Train on Synthetic - Test on Real: Toward Strain-Based Anomaly Detection on a Cantilever Beam

Philipp Conen M.Sc. - Philipp.Conen@dlr.de – Institute of Maintenance, Repair and Overhaul

Motivation

Machine learning can predict a certain output based on an input. However, large amounts of data are required for accurate predictions. In practice, however, **data is usually scarce or even unavailable**.

Related Work

In general, there are three methods against data scarcity: (1) data augmentation, (2) transferring data from other task areas (transfer learning) or (3) **generating synthetic data**. [1] Methods (1) and (2) are not suitable for all applications and are mostly used in the field of computer vision. Using synthetic data requires consideration of the absence of interference signals and environmental conditions. This challenge is also known as the **sim-to-real gap**.

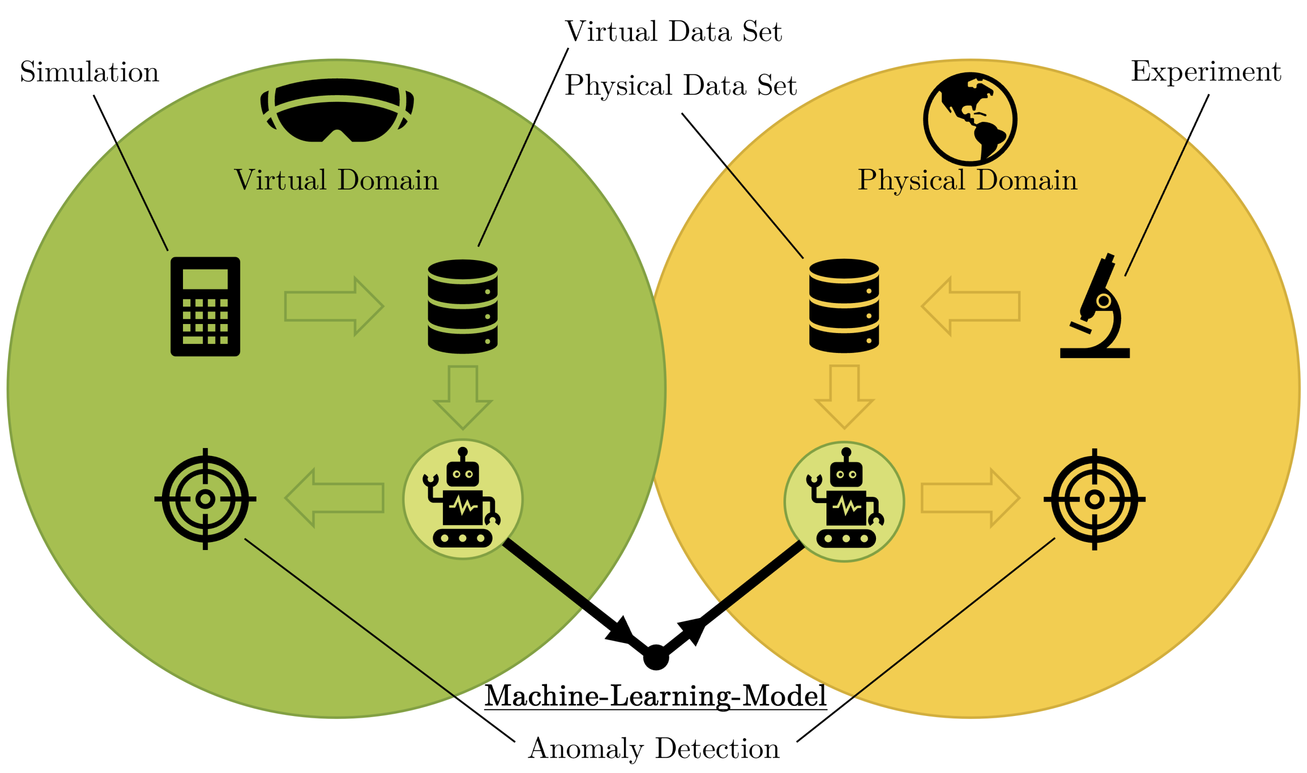


Figure 1: Proof of concept workflow for structural anomaly detection.

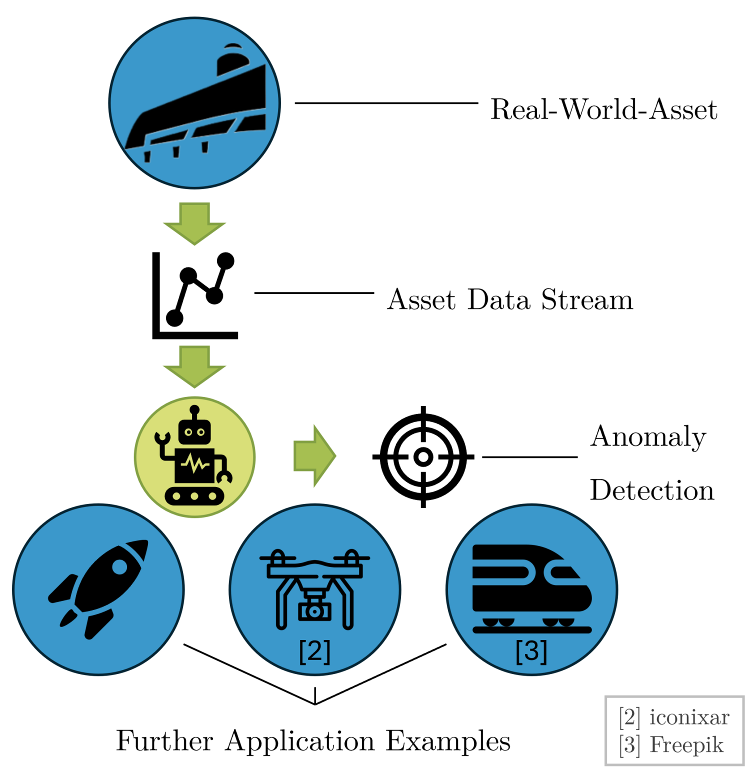


Figure 2: Operational workflow for structural anomaly detection.

Use Case

The use case considered here predicts structural anomalies under (initial) absence of real data. It applies the **synthetic data approach** (3).

Proof of Concept

A proof of concept ensures that the method works in principle. It requires an abstract initial **experimental investigation**. The experimental investigation distinguishes the proof of concept from the subsequent operation.

Operation

After the proof of concept, the approach monitors real structural assets **without physical trainng data**. For this, existing simulation models can generate a data set that trains a synthetic machine learning model.

Structural-Mechanical Scenario

The figure below shows the structural-mechanical scenario. It is equally valid for the synthetic and the real representation. The bending beam is clamped on the left-hand side and loaded with a force on the right-hand side. Strain gauges (90° rosette) are applied to the bending beam to measure pure **bending strain**. The stamp, which slides below the beam, interrupts the natural strain curve and thus represents a **structural anomaly**.

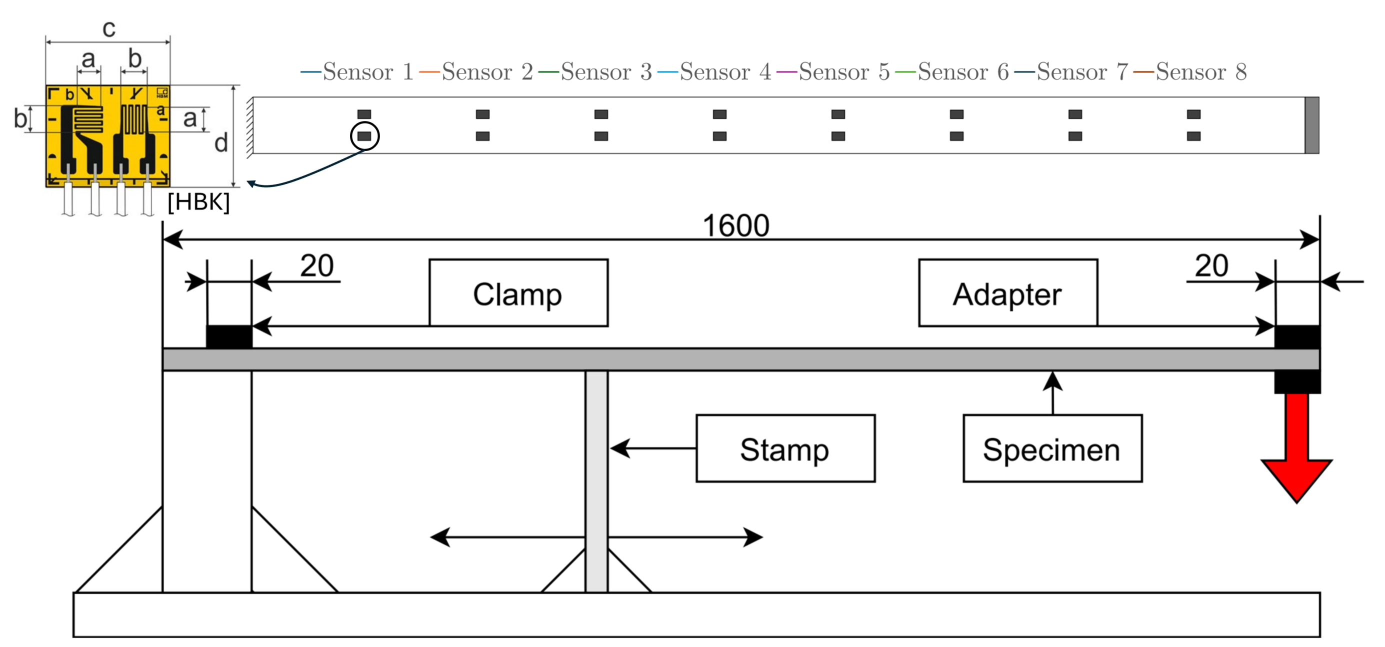
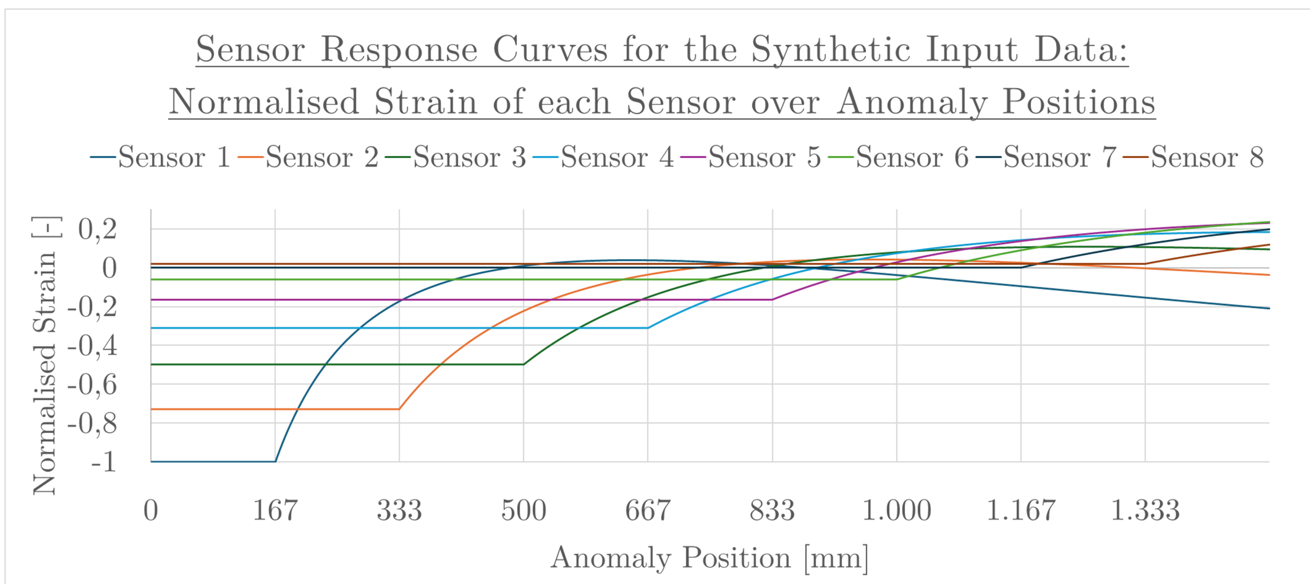


Figure 3: Synthetic and physical representation of the structural–mechanical scenario.

Machine Learning Model Investigation

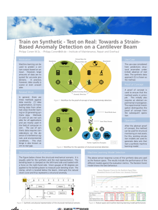
The figure below shows the sensor response curves of the synthetic data that are used as input. Furthermore, it shows the results, namely the performance of the models under investigation. The **Random Forest** model make the **best predictions**.

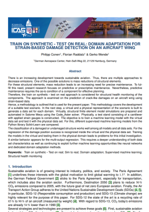


ML Regression Model	Linear Regression		Elastic Net		Polynomial Features		Random Forest	
Data Points	1499	440706	1499	440706	1499	440706	1499	440706
R^2 [-]	0,741	0,101	0,679	0,001	0,736	0,834	0,809	0,999
Mean-Squared-Error [mm]	4,474	0,643	5,885	0,719	4,132	0,119	3,512	0,001
Mean-Absolute-Error [mm]	1,233	0,679	1,544	0,734	0,989	0,190	0,809	0,001

Figure 4: Machine learning models: Input data and performance.

Deutsches Zentrum
für Luft- und Raumfahrt
German Aerospace Center





[1] L. Alzubaidi et al.:"Review of deep learning: concepts, CNN architectures, challenges, applications, future directions", Journal of big data, 2021