High-resolution Data Capture and Interpretation in Support of Port Infrastructure Maintenance

AHMED ALAMOURI[1], VANESSA DE ARRIBA LOPEZ[1], PEDRO ACHANCCARAY DÍAZ[1], JAN BACKHAUS[2] & MARKUS GERKE[1]

***Abstract**: German seaports feature approximately 3,000 km of quay walls and 2,500 facilities. Many of these facilities are in poor conditions, according to the Federal Waterways Engineering and Research Institute. As a result, it has become an urgent necessity to inspect such port infrastructures. In this paper, we present an inspection framework that deals with such kind of port infrastructures. It provides innovative methods and procedures that involve AI-technologies. This enables using structured datasets, namely drone-based images, in a fully digital process for inspection purposes. The framework emphasises automation, including the following steps: flight planning and image capturing, processing of image dataset, and evaluating visible damages on infrastructure surfaces. All steps and their experiments are being achieved within the joint research project port-AI funded by BMDV (Federal Ministry for Digital and Transport).*

1 Introduction

German seaports play a crucial role in securing economic power and prosperity in various industries and trades in Germany. In 2022, German seaports handled 279 million tons of goods on approximately 140 km of quay walls (STATISTISCHES BUNDESAMT, 2023). Overall, Germany has around 3,000 km of quay walls in ports and on waterways, and roughly 2,500 km of facilities on federal waterways, with gross fixed assets worth around €50 billion. However, about 80% of the locks and weirs are more than 50 years old, and 30% are more than 100 years old, with a theoretical service life of approximately 80-100 years. Studies conducted by the Federal Waterways Engineering and Research Institute BAW indicate that only about 70% of these facilities are in adequate condition (BAW 2018; BAW 2019). This implies that an increasing amount of personnel and technical resources will be required in the coming years to ensure the operation of such infrastructures. In addition, the demands for a safe and reliable management of seaport infrastructure facilities are constantly increasing due to globalized trade and changing shipping traffic. At the same time, personnel resources are becoming increasingly scarce. This situation led to a surge in questions and inquiries regarding the infrastructure's service life. These inquiries are related to the condition of the structure, its intended use, and the remaining service life. BAW has taken practical measures in recent years to meet the requirements for various services such as life and maintenance management systems. They have developed, for instance, a prioritization methodology for the maintenance strategy of hydraulic transportation structures, based on structural inspections, structural indicators and costs. However, despite these measures, there are still some difficulties in practical implementation, for instance in issues related to infrastructure inspection such as damage

[1] Technical University of Braunschweig, Institute of Geodesy and Photogrammetry, Bienroder Weg 81, D-38106 Braunschweig, Germany, E-Mail: [a.alamouri, vanessa.de-arriba-lopez, p.diaz, m.gerke@tu-braunschweig.de]

[2] Technical University of Braunschweig, Institute of Flight Guidance, Hermann-Blenk-Str. 27, D-38108 Braunschweig, Germany, E-Mail: [j.backhaus@tu-braunschweig.de]

detection and documentation (BAW 2019). Accordingly, we are working through the joint research project port-AI (TU BRAUNSCHWEIG 2021), funded by BMDV in the framework of the IHATEC-call, to find some effective solutions towards damage detection of seaport infrastructures. The project focuses on realising a fully digital twin and developing an inspection framework for seaport infrastructures with a focus on quay walls. Therefore, we aim in this paper to introduce the mentioned framework and its tests and experiments under development. The framework provides a fully digital process for image-based inspection using innovative methods and procedures that incorporate drone-based image acquisition and AI-technologies. Very high-resolution images from drones are processed to obtain a comprehensive insight into the structure's geometry and condition. Based on those data a methodology is employed to perform damage detection. In the following we first describe the overall framework, comprising image data capture, processing and damage detection. Section 3 presents experimental results while the paper closes with a discussion.

2 Related work and inspection framework

The framework is designed with a focus on automation and includes the following steps, as illustrated in Fig. 1: capturing of drone-based image dataset, photogrammetric process of the image data, and detecting of visible damages on infrastructure surfaces. The final step is dedicated to detecting cracks and spalling on quay walls, as these can directly impact the safety of facilities and their operation. Each of these steps will be discussed in detail in the following subsections.

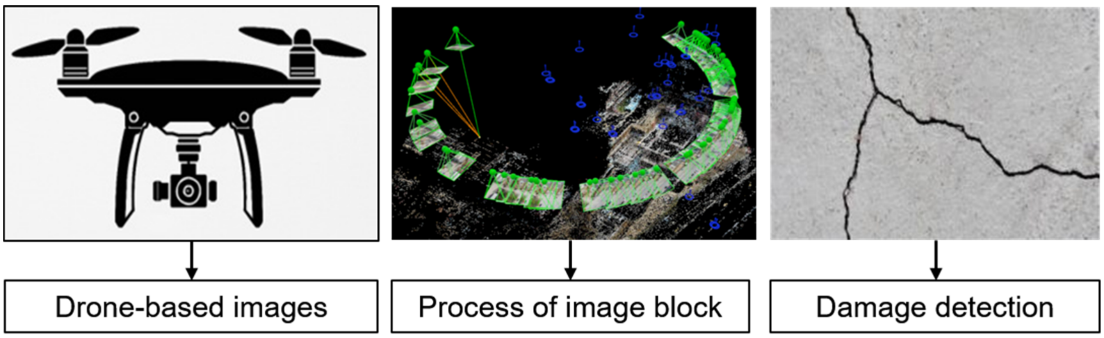

Fig. 1: Steps of the inspection framework

2.1 Acquisition of drone imagery

For some time now, drones have been a practical way of conveniently collecting image data of a complete infrastructure structure to be inspected. Developments in commercially available aircraft are leading to increasing reliability and user-friendliness, constantly increasing their usage in this field. However, automated use still poses a challenge in the context of critical infrastructure, as obstacles and disruptions often occur on site. For example, this could be moving cranes, which interfere with flight operations or block the recording of image data. Such disruptions require special attention to the local conditions and detailed planning of flight operations (GHASSOUN et al. 2021). Also, complexity of the object's geometry demands more viewpoints and different viewing angles, which have to be considered in the path planning.

The classic profile for inspection and mapping flights is the "lawnmower approach" (SANTOS et al. 2014; ASSMANN et al. 2018). Such a profile can be planned and automatically executed with almost any available flight control software. A further development is "terrain-following" profiles, in which the flight altitude is varied depending on the elevation of the terrain. Some commercial ground control stations for UAVs, such as "UgCS" from SGH Engineering, support such profiles.

Both profiles have in common that the parallel paths achieve a constant overlap in the resulting images. In most cases, overlap is a controlling parameter for path generation. Another is the Ground Sampling Distance (GSD), which specifies the distance between two points on the object that are mapped by neighbouring pixels. A lower value can be achieved by one of three factors or a combination of these: the resolution of the sensor, the field of view of the camera and the distance between the camera and the object. However, all three cannot be varied indefinitely. A higher resolution sensor is generally associated with a lower light yield per pixel and usually also results in a higher camera weight, which has a major impact on the flight time. A very narrow field of view is geometrically less favourable in the context of a three-dimensional reconstruction and can result in a lower accuracy in the Z-axis (MABOUDI et al. 2023; YUN 2020). The distance between the camera and the object cannot be reduced arbitrarily either, as it is associated with an increased risk of collision.

If the direct view is not blocked by obstacles, the most effective way to generate a data set that is suitable for the three-dimensional reconstruction of a building to be inspected is often to use a grid of images that are aligned perpendicular to the surface (nadir for horizontal planes, horizontal view for vertical planes). In addition, these images can be best used for automated damage detection even without reconstruction. In areas where a nadir image is not possible, e.g. due to obstacles, several oblique images must be taken. In general, such oblique images are also a way to increase the geometric quality and completeness of a three-dimensional reconstruction. However, if taken at every spot they always result in a multiple requirement for flight time, which is why a balance must always be found between the effort and benefit in the specific case.

Simple geometries such as quay walls can usually be approximated by combining several flat surfaces, whereby a nadir path can be planned in parallel paths for each one. Due to the feature-based mode of operation of the reconstruction algorithms, it must be ensured that points in the different perspectives are recognized as identical at the connection points of these surfaces. At an angle of 90 degrees, such as at the edge of a quay wall, this is generally not possible. Empirical tests carried out as part of the port_AI project show that the introduction of two paths in the transition can be sufficient for simple geometries and favourable conditions. This reduces the angle between two neighbouring perspectives to 30 degrees. MORGENTHAL et al. (2019) used a smaller angle of 22.5 degrees, which provides additional robustness against local geometric deviations and adverse optical effects in imaging. They also point out that a translation of the camera must be ensured between all images, which must be considered especially for concave corners of the structure.

The more complex the geometry of the structure to be inspected and the more disturbances occur in the immediate vicinity, the more important it is to visualize the flight in three dimensions for everyone involved. In addition to flight planning, this is also particularly relevant for flight execution, as the safety pilot is unable to adequately record progress along paths in three-dimensional space with a classic 2D map display. The project places particular emphasis on enabling the pilot to interact with the virtual 3D world shown on a display in a way that minimizes distraction.

2.2 Processing of image blocks

Many parts of the pipeline developed in the port-AI project benefit from a three-dimensional reconstruction of the structure from the image data. For this purpose, a pipeline consisting of feature extraction, structure-from-motion processing and dense matching is used. All these steps can be carried out with commercial software such as Agisoft Metashape or Pix4D.

Referencing the image block poses a challenge. Calibration of camera intrinsic is possible using only the features extracted from the images themselves, if a suitable flight planning was conducted. To determine the scale and global positions, external information is required. Several sources are available here. Firstly, the drone's position data at the time the image was taken is almost always available. These values are mostly GNSS-based and ideally, accuracies in the low centimetre range are possible with RTK systems. The high number of images largely compensates for large parts of the random and rapidly variable GNSS errors in relation to the referencing of the entire block. Systematic errors, such as the offset between the camera and the antenna centre, however, remain and can at best be corrected manually. Due to the usually quite soft connection between the camera and the drone, this offset is also dependent on acceleration.

In practice, an additional support of the block by independently measured targets as Ground Control Points (GCP) is therefore common. The use of scale bars as control parameters is not suitable for the application investigated in the port-AI project, but they are used as check parameters, as their introduction into the measurement area requires little effort.

One problem in many cases is that the GCPs cannot be applied evenly across the entire area. In the case of a quay wall, the positioning and measurement of targets on the upper surface is easy, but on the vertical surface of the wall this is only possible with considerable effort. However, our tests showed that a reliable reconstruction is still possible if the oblique paths in the transition between the surfaces are considered as described above. In this case, the referencing of the targets is also transferred to the vertical surface to a sufficient extent due to the relative orientation of the block.

Various products can also be derived from the resulting dense point cloud. In the port-AI project, textured meshes with a geometric resolution of a few millimetres or even below are generated. Firstly, these are used to support the building inspector during an on-site inspection by displaying them in an inspection app developed by project partners. Secondly, they are used to improve flight planning for future epochs by displaying them in the flight planning and inspection app.

By defining surfaces of high interest in the model, orthoimages for these areas can also be exported, which are used for crack detection (section 2.3). Unlike in the individual images, even long cracks are displayed uninterrupted in the orthoimage. They also represent the best possible vertical image of the corresponding point on the object at each point.

2.3 Damage detection – methods and workflow

A method has been developed to detect damages using high-resolution orthoimages. This method involves three steps: pretraining, fine-tuning, and geometric characteristics of detected damages. The proposed methodology employs deep learning (DL) models for semantic segmentation to detect cracks, and digital image processing techniques to identify the geometric characteristics of the cracks.

During the pretraining step, a semantic segmentation DL model is trained in a supervised manner using a public dataset containing images in the same domain as our application. This is performed to take advantage of the high number of annotated images, where the classes of interest (for instance cracks) is delineated, available in public datasets, and to reduce the number of images to be annotated from our dataset of acquired images. For this purpose, the DeepCrack (LIU et al. 2019) dataset was employed, which comprises 537 RGB-colour images (300 for training and 237 for testing), with 544×384 pixels, containing cracks (photos of real cracks and images collected from the internet) on asphalt and concrete with different background textures, as shown in Fig. 2. Then, a UNet (RONNEBERGER et al. 2015) architecture, with VGG19 (SIMONYAN & ZISSERMAN 2014) as the backbone, was trained on DeepCrack (images were resized to 300×300) with different hyperparameters to obtain the highest accuracy in the test set.

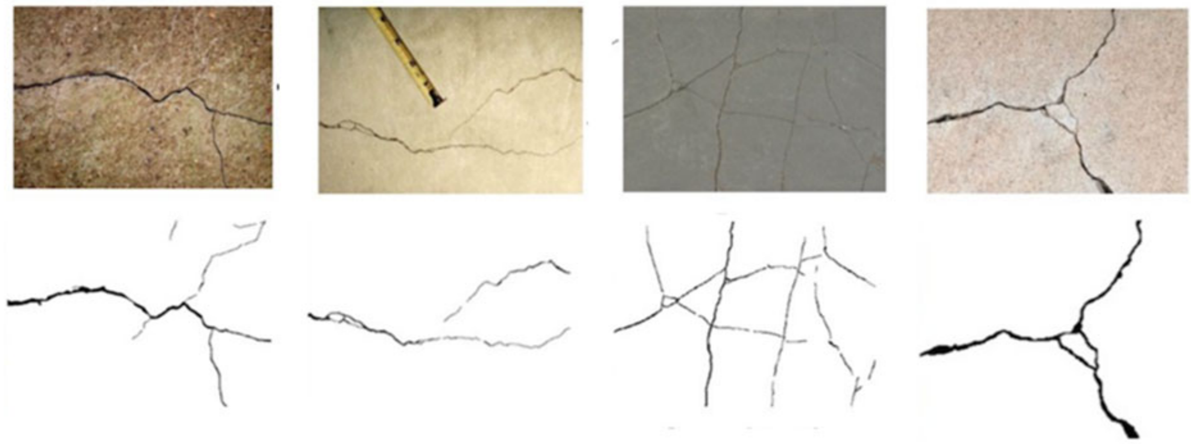

Fig. 2: Samples of images (top) and annotations (bottom) from the DeepCrack dataset

In the fine-tuning phase, we used our orthoimages and reference data to fine-tune the pretrained model with DeepCrack. The used orthoimages show quite different damage structure than the images from the public dataset and thus are manually annotated by visually inspecting and delineating the structural elements of interest (see Fig. 3): cracks (structural defects), joints (points of structural connection), steel (components), and concrete (structures). Two orthoimages - with sizes of 5272×1423 and 3285×1712 pixel - were annotated, and patches of 300×300 were extracted from each orthoimage to fine-tune the pretrained model. Damage detection was performed only on the "cracks" class, while the remaining classes were considered as background.

Fig. 3: Manually annotated orthoimage with different classes of structural elements

However, it is noted that long cracks have changes in their widths and orientation angles. For this, and to analyse and characterize such long cracks reasonably, we divided the orthoimage into a grid with cells of the same size (Fig. 4).

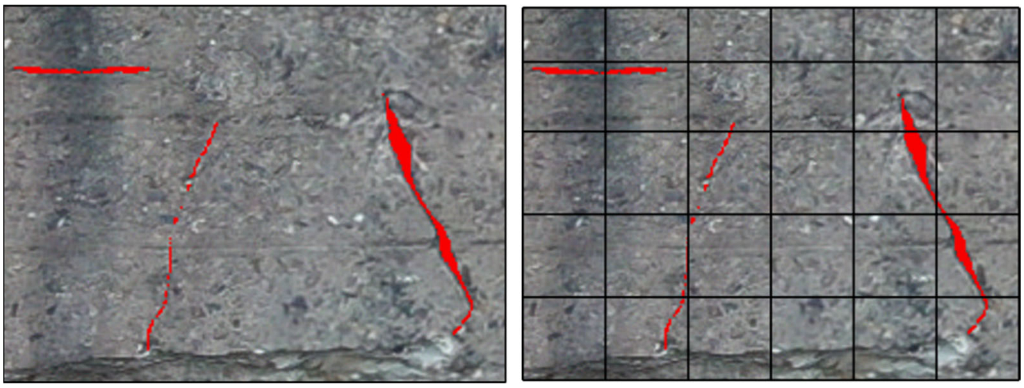

Fig. 4: An orthoimage with detected cracks (left), and the corresponding grid-image generated

In the last step of the above-mentioned method, we used digital image processing techniques to analyse each cell individually and determine the width and orientation angle of each crack-segment included in the entire cell (Fig. 5). First, the crack-segment area detected (in red) was optimized/refined using a specific threshold that excludes non-crack pixels. Next, a bounding box (in green) was determined for the valid crack-segment area. Depending on the orientation of the bounding box, the angle was estimated in the range of [-180° to 180°] taking the horizontal line as a reference. Afterwards, the midpoints (p1 and p2) between specific pairs of the bounding box were calculated, as far as they comprise the crack. The Euclidean distance (d) between these midpoints, which generate a perpendicular line across the crack, was utilized to calculate the crack width using the pixel size of the original orthoimage. Finally, the results, including the adjusted midpoints, width, and angle, were appended to a data array for further analysis.

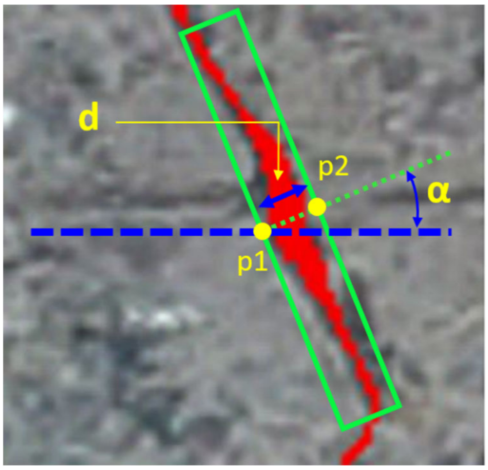

Fig. 5: Characterising a crack-segment including in an image cell: crack-segment area (red), bounding box (green), p1 and p2 are midpoints of the bounding box, (d) is the width and α is the orientation angle

3 Results

3.1 Image acquisition and processing

Experiments have shown that a pipeline such as the one described above is capable of reliably capturing maritime infrastructure of medium complexity and generating the image data required for associated applications such as crack detection. With suitable cameras such as the PhaseOne iXM-100, images with a GSD of under one millimetre are possible. In the experiment which generated the images below, a GSD of 1.53 millimetres was achieved.

Fig. 6 shows an example of a generated orthoimage of the front surface of a quay wall, as used for crack detection. Depending on the hardware, the processing time required to achieve these results ranges from a few hours to a few days and is largely automated.

Fig. 6: An orthoimage of a quay wall front with a cutout in full resolution. Visible cracks in the free concrete surface and behind the ladder

3.2 Crack detection

The results obtained in our experiments are summarized in Table 1, where different experiments were performed to measure the effect of data augmentation during training and fine-tuning using the annotated images. Two metrics are presented: Intersection over Union (IoU), which measures the level of agreement between the detected crack and annotation, and accuracy, which measures the number of correctly classified pixels for all classes (cracks and background). As each crack represents a very small portion of the image, the accuracy is expected to be higher than that of the IoU. The metrics were computed in the test set of the DeepCrack dataset because all manually annotated images were used for fine-tuning the pre-trained model.

Table 1: Results obtained for crack segmentation using data augmentation and fine-tuning

Model	Data Augmentation	Fine-tuning	Dataset train	IoU (Test)	Accuracy (Test)
UNet VGG19	No	No	DeepCrack	48.6	95.2
UNet VGG19	Yes	No	DeepCrack	81.7	97.9

We can see the positive effect of using data augmentation during training, where IoU and accuracy increased by 33.1% and 2.7%, respectively. Since all reference data was used to train the improved model, we did not yet employ the same kind of reference for testing the approach quantitatively.
However, the fine-tuned model was able to recognize cracks in the available orthoimages from our datasets, as shown in Fig. 7. Within current results, just a visual inspection has been done as we do not have enough annotated orthoimages to analyse the results which is scope of future works. However, Fig. 7 shows two sections of an orthoimage that was not used during the training with detected cracks (highlighted in red) provided by the fine-tuned model and the width and the slope angle computed for each crack.

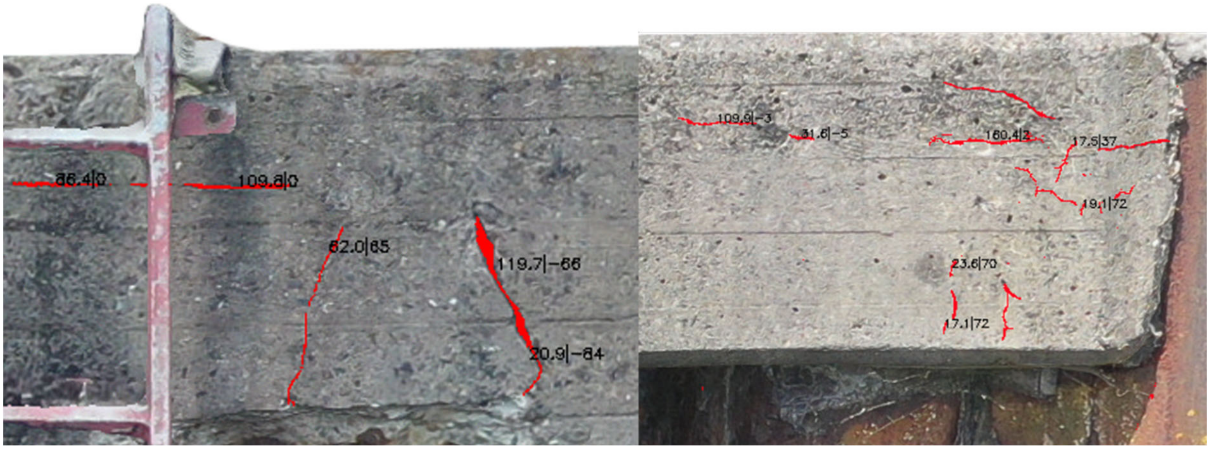

Fig. 7: Results obtained using the fine-tuned model for crack segmentation. Cracks (red) and geometric features (black text), namely the thickness and slope of each crack are represented

4 Summary and future work

Seaport infrastructures play a vital role in the maritime commercial sector in Germany. It is therefore crucial to monitor and maintain them regularly. To address this issue, we have presented an inspection framework in this paper that uses AI-technologies within the joint research project port-AI. The framework focuses on automation and includes three main steps: image capture using drone technology, photogrammetric evaluation of the image dataset, and detecting visible damages on infrastructure surfaces. The last step is especially important as it helps to identify cracks and spalling on quay walls that can directly affect the safety of the facilities and their operation.

The study focused on using drones for capturing images of seaport infrastructures that are used to detect cracks. The initial results showed that our inspection framework is reliable for medium complexity infrastructure. However, we plan to improve the automated flight execution by including moving obstacles like ships. In addition, minimizing a pilot distraction for safer and simpler navigation is a key future task. Furthermore, we will conduct experiments on other structures to investigate the relationship between the minimum detectable crack width, the GSD in the images and the resulting flight paths.

The current damage detection methods used for damage detection can identify cracks in the available orthoimages. However, we aim to enhance the proposed methodology by including the detection of additional damages such as spalling. To achieve this, we will use incremental learning methodologies (ZHOU et al. 2023) that allow us to add new annotations specific to diverse types of damages without having to retrain the developed model. This expansion will not only improve the versatility of our model but also increase its robustness by addressing a wider range of structural issues.

5 References

ASSMANN, J., KERBY, J., CUNLIFFE, A. & MYERS-SMITH, I., 2018: Vegetation monitoring using multispectral sensors - best practices and lessons learned from high latitudes. Journal of Unmanned Vehicle Systems, **7**(1), 54-75, https://doi.org/10.1101/334730.

BAW, 2018: Forschungsprogramm Verkehrswasserbau der Bundesanstalt für Wasserbau – Kompetenz für die Wasserstraßen, Heute und in Zukunft. https://www.baw.de/content/publikationen/pressemitteilungen/0/00221_2015.pdf, letzter Zugriff 27.01.2024.

BAW, 2019: Kompetenzzentrum für den Verkehrswasserbau. BAW-Strategie 2030. https://www.baw.de/content/publikationen/www-dokumente-oeffentlich/0/BAW_Kompetenzzentrum_Verkehrswasserbau_Web.pdf, letzter Zugriff 27.01.2024.

GHASSOUN, Y., GERKE, M., KHEDAR, Y., BACKHAUS, J., BOBBE, M., MEISSNER, H., TIWARY, P. K. & HEYEN, R., 2021: Implementation and Validation of a High Accuracy UAV-Photogrammetry Based Rail Track Inspection System. Remote Sens. 2021, **13**, 384. https://doi.org/10.3390/rs13030384.

FEDERAL MINISTRY FOR DIGITAL AND TRANSPORT, 2024: https://bmdv.bund.de/EN/Home/home.html, letzter Zugriff 28.01.2024.

LIU, Y., YAO, J., LU, X., XIE, R. & LI, L., 2019: DeepCrack-A deep hierarchical feature learning architecture for crack segmentation. Neurocomputing, **338**, 139-153.

MABOUDI, M., HOMAEI, M., SONG, S., MALIHI, S., SAADATSERESHT, M. & GERKE, M., 2023: A Review on Viewpoints and Path Planning for UAV-Based 3-D Reconstruction. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 5026-5048, https://doi.org/10.1109/JSTARS.2023.3276427.

MORGENTHAL, G., HALLERMANN, N., KERSTEN, J., TARABEN, J., DEBUS, P., HELMRICH, M. & RODEHORST, V., 2019: Framework for automated UAS-based structural condition assessment of bridges. Automation in Construction, **97**, 77-95, https://doi.org/10.1016/j.autcon.2018.10.006.

RONNEBERGER, O., FISCHER, P. & BROX, T., 2015: U-Net: Convolutional Networks for Biomedical Image Segmentation. Proceedings of 18th International Conference in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015, Lecture Notes in Computer Science, **9351**, Springer, Cham, 234-241, https://doi.org/10.1007/978-3-319-24574-4_28.

SANTOS, G. A. M., BARNES, Z., LO, E., RITOPER, B., NISHIZAKI, L., TEJEDA, X., KE, A., LIN, H., SCHURGERS, C., LIN, A. & KASTNER, R., 2014: Small Unmanned Aerial Vehicle System for Wildlife Radio Collar Tracking, 761-766, https://doi.org/10.1109/MASS.2014.48.

SIMONYAN, K. & ZISSERMAN, A., 2014: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

STATISTISCHES BUNDESAMT, 2023: Seeverkehr 2022: Güterumschlag 3,2% geringer als im Vorjahr. https://www.destatis.de/DE/Presse/Pressemitteilungen/2023/03/PD23_092_463.html, letzter Zugriff 28.01.2024.

TU BRAUNSCHWEIG, 2021: port_AI - A fully digital twin for port structures using IoT, 5G, BIM, AR and AI techniques to build smart building lifecycle management. https://www.tu-braunschweig.de/en/igp/projects/port-ai, letzter Zugriff 28.01.2024.

ZHOU, D. W., WANG, Q. W., QI, Z. H., YE, H. J., ZHAN, D. C. & LIU, Z., 2023: Deep class-incremental learning: A survey. arXiv preprint arXiv:2302.03648.

YUN, Z., 2020: Research on Aerial View Point Planning of Drone based on Multi-View. IOP Conference Series: Materials Science and Engineering, **739**, https://doi.org/10.1088/1757-899X/739/1/012055.