Enhanced Video-Level Anomaly Feature Detection for Nuclear Power Plant Component Inspections Using the Latency Mechanism

Zhouxiang Fei[1]*, Graeme M. West[1], Paul Murray[1], Gordon Dobie[1]

[1]Department of Electronic and Electrical Engineering, University of Strathclyde, Glasgow, U.K.

ABSTRACT

Inspections are regularly carried out at various nuclear power plant components to provide latest understanding of the facility condition. One common approach is via remote visual inspection which requires the engineers to watch a large volume of video footage and identify the anomalies in the video. This process is intensively manual-based as the video segments containing anomalies of interest is only a small part of the original video. For this reason, an automated anomaly detection tool is expected to ease the amount of human effort involved in the inspection process. Deep learning is a useful tool to autonomously detect anomalies in the inspection video, given that a well-prepared training dataset of anomaly types is available. However, the detection system may detect false-positives which can be difficult to remove without human intervention. To solve this problem, we introduce a new video-level detection workflow incorporating the latency mechanism to effectively reduce the false-positive detections in the inspection videos. In this workflow, each region in a frame is scanned using a convolutional neural network (CNN) trained for a specific anomaly type. The initial scanning results are then refined using the latency mechanism which flags a region as "anomaly" when the anomaly is detected in the current frame and in a series of previous consecutive frames. A case study of crack-like feature detection in superheaters is presented to demonstrate the efficacy of the proposed workflow. The results show that the false-positive detections seen in the initial scanning can be effectively reduced using the proposed latency mechanism. It is suggested that this workflow can be directly transferable to various anomaly detection tasks of nuclear plant facility inspection.

Keywords: Crack detection; deep learning; nuclear power plant inspection; remote visual inspection support

1. INTRODUCTION

Inspection of nuclear power plant facilities is routinely conducted to provide data which can be analysed to gain an understanding of its structural health monitoring. For each component under inspection, engineers need to identify if specific types of anomalies are embedded in the component, such as cracks in nuclear reactor cores [1]-[3] and corrosion in used nuclear fuel dry storage canisters [4]. Such understanding helps analyse the damage growth and predict the remaining life of components, and subsequent maintenance can be arranged if necessary. The traditional approach to understand the component conditions is mainly through manual-based inspection. The manual inspection may be performed to real-time videos, or reviewing of videos after being captured onsite may be conducted, depending on the inspection tools for the component. Despite the differences in the process of inspecting various facilities, one common challenging issue that the manual inspection faces is the requirement of a high level of constantly intensive

* Email address for the corresponding author: zhouxiang.fei@strath.ac.uk

concentration throughout a long inspection process and is time-consuming. As a result, the manual inspection can be an excessively lengthy and repetitive process.

Taking the superheater inspection [5] for example, an inspection video typically lasts around 2 hours, in which the crack-feature part may take only a few minutes long. The manual-labour cost can be significantly reduced if the engineer only needs to review the crack-feature parts (each lasting around 10~15 seconds and containing the suspected crack-feature contents) in the original video. To achieve this goal, automated detection of anomaly features is required to support the inspection process. Automated anomaly detection has drawn an intensive focus in the civil engineering domain to aid structural health monitoring. For example, image processing techniques were applied to detect cracks in bridge deck surfaces [6], steel beams [7] and concrete surfaces [8]. Recently the machine learning techniques have been widely applied to support the automated anomaly detection. For example, the Bag-of-Visual-Words technique was used to find the crack regions in the fuel channel images of nuclear reactor cores [3], and a classifier based on Haar-like features was designed to capture cracks in wind turbine blades [9]. Amongst various data-driven techniques, the deep learning technique family which deploys the convolutional neural network (CNN) becomes increasingly popular, as the anomaly feature types are automatically learnt by the network to make classifications. For example, the automated detection systems were developed for cracks in rails [10] and in concrete surfaces at patch level [11] and at pixel level [12]. The efficacy comparison of edge detector and CNN on concrete crack detection was investigated in [13]. Note that CNNs have had limited applications in the nuclear sector, compared to the civil engineering domain (e.g., pavement, road and bridge surfaces inspection). This could be partially due to the limited availability of anomaly types in the nuclear sector as well as its strict regulation on synthetic data. Currently some examples of CNN applications in the nuclear facility inspection aspect are the crack detection in mock-up reactor surfaces [14] and corrosion region identification in used nuclear fuel dry storage canisters [4].

The aforementioned research work ([4] [10] [11] [14]) mainly focused on anomaly detections at frame level. Though accurate image-level anomaly detection can be achieved, there are challenges remaining in progressing from frame-level detection to accurate and efficient video-level detection. For the video-level detection, it is of more interest to identify video segments containing the anomaly observations in the original whole video. In contrast, many existing works only approached accurate detection of anomalies in images, though spatio-temporal information between consecutive frames was used in [14] to enhance detection accuracy. With the video-level decision support, the inspection engineer only need to review a selection of summary video clips concentrating on anomaly features, instead of going through the entire video length. On the other hand, false-positive detections could occasionally occur. Taking superheater crack detection as example, the intensive edge features in frames may trigger false-positive detections, which is difficult to rectify only using the current frame under analysis. Video-level detection can be used to reduce false-positive detections by utilising the prior information between consecutive frames.

In this paper, we aim to improve the video-level detection by introducing a new video-level detection workflow incorporating the latency mechanism to effectively reduce the false-positive detections in inspection videos. In this workflow, the video frames are first sampled and each sampled frame in the video is divided into multiple regions. A full-grid scanning strategy is applied together with a trained CNN to check for anomaly features in each region in the sampled frames. After the initial scanning of all the sampled frames, the scanning results are refined using the latency mechanism. Specifically, a region in a frame is flagged as “anomaly feature” only when the both following conditions are met: (1) anomaly features are detected at this region in the current frame; and (2) the number of previous consecutive frames in which crack features are detected at the same region equals or is above a user-defined threshold. As a result, the false-positive detections seen in the initial scanning can be effectively reduced using the refinement of the proposed latency mechanism. As blurry frames may not produce useful detection information, a sliding window technique is applied to review the refined detection result and register the status of each frame as either “anomaly feature” or “non-anomaly”. As a result, enhanced detection of anomaly segments in the remote visual inspection videos of nuclear power plant components is achieved. In this paper, we demonstrate an application of the proposed video-level detection workflow to the video inspection of superheaters, which is a common structure within nuclear power plants. Similar approaches could also be applied to pipes and vessels in power plants. The remainder of this paper is organised as follows: Section 2 describes the proposed video-level detection framework. Section 3 presents a case study of the application to superheater inspection video and the result discussion. The conclusions are given in Section 4.

2. METHODOLOGY

The development of an automated anomaly system consists of the development of a trained classifier and the execution of the classifier for the video-level anomaly detection. Though various approaches (see [14] and [15] for example) can be used to obtain an anomaly detection classifier, we focus on the use of the deep learning technique in this paper.

2.1. Brief CNN Background

Deep learning is a popular technique for developing an automated anomaly detection system, on the basis of a well-prepared training dataset and accurate training process. In general, inside a typical CNN structure, the convolution and pooling layers are first deployed to extract the anomaly feature in the image, and the fully-connected and softmax layers are used to make classifications. Specifically, filters are used in the convolution layer to perform convolutions with different parts of the image to generate feature maps. The feature maps are passed through the pooling layer to reduce the size and then flattened into an array to feed into the fully-connected layer. The output of the fully-connected layer goes into the softmax layer to calculate the score of each class. The class with the highest score is predicted as the category of the image. As the implementation of the CNN classifier is only a subsection of the proposed video-level detection workflow, a very high-level description of the CNN operation is given herein. Please see [11] for a detailed introduction of CNNs.

The preparation of training dataset is an important step in the development of a CNN classifier, as this step allows the classifier to learn the correct features of interest for decision-making. On the other hand, preparing datasets is a lengthy and labour-intensive process if performed manually. This is because a large number of images are typically needed for each classification category. To solve this challenging task, an automated labelling technique was proposed in [5] to significantly reduce the manual-labour work spent in the dataset preparation step. The proposed labelling technique is versatile to generate automatically labelled images for various anomaly types. The GoogLeNet architecture [16] was originally developed to classify 1,000 different classes of objects in ImageNet [17] and is used in this paper as the classifier in the scanning phase of the proposed video-level detection workflow. Our previous work in [5] is advanced in this paper by using the CNN classifier to perform video-level anomaly detection. Please refer to [5] for a detailed introduction of the automated labelling technique. The workflow of the CNN classifier development is given in Fig. 1(a).

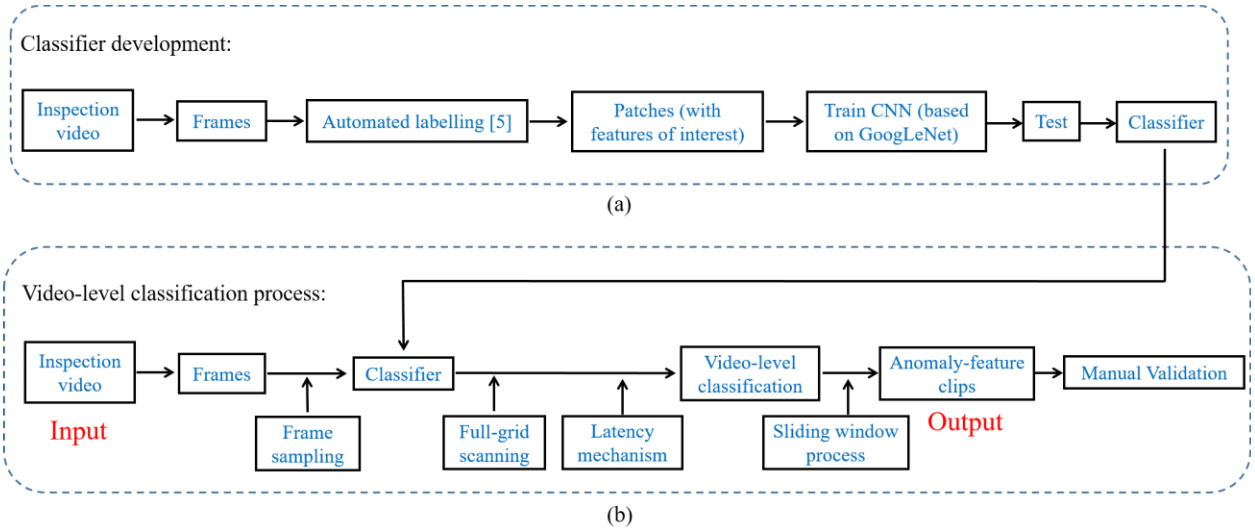


Figure 1. Overall architecture. (a) The development of the CNN classifier. (b) The video-level detection workflow.

2.2. Video-Level Detection

The frame-level detection can be performed using a full-grid scanning strategy via the developed classifier of the anomaly type. Taking the superheater crack-feature detection as example, the CNN classifier is trained to classify patches with the resolution of 224 × 224 × 3 (RGB channels) as "crack-feature" or "non-crack" classes. As shown in Fig. 2, all the non-overlapping patch locations in the frame are scanned and the patches detected as "crack-feature" are initially registered as potential anomaly region. A computationally-intensive video-level detection could be performed by repeating frame-level detection for each frame in the video. However, as adjacent frames may contain redundant information, frame sampling can be applied to reduce computational cost and speed-up the video-level detection process, without losing detected observations of the anomaly. As shown in Fig. 1(b), frame sampling is introduced in the first step of our proposed video-level detection workflow. In the context of this paper, the sampling rate is the number of frames from which a frame is used for analysis. For example, if *Frame Sampling Rate* is 3, then the 1st, 4th, 7th (and so forth) frames are analysed.

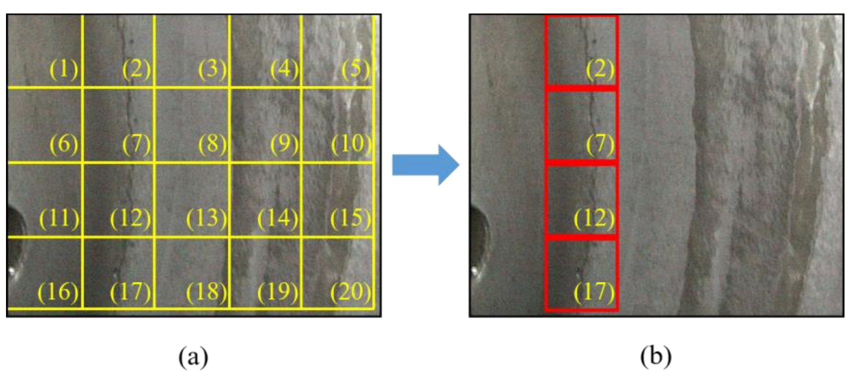


Figure 2. Full-grid anomaly scanning illustration. The footer of each patch denotes the region index. (a) An example of detecting crack features at all the non-overlapping patch locations (size: 224 × 224 × 3 in RGB channels) in a frame. (b) Detected regions as potential crack-feature regions by the classifier.

As shown in Fig. 1(b), after the sampled frames are scanned with full grid, the latency mechanism is applied to effectively remove false-positive detections. The advance of the latency mechanism is to leverage prior information about the nature of a video inspection (as opposed to a single frame) to improve the detection accuracy. During the latency mechanism step, the frame-level classification results are refined using a *Latency Threshold Value*. Specifically, for the analysed frames, a patch region in a frame is flagged as "anomaly" (in our case: "crack-feature") only when "anomaly" is detected in this frame and also at the same region in the previous (*Latency Threshold Value* − 1) frames. Taking Fig. 2(a) as example (assuming the frame index is 1000th in the video), when *Frame Sampling Rate* is 3 and *Latency Threshold Value* is 3, the patch region (index: 2) in the 1000th frame is registered as "crack-feature" only when "crack-feature" is detected at the patch region (index: 2) in the 994th, 997th and 1000th frames. To apply the latency mechanism, it is assumed that the crack features stay in the same patch region during the latency time window, which is usually true in most cases. Choosing the suitable *Latency Threshold Value* requires customised tuning. Generally speaking, the *Latency Threshold Value* is dependent on the moving speed of the camera to ensure that the potential crack features are mostly in the target region during this time window. A large *Latency Threshold Value* may cause missing out the observations of anomaly features which only last for a very short period of time in the inspection video.

At the end of the latency mechanism process, an analysed frame is registered as "anomaly" if any patch region in the frame is detected as "anomaly", otherwise the analysed frame is registered as "normal". Note that the video-level classification result at this step is still discrete, as some analysed frames may not provide useful information due to the blurriness in the frames. A sliding window technique is used following the latency mechanism to obtain "anomaly" video summary clips for the engineer to review. In the sliding window process, a sliding window with a specific window length (referred to as the *Sliding Window Length*) runs through the analysed frames. If the number of analysed frames registered as "anomaly" in a sliding window exceeds or equals a user-defined threshold (referred to as the *Sliding Window Threshold* in this paper), all the frames (both the analysed and unsampled frames) in the sliding window are registered as "anomaly". Then a 3-second period is added at the start and end of the sliding window to capture the manoeuvre of the camera approaching and leaving the anomaly region. This step may result in a large number of individual "anomaly" clips. In practice, it is highly likely that closely-adjacent clips may belong to the same anomaly observation. For this reason, the number of summary clips for the engineer to review can be significantly reduced by merging the adjacent clips (including the gap between the clips) if the adjacent clips are apart by less than 10 seconds. Finally, the resulting "anomaly" summary clips are reviewed by the engineer for cross-validation and report logging. In total, 4 tunable parameters are used in the video-level anomaly detection process, namely: the *Frame Sampling Rate*, *Latency Threshold Value*, *Sliding Window Length*, and *Sliding Window Threshold*.

3. CASE STUDY: CRACK-FEATURE DETECTION IN SUPERHEATERS

3.1. Superheater Inspection Background

Superheater inspection is used as a case study of the proposed video-level detection workflow. Superheater is a common component on the thermal aspect of nuclear power plants to overheat steam. Remote visual inspection of superheater is regularly performed. During the inspection process, the engineer sends an endoscope into the superheater via the air nozzle and carefully seek crack-like features at the bottom circumferential region (i.e., tube plate upper radius) of the superheater. An inspection video footage is recorded and reviewed to log the observation of crack features in the superheater. A schematic of the superheater structure is given in Fig. 3. Our aim is to use the proposed video-level detection workflow to automatically obtain short clips of crack features in the inspection video, without losing observation of crack features. As a result, engineers only need to watch a few summary clips instead of watching the entire video, thus to reduce the manual-labour cost involved.

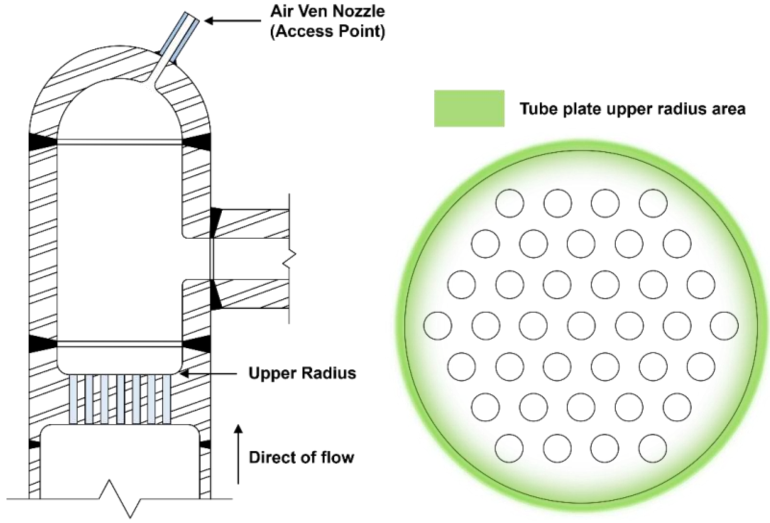


Figure 3. Superheater structure. (a) Cross-section view. (b) Top view of the tube plate.

3.2. Case Study Results

A CNN-based classifier is developed using the datasets generated via the labelling technique proposed in [5]. Specifically, the datasets consist of 800 "crack-feature" patches and 800 "non-crack" patches, and are divided into 60.5%, 15.5% and 24% portions for the training, validation and testing datasets, respectively. Each dataset is balanced and there is no data leakage across the three datasets (i.e., the data from a certain video is uniquely in one of the three datasets). A training process of 30 epochs using the GoogLeNet architecture via transfer learning [18] yields a recall of 91.15%, precision of 94.59%, F1-Score of 92.84% and total accuracy of 92.97% on the testing dataset.

Fig. 4 shows the efficacy of the latency mechanism on removing false-positive detections using a testing video. Specifically, the frames in the testing video are extracted and then classified using the full-grid scanning mechanism at *Frame Sampling Rate* = 1 (i.e., all the frames are analysed). Figs. 4(a)-(d) and (i)-(l) are classified without the latency mechanism, whereas Figs. 4(e)-(h) and (m)-(p) are the corresponding counterparts analysed using the latency mechanism with *Latency Threshold Value* = 3. It can be seen that the texture features in the frames can trigger the CNN classifier to make false-positive detections. Such false-positive detections can be difficult to remove if the prior information between adjacent frames is not used. On the other hand, the latency mechanism is an effective tool to reduce false-positive detections. Comparing pairs of Figs. 4(a) and (e), (i) and (m), (b) and (f), it is clear that the false-positive detections caused by strong edge features (due to discolouration and tube hole contours) are accurately removed using the latency mechanism, and the true-positive detections remain. More examples of removing false-positive detections caused by tube-hole edges, welding, strong lighting reflection and outlet contours are given in pairs of Figs. 4(j) and (n), (c) and (g), (k) and (o), (d) and (h), respectively. It is worth noting that the false-positive detections in blurry frames can also be rectified using the latency mechanism, as shown in Fig. 4(l) and (p).

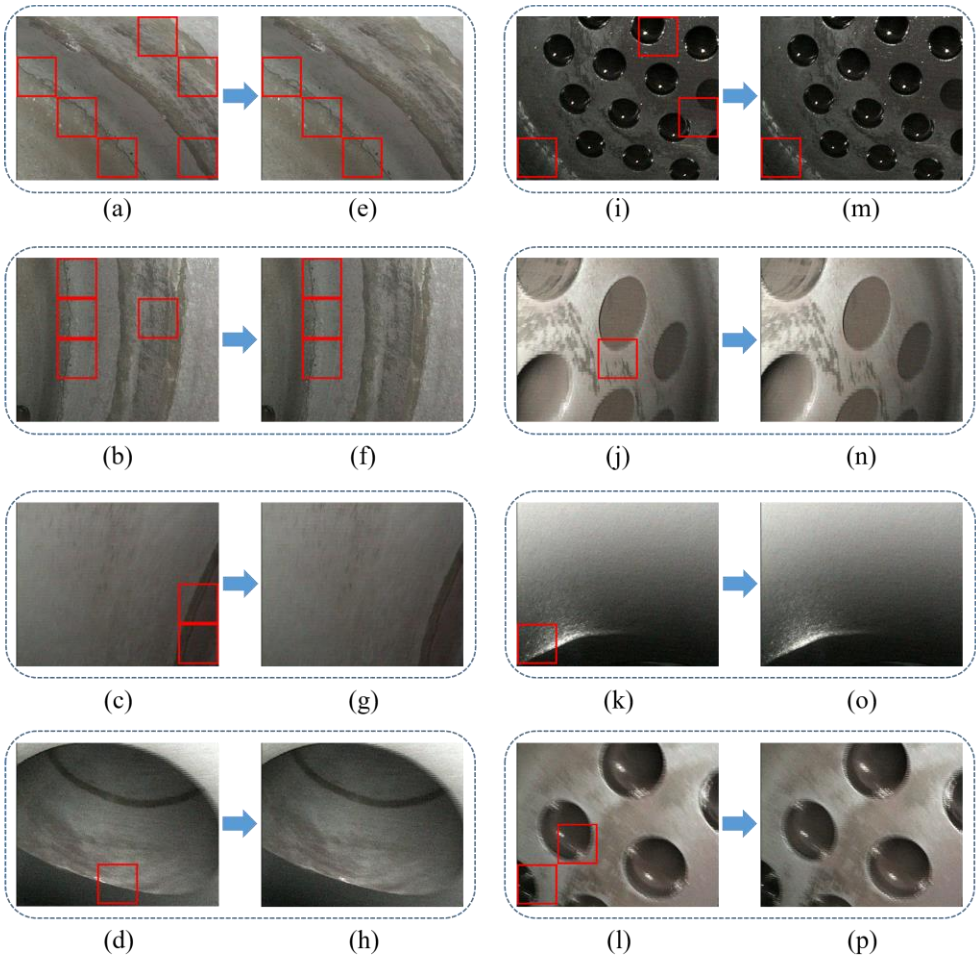
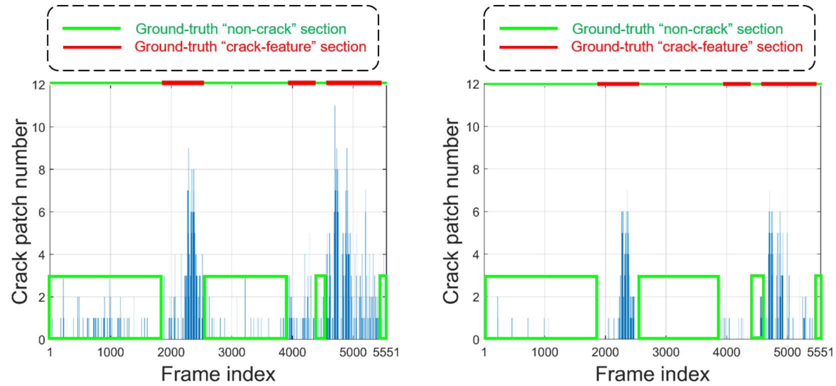


Figure 4. Exemplar classification results of removing false-positive detections using the latency mechanism at *Frame Sampling Rate* = 1 and *Latency Threshold Value* = 3. Figs. 4(a)-(d) and (i) -(l) demonstrate the false-positive detections classified by full-grid scanning without using the latency mechanism. The false-positive detections are mainly focused on the strong edge-feature regions caused by factors such as discolouration, tube hole and outlet contours, stains and strong lighting reflection. Such false-positive detections are effectively removed using the latency mechanism, as shown in Figs. 4(e)-(h) and (m)-(p).

The detected “crack-feature” patch number in each frame of the testing video (consisting of 5,551 frames, recorded at 25 frames per second) is shown in Fig. 5, aligned with the ground-truth contents at different times in the video. The ground-truth starting and ending frame indices for the “non-crack” and “crack-feature” segments are obtained by manually reviewing the video contents. Fig. 5(a) shows the number of “crack-feature” patch only using the full-grid scanning. As shown in Fig. 5(a), true-positive detections are accurately picked up in “crack-feature” segments, but false-positive detections could occur at the “non-crack” sections of the video. However, the classifier is not prone to making false-positive detections as the number of such misclassifications in the frames during non-crack sections is small. The classification result is then refined using the latency mechanism with *Latency Threshold Value* = 3 and is presented in Fig. 5(b). Comparing Fig. 5(b) to Fig. 5(a), it is clear that the false-positive detections at “non-crack” sections (bounded by green boxes) of the video are significantly reduced using the latency mechanism. Note that the blurriness of frames could hinder the classifier in making detections, resulting in no regions of interest found in some blurry frames.



(a) (b)

Figure 5. Efficacy of the latency mechanism on reducing false-positive detections in the testing video at *Frame Sampling Rate* = 1. (a) Detected "crack-feature" patch number in each frame of the video without using the latency mechanism. (b) Refined classification results using a *Latency Threshold Value* of 3.

Now the sliding window process introduced in Section 2.2 is performed to obtain "crack-feature" summary clips using the "crack-feature" registration status of the analysed frames refined by the latency mechanism. As shown in Fig. 6, three "crack-feature" summary clips are obtained using the proposed video-level detection workflow. The summary clips agree well with the ground-truth "crack-feature" contents in the video, with all the recorded crack features in the inspection report successfully captured.

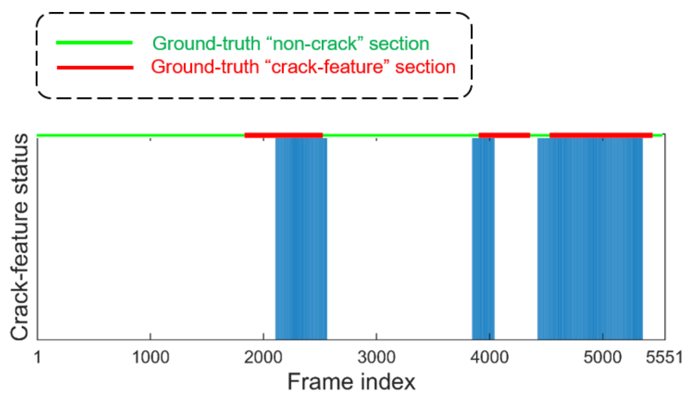


Figure 6. "Crack-feature" summary clips automatically obtained using the proposed video-level detection workflow, with *Frame Sampling Rate* = 3, *Latency Threshold Value* = 2, *Sliding Window Length* = 15, and *Sliding Window Threshold* = 6.

3.3. Discussion

Taking the crack-feature inspection in superheaters as an anomaly detection case study, the proposed video-level detection workflow has been demonstrated to be useful to accurately identify summary clips of anomaly contents from the inspection video in an efficient manner. Removing false-positive detections could be a challenging task if the priori information in adjacent frames is not used. Our proposed latency mechanism has been shown to be an effective tool to reduce false-positive detections. On this basis, the sliding window technique is used to convert discrete frame-level classification results into consecutive summary clips for review. The steps in the proposed workflow are general and directly transferrable to a wide range of anomaly inspection tasks in nuclear power plants. Choosing the parameters in the proposed workflow is related to the nature of the specific inspection task and therefore requires user-defined tailoring for the task of interest. As effective removal of false-positive detections is addressed in this study, part of future work will focus on developing techniques to reduce false-negative detections. Future work will also investigate tracking the movement of crack features across adjacent frames to further improve anomaly detection accuracy.

4. CONCLUSIONS

In this paper, a new video-level framework has been introduced to automatically detect anomalies in inspection videos accurately and efficiently. Supported by the proposed workflow, the associated manual-labour cost of the inspection task can be significantly reduced. The proposed workflow is general and is hoped to be applicable to a wide range of video-level anomaly inspections in nuclear power plants. Crack-like feature inspection in superheaters is used as a case study to demonstrate the performance of the proposed video-level detection workflow. The case study result has shown that the crack features of interest are accurately detected while false-positive detections are significantly reduced using the proposed latency mechanism. With the sliding window technique, the refined discrete frame-level classification results can be converted into short crack-feature summary clips for the engineer to review and make final decisions. Future work will focus on further enhancing anomaly detection accuracy using movement vector information across adjacent video frames.

ACKNOWLEDGMENTS

This work was carried out as part of a collaborative research project funded by the Advanced Nuclear Research Centre (ANRC) at the University of Strathclyde and supported by Altrad Babcock Ltd. The authors would like to thank Burkhard Zohm from Altrad Babcock Ltd, for his invaluable discussion.

REFERENCES

1. G. West, P. Murray, S. Marshall, and S. McArthur, "Improved visual inspection of advanced gas-cooled reactor fuel channels," *International Journal of Prognostics and Health Management*, **6** (3), pp.1-11 (2015).

2. P. Murray, G. West, S. Marshall, and S. McArthur, "Automated in-core image generation from video to aid visual inspection of nuclear power plant cores," *Nuclear Engineering and Design*, **300**, pp.57-66 (2016).

3. M. G. Devereux, P. Murray, and G. West, "A new approach for crack detection and sizing in nuclear reactor cores," *Nuclear Engineering and Design*, **359**, pp.1-11 (2019).

4. T. Papamarkou, H. Guy, B. Kroencke, et al., "Automated detection of corrosion in used nuclear fuel dry storage canisters using residual neural networks," *Nuclear Engineering and Technology*, **53** (2), pp.657-665 (2021).

5. Z. Fei, G. West, P. Murray, and G. Dobie, "Automated generation of training dataset for crack detection in nuclear power plant components," *Proceedings of the 12th Nuclear Plant Instrumentation, Control and Human-Machine Interface Technologies*, Illinois, USA, Jun. 2021, pp.251-258.

6. I. Abdel-Qader, O. Abudayyeh, and M. E. Kelly, "Analysis of edge-detection techniques for crack identification in bridges," *Journal of Computing in Civil Engineering*, **17** (4), pp.255-263 (2003).

7. C. M. Yeum and S. J. Dyke, "Vision-based automated crack detection for bridge inspection," *Computer-Aided Civil and Infrastructure Engineering*, **30**, pp.759-770 (2015).

8. T. Yamaguchi, S. Nakamura, R. Saegusa, et al., "Image-based crack detection for real concrete surfaces, *Transactions on Electrical and Electronic Engineering*, **3**, pp.128-135 (2008).

9. L. Wang and Z. Zhang, "Automatic detection of wind turbine blade surface cracks based on UAV-taken images," *IEEE Transactions on Industrial Electronics*, **64** (9), pp.7293-7309 (2017).
10. D. Li, Y. Wang, W. Yan, and W. Ren, "Acoustic emission wave classification for rail crack monitoring based on synchrosqueezed wavelet transform and multi-branch convolutional neural network," *Structural Health Monitoring*, **20** (4), pp.1563-1582 (2021).
11. Y. J. Cha, W. Choi, and O. Büyüköztürk, "Deep learning-based crack damage detection using convolutional neural networks," *Computer-Aided Civil and Infrastructure Engineering*, **32** (5), pp.361-378 (2017).
12. L. Zhang, J. Shen, and B. Zhu, "A research on an improved Unet-based concrete crack detection algorithm," *Structural Health Monitoring*, **20** (4), pp.1864-1879 (2021).
13. S. Dorafshan, R. J. Thomas, and M. Maguire, "Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete," *Construction and Building Materials*, **186**, pp.1031-1045 (2018).
14. F. Chen and M. R. Jahanshahi, "NB-CNN: Deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion," *IEEE Transactions on Industrial Electronics*, **65** (5), pp.4392-4400 (2018).
15. M. R. Jahanshahi and S. F. Masri, "Adaptive vision-based crack detection using 3D scene reconstruction for condition assessment of structures," *Automation in Construction*, **22**, pp.567-576 (2012).
16. C. Szegedy, W. Liu, P. Sermanet, et al., "Going deeper with convolutions," *Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition*, Boston, USA, Jun. 2015, pp.1-9.
17. J. Deng, W. Dong, R. Socher, et al., "ImageNet: a large-scale hierarchical image database," *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, USA, Jun. 2009, pp.248-255.
18. K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, "Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection," *Construction and Building Materials*, **157**, pp.322-330 (2017).