

# Lessons Learned in Developing a Convolutional Neural Network to Identify Damage Modes in Type-IV Composite Overwrapped Pressure Vessels

Finja Carlotta Rentzsch Holm, Tobias Schalm, Finn Eggers, Kai-Uwe Schröder

**Abstract.** Fiber composite structures are susceptible to various difficult-to-predict damage modes, including delamination, matrix cracking, and fiber breakage. The reliable identification of both, damage mode and location within the composite presents a distinctly complex challenge in Structural Health Monitoring. This work explores this challenge using Composite Overwrapped Pressure Vessels (COPVs) with surface-strain monitoring as a case study. Thereby, Convolutional Neural Networks (CNNs) are utilised for damage localisation and identification as they are particularly suited to process the emergent grid-like sensor data and to match a deviation in the surface-strain field to the corresponding damage states. The surface strain data used for supervised learning in this study is extracted from a finite element model for both an undamaged reference state and various damage conditions. These damage conditions consist of fibre breakage and matrix cracking, which are numerically modelled for different damage sizes and relative positions along the tank's main axis. The CNN used in this work is based on the U-Net architecture first introduced by Ronneberger et al. This study discusses the key lessons learned during the development of the CNN, including the choice of input data, suitable network architectures, training strategies, the impact of sensor spacing on performance, and challenges related to repeatability and generalisability.

**Key words:** Convolutional Neural Networks (CNNs), Smart Structures, Damage State Identification, Structural Health Monitoring (SHM), Hydrogen Pressure Vessels, Composite Overwrapped Pressure Vessels (COPVs)

## 1 INTRODUCTION

According to the European Green Deal [1], carbon neutrality within the EU shall be achieved by 2050. Renewable hydrogen is considered an essential energy vector towards decarbonisation of the transportation sector. In addition to production, transportation and utilization, hydrogen storage represents a critical link to the industrial chain for hydrogen energy production. Regarding hydrogen storage equipment and fuel cell vehicles, high-pressure storage tanks for gaseous

---

Finja Carlotta Rentzsch Holm & Tobias Schalm & Finn Eggers & Kai-Uwe Schröder  
Institute of Structural Mechanics and Lightweight Design, RWTH Aachen University, Wüllnerstraße 7, 52062 Aachen, Germany, e-mail: finja.carlotta.rentzsch.holm@rwth-aachen.de

hydrogen have prevailed over solid- and liquid-state based methods. In gaseous hydrogen storage, Type IV Composite Overwrapped Pressure Vessels (COPVs) continuously gain popularity due to superior mass storage efficiency of composite structures in comparison with those made of metallic alloys [2]. However, manufacturing costs will have to be reduced significantly to make hydrogen-powered road, water, and air vehicles competitive and to achieve the necessary economies of scale for broad market penetration. One cost driver is material costs, especially the carbon composite used in Type IV COPVs [3]. As a result of uncertainties related to the material properties, manufacturing and potentially in-service loads, COPVs tend to be overdimensioned. In this context, in-service Structural Health Monitoring (SHM) has the potential to greatly increase the confidence in COPVs, ultimately enabling lighter designs. Furthermore, SHM can help to transfer from a *Safe-Life (no damage during lifetime)* design philosophy to a *Fail Safe* and *Damage Tolerance* philosophy, improving efficiency while retaining or even improving reliability and safety when compared to a traditional Safe-Life philosophy. [4, 5, 6, 7]

Common SHM measurement concepts include strain-based, vibration-based and ultrasound-based methods. Strain-based methods are typically employed for static or quasi-static load conditions and based on the idea, that damage locally redirects the load path, leading to a measurable strain redistribution [5]. Common strain sensor technologies include electrical strain gauges and fiber optic sensors. For COPV monitoring, fiber optic sensors are particularly well suited as they can be easily embedded during the winding process. With multiple sensor positions integrated in a single optical fiber, fiber optic sensors also require significantly less cabling effort than strain gauges. Multiple works found in published literature have leveraged strain-sensing based on fiber optic sensors for health or load monitoring of Type IV COPVs [8, 9, 10, 11].

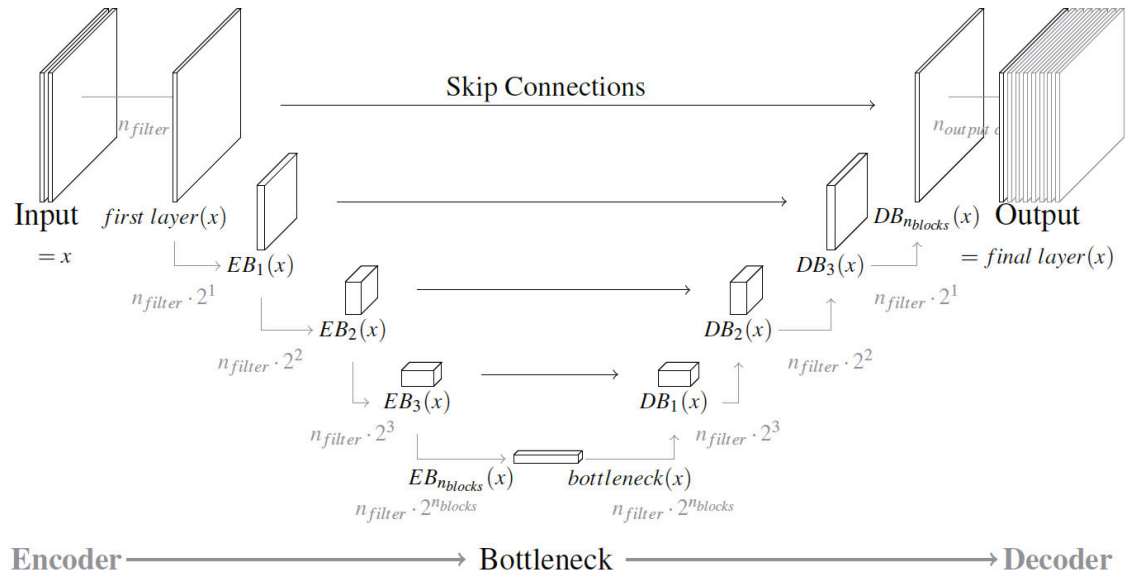
When analysing strain data, Artificial Neural Networks (ANNs) offer a potent tool to handle the complexity arising from the structure itself and its various possible damage modes, as well as uncertainties related to manufacturing, material properties, and loading conditions. This work describes an approach for damage detection, localization and identification based on Convolutional Neural Networks (CNNs) and the lessons learned during implementation.

## 2 THEORETICAL FOUNDATIONS

CNNs are a subset of ANNs and designed to process grid-structured input data with strong spatial dependencies (e.g. image data). For this work, CNNs with a U-Net architecture are utilized. The U-Net architecture was first presented by Ronneberger et al. [12] and originally developed for biomedical image analysis. The generalized architecture of a U-Net is depicted in Figure 1. It features a distinctive U-shaped structure comprising an encoder (contracting path), a bottleneck, and a decoder (expanding path). The encoder extracts hierarchical features through convolutional layers and max-pooling, reducing spatial dimensions while increasing feature depth. The bottleneck serves as a compressed representation of the input, capturing essential high-level features. The decoder then reconstructs the image using transposed convolutions (upsampling layers) to restore spatial resolution while refining segmentation details. A key feature of U-Net are its skip connections, which directly link corresponding layers in the

encoder and decoder, preserving fine-grained spatial information lost during downsampling.

U-Net architectures are particularly well-suited for classification tasks that require spatial localization within the output map—in this study, the precise position and extent of damage. Their ability to efficiently propagate features across layers while preserving spatial context makes them a powerful tool for image-based classification tasks. This work extends the application of U-Nets to pixel-like surface strain data, demonstrating their potential for accurate damage detection and localization in composite structures. [12]



**Figure 1:** Generalized U-Net architecture based on Ronneberger et al. [12].

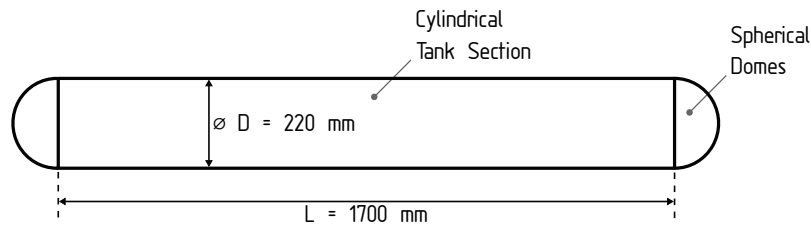
### 3 DEVELOPED METHOD

In this work, CNNs based on a U-Net-architecture are used for damage detection, localization and identification in Type IV hydrogen pressure vessels. Within this section, the input data, general methodology of the approach and constraints imposed by in-service applications are presented and discussed.

#### 3.1 Input data

To generate the input data used to train the CNN, a reference COPV is modelled using the Finite Element Method (FEM). It shall be noted, that the geometry, material and layup used in this study does not correspond to any validated tank design as information about commercially used designs generally may not be disclosed. Therefore, the authors have selected approximate outer dimensions as seen in tank designs for personal mobility and compressed hydrogen transport

applications, which are illustrated in Figure 2. The corresponding layup and material properties found in published literature are listed in Tables 1 and 2. The layer thickness provided in Table 2 is approximated using netting theory for a burst pressure of 1580 bar [13]. Fiber and matrix damage is modelled by reducing elastic material properties in the specified area according to Table 1. To reduce computational effort, it was assumed that the tank is symmetric in all three spacial directions (effectively modelling one-eighth of the tank) and only damage in the cylindrical section is considered. It is ensured that the area of influence of the stress redistribution caused by the investigated damage modes does not extend to the symmetry-boundaries in the FEM model.



**Figure 2:** Tank geometry used to generate the input data, based on tank designs found in personal mobility and compressed hydrogen transport applications.

**Table 1:** Material properties of composite materials for the reference state, matrix damage and fibre damage according to [14, 13]. (The Elastic Modulus  $E_i$  and the Shear Modulus  $G_i$  is provided in [MPa], the Poisson Ratio  $\nu_{ij}$  is dimensionless. The subscript 1 denotes the fiber direction.)

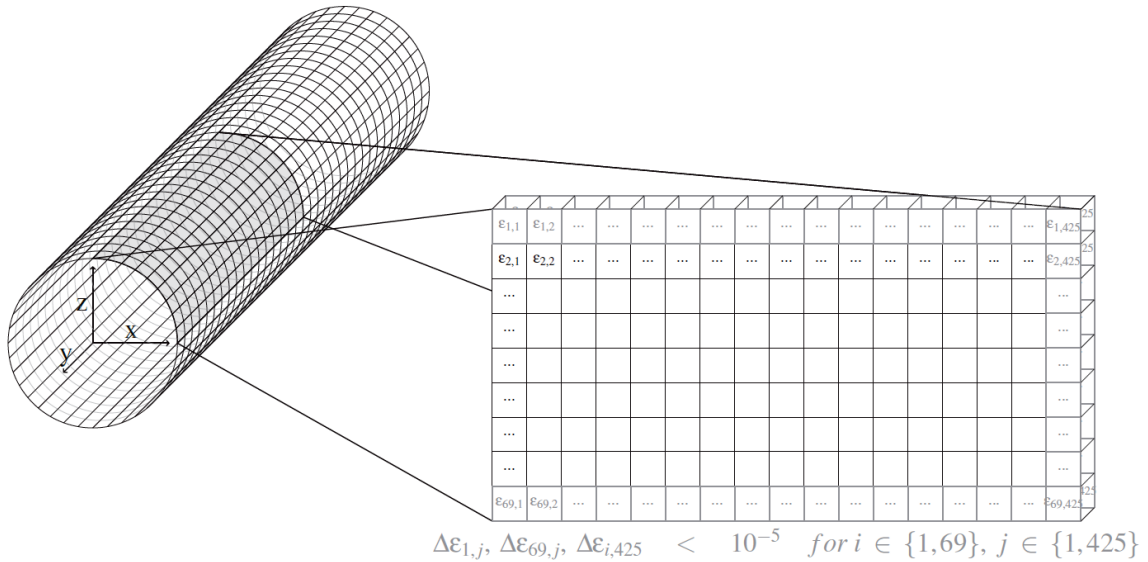
|                                        | $E_1$  | $E_2$ | $E_3$ | $\nu_{12}$ | $\nu_{13}$ | $\nu_{23}$ | $G_{12}$ | $G_{13}$ | $G_{23}$ |
|----------------------------------------|--------|-------|-------|------------|------------|------------|----------|----------|----------|
| Undamaged                              | 160000 | 7420  | 7420  | 0.28       | 0.28       | 0.3        | 3710     | 4790     | 3710     |
| Factor of reduction in case of damage: |        |       |       |            |            |            |          |          |          |
| Matrix damage                          | 1      | 0.1   | 0.1   | 1          | 1          | 0.1        | 1        | 1        | 0.1      |
| Fibre damage                           | 0.1    | 1     | 1     | 0.1        | 0.1        | 1          | 0.1      | 0.1      | 1        |

The surface-strain data of the cylindrical section of the eight-symmetric tank serves as input for training as illustrated in Figure 3. Matrix cracking or fibre breakage at different relative axial positions, layers and diameters are applied to each tank. The damage properties are generated randomly based on predefined boundaries. The surface strain data is read for 1000 damaged tank models as well as the undamaged reference tank model. A threshold value of  $\epsilon_{th} = 10^{-5}$  is introduced to accomodate the assumption that lower strains are not measurable in practical application. All strains below  $\epsilon_{th}$  are set to zero, effectively excluding these values and the

**Table 2:** Layer configuration of the composite hydrogen pressure vessel adapted from [15].

| layer          | 0    | 1          | 2          | 3          | 4          | 5    |
|----------------|------|------------|------------|------------|------------|------|
| thickness [mm] | 2.50 | 1.75       | 1.75       | 1.75       | 1.75       | 2.50 |
| angle [°]      | 89   | $\pm 22.1$ | $\pm 37.5$ | $\pm 22.1$ | $\pm 37.5$ | 89   |

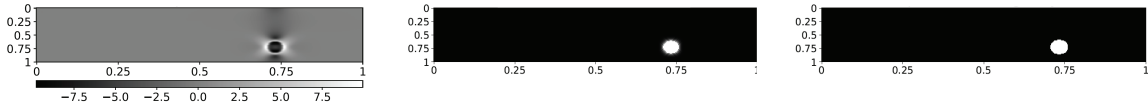
information contained within from the training process. To prevent overfitting, the dataset is augmented. Each damage is positioned along the circumferential centerline of the simulated tank segment. As the tanks are assumed to be rotationally symmetric, the input and target data is shifted in circumferential direction to generate augmented data. 100 random circumferential shifts between  $-45^\circ$  and  $45^\circ$  are performed on each damaged tank simulation resulting in a dataset of 53 100 input and target tensors.



**Figure 3:** Computation of input data by means of an FEM tank model. The cylindrical section of the FEM model is illustrated, whereby the surface strain data in longitudinal and circumferential direction is extracted for one eight-symmetric tank section.

Adequate data preprocessing can speed up the training process of ANNs, reduce network complexity and increase generalisability [16]. Data with varying scales can cause instabilities during training ANNs and must be adjusted to match the range of values expected by the input layer ( $[-1, 1]$  or  $[0, 1]$ ). For this work, z-score-normalisation is used for the input data. The z-score is a measure of how far a value deviates from the mean value of the data. Using z-index based standardization results is particularly suitable for normally distributed data. [16, 17]

The input data that is extracted from the numerical simulation is a tensor of dimension  $[2 \times 69 \times 425]$ : 2 strain values (deviation of the longitudinal and circumferential strain) are provided for all surface-elements of the modelled cylindrical tank section (69 in circumferential direction and 425 in longitudinal direction). The tank is discretized in 6 different layers, and it is assumed that fiber and matrix damage may occur in the inner five layers, thus focusing on damage not easily spotted in visual inspection. In order to detect both, fiber and matrix damage, in the 5 relevant layers, 10 output maps of size  $[69 \times 425]$  need to be predicted. This results in an output size of  $[10 \times 69 \times 425]$  with each output channel corresponding to a combination of damage type and layer. Damaged elements are labelled as 1 and undamaged elements are labelled as 0 in the channel corresponding to the respective damage type and layer. Label smoothing is used, adjusting the labels in close vicinity to the modelled damage. Label smoothing improves generalisability and accuracy of deep learning models [18]. Hard targets result in higher true logits and great differences between false logits. Therefore, using hard labels (i.e. assigning each element a binary value for damage / no damage), undamaged elements adjacent to damaged regions would be treated the same as undamaged elements far from damaged regions. Soft labels are introduced for target tensors between damaged and undamaged elements. Undamaged elements that are located at a distance of  $d$  ( $d \in [1, 16]$ ) from the damaged element are assigned values  $\frac{1}{2^d}$ . Exemplary input, target and output data is visualized in Figure 4 for matrix damage in layer 3.



**Figure 4:** Exemplary training data and results. Left: Input (Surface strain deviation in fiber direction  $\Delta\epsilon_1$ ), Middle: Target map (only for the channel corresponding to matrix damage in layer 3), Right: Output map (only for the channel corresponding to matrix damage in layer 3)

### 3.2 Approach for damage assessment in COPVs

During the first stage, training, validation and testing datasets are selected. The input and target dataset is randomly divided into training, validation and testing data with shares of 70 %, 10 % and 20 %, respectively. The next step is preprocessing of the input data and the creation of target data. This involves computing the strain deviation from the reference state, z-score-normalisation according to Equation ??, and the introduction of the aforementioned soft labels to the target data. Subsequently, different candidate topologies are generated and trained. The candidate topologies are created based on the general U-Net architecture illustrated in Figure 1. The encoder part is build up with  $n_{blocks}$  Encoder Blocks (EBs) that each contain  $n_{layer}$  convolutional layers and ReLU activation functions, batch normalisation, and max-pooling. The bottleneck has the highest degree of abstraction in relation to features of the data and incorporates multiple convolutional layers and ReLU activation functions. This is followed by the decoder block that is build up with  $n_{blocks}$  Decoder Blocks (DBs) that each contain  $n_{layer}$  ReLU

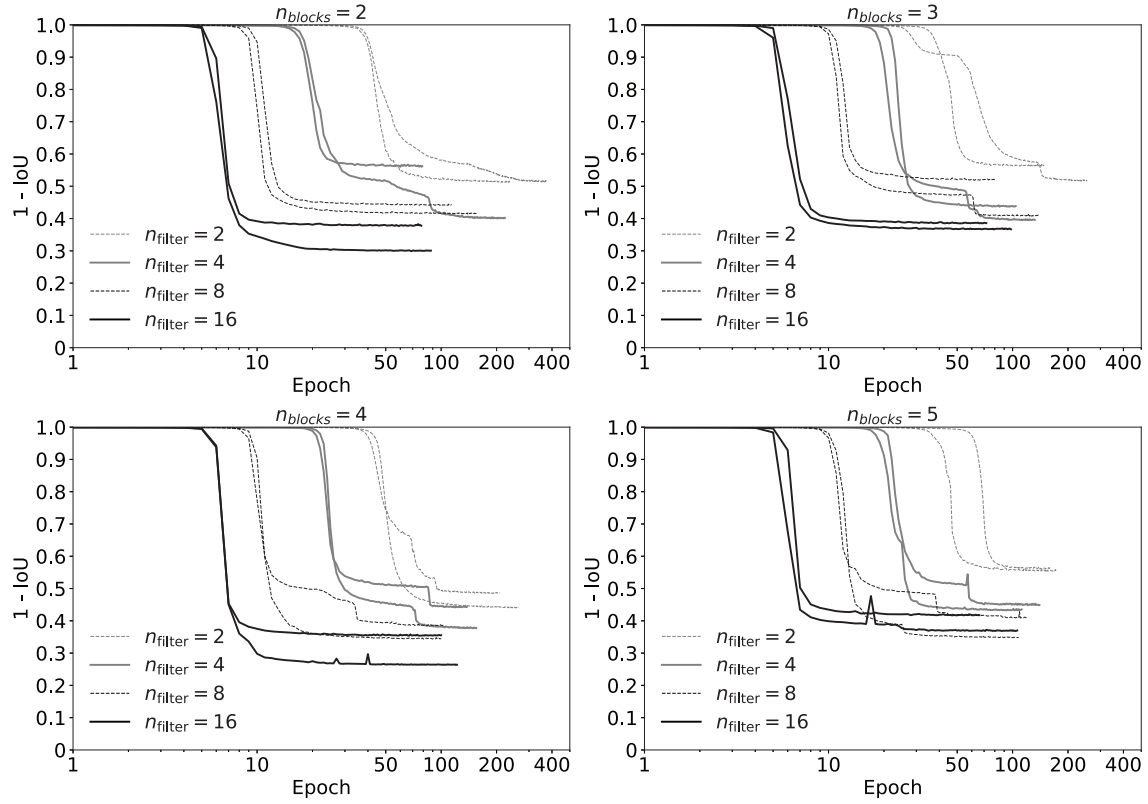
activation functions and convolutional layers, batch normalisation and ReLU activation analogous to the encoder. While the encoder reduces the dimensions of the data by applying maximum pooling, the decoder increases the dimensions to restore the original dimensions of the data by applying a transposed convolutional layer. The number of blocks  $n_{blocks}$  for the U-Net varies between 2 and 5 while the number of filters  $n_{filter}$  in the first block varies between 2 and 16. The number of EBs and DBs is determined by  $n_{blocks}$  and defines the degree to which the spatial dimensions of the input are compressed, indicating the level of abstraction applied to the input.  $n_{filter}$  indicates the range of features the network can learn. Skip connections are added between EBs and DBs to compensate for information loss during decoding and to improve feature integration [19]. Based on the network depth ( $n_{blocks}$ ), zero-padding is performed on the input and target data to ensure equal dimensions for the skip connections of the U-Net architecture. Therefore, the data is expanded with columns and rows of zeros to match the required dimensions. The best network topology is chosen based on the final validation error. The process of creating and training new candidate topologies is repeated until the validation error is sufficiently small [20]. For the loss function used in training, the Intersection over Union (IoU) is chosen. The IoU loss  $L_{IoU}$  is a suitable loss function for binary classification, in particular object detection or semantic segmentation. The IoU loss  $L_{IoU}$  is based on the overlap between the predicted segmentation mask and the ground truth mask and is calculated according to Equation 1 [21].  $TP$  amounts to the number of damaged elements predicted as damaged, while  $FP$  amounts to the number of undamaged elements predicted as damaged. Accordingly,  $FN$  describes the number of damaged elements that is predicted as undamaged.

$$L_{IoU} = 1 - \frac{TP}{TP + FP + FN} \quad (1)$$

During training, numerical instability is addressed by implementing ReLU activation, batch normalisation as well as gradient normalisation and gradient clipping at a threshold value of 1. Furthermore, the AdamW (Adam with decoupled weight decay) optimizer is applied showing superior performance for (over-parameterized) Deep Convolutional Neural Networks (DCNNs) compared to the Adam optimizer. Finally, learning rate scheduling is implemented based on a starting learning rate of  $10^{-4}$  that is reduced by 10 % if the validation loss fails to improve for five consecutive epochs until a minimum learning rate of  $10^{-6}$  is reached. High learning rates speed up the initial training process, whereas lower learning rates are favored in later stages to ensure stable model convergence. [22, 23, 24]

### 3.3 Initial results for damage detection and localization in COPVs based on a numerical test case

The progression of the validation loss function for different network topologies during training is depicted in Figure 5. The (expected) difference in performance after repeating the training for a model is referred to as *churn*. Churn can be attributed to different sources of randomness such as parameter initialisation and mini-batch ordering [25, 26]. On average, network performance improves with an increasing number of filters (see Figure 5). The IoU values calculated



**Figure 5:** Progression of the validation loss function over training epochs for different network topologies.

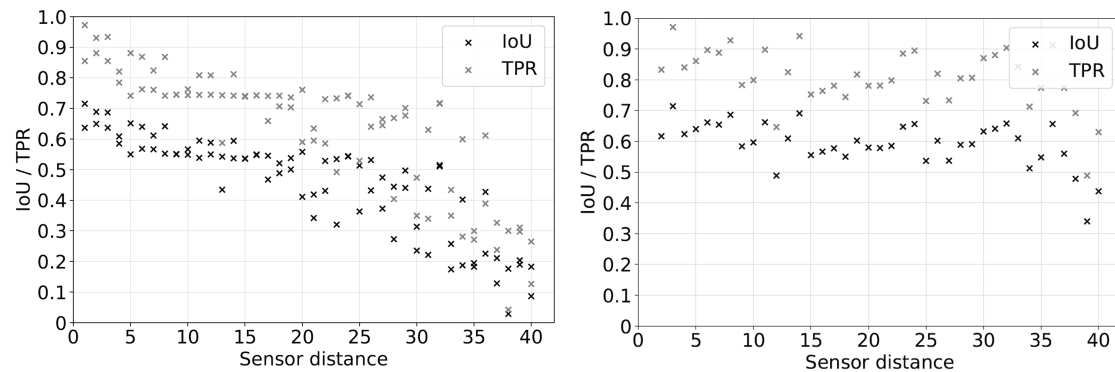
based on the validation and training data are sufficiently similar to conclude good generalisability and to indicate that the network was not overfitted on the training data. Measured by IoU, a network topology with  $n_{blocks} = 4$  and  $n_{filter} = 16$  represents the best model with a maximum IoU of 69.132 % on the validation data and 69.131 % on the training data. For this network topology, 93.198 % of the damaged elements are detected (corresponding to the Recall) and the Precision amounts to 74.815 %. Therefore, 25.185 % of the predicted damaged elements are actually undamaged. This data can be interpreted as follows: Generally, damage is detected, localized (position along the tank surface and damage depth), and identified (matrix or fiber damage) very well based on the numerical input data. However, a common overestimation of the damage size leads to a lower IoU value.

### 3.4 Investigating constraints imposed by in-service applications

So far, the presented proof-of-concept was conducted in a rather academic nature. Some steps have been undertaken to simulate a more realistic SHM scenario. It has already been considered that strain values below 10 microstrain cannot be accurately measured. When introducing random noise of  $\pm 10$  microstrain and using the network trained on ideal data for the

evaluation, the IoU drops to 58.4 % with a Recall (also referred to as True-Positive Rate (TPR)) of 84.4 % and a Precision of 78.6 %. This shows good generalizability of the algorithm. By including random noise in the training, it is expected that these values can be improved. Even if trained on only one damage mode at a time, the approach works well if multiple damage modes are present in the tank. 10260 modified tanks are created by randomly adding corresponding input and target values of 5 tanks from the complete input dataset. With multiple damage modes present in the investigated tank section, the network exhibits an IoU of 63.0 %, a Recall of 86.7 % and a Precision of 71.7 %.

Furthermore, it must be considered that the surface strain is most likely measured using FBGs. FBGs generally can only measure strain in one direction. However, including only the hoop strain of the outermost layer in the training process does not lead to a decrease of IoU, Recall or Precision. As such, the in-service health monitoring is viable if only relying on one strain direction. It must be further considered that sensors are spatially distributed on the tank surface with a larger spacing than the element size used in the FE-analysis (approx. 2 mm). This was modelled by reducing the size of the input data. As expected, the IoU decreases with increasing sensor distance. This decrease in performance can be partially circumvented by training a second CNN to reconstruct the missing strain data from numerical simulation and applying the CNN used for damage detection to the reconstructed data, see Figure 6.



**Figure 6:** IoU loss in dependence on sensor distance for validation data without data reconstruction (left) and with the dual-CNN data reconstruction approach (right).

#### 4 LESSONS LEARNED AND FUTURE WORK

The implementation of CNNs based on the U-Net architecture for damage detection, localization, and identification in Composite Overwrapped Pressure Vessels has provided several key insights. Firstly, the U-Net-based CNNs proved to be highly effective in accurately identifying and localizing damage within COPVs in a numerical test case. The network's ability to process spatially structured sensor data enables reliable mapping of deviations in the surface-strain field to underlying damage states. However, the damage size is commonly overestimated. Secondly,

it was observed that assessing only the longitudinal strain, which corresponds to utilizing a single sensor per sensing position in a real-world application, does not result in a significant reduction in performance for the numerical test-case. Furthermore, introducing a separate data reconstruction algorithm can enhance the CNN's performance when larger sensor distances are used. This technique mitigates the loss of spatial resolution due to sparse sensor placement, thereby preserving the network's ability to accurately infer damage characteristics even under suboptimal sensing conditions. Lastly, the results demonstrate good generalizability of the trained CNNs towards the presence of multiple damages. These findings underscore the potential of CNN-based approaches for damage identification in composite structures and provide valuable guidance for future experimental implementations and optimizations.

While the current study demonstrates promising results, several key areas for future research remain. Firstly, the repeatability of the CNN training shall be improved by adapting the loss-function and introducing soft labels encompassing multiple layers. Another crucial step towards improving the robustness of the CNN is the integration of manufacturing and load uncertainties as well as sensor noise within the numerical test case. Furthermore, a more realistic tank design will be used to generate training data. Additionally, extending the training dataset to include delaminations as a damage mode will provide a more comprehensive approach to SHM in COPVs. Lastly, the damage models shall be improved based on actual damage observed in CFRP-overwrapped pressure vessels during testing and service. It can be expected, that these steps greatly improve robustness and applicability of the approach.

## REFERENCES

- [1] E. Commission, "The european green deal. Communication from the commission to the european parliament, the european council, the council, the european economic and social committee and the committee of the regions," 2019.
- [2] G. Chi, S. Xu, D. Yu, Z. Wang, Z. He, K. Wang, and Q. Zhou, "A brief review of structural health monitoring based on flexible sensing technology for hydrogen storage tank," *International Journal of Hydrogen Energy*, vol. 80, pp. 980–998, 2024.
- [3] H. Yumiya, M. Kizaki, and H. Asai, "Toyota fuel cell system (TFCS)," *World Electric Vehicle Journal*, vol. 7, no. 1, pp. 85–92, 2015.
- [4] C. Viechtbauer and K.-U. Schröder, "Structural health control - a comprehensive concept for observation and assessment of damages," *13th Mechatronics Forum International Conference*, 2012.
- [5] A. Preisler, "Efficient damage detection and assessment based on structural damage indicators," Ph.D. dissertation, RWTH Aachen University, 2019.
- [6] D. F. Braga, S. Tavares, L. F. da Silva, P. Moreira, and P. M. de Castro, "Advanced design for lightweight structures: Review and prospects," *Progress in Aerospace Sciences*, vol. 69, pp. 29–39, 2014.

- [7] H. Rocha, P. Antunes, U. Lafont, and J. P. Nunes, “Processing and structural health monitoring of a composite overwrapped pressure vessel for hydrogen storage,” *Structural Health Monitoring*, vol. 23, no. 4, pp. 2391–2406, 2024.
- [8] G. Souza and Tarpani, “Using OBR for pressure monitoring and BVID detection in type IV composite overwrapped pressure vessels,” *Journal of Composite Materials*, vol. 55, no. 3, pp. 423–436, 2021.
- [9] P. Gasior, M. Malesa, J. Kaleta, M. Kujawińska, K. Malowany, and R. Rybczyński, “Application of complementary optical methods for strain investigation in composite high pressure vessel,” *Composite Structures*, vol. 203, no. 9, 2018.
- [10] I. G. Tapeinos, A. Rajabzadeh, D. S. Zarouchas, M. Stief, R. M. Groves, S. Koussios, and R. Benedictus, “Evaluation of the mechanical performance of a composite multi-cell tank for cryogenic storage: Part II – experimental assessment,” *International Journal of Hydrogen Energy*, vol. 44, no. 7, 2019.
- [11] W. Jiang, M. Liang, M. Schiebel, S. Zaremba, and K. Drechsler, “Development of machine learning based classifier for the pressure test result prediction of type IV composite overwrapped pressure vessels,” *International Journal of Hydrogen Energy*, vol. 58, no. 36, 2024.
- [12] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds. Springer, 2015, pp. 234–241.
- [13] Z. Hu, M. H. Chen, I. Zu, X. Jia, A. Shen, Q. Yang, and K. Xu, “Investigation on failure behaviors of 70 MPa type IV carbon fiber overwound hydrogen storage vessels,” *Composite Structures*, vol. 259, 2021.
- [14] W. Li, H. Lv, L. Zhang, P. He, and C. Zhang, “Experiment, simulation, optimization design, and damage detection of composite shell of hydrogen storage vessel - A review,” *Reinforced Plastics and Composites*, vol. 42, pp. 507–536, 2023.
- [15] C. Hopmann, N. Magura, R. Müller, D. Schneider, and K. Fischer, “Impact of winding parameters on the fiber bandwidth in the cylindrical area of a hydrogen pressure vessel for generating a digital twin,” *Polymer Composites*, pp. 1–13, 2022.
- [16] L. Yu, W. Shouyang, and K. K. Lai, “An integrated data preparation scheme for neural network data analysis,” *IEEE Transactions on Knowledge and data engineering*, vol. 18, no. 2, pp. 217–230, 2006.
- [17] C. Vercellis, *Business Intelligence: Data Mining and Optimization for Decision Making*. John Wiley and Sons Ltd, 2009.

- [18] R. Müller, S. Kornblith, and G. Hinton, “When does label smoothing help?” *33rd Conference on Neural Information Processing Systems*, 2019.
- [19] Y. Ji, H. Zhang, Z. Zhang, and M. Liu, “CNN-based encoder-decoder networks for salient object detection: A comprehensive review and recent advances,” *Information Sciences*, vol. 546, pp. 835–857, 2021.
- [20] I. Nunes da Silva, D. Hernane Spatti, R. Andrade Flauzino, L. H. Bartocci Liboni, and S. F. dos Reis Alves, *Artificial Neural Networks*. Springer, 2017.
- [21] J. R. Terven, D. M. Cordova-Esparza, A. Ramirez-Pedraza, E. A. Chavez-Urbiola, and J. A. Romero-Gonzalez, “Loss functions and metrics in deep learning,” *arXiv preprint*, 2023.
- [22] C. C. Aggarwal, “Neural networks and deep learning,” 2018.
- [23] M. A. Nielsen, “Neural networks and deep learning,” 2015.
- [24] S. Norouzi and M. R. Ebrahimi, “A survey on proposed methods to address adam optimizer deficiencies,” *Department of Electrical and Computer Engineering, University of Toronto*, 2019.
- [25] F. Renard, S. Guedria, N. de Palma, and N. Vuillerme, “Variability and reproducibility in deep learning for medical image segmentation,” *Scientific Reports, nature research*, vol. 10, 2020.
- [26] S. Bhojanapalli, K. Wilber, A. Veit, A. S. Rawat, S. Kim, A. Menon, and S. Kumar, “On the reproducibility of neural network predictions,” *arXiv*, 2021.

## ACKNOWLEDGEMENT

The project on which this contribution is based was funded by the Federal Ministry of Digital and Transport (Bundesministerium für Digitales und Verkehr) of the German government under the funding code 03B11024. The authors are responsible for the content of this publication. We are grateful for Battenberg Robotic GmbH’s partnership in this work.