

CONVOLUTIONAL NEURAL NETWORKS FOR DETECTING BRIDGE CROSSING EVENTS WITH GROUND-BASED INTERFEROMETRIC RADAR DATA

M. Arnold^a, M. Hoyer^b, S. Keller^c

^a ci-tec GmbH, 76137 Karlsruhe, Germany - m.arnold@ci-tec.de

^b ci-tec GmbH, 76137 Karlsruhe, Germany - m.hoyer@ci-tec.de

^c Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany - sina.keller@kit.edu

Commission I, WG I/3

KEY WORDS: Ground-based Interferometric Radar, Event Detection, CNN, Infrastructure Monitoring, Machine Learning, Time series classification, Field Campaign, UAV

ABSTRACT:

This study focuses on detecting vehicle crossings (events) with ground-based interferometric radar (GBR) time series data recorded at bridges in the course of critical infrastructure monitoring. To address the challenging event detection and time series classification task, we rely on a deep learning (DL) architecture. The GBR-displacement data originates from real-world measurements at two German bridges under normal traffic conditions. As preprocessing, we only apply a low-pass filter. We develop and evaluate a one-dimensional convolutional neural network (CNN) to achieve a solely data-driven event detection. As a baseline machine learning approach, we use a Random Forest (RF) with a selected feature-based input. Both models' performance is evaluated on two datasets by focusing on identifying events and pure bridge oscillations. Generally, the event classification results are promising, and the CNN outperforms the RF with an overall accuracy of 94.7% on the test subset. By relying on an entirely unknown second dataset, we focus on the models' performances regarding the distinction between events and decays. On this dataset, the CNN meets this challenge successfully, while the feature-based RF classifies the majority of non-event decays as events. To sum up, the presented results reveal the potential of a data-driven DL approach concerning the detection of bridge crossing events in GBR-based displacement time series data. Based on such an event detection, a prospective assessment of bridge conditions seems feasible as an extension to previous structural health monitoring approaches.

1. INTRODUCTION

Bridges are critical infrastructures as they play an essential role in transportation and traffic. Over time, many bridges' operation conditions have changed since, for example, the vehicle loads have increased. Most bridges have been designed for specifications different from the current traffic and load conditions. Therefore, close monitoring and recurrent inspection of bridge infrastructures are mandatory.

Thus structural health monitoring (SHM), particularly bridge monitoring, is an important topic in current research studies (Sun et al., 2020). SHM addresses the monitoring of infrastructures with the goal of condition assessment. For this purpose, bridges are usually equipped with measurement systems such as acceleration sensors or strain gauge sensors. The sensor installation, however, inflicts damage, and these measurement systems require regular maintenance. An alternative approach is to rely on non-invasive measurement systems such as ground-based interferometric radar (GBR). Furthermore, the GBR measurement setup is mobile and hence flexible in its use, since a single radar can be utilized for regular inspections at a large number of bridges. This makes it an economical alternative or a practical addition to a conventional sensor installation at each bridge. GBR measures the displacement of the bridge in line of sight (LOS), which then can be converted to vertical displacement (Pieraccini et al., 2006; Michel and Keller, 2020; Arnold and Keller, 2020). Thus the GBR can capture the bridge behavior in a displacement time series.

Approaches for assessing the bridge condition respectively

damage detection rely on such measured time series signals. One commonly applied approach is the Operational Modal Analysis (OMA). In OMA, the bridge's modal behavior is examined without prior knowledge about the system's input or the transfer function (Rainieri and Fabbrocino, 2014). Modal parameters such as the eigenfrequencies of bridges are investigated to achieve damage detection. Changes in these parameters can indicate deteriorations within the monitored bridge structures. When analyzing the structural conditions, strong bridge responses are necessary. Such responses are stimulated by crossing vehicles. SHM can (1) detect vehicle crossings and then, based upon that, (2) apply condition assessment using these crossings. This study focuses on extracting crossings in GBR signals, and we refer to these crossings as events.

Arnold and Keller (2020) introduce a feature-based machine learning (ML) approach to detect events in bridge displacement data measured with GBRs. Although the proposed approach performs well in general, decays and events occurring in GBR data are not distinguished reliably. Decays are strong bridge oscillations decaying over a period after an event if the bridge is heavily stimulated. To achieve a more profound event detection of bridge crossing events, we now propose a data-driven DL approach without applying any feature extraction beforehand. We train a Convolutional Neural Network (CNN) solely on the displacement data extracted from the GBR signal. The use of a CNN is motivated by the fact that the displacement data and the grayscale gradients in image edge detection are similar and image classification is often performed employing convolution layers (Lawrence et al., 1997; Wang et al., 2019).

As a baseline, we rely on the best performing feature-based approach of Arnold and Keller (2020). Labeled measurement data acquired during several measurement campaigns at two bridges in Germany is used to evaluate the CNN and feature-based approach. During these campaigns, we furthermore deploy an unmanned aerial vehicle (UAV) to gain ground-truth data about the vehicles on top of the bridge.

The novel contributions of this study are:

- a detailed discussion of the challenges in the event detection based on GBR displacement data and a demonstration of the potential of DL approaches to address some of these challenges;
- an in-depth description of the developed CNN architecture;
- an evaluation of the classification results accomplished by our purely data-driven CNN, thoroughly compared against the baseline of a feature-based ML approach;
- a comprehensive application of our approach to entirely unknown data.

In Section 2, we provide an overview of related work concerning GBR measurements on bridges, CNN time series signal processing, and SHM. Section 3 introduces the monitored bridges and the dataset acquired during GBR measurements. Our methodological approach is described in Section 4 concerning applied preprocessing steps and the developed CNN model. The event classification results are presented and evaluated in Section 5. We summarize our study in Section 6 and give an outlook on prospective research aspects.

2. RELATED WORK

In this section, we give a short overview of the relevant research studies. First, we summarize studies exploiting GBR for bridge monitoring. Second, selected applications of CNNs focussing on time series classification and anomaly detection are presented. Third, the focus is placed on the combination of ML approaches and bridge deformation data, such as GBR-measured displacement time series.

Pieraccini et al. (2007) deploy a GBR to measure a bridge in Florence, Italy, remotely. They evaluate the static and dynamic measurement capabilities of a GBR. Therefore the GBR scans the lower side of the bridge, measuring the phase of the reflected signal while the bridge is being loaded. Focussing on a static evaluation, they use a locomotive that slowly crosses the bridge, stopping in the middle for 6 minutes, while remotely and non-invasively monitoring the displacement along the lower side to acquire the maximum displacement of the bridge. For dynamic testing, a truck rapidly crosses the bridge. They are able to extract modal parameters, such as the eigenfrequency, from the structural dynamic response to this crossing. To verify the GBR measurement concept, Gentile and Bernardini (2008) compare measurement results of a GBR exploiting corner reflectors and conventional accelerations sensors. In the time domain, a direct comparison via the velocity is possible. Employing Singular Value Decomposition (SVD), the eigenfrequencies and mode shapes are examined. One finding of these studies is that it is possible to deploy a GBR to measure the bridge displacement. In principle, the GBR needs points with a high reflectivity for a high signal to noise ratio (SNR), which ensures precise displacement measurements. However, most of the bridges

are characterized by flat surfaces reducing the amount of signal backscattered towards the GBR. One option is to install corner reflectors, which provide a high reflectivity but may inflict damage to the bridge. Michel and Keller (2020) evaluate a non-invasive measurement setup based on GBR without mounting any targets on the bridge's lower side, called mirror mode. They place a reflector on the ground opposite of the GBR position at the bridge. The measurement setup is verified at a German bridge by simultaneously monitoring the bridge with two GBRs and a profile scanner. One GBR is set up in the mirror mode, whereas the other GBR measures the vertical displacement component directly. The GBR signal traverses a larger distance in the mirror mode, reducing the SNR compared to the standard GBR measurement setup. Nevertheless, the resulting displacement data is characterized by appropriate reflectivity, making the mirror mode valuable as an alternative, non-invasive measurement setup in GBR measurements. Dei et al. (2013) evaluate static GBR measurements at an 8-span bridge in Italy. Four corner reflectors are installed: two in the middle of the monitored span and two close to the piers. As up to 8 trucks are parked on the bridge, the two reflectors at the pier are elevated according to the GBR vertical measurements despite the bending of the bridge. This elevation is due to the position of the reflectors causing them to move vertically and horizontally. The superposition of both is measured by the GBR, leading to ambiguous results. The reflectors in the middle are not affected since they undergo no horizontal displacement. Miccinesi et al. (2021) propose a multiple input, multiple output interferometric (MIMO) GBR setup to acquire two independent displacement components to avoid equivocal measurements simultaneously. It is verified by monitoring a corner reflector oscillating under a controlled environment. The results are compared to a seismic sensor attached to the reflector. Both signals match in time and frequency domain. After verification, the MIMO GBR is exploited to measure the vertical and horizontal components at an Italian bridge.

In the context of GBR data, CNNs are applied for Synthetic Aperture Radar (SAR) image classification (e.g., Zhang et al., 2018). CNNs are mainly used as DL approaches in image classification tasks such as face recognition (Lawrence et al., 1997). When focusing on one-dimensional (1D) time series data, they are currently adopted to solve classification tasks such as electrocardiogram classification (Mahmud et al., 2020). Kenji Iwana and Uchida (2020) deploy CNN in time series classification on noisy time series to evaluate their capabilities. They rely on a toy dataset consisting of several simple waveforms such as sine waves superimposed with different noise levels. Additionally, they provide the network with multivariate time series signals and public real-world datasets during the training process. As a baseline, the CNN's classification results are compared with results of a Support Vector Machine (SVM) and Deep Belief Networks (DBN). For all datasets, CNNs obtain the best classification results due to their ability to extract suitable features themselves. In the study of Cook et al. (2020), another classification task is discussed involving event detection in a time series. Anomalies respectively events are, herein, deviations from a general pattern, which are described as outliers. Exploiting CNNs, anomaly detection is practiced in two steps: (1) CNNs predict a signal, and (2) a high difference between prediction and measurement indicates an anomaly. This specific approach is based on a kind of predictability, for example, a periodicity, in the signal, which a CNN can reproduce.

Several studies include ML approaches and bridge deformation data. Arnold and Keller (2020) use GBR-measured data and introduce a feature-based ML approach to detect events in displacement time series. GBR-measured data of three German bridges is presented. Selected ML approaches are applied to detect and classify events of three different vehicle classes. One central aspect of the applied preprocessing is to reduce the influence of time series drifts caused by, for example, environmental conditions. Therefore, features that are independent of such a drift are investigated and selected. As a result, events are detected with an accuracy of up to 83.8%. Such an event classification can be the basis for damage detection on bridges using displacement data recorded during normal traffic conditions. Another study conducted by Döring et al. (2020) exploited ratio-based features of single vehicle crossings to detect damage in a two-span German bridge. Several damages are inflicted at two locations during the measurements by cutting tendon cables to different degrees. A single-vehicle then crosses the bridge for each damage type while the strain is measured at 18 different positions at the bridge. Features, independent of vehicle weight and velocity, are extracted from the strain time series and used as input data for several ML approaches. Accuracy of up to 95.1% for damage classification is achieved.

3. GBR-BASED MEASUREMENTS AT BRIDGES AND RESULTING DATASETS

In this section, we give a brief overview of the measurement setup used to record the data at the monitored bridges. A detailed description of the measured data is given in Arnold and Keller (2020). Subsequently, we describe the nature of the displacement data while highlighting the challenges concerning decays. For verification purposes, we also rely on RGB images acquired by an unmanned aerial vehicle (UAV).

The GBR can monitor multiple points in the line of sight (LOS) simultaneously with a sampling rate of up to 200 Hz using modulation. For this purpose, the signal reflection of every 0.75 m is accumulated to one single range bin (Pieraccini, 2013). In sum, the GBR returns one time series every 0.75 m. Therefore, it is advantageous to have precisely one highly reflective location in each range bin. Otherwise, signal superposition and thus ambiguities occur. Since the GBR uses the interferometric principle, it measures the phase of every reflected signal within one range bin. The phase difference can be converted to the bridge displacement in LOS between two consecutive samples. Concerning the height difference between the GBR and the bridge, the vertical displacement Δz can be deduced (Rödelsperger et al., 2010). Dei et al. (2013) state that the GBR measures several displacement components of the bridge depending on the position of the corner reflectors. Since our reflectors are relatively centered, we speak of vertical displacement for convenience only.

Figure 1 schematically shows the general idea of the GBR measurement setup. The colored triangles in Figure 1a respective circles in Figure 1b represent reflectors that have been attached to the bridge to provide points of strong local backscattering despite the bridges' flat surface. As the GBR is positioned with the LOS parallel to the lanes, the reflectors are installed with an offset in LOS to be in unique range bins. Additionally, a vertical offset has been added to extract further information, for example, the driving side of a crossing vehicle.

In this study, we consider two bridges, A and B, GBR-monitored during several measurement campaigns covering an

appropriate range of environmental conditions. Table 1 summarizes relevant details concerning these two bridges. They are equipped with reflectors to obtain specific reflection points. Since the first natural frequency of the two bridges is close together, similar decay processes occur.

We define an *event* as the time during which a vehicle crosses the monitored field. In the case of bridge A, this means that only one field, not the whole bridge, is considered. Figure 2 presents bridge A's vertical displacement during three events for all four range bins (cf. Figure 1). For visualization purposes, we have removed the offset of each time series caused by environmental conditions. An UAV was used to acquire the images to gain ground-truth data of events. The difference in displacement at different parts of the bridge indicates the vehicle's driving side: range bins 21 and 22 show a greater deflection than range bins 23 and 24, which indicates that all vehicles drive from right to left (see also the reflector positions in Figure 1).

The depicted situation involves several cars queueing behind a slower truck, which is commonplace. The truck's heavy weight causes a strong oscillation of the bridge, which slowly decays over several seconds. Vehicles entering during this decay still produce a bend of the bridge. However, the deformation is superimposed by vibration. Nevertheless, we only want to detect the bending process as an event, regardless of the general bridge oscillation, which is classified as a non-event.

Table 2 summarizes information about events recorded during five measurement campaigns conducted during different days, including the event figures per day. As an event takes approximately 1.35 s on average, the number of non-event incidents is much larger considering each campaign's duration. The number of range bins refers to how many range bins we used for training. They have been selected following their respective SNR to ensure low noise. Since the bridge's eigenfrequency generally varies depending on the air temperature (Mahowald et al., 2014), we ensure a wide temperature range when selecting the study's dataset. The rows 1 to 4 in Table 2 are combined into a **dataset I**, which is partially used for the training of the models. Besides, row 5 serves as an entirely unknown dataset (**dataset II**) to specifically evaluate the performance in distinguishing decays and events.

4. METHODOLOGY

In this section, we present our methodological approach. First, we provide information about all preprocessing steps concerning the GBR time series in Section 4.1. Second, the applied ML models are described focusing mainly on the CNN architecture (see Section 4.2).

4.1 GBR Time Series Preprocessing

As an objective of this study, we apply as little data preprocessing as possible on the GBR-measured time series data described in Section 3. Based on this precaution, we will be able to deploy the proposed DL approach as an online tool to previously unmeasured bridges someday. Therefore, no normalization is applied before the training process of the ML models.

Since bridges usually vibrate in a lower frequency range (Mehlhorn and Curbach, 2014) than the GBR's Nyquist frequency of 100 Hz, we use a low-pass Butterworth filter to remove high-frequency noise (Bianchi and Sorrentino, 2007) as the only signal preprocessing step. Additionally, we avoid removing the

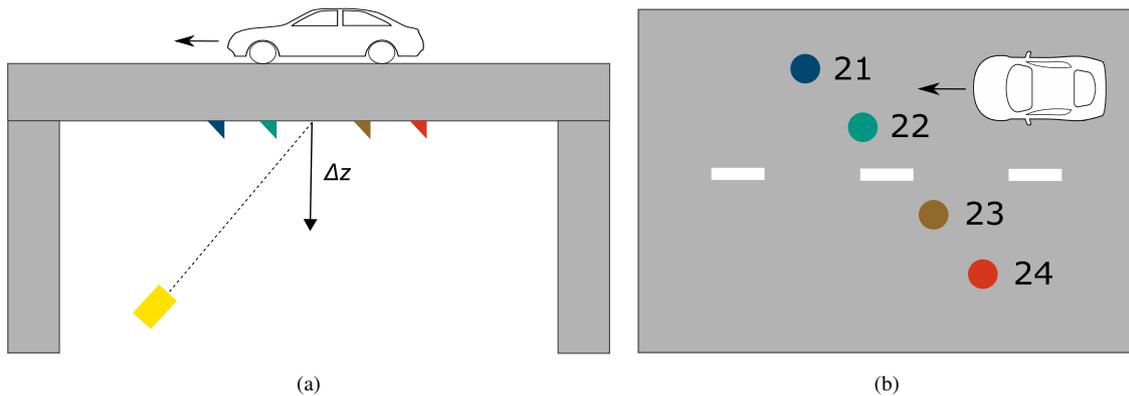


Figure 1. (a) Lateral view of a prototypical bridge during a GBR measurement. The GBR (yellow) measures the displacement in line of sight (LOS) visualized as a dotted line. This LOS measurement, then, is transformed into vertical displacement Δz . The colored triangles represent corner reflectors, which have been installed to achieve a higher reflectivity. (b) Top view of the same measurement setup as in (a). It shows the two lanes on top of the bridge as well as the offset of the range bins 21 to 24.

Table 1. Detailed information on the selected bridges.

Structure	Type	Fields	Length	Width	Natural frequency
Bridge A	Beam/plate mixing system	2	57.00 m	13.60 m	3.66 Hz
Bridge B	Plate girder bridge/Girder grid bridge	1	26.36 m	11.69 m	3.75 Hz

long-term drift to ensure the models' robustness against environmental impacts.

The filtered time series data are automatically split based on non-overlapping windows of 100 samples, which means 0.5 s, automatically. Subsequently, the data of each window is compared to a labeled event. If at least 40 % of the 100 samples window lie within the displacement caused by an event, the entire window is labeled as an event. This step is necessary for the preparation of the supervised classification task. Based on the defined window size of 100 samples and the ratio of 40 %, we obtain a satisfying detection of the event's start- and endpoint. Besides, this window size ensures that we capture more than one period of the bridge vibration.

The amount of events is significantly lower than the number of non-event incidences. This imbalance can influence the classification performance of the applied ML models since the entire dataset is biased towards non-events. Therefore, we apply a random undersampling by deleting samples of the majority class, which, in our case, is called a *no event*.

As a last preprocessing step, we randomly split the dataset I with the ratio 70 : 15 : 15 for the training, validation, and test subset. The training subset contains 47710 incidences overall with 23719 *event* and 23991 *no event*. Dataset II (see Table 2) is used as an entirely unknown test subset for the ML models. The *no event* class of dataset II solely consists of decays, which allows us a more thorough analysis concerning events and decays. It is composed of 236 events and 24 undisturbed decays, respectively 516 and 249 incidents, after splitting them into windows of 100 samples.

4.2 Machine Learning Models

We evaluate a solely data-driven DL approach for detecting events in GBR displacement data. As a baseline for the CNN classification performance, we rely on a feature-based ML approach, a Random Forest (RF) model (see Arnold and Keller

(2020)). The RF is applied on features extracted from the low-pass filtered data such as the variance and implemented using the widely-used scikit-learn package (Pedregosa et al., 2011). The implementation of the introduced CNN is based on TensorFlow (Abadi et al., 2016). In the following, we describe the architecture of the (1D) CNN. As described in Section 2, CNNs are mainly popular in image classification but are not often used in the context of 1D GBR time series data. The displacement time series of an event has similarities to the 1D grayscale gradient of an edge. Since edge detection is commonly performed by convoluting a filter, we exploit convolutional layers in our model. The CNN architecture is visualized in Figure 3. 100 bridge displacement samples of one single range bin represent the input. The next three parts consist of three 1D convolutional layers (CONV) with different filters and filter sizes, as summarized in Table 3. Each CONV layer is followed by a max-pooling layer (MaxPooling). The CONV1 layer has *tanh* as an activation function, in contrast to the other CONV layers. The main reason for this activation function is that the input data is not normalized to ensure the transferability to other, unseen bridge data. Based on the *tanh* activation, the first input features are mapped to the range -1 to 1 . The additional CONV layers, therefore, have scaled input features. Relatively large *kernel_sizes* characterize CONV1 and CONV2, aiming at capturing gradients over several samples. As the GBR measures with a frequency of 200 Hz, the displacement change between consecutive samples is small. A large *kernel_size* allows the filter to register more significant changes within the signal. Additionally, the vibration of the signal is easier to detect, which ensures a more profound classification of pure decays and events superimposed with such vibrations. A flatten layer reshapes the CONV output to be used as an input for fully-connected (FC) layers. FC1 reduces the vector dimension, thus condensing on important features. Besides, FC2 expands, allowing for more complex evaluations based on the results of FC1. FC3 and FC4 then compress the features vector for a SoftMax, which returns the final classification results. The RF is implemented with the default hyperparameters except for the

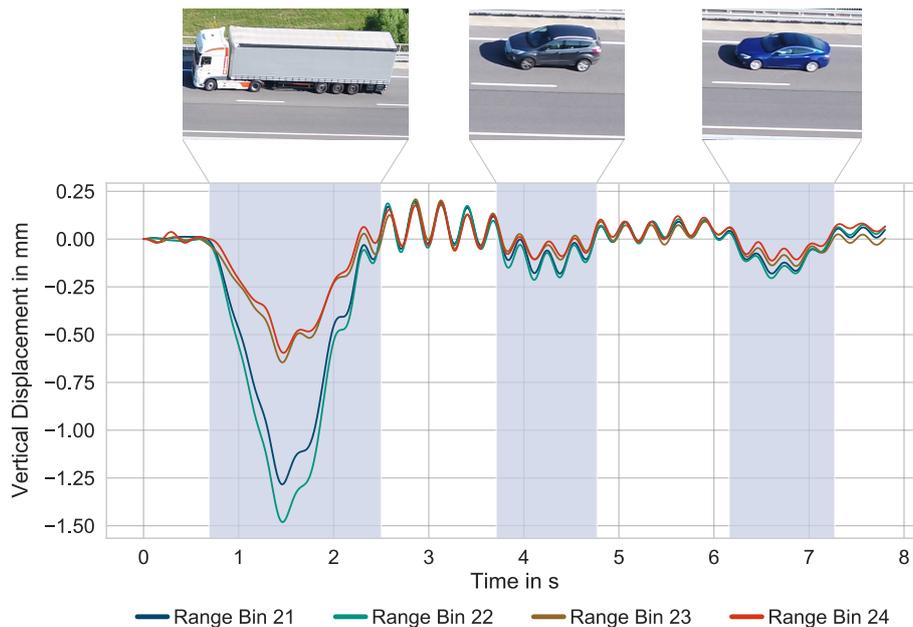


Figure 2. The displacement time series of three consecutive events for four range bins on bridge A. All events are highlighted and pictures of the corresponding vehicles, acquired with a UAV, are attached. The leading truck causes the bridge to vibrate even after leaving the bridge deck. Both following cars lead to a bending of the bridge which is superimposed with the oscillation.

Table 2. Overview of the measurement campaigns at the bridges conducted for this study. The duration refers to the period of measurements at the respective bridge, number of events to the number of crossing vehicles, and number of range bins to the number of bins selected for analysis. The temperature range covers the air temperature during the measurements. Dataset I consists of the training, validation, and test subsets necessary for the ML approaches; while dataset II is an extra test subset, which is entirely unknown to the ML approaches.

Number	Bridge	Duration	Number of events	Number of range bins	Temperature range	Dataset
1	A	4h30min	1007	6	11.20 °C to 12.35 °C	I
2	A	1h45min	395	5	8.90 °C to 10.0 °C	I
3	A	30min	119	2	8.10 °C to 8.90 °C	I
4	B	3h30min	1482	2	14.70 °C to 15.50 °C	I
5	A	1h20min	236	1	18.10 °C to 19.40 °C	II

following parameters: `max_depth = 2`, `max_leaf_nodes = 6`, `min_samples_split = 0.6` and `n_estimators = 200`. They are fine-tuned using GridSearch.

5. RESULTS AND DISCUSSION

The study's objective is to investigate the potential of a DL approach in the context of detecting events in GBR time series data without any feature extraction. The classification performances of the RF with selected input features and of the CNN are contrasted in Figure 4 and Table 4. Figure 4 shows the normalized confusion matrix for the RF and the CNN model, while Table 4 states the Overall Accuracy (OA), Precision (P), and Recall (RC). Both models are evaluated on the test subset of dataset I as well as on the entirely unknown dataset II (see Table 2).

In the case of **dataset I**, the RF and the CNN achieve high scores with values above 90 % (see Table 2 and Figure 4a). Besides, the CNN achieves a significantly better classification with an OA of 94.7 %, which outperforms the RFs OA by 3.2 percentage points (p.p.). As dataset I contains events measured at two distinct bridges, this overall good performance reveals that the ML models can cope with heterogeneous GBR time series

data. Therefore, the results indicate an appropriate transferability between data measured at different bridges (under specific prerequisites). When focusing on the correctly classified events concerning all events, the RC-score, the CNN, and the RF achieve almost equal scores. The CNN performance is only 0.8 p.p. better than the RF. However, the P-score differs considerably. The P-score indicates how many of the incidence classified as events are actual events. According to this score, the CNN (95.3 %) performs by 5 p.p. better. The main reason is that the RF experiences difficulties in distinguishing between decays and events, as shown in (Arnold and Keller, 2020). We have to consider that no information about the number of decay incidents of no events in the dataset I is given, which impedes a further and more detailed analysis.

Therefore, we focus mainly on the distinction between decays as no events and events in **dataset II**. The classification results based on this dataset reveal that the CNN outperforms the feature-based RF (see Figure 4b and Table 4). Although the OA decreases compared to the results achieved with dataset I, all metrics scores are still above 90 %, with the RC for the CNN even at 99.4 %. When focusing on the RF's performance, it strikes that the RF model achieves an almost equally high RC-score, but the P-score and especially the OA are relatively low. According to the confusion matrix in Figure 4b, the RF

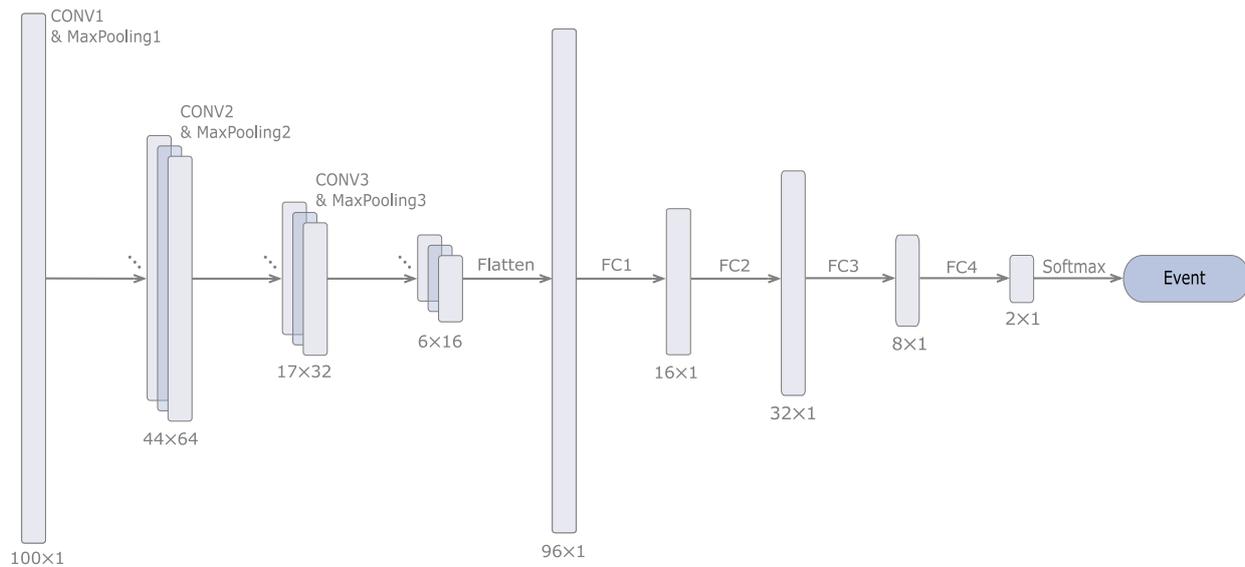


Figure 3. Flowchart of our developed CNN. The DL network combines convolutional layers (CONV), fully-connected (FC), and max-pooling (MaxPooling) layers. The output of each layer is presented as a vertical bar annotated with its corresponding shape. Details on the parameters of individual layers are given in Table 3.

Table 3. Hyperparameters of the applied CNN visualized in Figure 3.

Layer	Hyperparameters
CONV1	filters = 64 kernel_size = 13 activation = tanh
MaxPooling1	pool_size = 2
CONV2	filters = 32 kernel_size = 11 activation = ReLU
MaxPooling2	pool_size = 2
CONV3	filters = 16 kernel_size = 5 activation = ReLU
MaxPooling3	pool_size = 2
FC1	units = 16
FC2	units = 32
FC3	units = 8

Table 4. Results of the applied ML models, RF and CNN, for the dataset I and II on the classification task event vs. no event.

Model	Dataset I			Dataset II		
	OA in %	P in %	RC in %	OA in %	P in %	RC in %
RF	91.5	90.3	93.5	64.5	74.6	98.1
CNN	94.7	95.3	94.3	92.7	93.6	99.4

classifies the majority of no event decay (69 %) as an event. This is again due to the challenges of distinguishing decays and events. Since in dataset II all no events consist of decays, the RF misclassifies many incidents. Taking the deviating atmospheric conditions included in dataset II from dataset I, we recognize that the CNN can handle the deviation in the air temperature range and the resulting drift in the GBR-based time series. Thus, our approach with a minimized preprocessing of the GBR time series data seems appropriate in applying a CNN. This allows us to exploit our model during measurement campaigns during different environmental conditions.

Figure 5 shows a classification example of a 60 s period of both the CNN and the RF model. Note that the shown period originates from the dataset II, which is entirely unknown to the ML models. The CNN can detect and distinguish events and decays. It only classifies a small proportion of the given excerpt falsely. One reason for this finding could be that the example excerpt, including an event's end, extends into the decay. As a

solution, we can, for example, reduce the step size for the prediction window. The RF, on the other hand, fails in this case since it classifies the complete decay as an event after the first vehicle's crossing at around 3 s.

The window highlighted in pale blue corresponds to the vertical displacement situation visualized in Figure 2. In this situation, two cars cross the bridge shortly after a truck caused the bridge to vibrate. Thus, the bending caused by the two cars is superimposed by the bridge oscillation. It sticks out that the CNN performs better in extracting only the events regardless of the decay. This finding implies that the CNN learns to filter out the oscillation and considers the actual bending of the bridge. Based solely on the selected input features, the RF is incapable of classifying the entire excerpt correctly and considers the decay as an event. Both ML models have in common that they detect the single, undisturbed events at approximately 36 s and 52 s successfully. Although a small offset of around 0.25 mm caused, for example, by temperature changes is present, the ML models' performances are not affected.

In summary, our results show that a data-driven CNN is superior in event detection based on displacement time series as compared to feature-based models. While both approaches achieve appropriate results in general, the main difference is the challenge of distinguishing events and decays. In this case, the CNN performs well, while the RF classifies a majority of the incidences incorrectly.

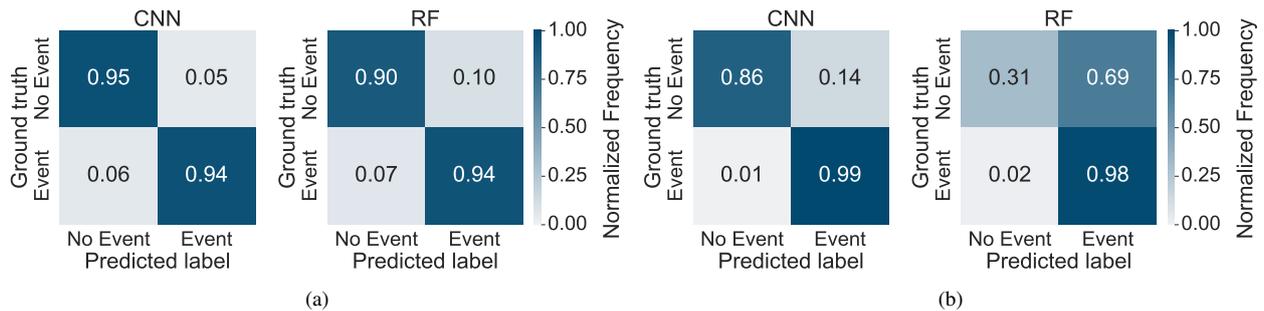


Figure 4. Confusion matrices of the CNN model (left) and the RF model (right) for (a) dataset I and (b) dataset II.

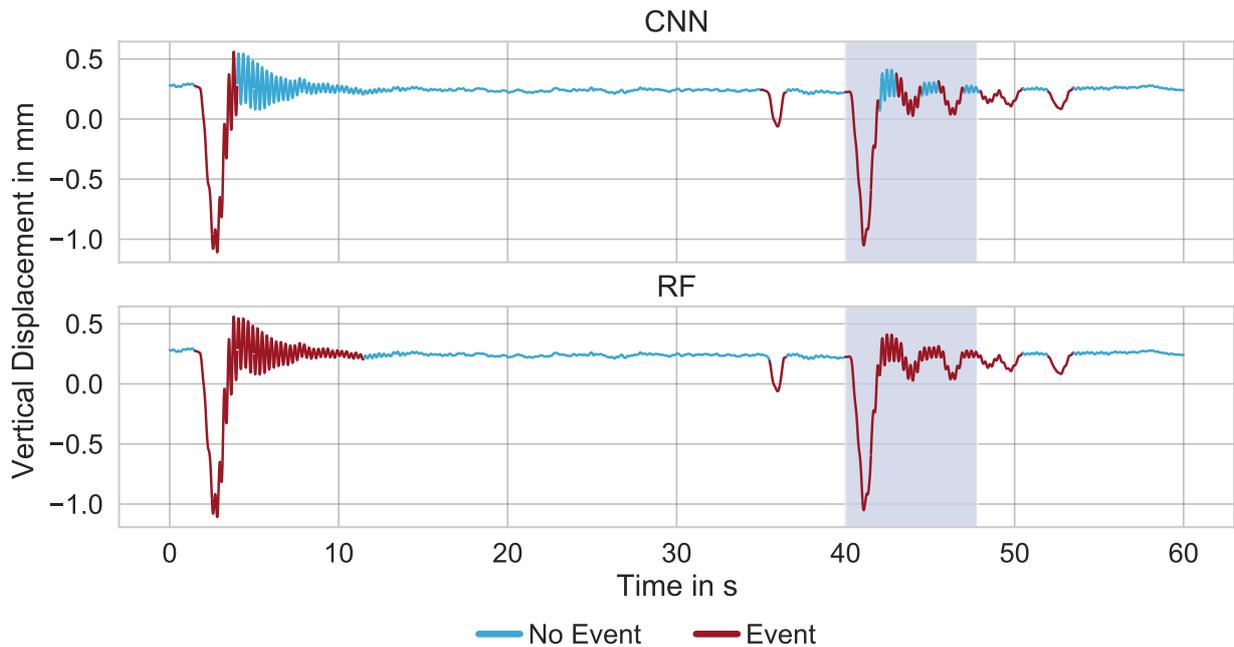


Figure 5. Classification results of an exemplary, 60 s long section of bridge A for CNN (upper) and RF (lower). The measurement campaign from which this section originates is unknown to the model. The highlighted area corresponds to the excerpt depicted in Figure 2.

6. CONCLUSION AND OUTLOOK

In this paper, we introduce and investigate a data-driven DL approach exploiting convolutional layers for event detection in 1D GBR-displacement time series data. Events refer to vehicles crossing the bridge during the GBR measurements. The data originates from real-world measurement campaigns at two German bridges. Our proposed DL approach addresses the current challenges, especially regarding the differentiation of crossing events and subsequent decays, present in state-of-the-art feature-based ML approaches based on shallow learners such as an RF model. The CNN detects events successfully and reliably; even in distinguishing decays and events, it performs satisfyingly. As a shallow baseline learner, we rely on a RF model. Overall, the CNN outperforms the RF in detecting only the bridge crossing while omitting the decay. In two distinct datasets, the ML models' performances are evaluated concerning, among others, the decays. The first dataset is used for training, validation, and test of the ML models. In contrast, the second dataset remains unknown in order to investigate the applied models' transferability and reproducibility.

Our CNN achieves higher overall accuracy, precision, and recall

on both datasets compared to the baseline model. The shortcomings of a feature-based, commonly applied ML model are revealed, especially in the second test on an unknown dataset. As presented, the CNN achieves an overall accuracy of 92.7% and a precision of 93.6% in classifying events and decays (non-events). The RF falls short with an overall accuracy of 64.5% and a precision of 74.5%. Both models' recall-score is similar, implying that they rarely classify events as non-events. Besides, we compare the classification results exploiting a time series excerpt from a measurement campaign unexploited in the ML models' training. Both the CNN and the RF reveal robustness against environmental impacts since a long-term time series drift does not affect the classification performances.

In sum, our study shows promising results for relying on a solely data-driven CNN as DL approach in the context of event detection with GBR-based time series data without any extensive preprocessing. Focusing on the further improvement of the CNN performance and a detailed evaluation of its limits and opportunities, we need to increase the number of decays in the training subset. Therefore, we will conduct further measurement campaigns at the presented bridges as well as add new

bridges to the portfolio. Based on this new data, we can enhance the generalization abilities of our proposed DL approach. Besides, we need to consider including bridges with strongly different eigenfrequencies in our measurement campaigns to investigate the CNN performance concerning varying decay oscillations. Based on the presented DL classifiers, a bridge damage assessment could be established which functions under real-world conditions with non-invasive GBR time series data and stimulations caused by vehicle crossings. Such a prospective assessment could combine a data-driven event detection and subsequently an event classification to extract the input causing the measured bridge displacement. This combination might be a first step towards determining changes in the dynamics of bridges concerning a more profound structural health monitoring.

ACKNOWLEDGEMENTS

This research study is part of the ZEBBRA project funded by the German Federal Ministry of Education and Research (BMBF). We especially thank our partners, the Institute for Automation and Applied Informatics (IAI) at the Karlsruhe Institute of Technology and the company Büro für Strukturmechanik, Coburg, for their support during the measurement campaigns.

References

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. et al., 2016. Tensorflow: A system for large-scale machine learning. *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, 265–283.
- Arnold, M., Keller, S., 2020. Detection and classification of bridge crossing Events with ground-based interferometric radar data and machine learning approaches. *ISPRS annals*, V-1-2020, 109–116.
- Bianchi, G., Sorrentino, R., 2007. *Electronic Filter Simulation & Design*. McGraw-Hill Education.
- Cook, A. A., Mısırlı, G., Fan, Z., 2020. Anomaly Detection for IoT Time-Series Data: A Survey. *IEEE Internet of Things Journal*, 7(7), 6481–6494.
- Dei, D., Mecatti, D., Pieraccini, M., 2013. Static Testing of a Bridge Using an Interferometric Radar: The Case Study of “Ponte degli Alpini,” Belluno, Italy. *The Scientific World Journal*, 2013.
- Döring, A., Waibel, P., Matthes, J., Scherer, O., Keller, H. B., Keller, S., Müller, J., Schneider, O., 2020. Ratio-based features for data-driven bridge monitoring and damage detection. Accepted.
- Gentile, C., Bernardini, G., 2008. Output-only modal identification of a reinforced concrete bridge from radar-based measurements. *NDT & E International*, 41(7), 544 - 553.
- Kenji Iwana, B., Uchida, S., 2020. Time series classification using local distance-based features in multi-modal fusion networks. *Pattern Recognition*, 97, 107024.
- Lawrence, S., Giles, C. L., Tsoi, A. C., Back, A. D., 1997. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*, 8(1), 98–113.
- Mahmud, T., Fattah, S. A., Saquib, M., 2020. Deeparnet: An efficient deep cnn architecture for automatic arrhythmia detection and classification from denoised ecg beats. *IEEE Access*, 8, 104788–104800.
- Mahowald, J., Maas, S., Nguyen, V. H., Waldmann, D., Zuerbes, A., 2014. Some conclusions from the measurements of temperatures and their gradients on eigenfrequencies of bridges.
- Mehlhorn, G., Curbach, M., 2014. *Handbuch Brücken*. Springer.
- Miccinesi, L., Beni, A., Pieraccini, M., 2021. Multi-Monostatic Interferometric Radar for Bridge Monitoring. *Electronics*, 10(3), 247.
- Michel, C., Keller, S., 2020. Introducing a non-invasive monitoring approach for bridge infrastructure with ground-based interferometric radar. Accepted.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. et al., 2011. Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825–2830.
- Pieraccini, M., 2013. Monitoring of civil infrastructures by interferometric radar: A review. *The Scientific World Journal*, 2013.
- Pieraccini, M., Fratini, M., Parrini, F., Atzeni, C., 2006. Dynamic Monitoring of Bridges Using a High-Speed Coherent Radar. *IEEE Transactions on Geoscience and Remote Sensing*, 44(11), 3284–3288.
- Pieraccini, M., Parrini, F., Fratini, M., Atzeni, C., Spinelli, P., Micheloni, M., 2007. Static and dynamic testing of bridges through microwave interferometry. *Ndt & E International*, 40(3), 208–214.
- Rainieri, C., Fabbrocino, G., 2014. Operational modal analysis of civil engineering structures. *Springer, New York*, 142, 143.
- Rödelsperger, S., Läufer, G., Gerstenecker, C., Becker, M., 2010. Monitoring of displacements with ground-based microwave interferometry: IBIS-S and IBIS-L. *Journal of Applied Geodesy*, 4(1), 41–54.
- Sun, L., Shang, Z., Xia, Y., Bhowmick, S., Nagarajaiah, S., 2020. Review of bridge structural health monitoring aided by big data and artificial intelligence: from condition assessment to damage detection. *Journal of Structural Engineering*, 146(5), 04020073.
- Wang, W., Yang, Y., Wang, X., Wang, W., Li, J., 2019. Development of convolutional neural network and its application in image classification: a survey. *Optical Engineering*, 58(4), 040901.
- Zhang, C., Pan, X., Li, H., Gardiner, A., Sargent, I., Hare, J., Atkinson, P. M., 2018. A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 140, 133–144.