In case of classification, CNN models are trained for two, three and five classes of residual capacities, presented in Fig. 7a-c. Since the problem is a multi-class classification task, the classes of the data set are one-hot encoded to prevent possible ordinal relationships. The data set is shuffled and split into 64% training, 16% validation and 20% testing data.

Classification CNN models are built by using the TensorFlow [42] deep learning library. The first layer of the CNN is a rescaling layer, which rescales the input data of the RGB color in a range between 0 – 1.

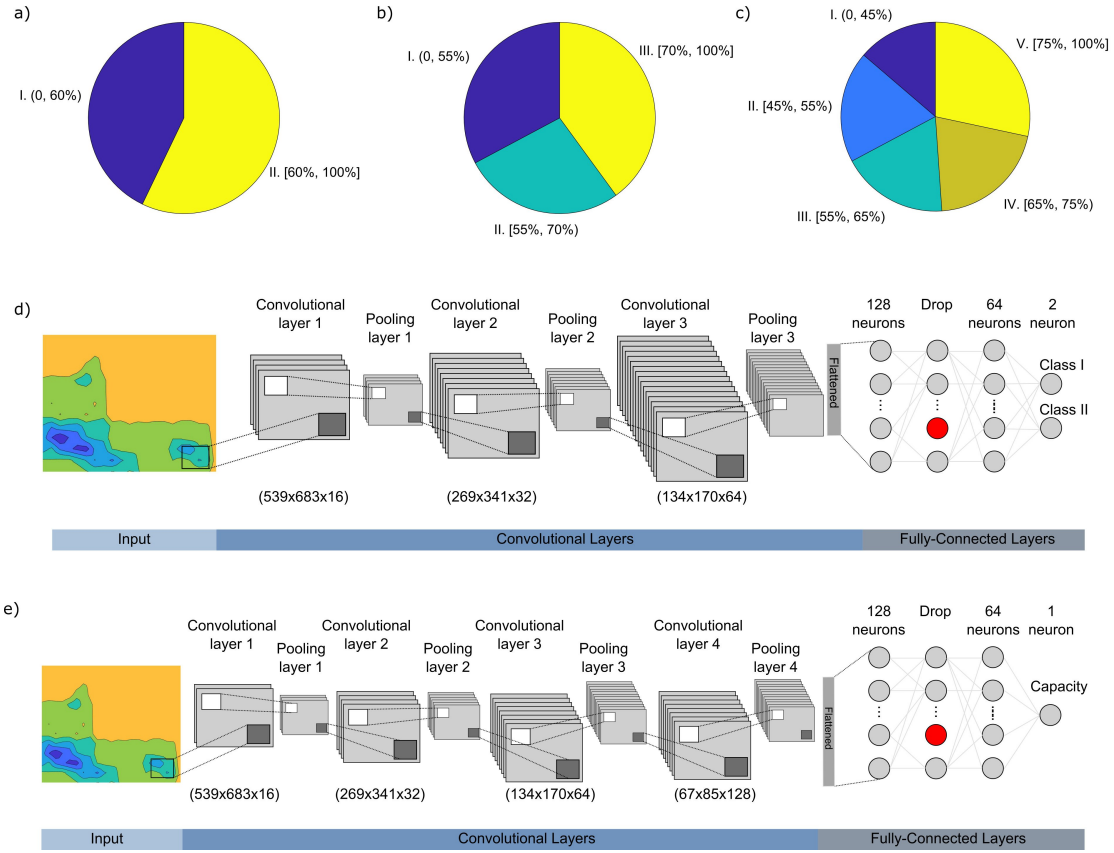


Figure 7: Training the CNN evaluation tools. For classification models, data is distributed into **a)** two, **b)** three, and **c)** five classes, while the regression model is trained on the entire data sample. The convolution neural network architecture for **d)** the classification model with two classes, and **e)** the regression model. In the dropout layer, the red color depicts neurons that are intentionally ignored to prevent overfitting on the training data.

The following stacked layers consist of multiple Conv2D layers that create feature maps with subsequent MaxPooling2D layers that downsample the feature maps. After the reduction of the feature maps with the stack of convolutional and pooling layers, three fully connected dense layers with a Dropout layer in between are followed.

The ReLU (rectified linear unit) activation function is used for the convolutional layers as well as the first dense layer due to the efficient computation. The last dense layers depending on the classification target consist of two, three or five neurons representing the different classes. Therefore, the last dense layer uses a softmax activation function that assigns the prediction probability to the corresponding class. The architecture of the model is shown in Fig. 7d. The hyperparameters of the CNN layers (Supplementary Table 1) are determined by conducting an optimization with a random grid search from a predefined hyperparameter set. Due to the high computational effort and memory allocation during training of the CNNs, the hyperparameters are chosen in such a way that the computational effort is limited and the CNN is trainable regarding memory consumption of the computer, as shown in Supplementary Table 2.

The parameter depth represents the number of paired convolutional and pooling layers of the CNN model. A higher depth results in a deeper and more complex model. The optimization iterated 100 trials and the best-performing models are saved. In each iteration, the CNN model is trained with a set of randomly chosen parameters from the hyperparameter set. The CNN is trained by using the Adam optimizer and a batch size of 32 samples. During the training of the CNN, an EarlyStopping callback function is used to monitor the validation loss and stops the training when the CNN starts overfitting. As performance measure, the accuracy metric is used with the categorical cross-entropy loss function.

In addition to classification, a Convolutional Neural Network (CNN) is trained on regression to predict the remaining capacity as a numerical value. The same data and dataset splits (train, validation, test) used for classification are also employed to train the CNN. The architecture remains identical except for an additional Conv2D layer and the absence of an activation function in the output layer. Since the task is to predict a single numerical value, the output layer consists of just one neuron, Fig. 7e. The Adam optimizer is used in conjunction with a batch size of 32 and an EarlyStopping callback function. During training, the loss is computed using the Mean Squared Error (MSE), and performance is evaluated using the Root Mean Squared Error (RMSE).

Data availability

The point clouds and the database that support the findings of this study are available from the corresponding author upon reasonable request.

Code availability

The scripts developed for data generation are available from the corresponding author upon reasonable request.

Acknowledgements

This work was supported by the Massachusetts Department of Transportation (grants no. S14000000049794, S13810000051537, S13810000054013) and the New England Transportation Consortium (grant no. S14000000049794). This study was undertaken with the support by Graduate Academy of the Technische Universität Dresden, supported by Federal and State Funds.

Competing interests

The authors declare no competing interests.

References

- [1] ASCE, Infrastructure report card (2021).
URL <https://www.infrastructurereportcard.org/>
- [2] FHWA, LTBP Infobridge DATA, FHWA, US DOT (2023).
- [3] National bridge inspection standards (2022).
URL <https://www.fhwa.dot.gov/bridge/nbis2022.cfm>

- [4] PocketMIKE, Operating Manual, GE Inspection Technologies, Lewistown, PA (2004).
- [5] B. M. Phares, G. A. Washer, D. D. Rolander, B. A. Graybeal, M. Moore, Routine highway bridge inspection condition documentation accuracy and reliability, *Journal of Bridge Engineering* 9 (4) (2004) 403–413. doi:10.1061/(ASCE)1084-0702(2004)9:4(403).
- [6] R. J. Ross, B. K. Brashaw, S. J. Anderson, Use of laser scanning technology to obtain as-built records of historic covered bridges, USDA Forest Service, Forest Products Laboratory, Report Paper, FPL-RP-669, 2012; 21 p. 669 (2012) 1–21.
- [7] I. Lubowiecka, J. Armesto, P. Arias, H. Lorenzo, Historic bridge modelling using laser scanning, ground penetrating radar and finite element methods in the context of structural dynamics, *Engineering Structures* 31 (11) (2009) 2667–2676. doi:https://doi.org/10.1016/j.engstruct.2009.06.018.
- [8] M. Minehane, R. O'Donovan, K. Ruane, B. O'Keeffe, The use of 3d laser scanning technology for bridge inspection and assessment, *Structural Health Monitoring (SHM)* 13 (2014) 14.
- [9] C. Popescu, B. Täljsten, T. Blanksvärd, L. Elfgren, 3d reconstruction of existing concrete bridges using optical methods, *Structure and Infrastructure Engineering* 15 (7) (2019) 912–924. doi:10.1080/15732479.2019.1594315.
- [10] H. Kim, Y. Narazaki, B. F. Spencer Jr., Automated bridge component recognition using close-range images from unmanned aerial vehicles, *Engineering Structures* 274 (2023) 115184. doi:https://doi.org/10.1016/j.engstruct.2022.115184.
- [11] B. Conde-Carnero, B. Riveiro, P. Arias, J. C. Caamaño, Exploitation of geometric data provided by laser scanning to create fem structural models of bridges, *Journal of Performance of Constructed Facilities* 30 (3) (2016) 04015053. doi:10.1061/(ASCE)CF.1943-5509.0000807.
- [12] N. Gyetvai, L. Truong-Hong, D. F. Laefer, Laser scan-based structural assessment of wrought iron bridges: Guinness bridge, Ireland, *Proceedings of the Institution of Civil Engineers-Engineering History and Heritage* 171 (2) (2018) 76–89.
- [13] L. Truong-Hong, R. Lindenbergh, Automatically extracting surfaces of reinforced concrete bridges from terrestrial laser scanning point clouds, *Automation in Construction* 135 (2022) 104127. doi:https://doi.org/10.1016/j.autcon.2021.104127.
- [14] A. Bautista-De Castro, L. J. Sánchez-Aparicio, L. F. Ramos, J. Sena-Cruz, D. González-Aguilera, Integrating geomatic approaches, operational modal analysis, advanced numerical and updating methods to evaluate the current safety conditions of the historical bôco bridge, *Construction and Building Materials* 158 (2018) 961–984. doi:https://doi.org/10.1016/j.conbuildmat.2017.10.084.
- [15] M.-K. Kim, H. Sohn, C.-C. Chang, Localization and quantification of concrete spalling defects using terrestrial laser scanning, *Journal of Computing in Civil Engineering* 29 (6) (2015) 04014086. doi:10.1061/(ASCE)CP.1943-5487.0000415.
- [16] V. Kasireddy, B. Akinci, Assessing the impact of 3d point neighborhood size selection on unsupervised spall classification with 3d bridge point clouds, *Advanced Engineering Informatics* 52 (2022) 101624. doi:https://doi.org/10.1016/j.aei.2022.101624.

- [17] G. Teza, A. Galgaro, F. Moro, Contactless recognition of concrete surface damage from laser scanning and curvature computation, *NDT & E International* 42 (4) (2009) 240–249. doi:<https://doi.org/10.1016/j.ndteint.2008.10.009>.
- [18] N. Bolourian, A. Hammad, Lidar-equipped uav path planning considering potential locations of defects for bridge inspection, *Automation in Construction* 117 (2020) 103250. doi:<https://doi.org/10.1016/j.autcon.2020.103250>.
- [19] J. Valença, I. Puente, E. Júlio, H. González-Jorge, P. Arias-Sánchez, Assessment of cracks on concrete bridges using image processing supported by laser scanning survey, *Construction and Building Materials* 146 (2017) 668–678. doi:<https://doi.org/10.1016/j.conbuildmat.2017.04.096>.
- [20] M. Rabah, A. Elhattab, A. Fayad, Automatic concrete cracks detection and mapping of terrestrial laser scan data, *NRIAG Journal of Astronomy and Geophysics* 2 (2) (2013) 250–255. doi:[10.1016/j.nrjag.2013.12.002](https://doi.org/10.1016/j.nrjag.2013.12.002).
- [21] X. Chen, J. Li, S. Huang, H. Cui, P. Liu, Q. Sun, An automatic concrete crack-detection method fusing point clouds and images based on improved otsu’s algorithm, *Sensors* 21 (5) (2021). doi:[10.3390/s21051581](https://doi.org/10.3390/s21051581).
- [22] M. Omer, L. Margetts, M. H. Mosleh, S. Hewitt, M. Parwaiz, Use of gaming technology to bring bridge inspection to the office, *Structure and Infrastructure Engineering* 15 (10) (2019) 1292–1307. doi:[10.1080/15732479.2019.1615962](https://doi.org/10.1080/15732479.2019.1615962).
- [23] L. Truong-Hong, H. Falter, D. Lennon, D. F. Laefer, Framework for bridge inspection with laser scanning (2016).
- [24] S. K. P. Kushwaha, S. Raghavendra, H. Pande, S. Agrawal, Analysis and integration of surface and subsurface information of different bridges, *Journal of the Indian Society of Remote Sensing* 48 (2) (2020) 315–331.
- [25] H.-K. Shen, P.-H. Chen, L.-M. Chang, Automated steel bridge coating rust defect recognition method based on color and texture feature, *Automation in Construction* 31 (2013) 338–356. doi:<https://doi.org/10.1016/j.autcon.2012.11.003>.
- [26] E. M. Javier, Methods for evaluation of the remaining strength in steel bridge beams with section losses due to corrosion damage, Master’s thesis, Virginia Tech (2021).
- [27] Y. Wang, S. Xu, A. Li, Flexural performance evaluation of corroded steel beams based on 3d corrosion morphology, *Structure and Infrastructure Engineering* 16 (11) (2020) 1562–1577. doi:[10.1080/15732479.2020.1713169](https://doi.org/10.1080/15732479.2020.1713169).
- [28] A. Hain, T. Zhang, A. Zaghi, Bridge Maintenance, Safety, Management, Life-Cycle Sustainability and Innovations, Taylor and Francis Group, 2021, Ch. Estimation of the residual bearing capacity of corrosion damaged bridge beams using 3D scanning and finite element analysis.
- [29] G. Tzortzinis, C. Ai, S. F. Breña, S. Gerasimidis, Using 3d laser scanning for estimating the capacity of corroded steel bridge girders: Experiments, computations and analytical solutions, *Engineering Structures* 265 (2022) 114407. doi:<https://doi.org/10.1016/j.engstruct.2022.114407>.

- [30] G. Tzortzinis, S. Gerasimidis, S. F. Breña, Improved load rating procedures for deteriorated steel beam ends with deteriorated stiffeners, Tech. Rep. 21-024, Massachusetts Department of Transportation, Office of Transportation Planning, Boston, MA (2021).
- [31] T. Zhang, A. E. Zaghi, Estimation of the residual bearing strength of corroded bridge girders using 3d scan data, *Thin-Walled Structures* 188 (2023) 110798. doi:<https://doi.org/10.1016/j.tws.2023.110798>.
- [32] V. V. Degtyarev, K. D. Tsavdaridis, Buckling and ultimate load prediction models for perforated steel beams using machine learning algorithms, *Journal of Building Engineering* 51 (2022) 104316. doi:<https://doi.org/10.1016/j.jobbe.2022.104316>.
- [33] V. Degtyarev, M. Naser, Boosting machines for predicting shear strength of cfs channels with staggered web perforations, *Structures* 34 (2021) 3391–3403. doi:<https://doi.org/10.1016/j.istruc.2021.09.060>.
- [34] V.-H. Truong, H.-A. Pham, T. Huynh Van, S. Tangaramvong, Evaluation of machine learning models for load-carrying capacity assessment of semi-rigid steel structures, *Engineering Structures* 273 (2022) 115001. doi:<https://doi.org/10.1016/j.engstruct.2022.115001>.
- [35] S. M. Mojtabaei, J. Becque, I. Hajirasouliha, R. Khandan, Predicting the buckling behaviour of thin-walled structural elements using machine learning methods, *Thin-Walled Structures* 184 (2023) 110518. doi:<https://doi.org/10.1016/j.tws.2022.110518>.
- [36] S. Mangalathu, S.-H. Hwang, E. Choi, J.-S. Jeon, Rapid seismic damage evaluation of bridge portfolios using machine learning techniques, *Engineering Structures* 201 (2019) 109785. doi:<https://doi.org/10.1016/j.engstruct.2019.109785>.
- [37] G. Tzortzinis, B. T. Knickle, A. Bardow, S. F. Breña, S. Gerasimidis, Strength evaluation of deteriorated girder ends. I: Experimental study on naturally corroded I-beams, *Thin-Walled Struct.* 159 (2021) 107220.
- [38] G. Tzortzinis, B. T. Knickle, A. Bardow, S. F. Breña, S. Gerasimidis, Strength evaluation of deteriorated girder ends. II: Numerical study on corroded I-beams, *Thin-Walled Structures* 159 (2021) 107216. doi:<https://doi.org/10.1016/j.tws.2020.107216>.
- [39] MassDOT, Draft LRFD Bridge Manual Part I, Boston, MA (2020).
- [40] SIMULIA, Providence, RI, Abaqus user's manual (2014).
- [41] K. O'Shea, R. Nash, An introduction to convolutional neural networks, ArXiv e-prints (11 2015).
- [42] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems, software available from tensorflow.org (2015).
URL <https://www.tensorflow.org/>

Supplementary Files

Layer	Filters/Neurons	Activation	Kernel	Pool size	Dropout rate
Classification-cov1	16	ReLU	3	-	-
Classification-cov2	32	ReLU	3	-	-
Classification-cov3	64	ReLU	3	-	-
Classification-dense1	128	ReLU	-	-	-
Classification-dense2	64	ReLU	-	-	-
Classification-dense3	5	softmax	-	-	-
Classification-pool	-	-	-	2	-
Classification-drop	-	-	-	-	0.2
Regression-cov1	16	ReLU	3	-	-
Regression-cov2	32	ReLU	3	-	-
Regression-cov3	64	ReLU	3	-	-
Regression-cov4	128	ReLU	3	-	-
Regression-dense1	128	ReLU	-	-	-
Regression-dense2	64	ReLU	-	-	-
Regression-dense3	1	-	-	-	-
Regression-pool	-	-	-	2	-
Regression-drop	-	-	-	-	0.2

Supplementary Table 1: Hyperparameters of the convolution neural network layers for the classification and regression models.
The Hyperparameters of the CNN layers are determined by conducting optimization with a random grid search from a predefined hyperparameter set, shown in Supplementary Table 2.

Entity	Values
Convolutional layer	Filters
cov1	4, 8, 16, 32
cov2	8, 16, 32, 64
cov3	16, 32, 64, 128
Dense layer	Neurons
dense1	16, 32, 64, 128
dense2	64, 128, 256
Others	
Kernel	3, 5
Learning rate	1e-4, 5e-4, 1e-3
Dropout rate	0.1, 0.2, 0.3
Depth	2, 3

Supplementary Table 2: Hyperparameter set. Initial set for optimization.