Tunnelling and Underground Space Technology 140 (2023) 105310

Contents lists available at ScienceDirect

Tunnelling and Underground Space Technology incorporating Trenchless Technology Research

journal homepage: www.elsevier.com/locate/tust

Automated pixel-level crack monitoring system for large-scale underground infrastructure – A case study at CERN

Aohui Ouyang [a,b], Vanessa Di Murro [b], Martin Cull [b], Roddy Cunningham [b], John Andrew Osborne [b], Zili Li [a,*]

[a] Civil, Structural & Environmental Engineering, University College Cork, Cork, Ireland
[b] SCE-SAM-FS, European Organization for Nuclear Research, CERN, Geneva, Switzerland

ARTICLE INFO

Keywords:
Image-based deep learning
Crack segmentation
Crack density
Crack spatial distribution
Tunnel deformation mode

ABSTRACT

In concrete tunnel linings, cracks usually appear and develop as an early sign of structural degradation prior to severe intolerable serviceability damage. The monitoring and assessment of crack spatial distribution can highlight long-term tunnel structural behavior and facilitate tunnel maintenance. This study describes a remote and automated system for conducting crack monitoring at a pixel-level scale using robot-mounted imaging technology. This system collects crack images remotely and stitches them together to create a panorama image of the tunnel surface. Employing transfer learning, this study fine-tunes and improves the state-of-the-art semantic segmentation model with a lightweight backbone, DeepLab V3plus, to detect cracks automatically. A novel smooth blending prediction method is implemented on the panorama to present long-distance tunnel crack distribution, alleviating misclassification problems encountered in high-resolution image inference. In addition, transfer learning, tailored loss functions, and regularization techniques have been developed based on the CERN tunnel crack database characteristics to maintain high performance and generalization of the proposed method.

Field trials conducted in tunnels at CERN demonstrate the feasibility of the proposed crack monitoring system. Results show that the proposed system allows the identification of severe crack-damaged tunnel sections and specific crack patterns, which can be related to the structural behavior of the tunnel lining.

1. Introduction

CERN, the European Organization for Nuclear Research, is the largest particle physics laboratory in the world, with more than 83 km of underground infrastructure. The time-related structural degradation of CERN underground infrastructure, combined with complex geological conditions, has compromised the long-term serviceability of some tunnel sections. As an integral part of the CERN tunnel asset management (TAM) program, crack monitoring is used to improve the understanding of long-term tunnel behavior and facilitate decisions on mitigation measures. However, crack monitoring in tunnels at CERN is difficult to perform due to the large scale of the tunnel infrastructure and the presence of radioactive areas. Remote and automated crack monitoring methods are, therefore essential for CERN tunnel structural assessments.

It has been common practice to perform periodic visual crack inspections by well-trained engineers during the shutdown of accelerators. However, manual visual inspections are inefficient, subjective, and labor-intensive (Huang et al., 2017). Recently, digital image-based crack detection methods have been developed for remote structural inspection (Jahanshahi & Masri, 2012; Jahanshahi & Masri, 2013). Traditional image-based tunnel crack detection techniques rely on the difference in pixel values between backgrounds and cracks (Yu et al., 2007; Huang, Sun et al., 2017; Lei et al., 2021). To improve the accuracy, machine learning techniques, including KNN (Zhang et al., 2014) or Support Vector Machine (Hadjidemetriou et al., 2018; Sari et al., 2019), were introduced as additional classifiers to distinguish crack pixels from unremovable noise pixels left even after thresholding.

Advances in Convolution Neural Networks (CNN) offer a more robust and efficient technique for crack detection. A simple CNN model with several layers developed by Cha et al. (2017) outperforms a machine learning-based classifier in concrete crack detection. In the case of patch-level crack detection, the VGG16 was employed to detect distressed pavement cracks (Gopalakrishnan et al., 2017). Googlenet was adopted for crack detection in tunnels which are more complex

* Corresponding author.
E-mail addresses: aohui.ouyang@cern.ch (A. Ouyang), vanessa.di.murro@cern.ch (V. Di Murro), martin.cull@cern.ch (M. Cull), roddy.cunningham@cern.ch (R. Cunningham), john.andrew.osborne@cern.ch (J.A. Osborne), zili.li@ucc.ie (Z. Li).

https://doi.org/10.1016/j.tust.2023.105310
Received 6 February 2023; Received in revised form 30 June 2023; Accepted 2 July 2023
Available online 14 July 2023

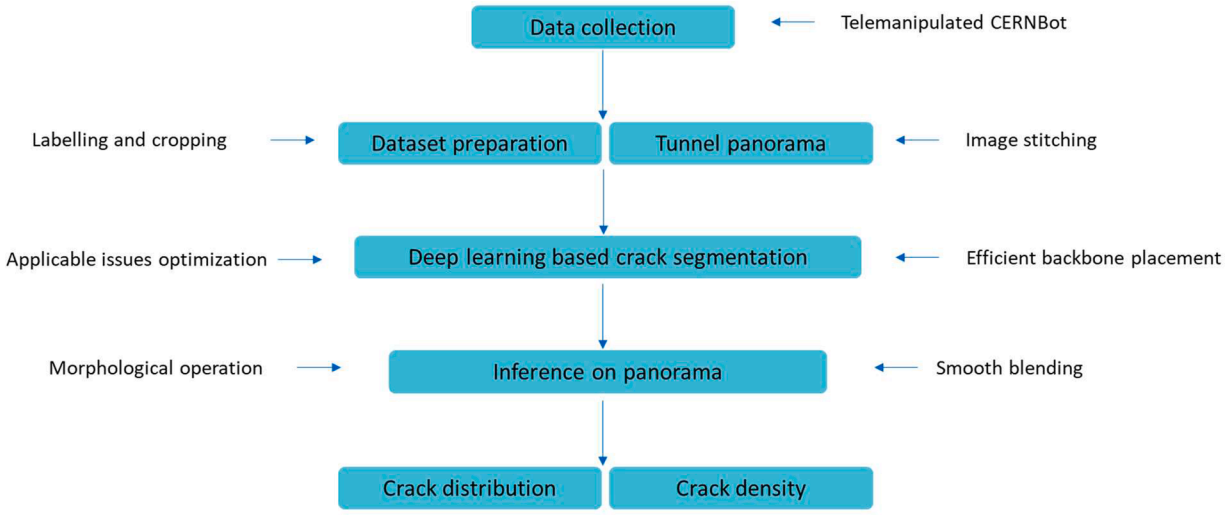


Fig. 1. Flow chart of the crack monitoring system.

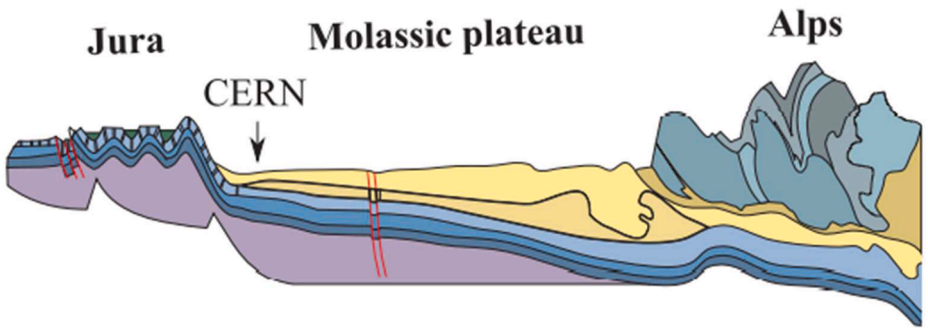


Fig. 2. Location of CERN.

environments with less illumination and curved surfaces (Gong et al., 2018). Moreover, residual blocks, facilitating CNN modes to go deeper without gradient vanishing, improves the performance of crack detection (Maeda et al., 2018; Ni et al., 2019).

Semantic segmentation models are being developed to achieve pixel-to-pixel detection by removing fully connected layers within CNN models. Huang et al. (2018) introduced the Fully Convolutional Network (FCN) (Long et al., 2015) in pixel-level detections of cracks and water leakage in segmentally lined tunnels. In semantic segmentation models, CNN models play vital roles in prediction accuracy as they are embedded as backbone structures for feature extraction. Dung (2019) adopted FCN for crack segmentation and explored the contributions of variant pre-trained backbones. The U-net (Ronneberger et al., 2015) showed robust performance in variant scenarios such as cracks in tunnel lining (Dang et al., 2022), desiccated clay (Xu et al., 2022), or bridges (Zhang et al., 2021). Other benchmark models like SegNet (Badrinarayanan et al., 2015) and Feature Pyramid Network (FPN) (Lin et al., 2017) were fine-tuned in crack segmentation of roads (Chen et al., 2020) or masonry structures (Dais et al., 2021). Mask R-CNN (He et al., 2017) was used for instance level crack detection and water detection in segmentally lined tunnels (Zhao et al., 2020) and s(Huang et al., 2022). The same approach was employed in the CERN tunnel crack detection (Attard et al., 2019); Attard (2020).

Despite previous successes of crack segmentation in tunnel engineering, gaps remain in their practical applications in underground tunnel networks that extend over many kilometers. Analyzing a long tunnel panorama over hundreds of meters in this context involves image stitching and robust inference of deep networks on high-resolution images. Therefore, an automated crack monitoring system is proposed as shown in Fig. 1. First, the telemanipulated data acquisition platform CERNbot is employed for remote image data collection. Collected images are reorganized by image stitching to create a tunnel panorama for a large field of inspection view. The state-of-the-art semantic segmentation deep network, DeepLab V3 plus, is fine-tuned on the collected CERN crack database and improved with a lightweight backbone to achieve efficient automatic crack segmentation. Optimization methods, including a tailored loss function, transfer learning, and regularization techniques, are proposed for applicable issues of the customized database. Finally, a smooth blending method is proposed to achieve the prediction of the panorama for crack distribution and statistics. Field trials have been conducted in two typical sections of tunnels at CERN to test the feasibility of the proposed crack monitoring system. The panorama image provides a large field of view of tunnel surfaces. Monitoring results present a clear crack profile with specific crack patterns and identify the crack-damaged tunnel chainages, facilitating the understanding of long-term tunnel deformation patterns.

2. Automatic image data collection in the CERN tunnels

2.1. CERN tunnel background

Founded in 1954, CERN houses a complex system of underground structures. Accelerators and other auxiliary machinery for high-energy physics research are mainly housed in tunnels and caverns, with access via shafts. There are 83 km of underground structures at CERN where the main two components are a 7 km circular tunnel, the *Super Proton Synchrotron* (SPS), built in 1974 and a 27 km circular tunnel, the *Large Hadron Collider* (LHC), built in 1987. The aging and deterioration

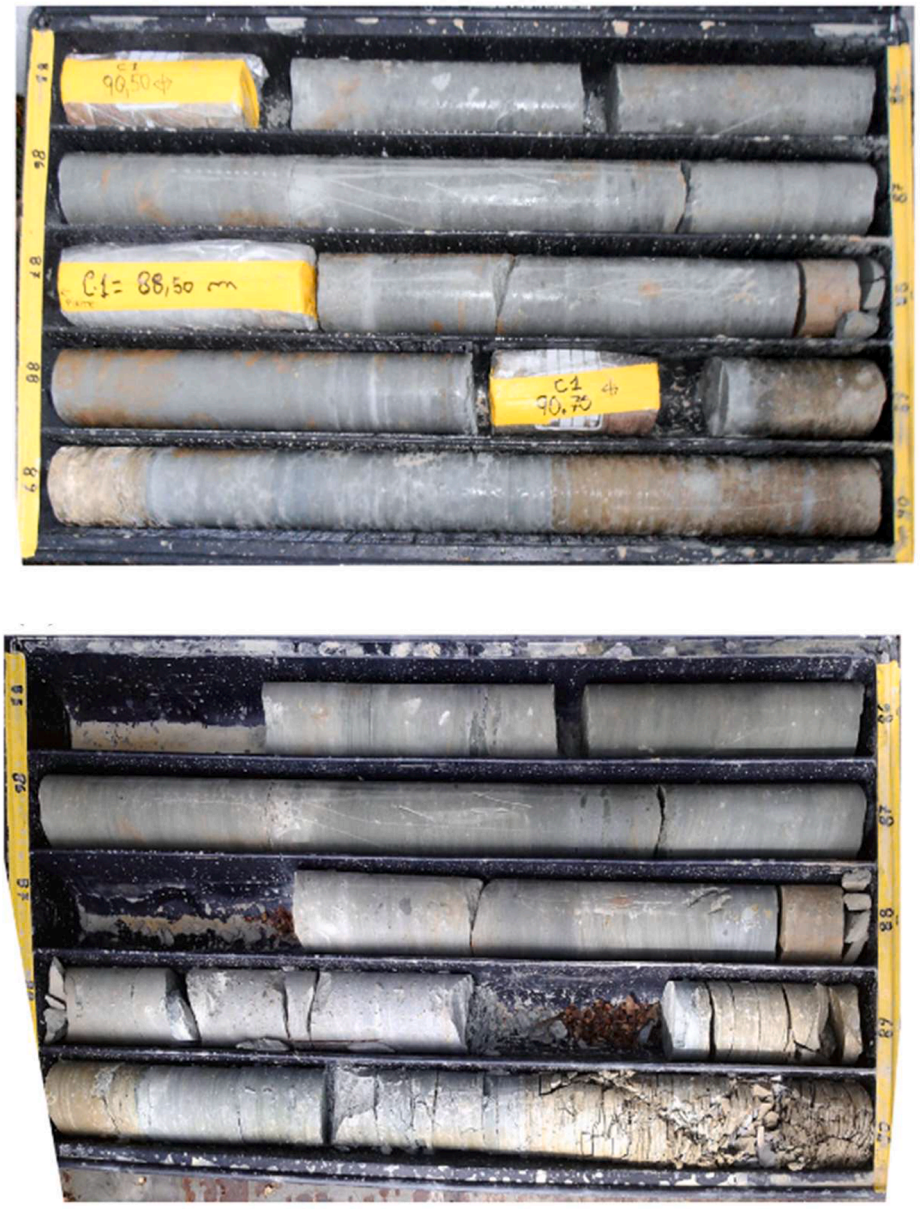


Fig. 3. Integrity comparison of cores from borehole C1 revealed at ATLAS (a) Photo Oct. 2015; (b) Photo Mar. 2017 (Fern, Di Murro et al. 2018).

of the CERN underground complex present challenges in terms of the future exploitation of the site infrastructure.

With regards to geological conditions, CERN tunnels sit astride the France-Switzerland border, located on the Molassic plateau, as shown in Fig. 2. Due to the compression caused by tectonic thrust between the Alps and Jura ranges, the underground structures suffer from an imbalance of the ground stress. In-situ dilatometer tests at the ATLAS cavern of LHC, revealed anisotropic stress ratios of horizontal and vertical stresses varying from 1.29 to 2 (Fern et al., 2018). In addition, the very weak marl layer, one constituent of the widespread red molasse, swells, slakes and spalls upon contact with moisture, Fig. 3 illustrates the weathering behavior of the red molasse, extracted from the borehole C1 taken near ATLAS, has already been weathered and disintegrated after 18 months of air exposure.

CERN has recorded several cases of lining defects including tunnel crown failure in 1990, as shown in Fig. 4. Previous research showed that a concrete lined tunnel was experiencing vertical tunnel elongation with time (Di Murro, 2019; Di Murro et al., 2019). The observed tunnel deformation mode was detected by using the novel distributed fibre optic sensors (DFOS), a remote monitoring method shown in Fig. 5. The development of compressive and tensile cracks spreading along tunnel crown and sidewalls respectively, as shown in Fig. 6, resulted in the the abovementioned tunnel lining mechanism of deformation.

Tunnel inspections at CERN are challenging due to the ionizing radiation generated from the operations of the accelerators and their residual radiation after operation. Inspection campaigns can only be carried out during annual technical stops or the long-term shutdown of accelerators (occurring every 3–4 years). Engineers need to take mandatory training in radiation protection, evacuation and wear specialized personal protection equipment (e.g., the radiation dosimeter and self-rescue mask) before entering tunnels. Considering tunnel safety & serviceability, a remote, automatic, and robust crack monitoring system is desired for routine tunnel inspection and assessment, as this enables the tunnels to be inspected throughout the year.

2.2. *Image data collection and tunnel panorama stitching*

To achieve remote tunnel inspection in CERN, a data collection tool

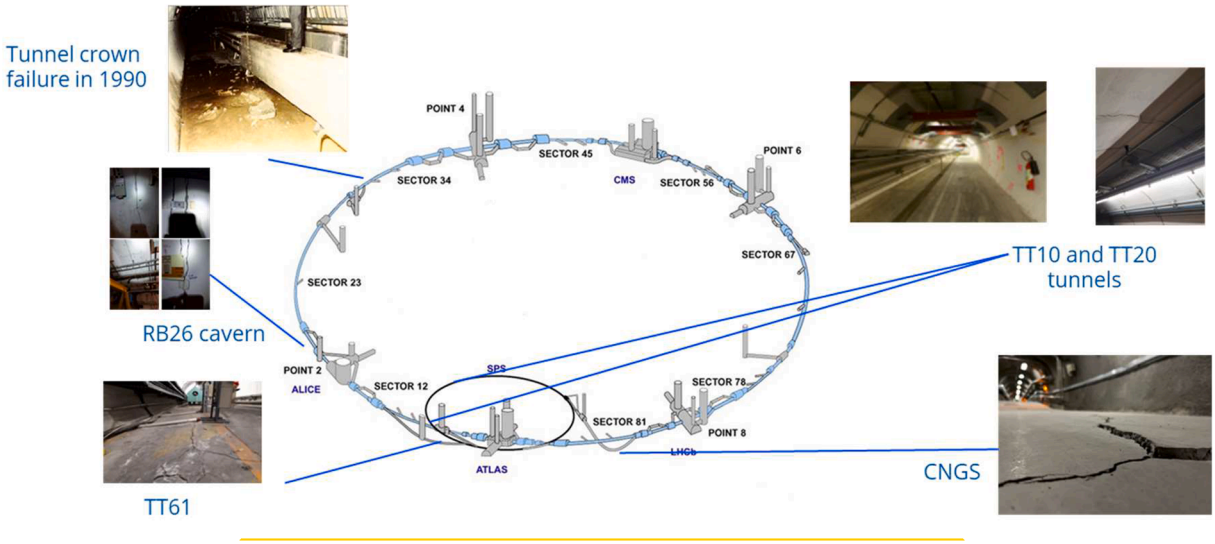


Fig. 4. Tunnels at CERN and their recorded defects.

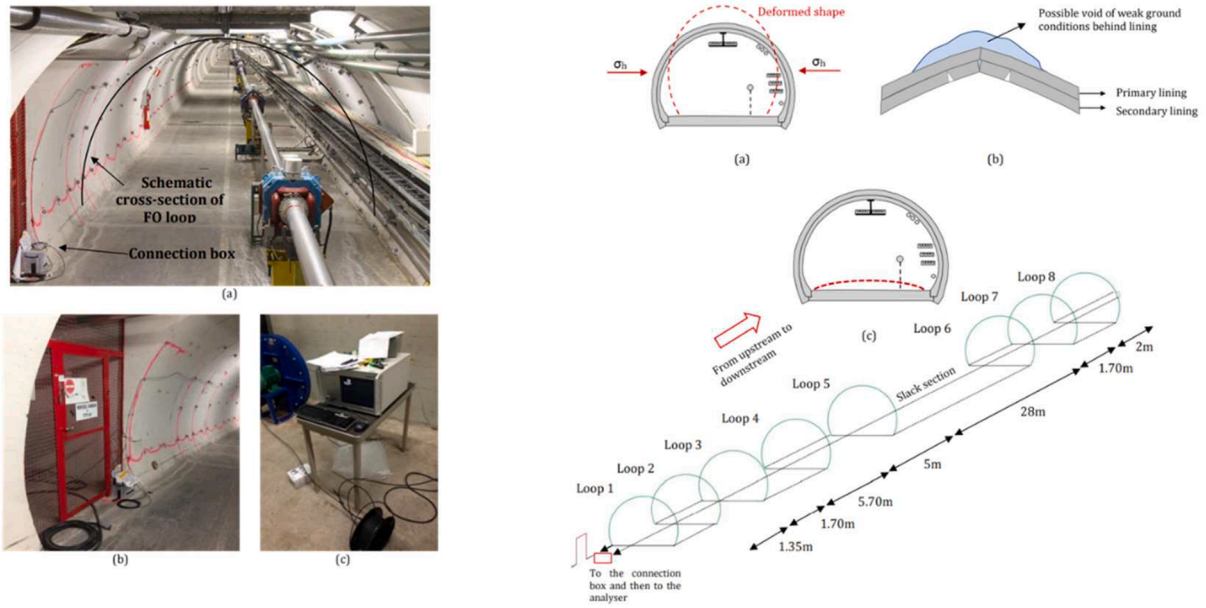


Fig. 5. Long-term tunnel monitoring at CERN by the distributed fiber optical sensors (Di Murro 2019).

that enables telemanipulation underground with robust performance in harsh environments is essential. To provide intervention on beam facilities without the requirement of physical personnel underground, the BE-CEM-MRO section of CERN has developed and designed a series of robots. In this scenario, the CERNbot is employed to capture image data for tunnel crack distribution monitoring. The CERNbot is a Remotely Operated Vehicle (ROV) that can connect to mobile control devices or PC through the underground coverage of internal telecommunication networks. It is a supervised telemanipulation system, consisting of 5 layers as shown in Table 1, and enables real-time compliant control through a human-robot interface developed by (Di Castro et al., 2017; Di Castro, 2019; Lunghi et al., 2019). In the Graphical User Interface, operators can pick available robots with different configurations on-site to perform tasks (Lunghi et al., 2016). The CERNbot is also a modular robot, comprising modifiable and compatible subsystems (Di Castro, 2019). Subsystems can work individually and merge for various mechanical repair tasks like cutting, welding, or screwing by adopting different configurations, as shown in Fig. 7. For tunnel inspections, a steel framework and three Digital Single Lens Reflex cameras are mounted on the upper part of the CERNbot, shown in Fig. 8. The software development kit enables automatic image collection and storage during the inspection process.

Training of deep networks for automated crack segmentation needs sufficient labeled image data. The crack database for deep model training consists of 517 RBG images collected within the LHC, TT10, and TT1 tunnels, with an average resolution of 1920x1080. 80 of them, captured by the CERNbot, are the subset of raw images which are later employed for panorama stitching. The other training data are collected during previous manual inspections. Fig. 9 illustrated the dataset preparation workflow in which raw images are cropped into around 4000 image patches, with size of 512x512 and containing cracks only. Not like the like smooth blending, raw images were cropped without any overlaps. Following that, crack pixels are annotated by well-trained engineers manually into truth masks. The dataset for deep learning training, consisting of image patches and corresponding masks, is split with ratio of 4:1 into training and test datasets, then. The unique

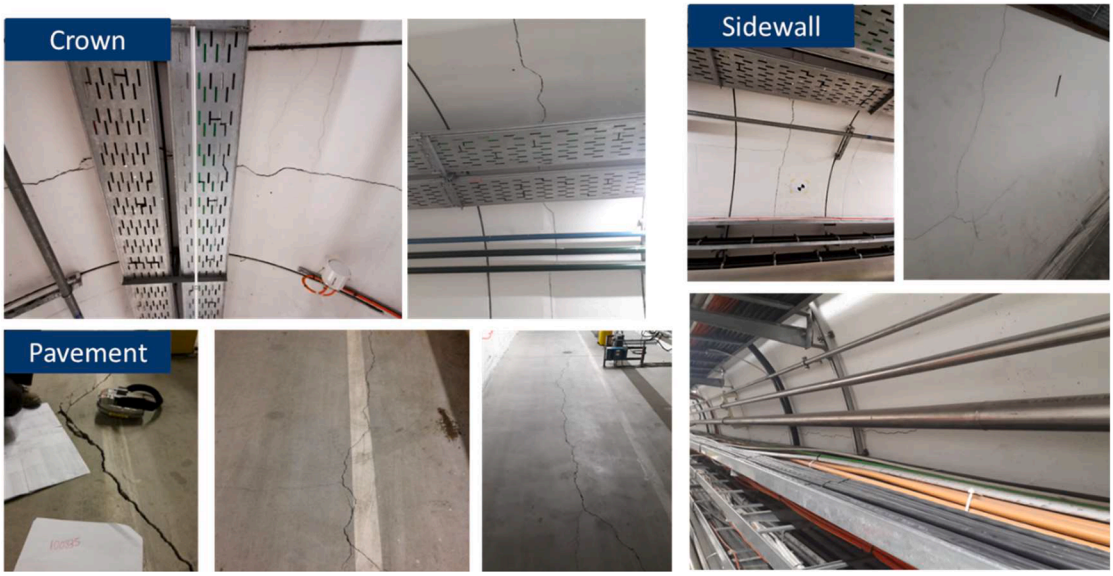


Fig. 6. Crack photos in TT10 tunnel of CERN.

Table 1
Layers of the telemanipulation system of CERNbot.

Layers	Function
Hardware Abstraction layer	Coding compiler for hardware
Control layer	Offering control strategies for robotic arms, traction platform, and conducting sensor fusion
Supervision layer	Responsible for robot localization, environmental reconstruction, energy optimization, and assisted grasping.
Application layer	Responsible for robot operations such as assisted navigation, object alignment, or sequencer of multiple operations.
OS layer	Human-robot interface

characteristics of CERN tunnel crack image data are summarized in Fig. 10: a) images taken in insufficient or variant illumination conditions, b) crack split by beam facilities, c) crack with complex pattern or bifurcation, and d) non-defined marks with similar characteristics to that of cracks.

The image stitching allows the alignment of hundreds of images in spatial order and offers a large field of view than separate images. The stitched panorama also enables engineers to better understand the spatial layouts of the cracks. This is important as it is challenging to link cracks taken from separate images with the exact tunnel locations. Not to mention that retrieving spatial correspondences of cracks from isolated frames it is also troublesome. These spatial layouts can offer a thorough understanding of long-term structural performances. The shooting scheme is proposed in Fig. 11 to facilitate the image stitching process. Separate images captured by the coordinated camera set in consecutive inspection spots are utilized for the creation of the panorama. The overlapping of images can ensure that there are no sections missed, and it offers common features among different frames. In the image stitching process, SURF detectors (Bay et al., 2008) are implemented in adjacent image pairs to extract key points invariant to change of viewpoint. Then, gradient orientations of key points are computed to describe their local appearances. Transformation matrixes among

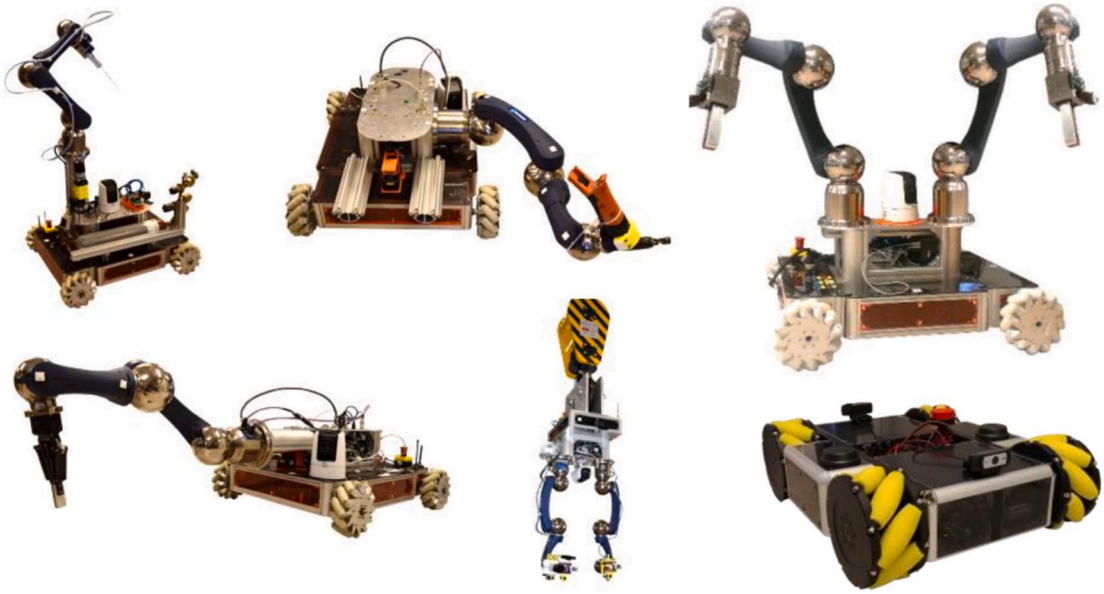


Fig. 7. Different CERNbots at BE-CEM-MRO department in CERN (Di Castro 2019).

Fig. 8. The customized CERNbot for tunnel lining inspection.

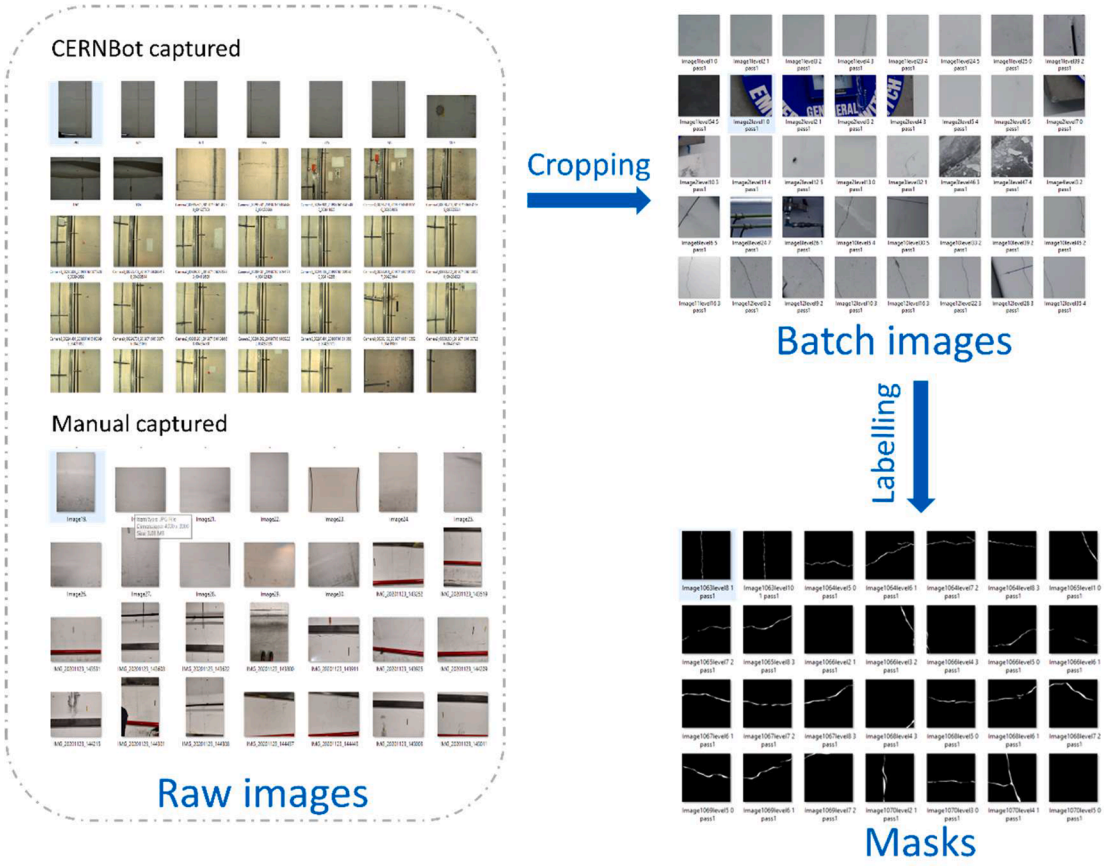


Fig. 9. Deep learning dataset preparation.

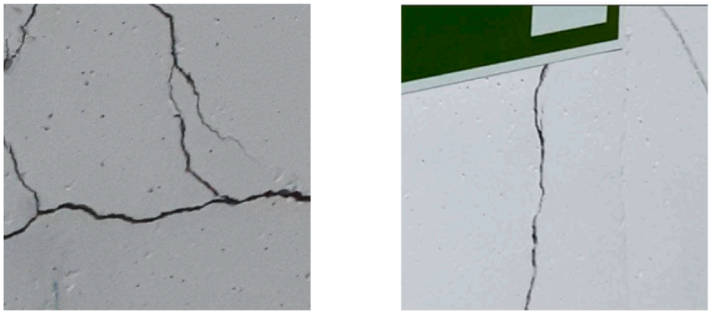
images are derived by matching these common features in the overlapping areas. In addition, to improve stitching quality, the RANSAC algorithm(Fischler & Bolles, 1981) is also applied to refine the feature-matching process by removing outliers. Finally, consecutive images are transformed and merged into a panorama.

3. Methodology in the image data processing

3.1. The model architecture of tunnel crack segmentation based on Atrous convolution

The benchmark semantic segmentation DeepLab V3plus model (Chen et al., 2018) is employed for pixel-level tunnel crack detection in

a) Insufficient and variant illumination b) Crack covered by beam facilities



c) Crack with a complex pattern or bifurcations d) Non-defined mark similar to crack

Figure.10 Characteristics of CERN tunnel crack data

Fig. 10. Characteristics of CERN tunnel crack data.

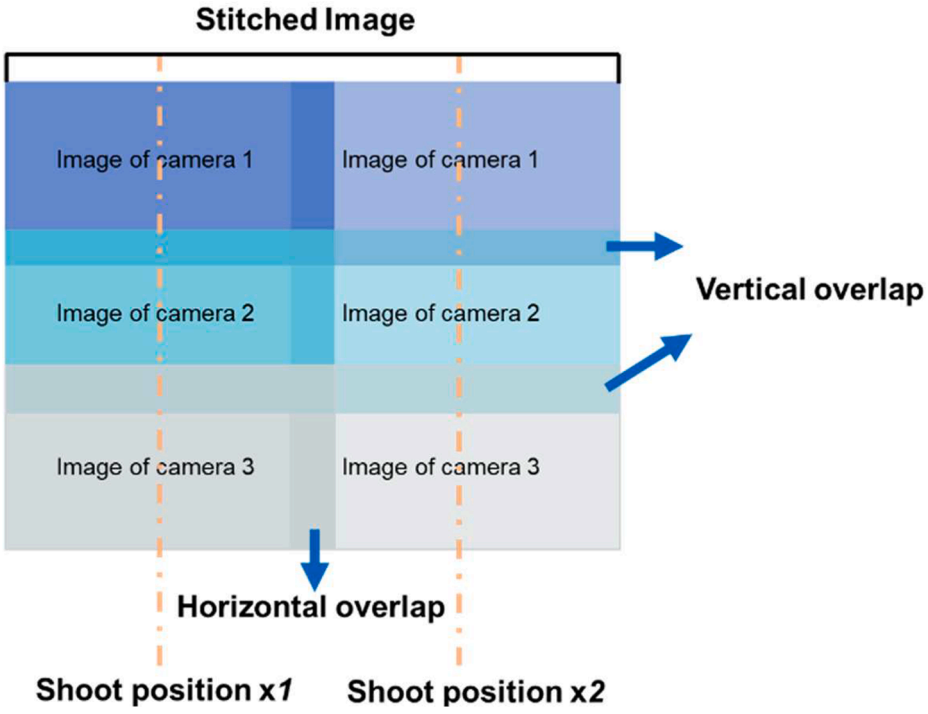


Fig. 11. Shooting scheme.

this study. The DeepLab V3plus model incorporates capabilities of previous CNN models, including the encoder-decoder conception of FCN and SegNet, the skip connection of Unet, the pyramid pooling structure of PSPNet (Zhao et al., 2017), and the Atrous convolution from previous versions of DeepLab models. Fig. 12 illustrates the architecture of the DeepLab V3plus Model. The backbone structure extracts coarse crack features from input RGB images and passes them to the encoder and decoder paths. In the decoder path, the low-level features, after upsampling, are fused together with the high-level features from skip connections for further convolution operations. This kind of operation is

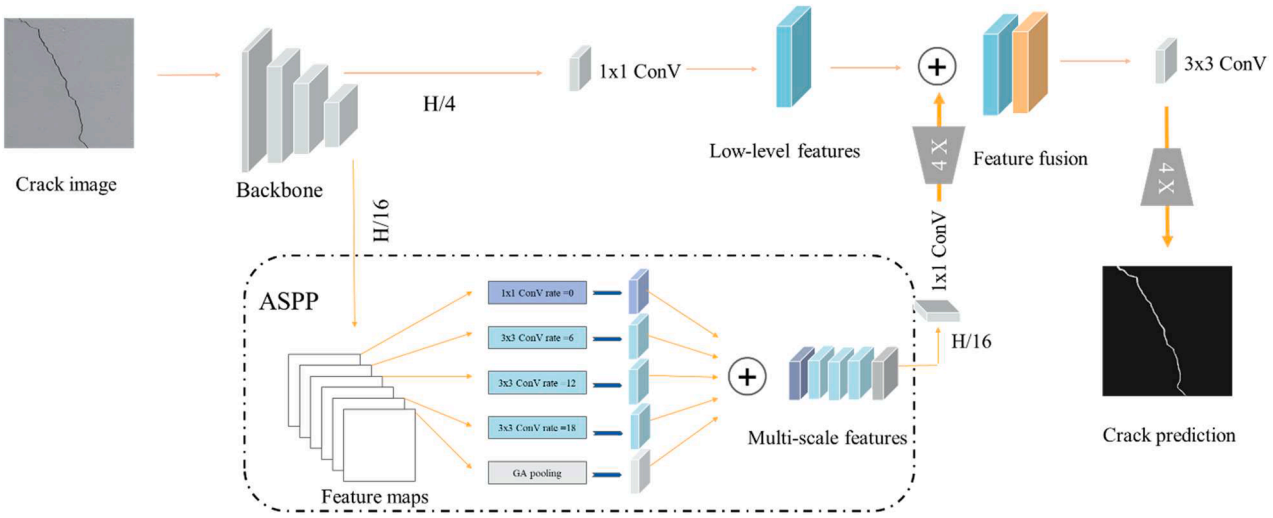


Fig. 12. Schematic of DeepLab V3plus architecture.

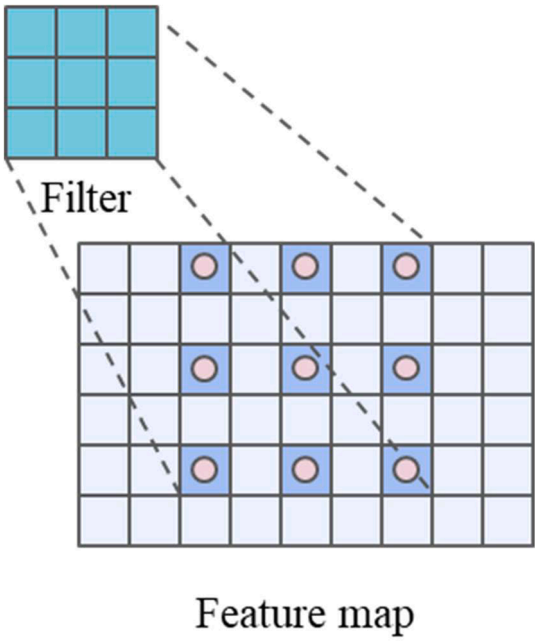


Fig. 13. Receptive field of Atrous filter.

similar to the idea of multi-view image fusion (Kuhn et al., 2013), facilitating detaching sharp crack boundaries from complex underground backgrounds. Finally, fused feature maps are upsampled again to restore the final crack pixel detection and crack spatial layout information.

In the DeepLab V3plus model, the Atrous Spatial Pyramid Pooling (ASPP) block is essential for extracting multi-scale image features. Feature maps from the backbone are processed in parallel by different operations in the ASPP block. A pooling layer and Atrous convolution layers with dilation rates of 0, 6, 12, and 18 produce five kinds of output features representing different scales of crack information. This pyramid-style structure enables the network to focus on different scales of instances on input images, thus preventing the network from failing to predict small instances or fine details of large instances. On top of that, compared to normal convolution kernels, kernels of the Atrous convolutions enable the network to enlarge the receptive field, providing contextual information from a larger area, as shown in Fig. 13.

The backbone structure plays a vital role in the performance of semantic segmentation models, serving as feature extractors. Its convolution blocks can condense essential information on input images into feature maps storing different classification knowledge on each channel. The original DeepLab V3plus model is built upon Xception. Although these deep CNN backbones could achieve higher accuracy, the required training data and time would increase (Dais,Bal et al., 2021). Therefore, the DeepLab V3plus model is improved with a trimmed version of the MobileNet V2 backbone (Sandler et al., 2018). The main objectives for this improvement is two-fold. The first reason is the trade-off between the model complexity and computational demand. More specifically, deep learning networks get millions of parameters and weights to train

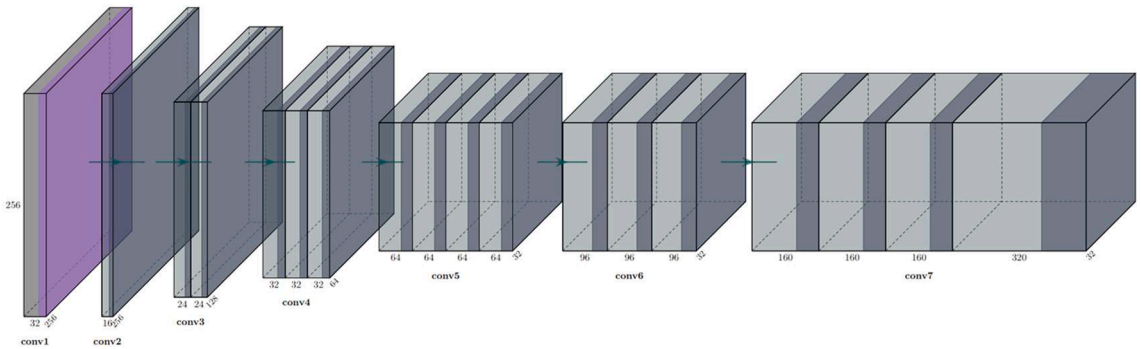


Fig. 14. Sketch of the trimmed MobileNet V2 backbone.

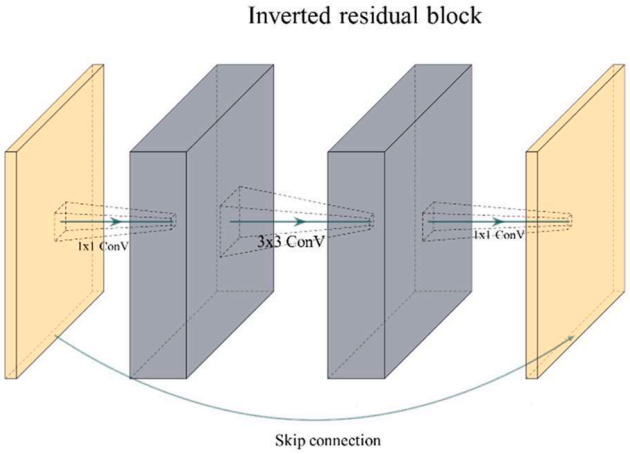


Fig. 15. Sketch of the inverted residual block.

Table 2
Training configurations.

Training configuration	Hyper-parameters
Backbone type	Xception and MobileNet V2
Pretrained dataset	PASCAL VOC
Image Size	[512,512]
Batch size	4
Learning rate	From 5e to 4 decaying to 5e-7 by step decay strategy
Optimizer	Adam
Epochs	140 epochs
Output stride	16
Momentum	0.9
Atrous rate of ASPP	[6,12,18]
Data shuffle	True
Regularizations	L2 Norm with weight decay rate = 2e-4
Data Augmentation strategy	Horizontal flip; Vertical flip; Gaussian blur; Random rotation (-10 to 10 degrees); Color space convert
Framework	Keras
GPU	NVIDIA-A100-SXM4-40 GB

and update, thus requiring tremendous database scales (Xue et al., 2022). The second objective is to on-device crack segmentation for further research of CERN TAM tasks. Combining knowledge distill and parameter pruning techniques, a lightweight semantic segmentation model might be deployed on mobile or edge devices like drones or the CERNbot. The proposed trimmed version of MobileNetV2 consists of only seven inverted residual convolution blocks. Fig. 14 depicts the dimensionality shrinkage of feature map size and expansion of their channels of the trimmed MobileNetV2 backbone. The depth wise separable convolution can considerably reduce trainable parameters by breaking down traditional convolutions into two separate procedures: depth wise convolution and pointwise convolution(Howard et al., 2017). In addition, the inverted residual block structure shown in Fig. 15 can also prevent deep networks from gradient or exploding phenomena vanishing by skip connections between input and output features.

3.2. Training strategies for applicable issues

After data collection and labeling, cropped patches are fed into the DeepLab V3plus model for fine tuning. Table 2 summarizes training configurations. However, some practical problems remain to be solved when applying deep networks on tunnel crack segmentation tasks at CERN. The first issue is the class imbalance as tunnel cracks represent only a minor part of the collected images. Deep networks tend to create bias towards the major class, the non-crack pixels, resulting in underperformed misclassification issues. To fix this, a compound loss function is proposed, comprising weighted dice loss (Sudre et al., 2017) and focal loss (Lin et al., 2017). Differing from the cross-entropy loss function utilized by (Dang,Wang et al., 2022), the weighted dice loss shown as eq. (1) is a region-based loss function, and the focus of deep network on cracked pixels can be regulated by increasing corresponding weights. Additionally, the focal loss shown in eq.(2) enables the model to focus on hard-to-learn classes, crack pixels, by limiting the contribution of non-crack pixels on the loss. Combining these two strategies, the compound loss function shown in eq.(3) is therefore proposed to optimize deep networks on the unbalanced crack dataset of the CERN tunnel.

$$L_{wdice} = 1 - 2\frac{\sum_{C=1}^{2} w_c \sum_x p_{cx} g_{cx} + s}{\sum_{C=1}^{2} w_c \sum_x (p_{cx} + g_{cx}) + s} \tag{1}$$

Where L_{wdice} is the weighted dice loss; C ranges over classes; w_c is class weights; p_{cx} and g_{cx} represents predicted tensors and the ground truths of position$x$; s is the coefficient added to prevent undefined situation dividing by 0.

$$L_{focal} = -(1-pt)^{\gamma}\log(pt)$$

$$pt = \begin{cases} 1-p, & \text{otherwise} \\ p, & \text{if } y = 1 \end{cases} \tag{2}$$

Where p is the probability and the modulating factor γ is added to regulate the loss function keeping the balance between misclassified and well-classified pixels.

$$L_{compound} = L_{focal} + L_{wdice} \tag{3}$$

The second issue is the underperformance of deep networks when training from scratch. Therefore, transfer learning is introduced here to accelerate the training and to reduce the required data size, noted its efficiency in CERN's previous research (O'Brien et al., 2023). The fundamental point of transfer learning is to employ common knowledge from other datasets to reduce the requisite converge time and data. Although data resemblance is limited, primary knowledge like edge detectors, color detectors, shape detectors, etc., were transferred into the new models to reduce repetitive training. The initial training is conducted on the PASCAL VOC database first (Everingham et al., 2015), and pre-trained weights are then transferred to initialize the improved DeepLab V3plus model. Due to the different classes of these two datasets, the last layer, the classifier, is removed and retrained during the transfer learning process. When image sizes between the crack database of CERN tunnels and the pre-trained database do not match, one extra convolution layer is added to decrease the input size or pad the input image to enlarge their sizes.

The last one is the overfitting issue of deep learning models, which is quite common in civil engineering defect detection applications (Liu et al., 2020; Xue,Jia et al., 2022). Overfitting would result in the underperformance of deep networks on the test dataset, as trivial and misleading details in training set are learned as key features. Two regularization methods are adopted to address this issue: data augmentation and model complexity reduction. The data scale is enlarged by augmentation methods shown in Table 2. Additionally, the L2 normalization is implemented for the penalization of model complexity by adding extra losses in convolution layers with higher weights, as shown in Eq(4).

$$L_{\text{total}} = L_{compound} + \eta \sum_{l=1}^{L} \sum_{i=1}^{n} |w_{l,i}|^2 \tag{4}$$

in which L_{total} denotes the total loss; the η is the weight decay rates; the L and n represent the number of layers and nodes.

3.3. The fundamental of crack analysis in stitched images

The final step of the crack monitoring system is to scan the stitched panorama with deep networks for the large-scale tunnel crack distribution. However, this procedure could be very challenging due to the

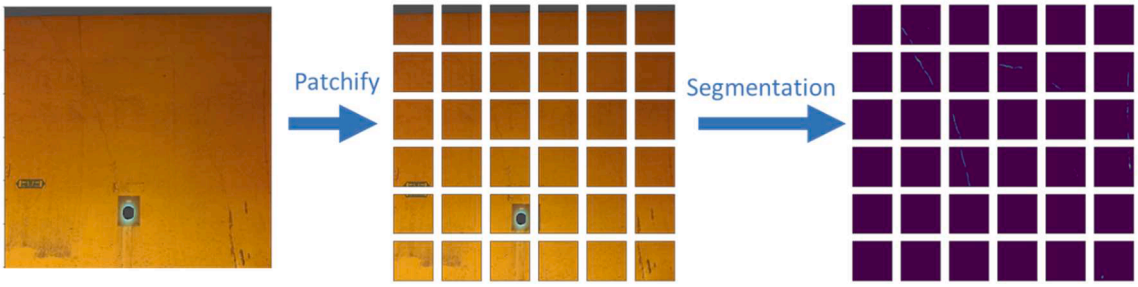


Fig. 16. Workflow of the patch-wise prediction.

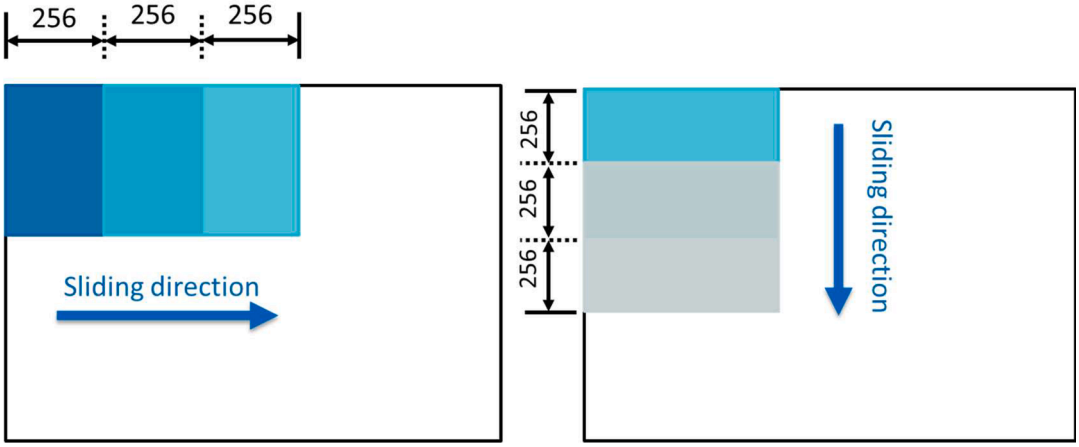


Fig. 17. Sliding window path of the smooth blending method.

Table 3
Comparison of model performance on test sets.

Model name	mIoU	mPA	Trainable parameters (Million)
DeepLab V3plus with Xception	83.92%	92.64%	41.2
DeepLab V3plus improved with trimmed MobileNet V2 (Tailored loss)	82.81%	92.20%	2.8
DeepLab V3plus improved with trimmed MobileNet V2 (Cross-entropy loss)	80.68%	87.1%	
U-net with ResNet 152	82.86%	88.71%	67.3
FPN with Googlenet (Tailored loss)	82.83%	89.85%	25
FPN with Googlenet (Cross-entropy loss)	81.54%	88.44%	

high resolution of the panorama stitched from hundreds of images. Loading the stitched panorama into deep networks directly for crack detachment would end with a GPU crash and graph extrusion error. Therefore, it has been a common practice in segmentation tasks of remote satellite sensing communities (Paisitkriangkrai et al., 2015; Liu et al., 2018) to employ the resizing strategy or Patch-wise prediction methods to deal with the overly high-resolution image.

The resizing strategy is to resize a high-resolution image into suitable size for segmentation. However, image details would sharply decrease when the resizing ratio is too large. This might be troublesome because the observed tunnel crack pixels occupy less than 1% of the customized crack database. Therefore, there is a high risk of losing crack pixels when predicting the stitched panorama. The workflow of the patch-wise prediction method is graphicly illustrated in Fig. 16. In this method, the high-resolution image is divided into small patches suitable for the input of deep networks. Then, each sub-patch is predicted parallelly by the deep network, and all sub-patches are stitched to produce the final crack map. The inherent drawbacks of this method are that cracks on each sub-patch are predicted separately and all pixels are only reviewed once. This failure to employ contextual information from surrounding patches can result in jagged prediction issues like crack discontinuity.

To address this high-resolution related prediction issue, the smooth blending prediction method proposed by (Chevalier, 2017) is introduced here, noting its success in cell segmentation (Greenwald et al., 2022). In this method, a sliding window is employed to crop high-resolution images into patches with overlaps in an order shown in Fig. 17. On the one hand, pixels in the panorama are reviewed in prediction windows four times, compared with the patch-wise prediction method. Increasing prediction possibilities can produce more robust results. On the other hand, the new order of sliding window can relocate crack pixels close to the fringes at the image center. This is important because the prediction accuracy between image borders and the center is different in semantic segmentation, due to variant contextual knowledge (Brackenbury, 2022). This relocation allows cracked pixels to be predicted at different positions in an image patch and helps deep networks to employ more contextual knowledge for more precise inferences. In addition, each cropped patch would be predicted several times after being rotated and mirrored like dihedral group D_4. This multi-angle view facilitates deep networks to get a stable prediction. Finally, the prediction windows are merged togther by interpolating overlap areas with the simplified spline window function for smooth blending. Morphological operations, including opening and closing, are also adopted to improve the quality of the final crack map by a python script. The morphological opening removes isolated noise pixels misclassified by deep networks. On the contrary, connections between adjacent cracks are bridged by the morphological closing operation.

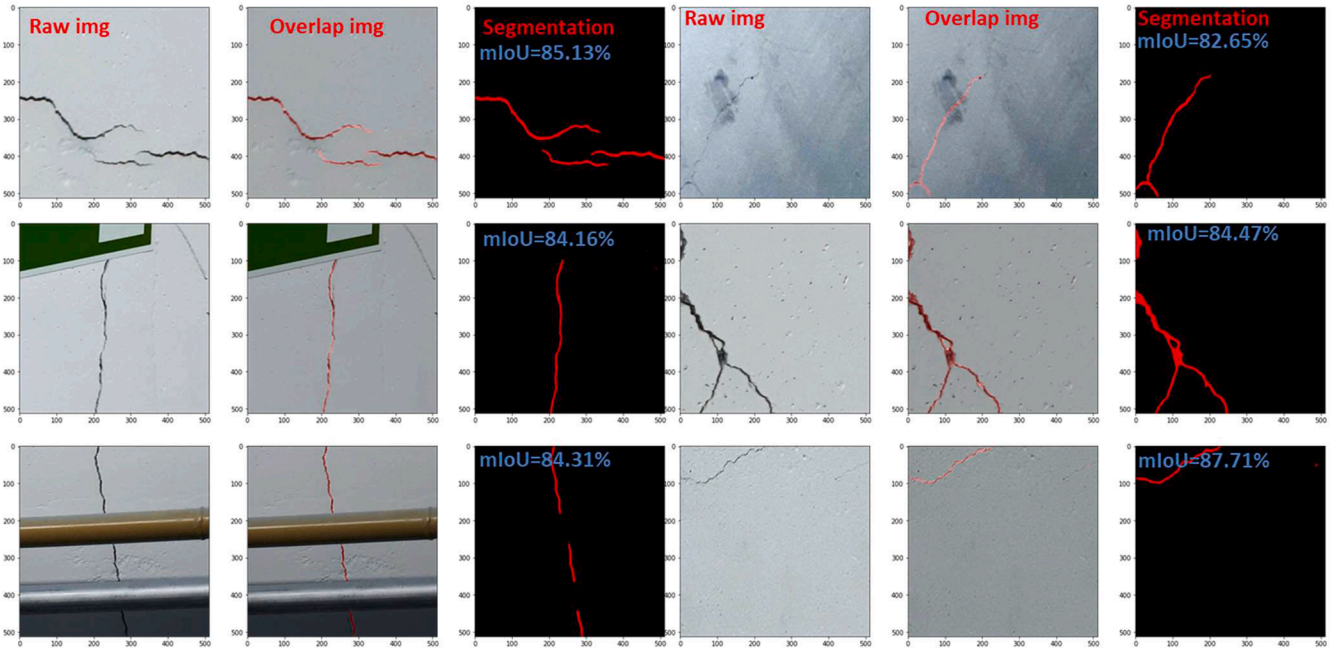


Fig. 18. Pixel-level crack segmentation results.

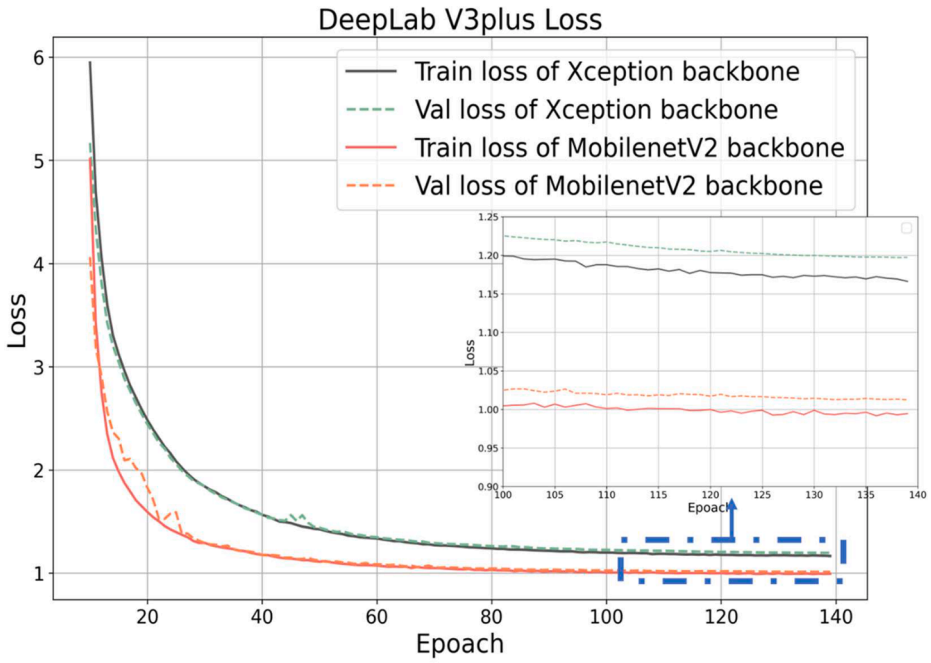


Fig. 19. Losses in the training and validation stage.

4. Case study of CERN tunnel cracks using the proposed system

4.1. Performance of deep learning model

Table 3 compares the performances and complexities of different models on the customized CERN datasets. The modified U-net with ResNet backbone for soil densification (Xu,Zhang et al., 2022) crack or tunnel crack segmentation (Dang,Wang et al., 2022) and the FPN improved with GoogleNet backbone for concrete crack segmentation (Ni et al., 2019) are introduced for comparison. In terms of performance, all deep networks achieve mIoU scores above 80%. The original DeepLab V3plus model outperforms the others with a mIoU score of 83.92% and an mPA score of 92.64%. In addition, the improved DeepLab V3plus model achieves an mPA score of 92.20%, higher than Unet and FPN. Regarding model complexity, the required training time and computational demands increase with piled-up trainable parameters, as stated by Dais,Bal et al. (2021). The trimmed MobileNet V2 efficiently reduces trainable parameters of the DeepLab V3plus model into 5%, thus accelerating the training. On top of that, compared with Unet and FPN, the proposed DeepLab V3 plus model achieves a higher MPA with less model complexity. Moreover, to validate the performance improvement of the tailored loss functions, the FPN improved with GoogleNet backbone and DeepLab V3plus improved with trimmed MobileNet V2 were trained again with the Cross-entropy loss function. Comparisons in

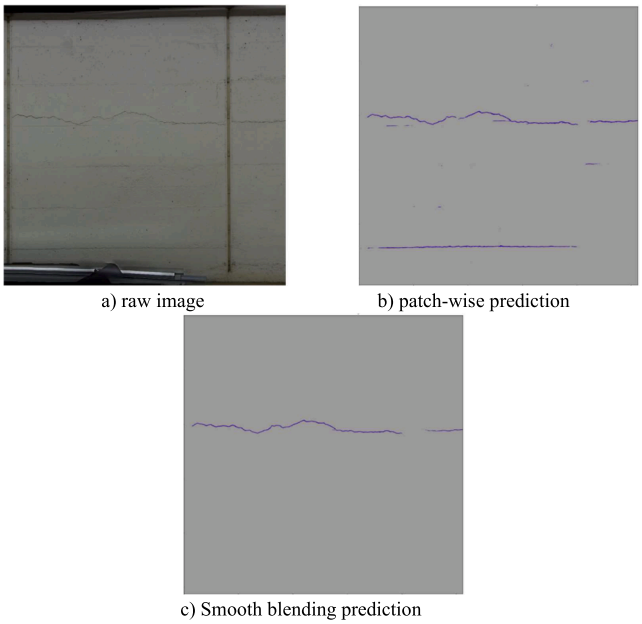


Fig. 20. Comparison of patch-wise and smooth blending prediction method.

Table 3 show a 1–2% improvements in mIoU scores and more than 2% improvements in mPA, for both models.

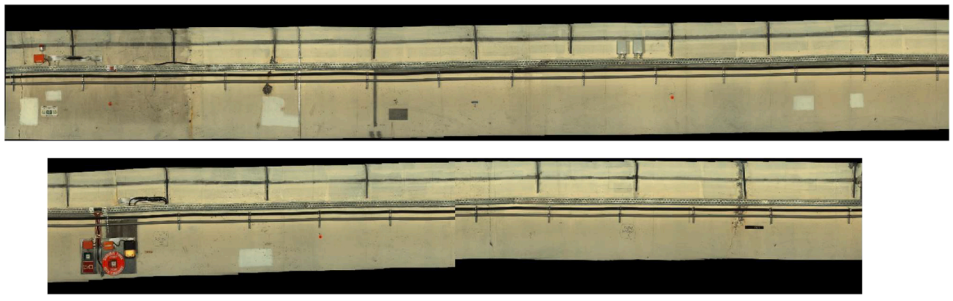
Fig. 18 enumerates some segmentation results on image patches with a size of 512 by 512, inferred by the proposed Deeplab V3plus. The prediction results consisted of the original images, segmentation results, and overlaps. The capability of the proposed crack segmentation model on harsh tunnel environments is demonstrated by detecting slender fissures, differentiating cracks from crack-similar marks, and segmenting small branches of complex cracks with fair accuracy.

Fig. 19 plots the losses descending of the improved and original DeepLab V3plus Models in the training and validation stages. Both losses drop sharply in the first 60 epochs and reach a plateau after 100. The overlap of loss curves of the training and validation phases in Fig. 19 validates the effectiveness of regularization methods, including data augmentation and L2 normalization, in addressing overfitting. The maximum difference between the validation loss and the training loss is less than 0.022, indicating that the proposed DeepLab V3plus model can effectively detect tunnel cracks on both the training and unseen detection images.

Performances of a deep learning model on high resolution images could change with different applied prediction strategies. To corroborate this, a raw image, covering both cracks and crack-like indentations, is sifted to test the performances of the patch-wise and the proposed smooth blending prediction method. The raw image in Fig. 20a) is a stitched photo taken from the TT1 tunnel with a resolution of 5120 by 5120 pixels. The mIoU scores of the smooth blending and the patch-wise predictions are 81.76% and 66.22%, accordingly. It is worth to mention that these two mIoU scores are computed directly from the ground truth and the prediction mask in full sizes, instead of averaging the mIoU values of all sub image patches. The better performance of the smooth blending method could be related to its better employment of more contextual information from surrounding patches. To provide better understanding, comparisons of visual results are illustrated in Fig. 20. In patch-wise prediction, the deep network tends to misclassify longitudinal indentations on the sidewall as tunnel cracks due to their high resemblance. To the best of the authors' knowledge, indentations on the tunnel wall are probably due to concrete defects subjected to poor concrete construction procedures, which should not be labeled as cracks. In our context, however, tunnel cracks should be concrete fractures triggered by considerable tunnel lining deformation. The misclassification is related to the small, cropped size: longitudinal indentations resemble horizontal cracks more on small image patches. Therefore, more contextual information from surrounding patches in the proposed sliding-blending method enables deep networks to avoid such misclassification.

4.2. Crack distribution of CERN tunnel

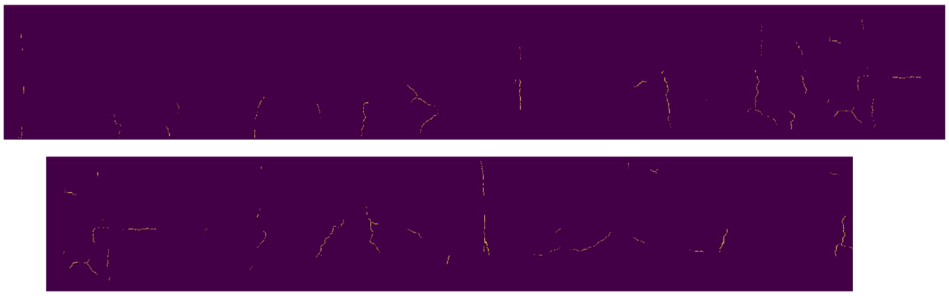
Inspections were implemented in two separate sections of the CERN



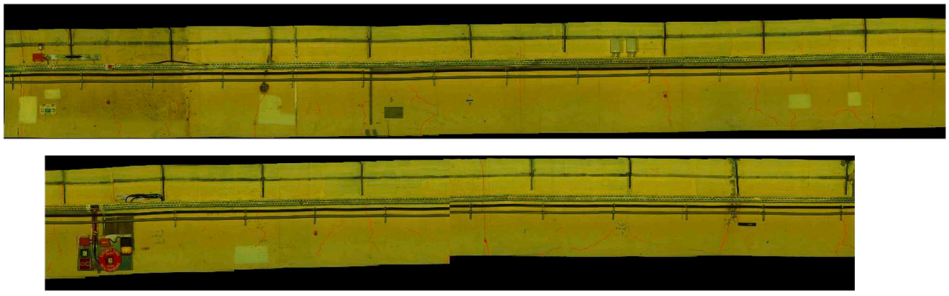
a) the 80-meter panorama of LHC tunnel with resolution of 3479 by 49152 pixels

b) 45-meter panorama of TT1 tunnel with resolution of 2048 by 26309 pixels

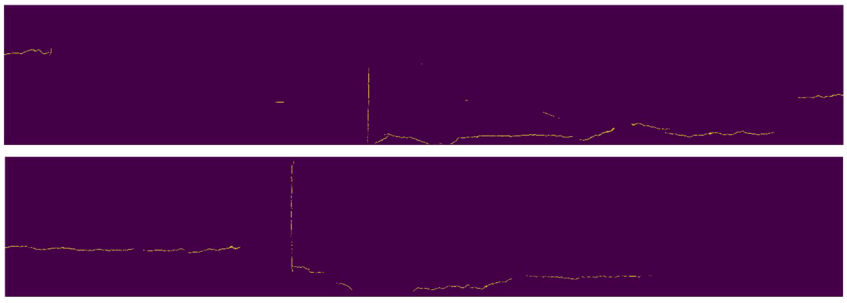
Fig. 21. Stitched tunnel panorama of two typical tunnel sections at CERN.



a) Segmentation results of 80-meter LHC tunnel



b) Overlay results on panorama created from 80-meter LHC tunnel



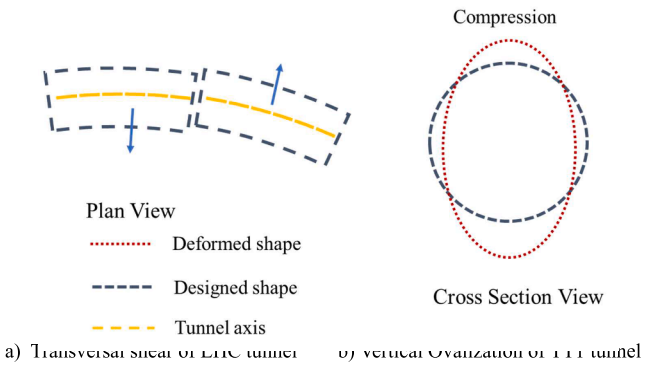
c) Segmentation results of 50-meter TT1 tunnel

d) Overlay results on panorama created fa rom 50-meter TT1 tunnel

Fig. 22. Crack distribution of inspected tunnel chainages at CERN.

tunnels: a section of approximately 80 m along the curved LHC tunnel and a section of around 45 m along the straight TT1 tunnel. During the inspections, cameras of the CERNbot, set at an approximate distance of 1.5 m away from the tunnel wall, acquire tunnel images while moving at a speed of 0.2 m/s. Fig. 21 presents the stitched panorama with the corresponding image resolution attached for both tunnels.

Segmentation and overlay results illustrate the crack distribution of these two experimental tunnel sections in Fig. 22, where cracks are positioned in precise areas with recognizable patterns. In this way, the proposed pixel-level crack monitoring system provides an accurate depiction of the condition of the tunnel and contributes to the comprehensive understanding of the long-term behavior of the tunnels. Cracks observed in the LHC tunnel are more dispersed with a circumferential oriental pattern. This kind of crack pattern indicates an uneven tunnel shearing deformation along the transversal direction described in Fig. 23 (a). In contrast, the straight TT1 tunnel section encompasses a few major longitudinal cracks spread along the sidewall. This kind of crack distribution reveals insights into the tunnel cross-section bending deformation mode of the inspected tunnel chainage as illustrated in the tunnel cross-section of Fig. 23 (b); the tunnel crown is experiencing compression while the sidewalls are in tension, resulting in continuous cracks along the longitudinal tunnel direction. Such tunnel deformation mode is also recorded at the straight TT10 tunnel using field monitoring tools, as described by (Di Murro, 2019). In addition, two circumferential cracks also appear in the 45 m TT1 tunnel section, located near the construction joints of the seconding lining, from which



a) Transversal shear of LHC tunnel b) Vertical Ovalization of TT1 tunnel

Fig. 23. The proposed deformation mechanism of LHC and TT1 tunnel.

adjacent tunnel cross sections tend to deform unevenly.

4.3. Crack density of CERN tunnel

Crack density can statistically quantify structural conditions of tunnels and facilitate asset management. Here, the global crack density (denoting the ratio of crack pixels to backgrounds) is employed to quantify the severity of crack defect. For instance, the global crack densities of the inspected two tunnel sections are 0.2184% of the 80 m LHC tunnel and 0.3471% of the 45 m TT1 tunnel. This result indicates that TT1 tunnel is probably developing a greater deformation in the long term, subject to surrounding soft rock or tunnel deterioration. In addition, the sectional crack percentage denotes the ratio of crack pixels within the designated chainage to the counterpart of the whole inspected tunnel section. This index is introduced to reflect the uneven crack distribution from chainages to chainages and identify the most severe crack-damaged chainages. The supplementary sectional summary, plotted in Fig. 24, illustrates that the worst part in the inspected areas appears at the 6th −8th meter of the TT1 tunnel with a peak value of 12.57%. Such higher crack density is probably attributed to the complex interaction between layered ground and inclined tunnel subject to time-varying water pressure at TT1 tunnel section, which deserves specific investigation in future studies.

In addition, this sectional summary also shows a more evenly distributed crack pattern in the LHC tunnel than in the TT1 tunnel, as mentioned before.

5. Conclusion

This study proposes a remote and automated tunnel crack monitoring system. The system consists of remote data acquisition, data processing for a long panorama, deep learning methods for crack segmentation, and high-resolution image prediction for crack density and distribution. Field trials have been conducted in two typical CERN tunnel sections. The major conclusions are summarized as follows:

- The proposed tunnel crack monitoring system successfully carries out TAM tasks of CERN underground facilities and facilitates the understanding of long-term tunnel behavior. Results demonstrate that the panorama image enlarges the field of inspection view and renders a realistic tunnel surface condition. The crack density distribution allows identification of crack-damaged tunnel sections. Crack pattern results reveal long-term deformation mechanisms of different tunnel sections at CERN: the transversal shearing mode is prevalent in LHC tunnel section, and the vertical elongation of the cross section is the dominant deformation mode in TT1 tunnel.
- The crack monitoring results show that longitudinal cracks are dominant in the inspected TT1 tunnel, whereas sparse circumferential cracks are more often encountered in the LHC tunnel. In addition, compared to LHC tunnel with a crack density of 0.22%, a much higher crack density of 0.35% appears in the inspected TT1 tunnel section.
- A crack database of tunnels at CERN has been established during this study. Incorporating regularization methods, tailored loss function, and transfer learning, the DeepLab V3plus models show good performance in segmenting complex CERN tunnel cracks in harsh tunnel environments. Training results on the customized database show that the fine-tuned DeepLab V3plus outperforms extant deep networks, with mIoU of 83.92% and mPA of 92.64%, and the improved lightweight DeepLab V3plus model is efficient with satisfactory accuracy. In addition, the proposed smooth blending inference method can improve the performance of deep networks on high-resolution tunnel panorama.

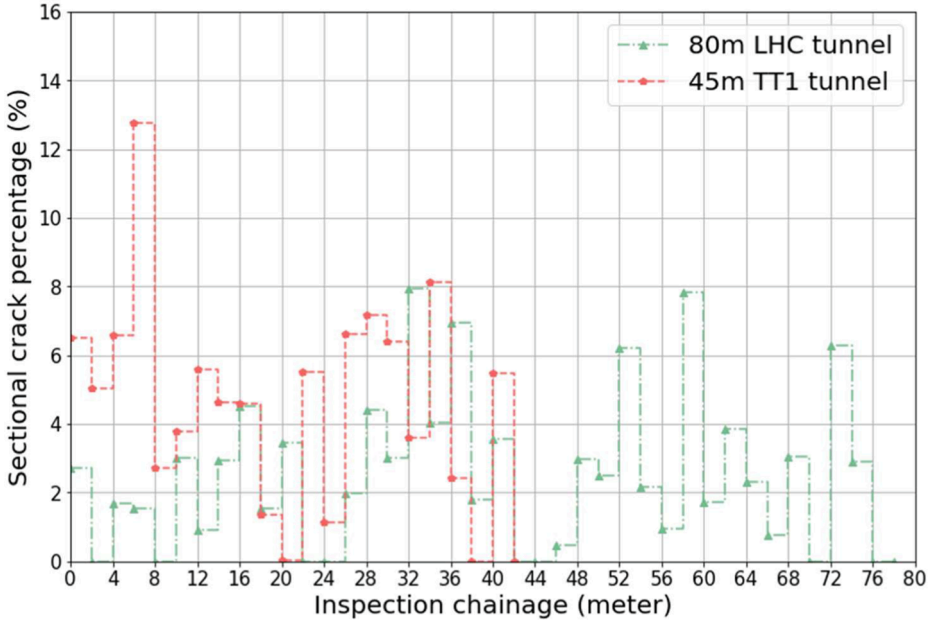


Fig. 24. Sectional crack percentages of two inspected chainages.

CRediT authorship contribution statement

Aohui Ouyang: Conceptualization, Methodology, Formal analysis, Software, Writing – original draft. **Vanessa Di Murro:** Investigation, Writing – review & editing. **Martin Cull:** Investigation, Writing – review & editing. **Roddy Cunningham:** Investigation, Writing – review & editing. **John Andrew Osborne:** Investigation, Resources, Supervision, Project administration. **Zili Li:** Conceptualization, Supervision, Project administration.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

This research is supported by a grant from the European Organization for Nuclear Research (CERN), the Irish Centre for Applied Geoscience (iCRAG), and the Science Foundation Ireland (SFI) under Grant Number 13/RC/2092_P2. This funding is greatly acknowledged. The authors also appreciate the image data collected and shared by the BE-CEM-MRO section at CERN.

References

Attard, L., Debono, C.J., Valentino, G., Di Castro, M., Masi, A., Scibile, L., 2019. Automatic crack detection using mask R-CNN. 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA).

Attard, L., 2020. A tunnel structural health monitoring solution using computer vision and data fusion.

Badrinarayanan, V., Handa, A. and Cipolla, R., 2015. Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. arXiv preprint arXiv:1505.07293.

Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., 2008. Speeded-up robust features (SURF). Comput. Vis. Image Underst. 110 (3), 346–359.

Brackenbury, D., 2022. Automated Image-Based Inspection of Masonry Arch Bridges. University of Cambridge.

Cha, Y.J., Choi, W., Büyüköztürk, O., 2017. Deep learning-based crack damage detection using convolutional neural networks. Comput. Aided Civ. Inf. Eng. 32 (5), 361–378.

Chen, T., Cai, Z., Zhao, X., Chen, C., Liang, X., Zou, T., Wang, P., 2020. Pavement crack detection and recognition using the architecture of segNet. J. Ind. Inf. Integr. 18, 100144.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. Proceedings of the European Conference on Computer Vision (ECCV).

Chevalier, G., 2017. Smoothly Blend Image Patches. GitHub. <https://github. com/Vooban/Smoothly-Blend-Image-Patches>.

Dais, D., Bal, I.E., Smyrou, E., Sarhosis, V., 2021. Automatic crack classification and segmentation on masonry surfaces using convolutional neural networks and transfer learning. Autom. Constr. 125, 103606.

Dang, L.M., Wang, H., Li, Y., Park, Y., Oh, C., Nguyen, T.N., Moon, H., 2022. Automatic tunnel lining crack evaluation and measurement using deep learning. Tunn. Undergr. Space Technol. 124, 104472.

Di Castro, M., Buonocore, L. R., Ferre, M., Gilardoni, S., Losito, R., Lunghi, G. and Masi, A., 2017. A dual arms robotic platform control for navigation, inspection and telemanipulation. Proceedings of the 16th International Conference on Accelerator and Large Experimental Control Systems (ICALEPCS'17), Barcelona, Spain.

Di Castro, M., 2019. A novel robotic framework for safe inspection and telemanipulation in hazardous and unstructured environments, Industriales.

Di Murro, V., Pelecanos, L., Soga, K., Kechavarzi, C., Morton, R. F. and Scibile, L., 2019. Long-term deformation monitoring of CERN concrete-lined tunnels using distributed fibre-optic sensing. Geotechnical Engineering Journal of the SEAGS & AGSSEA 50 (1).

Di Murro, V., 2019. Long-term performance of a concrete-lined tunnel at CERN, Cambridge U.

Dung, C.V., 2019. Autonomous concrete crack detection using deep fully convolutional neural network. Autom. Constr. 99, 52–58.

Everingham, M., Eslami, S., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A., 2015. The pascal visual object classes challenge: a retrospective. Int. J. Comput. Vis. 111 (1), 98–136.

Fern, E.J., Di Murro, V., Soga, K., Li, Z., Scibile, L., Osborne, J.A., 2018. Geotechnical characterisation of a weak sedimentary rock mass at CERN, Geneva. Tunn. Undergr. Space Technol. 77, 249–260.

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24 (6), 381–395.

Gong, Q., Wang, Y., Yu, Z., Zhu, L., Shi, H., Huang, H., 2018. A Tunnel Crack Identification Algorithm with Convolutional Neural Networks. 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC).

Gopalakrishnan, K., Khaitan, S.K., Choudhary, A., Agrawal, A., 2017. Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. Constr. Build. Mater. 157, 322–330.

Greenwald, N.F., Miller, G., Moen, E., Kong, A., Kagel, A., Dougherty, T., Fullaway, C.C., McIntosh, B.J., Leow, K.X., Schwartz, M.S., 2022. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. Nat. Biotechnol. 40 (4), 555–565.

Hadjidemetriou, G.M., Vela, P.A., Christodoulou, S.E., 2018. Automated pavement patch detection and quantification using support vector machines. J. Comput. Civ. Eng. 32 (1), 04017073.

He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask r-cnn. Proceedings of the IEEE International Conference on Computer Vision.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H., 2017. Mobilenets: efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

Huang, H.-W., Li, Q.-T., Zhang, D.-M., 2018. Deep learning based image recognition for crack and leakage defects of metro shield tunnel. Tunn. Undergr. Space Technol. 77, 166–176.

Huang, H., Sun, Y., Xue, Y., Wang, F., 2017. Inspection equipment study for subway tunnel defects by grey-scale image processing. Adv. Eng. Inf. 32, 188–201.

Huang, H., Zhao, S., Zhang, D., Chen, J., 2022. Deep learning-based instance segmentation of cracks from shield tunnel lining images. Struct. Infrastruct. Eng. 18 (2), 183–196.

Jahanshahi, M.R., Masri, S.F., 2012. Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures. Autom. Constr. 22, 567–576.

Jahanshahi, M.R., Masri, S.F., 2013. A new methodology for non-contact accurate crack width measurement through photogrammetry for automated structural safety evaluation. Smart Mater. Struct. 22 (3), 035019.

Kuhn, A., Hirschmüller, H. and Mayer, H., 2013. Multi-resolution range data fusion for multi-view stereo reconstruction. German Conference on Pattern Recognition.

Lei, M., Liu, L., Shi, C., Tan, Y., Lin, Y., Wang, W., 2021. A novel tunnel-lining crack recognition system based on digital image technology. Tunn. Undergr. Space Technol. 108, 103724.

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S., 2017a. Feature pyramid networks for object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017b. Focal loss for dense object detection. Proceedings of the IEEE International Conference on Computer Vision.

Liu, Y., Ren, Q., Geng, J., Ding, M., Li, J., 2018. Efficient patch-wise semantic segmentation for large-scale remote sensing images. Sensors 18 (10), 3232.

Liu, J., Yang, X., Lau, S., Wang, X., Luo, S., Lee, V.C.S., Ding, L., 2020. Automated pavement crack detection and segmentation based on two-step convolutional neural network. Comput. Aided Civ. Inf. Eng. 35 (11), 1291–1305.

Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Lunghi, G., Prades, R. M. and Di Castro, M., 2016. An advanced, adaptive and multimodal graphical user interface for human-robot teleoperation in radioactive scenarios. International Conference on Informatics in Control, Automation and Robotics.

Lunghi, G., Marin, R., Di Castro, M., Masi, A., Sanz, P.J., 2019. Multimodal human-robot interface for accessible remote robotic interventions in hazardous environments. IEEE Access 7, 127290–127319.

Maeda, H., Sekimoto, Y., Seto, T., Kashiyama, T., Omata, H., 2018. Road damage detection and classification using deep neural networks with smartphone images. Comput. Aided Civ. Inf. Eng. 33 (12), 1127–1141.

Ni, F., Zhang, J., Chen, Z., 2019a. Pixel-level crack delineation in images with convolutional feature fusion. Struct. Control Health Monit. 26 (1), e2286.

Ni, F., Zhang, J., Chen, Z., 2019b. Zernike-moment measurement of thin-crack width in images enabled by dual-scale deep learning. Comput. Aid. Civ. Inf. Eng. 34 (5), 367–384.

O'Brien, D., Osborne, J.A., Perez-Duenas, E., Cunningham, R., Li, Z., 2023. Automated crack classification for the CERN underground tunnel infrastructure using deep learning. Tunn. Undergr. Space Technol. 131, 104668.

Paisitkriangkrai, S., Sherrah, J., Janney, P., Hengel, V.-D., 2015. Effective semantic pixel labelling with convolutional networks and conditional random fields. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C., 2018. Mobilenetv 2: inverted residuals and linear bottlenecks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Sari, Y., Prakoso, P.B., Baskara, A.R., 2019. Road crack detection using support vector machine (SVM) and OTSU algorithm. 2019 6th International Conference on Electric Vehicular Technology (ICEVT).

Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Jorge Cardoso, M., 2017. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support.

Xu, J.-J., Zhang, H., Tang, C.-S., Cheng, Q., Liu, B., Shi, B., 2022. Automatic soil desiccation crack recognition using deep learning. Geotechnique 72 (4), 337–349.

Xue, Y., Jia, F., Cai, X., Shadabfar, M., Huang, H., 2022. An optimization strategy to improve the deep learning-based recognition model of leakage in shield tunnels. Comput. Aided Civ. Inf. Eng. 37 (3), 386–402.

Yu, S.-N., Jang, J.-H., Han, C.-S., 2007. Auto inspection system using a mobile robot for detecting concrete cracks in a tunnel. Autom. Constr. 16 (3), 255–261.

Zhang, L., Shen, J., Zhu, B., 2021. A research on an improved Unet-based concrete crack detection algorithm. Struct. Health Monit. 20 (4), 1864–1879.

Zhang, W., Zhang, Z., Qi, D., Liu, Y., 2014. Automatic crack detection and classification method for subway tunnel safety monitoring. Sensors 14 (10), 19307–19328.

Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2017. Pyramid scene parsing network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

Zhao, S., Zhang, D.M., Huang, H.W., 2020. Deep learning–based image instance segmentation for moisture marks of shield tunnel lining. Tunn. Undergr. Space Technol. 95, 103156.