

Multi-modal robotic visual-tactile localisation and detection of surface cracks

Francesca Palermo^{1†}, Liz Rincon-Ardila², Changjae Oh¹, Kaspar Althoefer^{1,3},
Stefan Poslad¹, Gentiane Venture², Ildar Farkhatdinov^{1,4†}

Abstract—We present and validate a method to detect surface cracks with visual and tactile sensing. The proposed algorithm localises cracks in remote environments through videos/photos taken by an on-board robot camera. The identified areas of interest are then explored by a robot with a tactile sensor. Faster R-CNN object detection is used for identifying the location of potential cracks. Random forest classifier is used for tactile identification of the cracks to confirm their presences. Offline and online experiments to compare vision only and combined vision and tactile based crack detection are demonstrated. Two experiments are developed to test the efficiency of the multi-modal approach: online accuracy detection and time required to explore a surface and localise a crack. Exploring a cracked surface using combined visual and tactile modalities required four times less time than using the tactile modality only. The accuracy of detection was also improved with the combination of the two modalities. This approach may be implemented also in extreme environments since gamma radiation does not interfere with the sensing mechanism of fibre optic-based sensors.

I. INTRODUCTION

Localising and recognising the presence of mechanical fractures is an important task necessary in hazardous environments during waste decommissioning. It is especially useful to avoid spillage from containers keeping chemical and radioactive waste or to identify concrete fractures at early stages, to prevent their growth which may lead to larger macro-scale catastrophic failures. Crack detection and localisation have been already investigated in the past and multiple techniques exist which rely on visual analysis of the analysed segment [1], eddy current measuring devices [2], ultrasonic techniques [3] or X-ray scanning [4]. Adhikari et al. [5] proposed an automated model to detect and evaluate cracks with a supervised neural network approach and 3D visualisation model. The proposed method automatically defines crack segments by subtracting branch points from the crack skeleton. This approach avoids searching for crack connectivity and is efficient in terms of time and cost. Jahanshahi et al. [6] developed a contact-less remote sensing crack detection and quantification method based on 3D scene reconstruction. They utilised depth perception to detect

cracks and quantify their width. This feature is especially useful for incorporating mobile systems into structural inspection methods since it allows inaccessible regions to be properly inspected for cracks. Nguyen et al. [7] proposed a method to recognise concrete fractures from 2D images implementing edge detection. Based on their symmetric and line-like properties, cracks were distinguished from other objects. The centre lines of the fractures were estimated and fitted by cubic splines to detect the edge points which were linked together to represent the fracture. Pereira et al. [8] implemented an edge detection algorithm based on Sobel filter to automatically localise fractures in civil construction using Unmanned Aerial Vehicles. Improved crack detection performance can be achieved with edge detection and image segmentation methods applied in supervised and well-structured environments when the crack has clear continuity and its image has high contrast [1]. However, cracks are usually found in noisy backgrounds which lead to poor continuity, low contrast and impact the acquired imaging quality. Recently, deep learning-based methods have been developed to localise and classify cracks [9]. Chen et al. [10] proposed a fusion deep learning framework called NB-CNN (Naïve Bayes - Convolutional Neural Network). It analyses individual video frames for crack detection and detects crack patches in each video frame. Zou et al. [11] proposed DeepCrack, a deep convolutional neural network for automatic crack detection. It recognises the line structures by employing multi-scale deep convolutional features learned at hierarchical convolutional stages. In contrast to the visual modality, tactile and force sensing combined with proximity sensing can provide important information on material properties such as shape, texture and hardness [12], [13]. The stiffness of objects has been investigated [14] implementing a hybrid force and proximity finger-shaped sensor achieving 87% classification accuracy on a set of household objects. Liu et al. [15] developed a contact-sensing fingertip sensor to estimate the direction and the magnitude of the friction, normal forces and the local torque generated at the surface of explored objects with limitation of exploration for lightweight objects.

Proposed scenario. We consider a mechanical fracture detection task performed by a remotely operated robotic manipulator equipped with a vision and tactile sensing system. A possible remote inspection scenario may include a robot entering a space to be inspected, and performing a visual scan of the environment. The captured visual images are analysed by the algorithm proposed in this work to identify the areas

¹ School of Electronic Engineering and Computer Science, Queen Mary University of London, London, United Kingdom.

² Department of Mechanical Systems Engineering, Tokyo University of Agriculture and Technology, Tokyo, Japan.

³ School of Engineering and Materials Science, Queen Mary University of London, London, United Kingdom.

⁴ Department of Bioengineering, Imperial College of Science, Technology and Medicine, London, United Kingdom.

† {f.palermo,i.farkhatdinov}@qmul.ac.uk

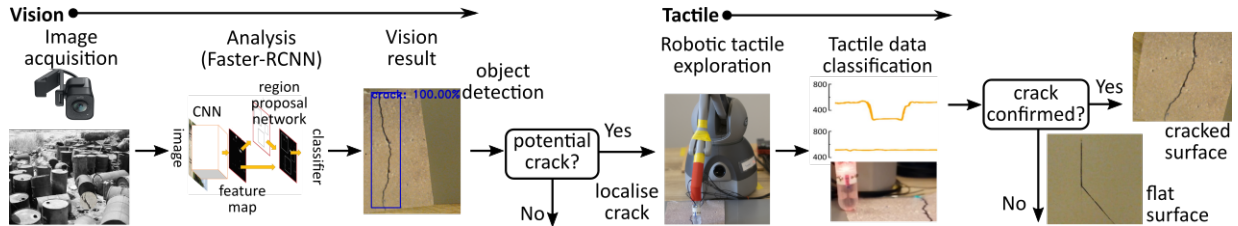


Fig. 1. Algorithm of the proposed multi-modal approach combining both visual and tactile and proximity data for fracture detection and localisation. The camera acquires a video of the scene. One frame is extracted from the video and object detection is applied for crack detection. The positions of the detected cracks are sent to the manipulator which re-positions itself and starts exploration. The potential crack is confirmed or rejected as a flat surface.

TABLE I
FRACTURE DETECTION RESULTS OF THE TRAINED MODEL ON THE TEST DATASET

Model	Training	Testing	Steps	mAP (%)	mAP 50IOU (%)	AR@100 (%)
Faster R-CNN Inception V2	3500	1500	200'000	26.90	43.82	38.76
Faster R-CNN Inception V2 - Augmented	3500	1500	200'000	65.43	93.03	71.92

which are likely to contain mechanical fractures. Following the identification of the area of interest, the coordinates of the location of the region of interest in respect to the camera are transformed to the robot's coordinates and sent to it. The robot moves closer to the area to be inspected and uses the on-board manipulator equipped with tactile sensors to physically explore the surface for further characterisation. The overview of our approach is shown in Fig. 1. Improving our previous work [16], [17], the proposed method presents a multi-modal approach with both vision and force sensing for crack localisation and recognition. A fibre optic sensor has been implemented for data acquisition because of the reduced dimensions ($\sim 55\text{mm}$), weight ($\sim 200\text{g}$), low cost, the strong immunity to electromagnetic interference and the improved environmental resistance. This approach may be implemented also in different extreme environments (e.g. in nuclear plants), since gamma radiation does not interfere with the basic sensing mechanism of fibre optic-based sensors [18]. Further, the nylon component of the implemented sensor can be used in irradiated conditions with limitations as [19] presented.

II. CRACK DETECTION METHOD AND SETUP

Fig. 1 presents the overview of the proposed crack detection method based on visual and tactile sensing modalities. The experimental setup with the video camera and the tactile sensor mounted at the end-effector of a desktop robotic manipulator is shown in Fig. 2.

A. Vision-based Crack Detection

For the vision feedback part, a Logitech StreamCam is used to scan the surface and localise possible fractures. We adopt a Faster Region-based Convolutional Neural Network (Faster R-CNN) [20] with Inception v2 architecture [21], pre-trained on Common Objects in Context (COCO) dataset [22]. The model is trained on Windows OS with the TensorFlow Object Detection API [23] with GeForce GTX 850M 4GB. The dataset used to train and validate the model consists of a subset of 3000 RGB images with 227×227 size, representing

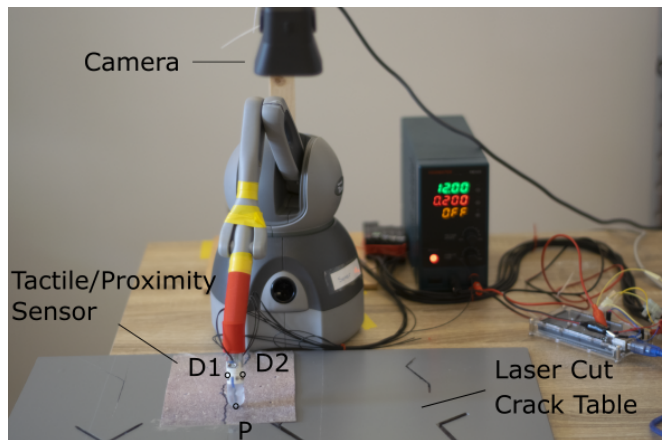


Fig. 2. The experimental setup for the surface exploration multi-modal approach with vision, force and proximity sensor.

fractures in concrete extracted from [24]. All the images were manually labelled with LabelImg¹, a graphical image annotation tool. The images available in the implemented dataset were similar to each other and did not have any occlusions or other objects besides the crack. This is not optimal when training an object detection model. Thus, multiple data augmentation processes were implemented to make the model more robust. First, the patches with crack (one or two at a time) were randomly pasted in a randomly chosen larger background image. To avoid the model to only focus on the patch, a patch, without the crack, is additionally extracted from the concrete image and pasted in the background image. Since the images were missing any level of occlusion, to each of the resulting and original images, a random number of different screws (between 1 and 2) was added to random positions. Screws have been chosen since they are part of a set of objects which may be easily found in extreme and nuclear environments. Finally, a

¹<https://github.com/tzutalin/labelImg>

random set of data augmentation techniques were applied. The complete set of augmentations consists of changing the HSV (hue, saturation, lightness), horizontal flip, scale, translate, rotate and shear. For each image, the script decided on a random set of the previously described augmentation and applied that to augment the image. The labels of the original dataset were adopted accordingly to the various modifications during the augmentation process. After the augmentation steps, a total of (approx.) 24000 images were created. Considering the limited GPU available, a total of 5000 images (3500 for training, 1500 for validation and testing) were randomly chosen for training and validating the model. Fig. 3 shows the workflow of the augmentation process.

B. Crack Detection with Tactile Sensing

An integrated force and proximity finger-shaped sensor, described in [14], is used for automatic crack detection. The sensor employs three pairs of optical fibre cables (D1, D2, D3) to measure the sensor's body deformation of the flexible middle part based on the changes in reflected light intensity. The fourth pair of optical fibre cables (P) is used to sense the proximity between external objects and the tip of the finger. The sensor is attached to the end effector of a Geomagic (old name Phantom Omni) desktop haptic interface that was used as a robot-manipulator. The end-effector was programmed to follow predefined tactile sensing paths as described in [16], [17]. The Geomagic robot moved the tactile sensor across the sample objects: the periodic sliding has a magnitude of 1.6 cm and a frequency of 0.25 Hz. The average sliding velocity was 3.89 mm/s. The initial position of the tactile sensors was not controlled and varied from trial to trial at approximately 5-10 mm from the crack edge. No normal force was applied by the sensor to the sampled surfaces except the force caused by the sensor weight (approx 200 g). The force and proximity data were recorded and feature extraction was performed; the resulting output was used as input for the classification algorithm. Feature extraction (Mean Absolute Value and Root Mean Square) was performed on each consequent 25 ms long time window with an increment of 5 ms. A Random Forest (100 trees) classifier was implemented to determine the surface pattern of the examined material. Considering the dependency of the fibre optics to different colours of the explored materials, the Random Forest classifier implemented in [16], [17] had to be adapted. For this purpose, data were acquired for each of the 3D printed surfaces (Ultimaker III 3D printer, 0.2 mm layer height, 0.4 mm nozzle diameter) shown in [17]. In addition, since the proximity data is dependent on the explored colour, two paper printed cracks were also included in the acquisitions. These printed cracks were not used for the online experiment described in section III-C. To recognise the surface of the material, the classification labels were set to: 'no crack', 'crack'. Tactile and proximity signals were recorded for 12 repeated continuous sliding movements. This continuous recording was repeated five times. First, raw, MAV and RMS data were classified using only the

proximity data (P) or the deformation signals (D1, D2, D3). Then, combinations of the above-mentioned features were conducted. A set of observations consisted of no-crack, crack 1mm, crack 2mm, crack 5mm, crack 8mm. Each set of observations was trained on itself and tested against the rest of the set one at a time for intersession investigation. Each observation was split 70% for training test and 30% for testing. In total, 20 results for each analysed feature were obtained. The software to control the Geomagic has been developed in C++ on an Ubuntu machine while the classification software has been developed on Ubuntu with Python 3.7.

III. DATA ANALYSIS AND RESULTS

The results are divided into crack detection through visual inspection and classification of fractures via tactile (force and proximity) data. We present the results of the offline tests on the recorded data (vision and tactile data analysed independently), and the performance comparison from the online experiment with the robot in which visual and tactile data are used cooperatively.

A. Testing Vision-based Crack Detection

We investigate the effect of using augmented images for training a model for fracture detection by comparing the performance of the trained model with a model trained without using augmented images. Each model was trained on 3500 images for 200k iterations. As shown in Table I and Fig. 4 a), the model trained with augmented images outperforms the baseline model with mean average precision (mAP) of ~65.2%, mean average precision at 50% intersection over union (mAP 50IOU) of 93.0% and recall with 100 detections (AR@100) of 71.9%. It is worth noting that the performance increase on mAP 50IOU are important as well as mAP since the detected locations will be further inspected by the tactile sensor. Fig. 5(a-b) shows examples of the comparison between ground-truth and object detection results. We then tested the model on the real scenario using the webcam on a laser-cut surface with cracks from 0.5mm to 5mm. Fig. 5 c) shows the results of the object detection model on a laser-cut surface. The network can generally detect and localise cracks while the performance decreases on fine cracks.

B. Testing Tactile based Crack Detection

For the data analysis of the acquired proximity and deformation data, the automated preprocessing method described in [17] has been implemented. Fig. 4(b) shows the results for the above-described offline analysis. The Random Forest classifier achieves the best classification accuracy of 89.56% when implementing MAV or RMS features with force and proximity data combined. For additional information on preprocessing and analysis please refer to [16], [17].

C. Online Experiment with Robotic System

Crack detection performance. We compared the crack detection score for vision-only and combined vision-tactile approaches. The robot's workspace was divided into two left

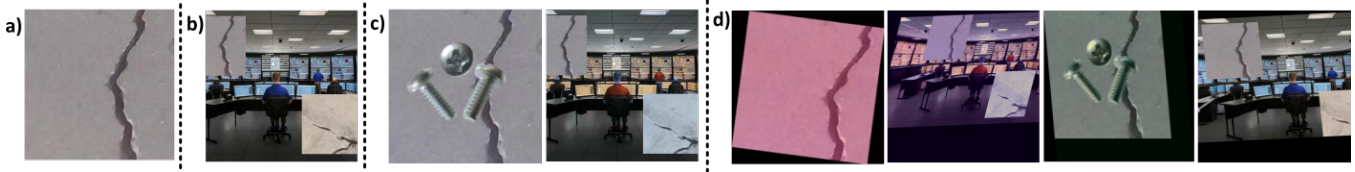


Fig. 3. Examples of images used for the crack localisation and detection. To note: the following figures have been resized and cropped to make them more visible in the paper. a) Original images of the dataset; b) Original images added to various backgrounds. A random number of cracks and flat surfaces included; c) a random number of screws added to the previous figures to create occlusion. d) Data augmentation applied to all the previous figures.

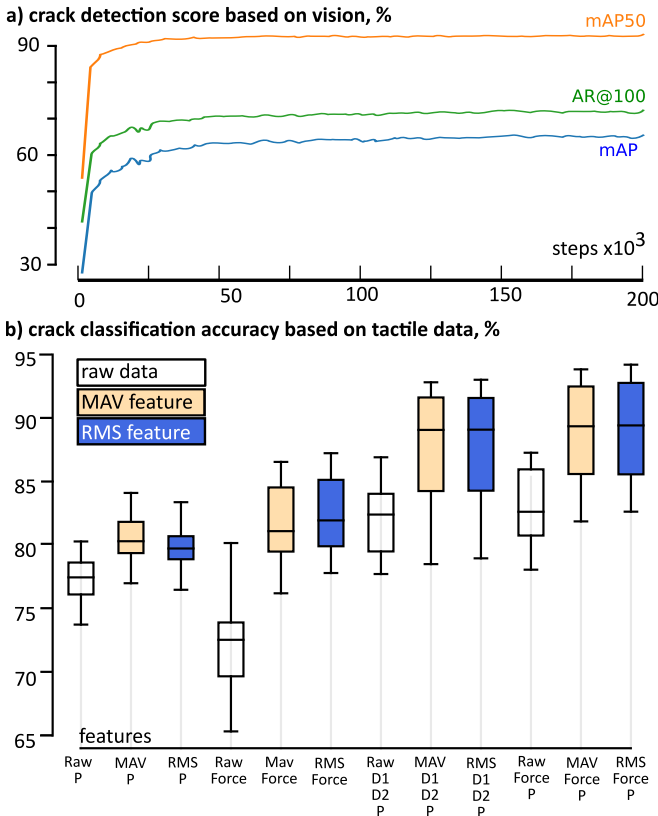


Fig. 4. Offline results for crack localisation and recognition. a) Object detection analysis b) Classification accuracy for proximity and deformation data. Each box extends from the lower to upper quartile values of the data, with a line at the median.

(L) and right (R) areas. The camera and the proposed R-CNN model were used to localise the area (L or R) with possible cracks and send their positions to the robot manipulator's controller (via the UDP channel). Once the positions were received, the robot started the exploration of each of the cracks, followed by tactile data analysis using the random forest classifier that recognised the crack or stated a false-positive result from vision.

A total of 10 trials have been performed to test the system. Fig. 6 a) shows two examples of the online multi-modal trials. In the first row, the frames captured by the webcam. In the second row, the object detection results on the previously acquired frames. In the third row, the tactile results are

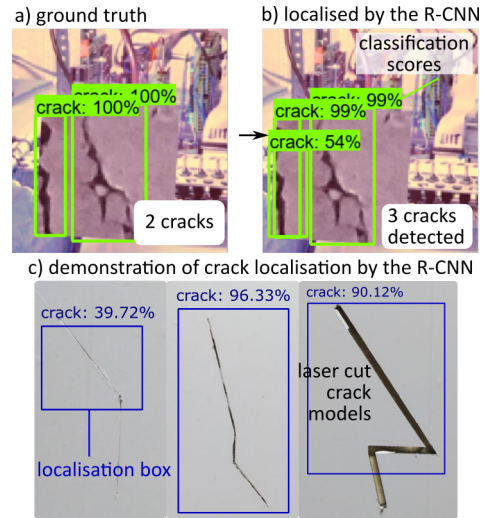


Fig. 5. Offline results of the object detection faster-RCNN Inception V2 model. In a) the manually labelled images. b) Object detection applied to the images. c) Example of detection on the laser-cut crack table that will later be used in the online experiments.

shown. Trial 4 shows an example of False Positive from the vision algorithm which detects as crack a printed crack. This is then rectified by the tactile sensor. During the trials, the tactile classification failed once in case of 0.5 mm crack. This could be due to the limited size of the crack which didn't create enough lateral deformations in the sensor. Making the proximity sensor more robust to lighting conditions and recalibrating the sensor to be more sensitive to deformations should solve this problem. From the vision part, there was a false negative for a fracture with double occlusion. Manually inspecting the online video with object detection captured by the camera, showed correct recognition of the fracture which was not found with the analysed frame. This might be due to luminosity or other factors which may influence the model. Using multiple frames on the vision side may make the model more robust and more precise to fracture detection. Fig. 7 (a) shows the correct scores for crack detection when using only vision and the two modalities together. In the 10 trials, vision detected 14 cracks plus a false negative during trial 10, for a total of 15. Of this, seven were false positive (paper printed crack), one false negative (crack not detected), seven correctly detected cracks. During the 10 trials, tactile received from the vision algorithm in total 14 possible cracks locations to inspect. The 15-th false negative

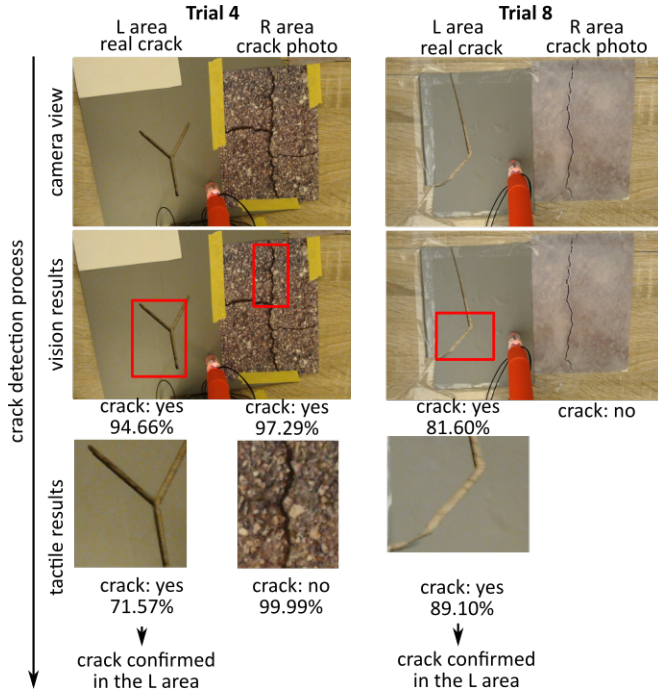


Fig. 6. Examples of online trails with multi-modal vision and tactile features for crack exploration. A total of 10 trials were performed. In the first row, the frames captured by the webcam. In the second row, the object detection results on the previously acquired frames. In the third row, the tactile results are shown.

TABLE II
TIME REQUIRED TO EXPLORE A SURFACE WITH TACTILE (T) AND
TACTILE+VISUAL (T+V) MODIFICATION

Modality	γ	N_s	Time \pm st.dev, s
T	0.25	13	519.40 \pm 24.90
T	0.50	7	278.61 \pm 16.58
T	0.75	5	198.56 \pm 8.06
T+V	N/A	N/A	31.85 \pm 1.91

from vision was never detected. Of the total 14 cracks, the tactile failed in one case when exploring the 0.5 mm fracture. In summary, the model was able to correctly detect 92.85% of the cracks when both visual and tactile modalities were used, and the detection performance decreased to 46.66% when only vision was used. Further experiments will include different lighting conditions, dirty backgrounds and other types of marks instead of paper printed cracks.

Task completion time. Detecting fractures, using only tactile data would be time-consuming as the robot would require to perform several surface scanning operations to find and characterise a crack. Using a combination of tactile and vision would reduce this time. For this purpose, an experiment has been carried out to compare the time required when exploring a surface using only tactile and when using tactile and vision combined. For tactile exploration, several scanning movements on the surface are required to localise and detect a crack. The number of required scanning movements can be estimated based on assumptions we have on the

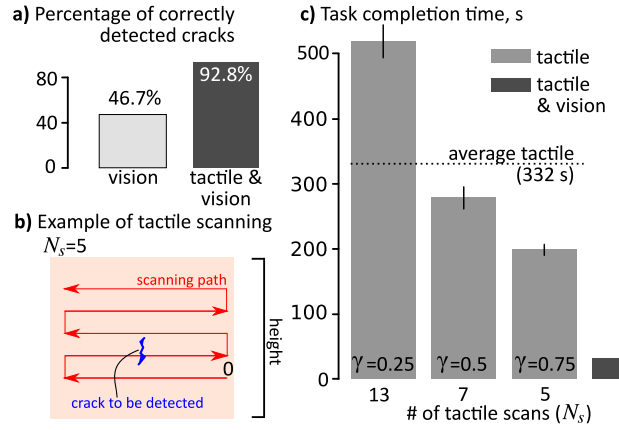


Fig. 7. a) Percentage of correct fracture detection with only vision and vision + tactile. b) Scanning example of $N_s = 5$ for the tactile exploration only. c) Time required for fracture detection comparison with tactile and tactile + vision modalities. Multiple analysis with different γ have been performed. Mean and standard deviation are shown.

crack geometry (length, shape) and their spatial frequency. We quantified these assumptions based on the dataset of the crack implemented for object detection. Each of the images was converted to binary mode (black and white pixels). We assumed and verified that black pixels represented the cracks. A subset of randomly selected 30 images was used to calculate the number of black pixels for each of the images. Then, the average number of black pixels was calculated. Inspecting each of the image, the width of the crack was calculated in pixels. With these two measurements, it was possible to determine the approximate average length of a crack in pixels, assuming that the crack is uniform and can be represented with multiple lines. The above is summarised, as follows:

$$L_i = \frac{A_i}{\bar{w}}, \quad \bar{L} = \frac{\sum L_i}{N_c} \quad (1)$$

with the length, L_i , of the i^{th} -crack selected from the i^{th} -image; the area of black pixels, A_i , in the i^{th} image (number of black pixels describing the size of the crack); the average width of the cracks, \bar{w} ; the estimated average length of the cracks in pixels, \bar{L} ; the total number of cracks used in the parameters estimation, N_c . The estimated cracks' length was then used to calculate the required number of tactile scans required by the robot's end-effector to detect a crack using only tactile sensor, as follows

$$N_s = \frac{h}{\gamma \cdot \bar{L}} + 1 \quad (2)$$

with the minimal number of scanning movements, N_s , required by the sensor to localise a crack; the height of the image captured by the camera in pixels, h ; a design parameter, $0 < \gamma < 1$, introduced to vary the length of the crack as we used the estimated length. In other words, we are calculating how many cracks of a given length can be located in a single image aligned in one direction. The above conservative estimation of the number of tactile scanning cycles is based on several assumptions: 1) a crack can

be located anywhere in the workspace of the robot and its location is unknown to the robot; 2) a crack can be represented with a combination of simple line geometry; 3) the robot will continue to scan the surface even if the crack was detected. The parameter γ can be adjusted to make the estimation more or less conservative. An example of the tactile scanning movement performed by the robot with $N_s=5$ is shown in Fig. 7b. The following parameter and estimation values were used: $\bar{A} = 2961$ black pixels on a average total of 51529 pixels for image, $\bar{w} = 20$ pixels, $\bar{L} = 148.06$ pixels, $N_c = 30$ images, $h = 480$ pixels. The parameter γ was set to 0.25, 0.5 and 0.75 in the testing trials.

Using the number of scans required to detect a crack with tactile modality only we conducted an experiment to measure the task completion time and compared it with combined vision and tactile modality. Table II and Fig. 7c show the results of the trials. Exploring a 200x100mm surface, with only tactile, requires a minimum of 199.34 seconds (~3 minutes). On the other hand, when using both vision and tactile, that time is reduced to 31.87 seconds (~ half a minute). Since the object detection phase is the part requiring more time during the multi-modal experiment, this time can be additionally shortened by implementing a stronger GPU.

IV. CONCLUSION

In this paper, a multi-modal approach for crack detection is presented. The proposed algorithm implements both visual and tactile information to classify fractures. The method uses object detection to localise fractures on surfaces which are then analysed and confirmed based on fibre optical proximity signals which are recorded during physical interaction between a custom-designed robotic finger and the remote environment. Two experiments have been developed to test the efficiency of the multi-modal approach: online accuracy detection and time required to explore a surface and localise a crack. A total of 10 trials have been implemented to compare the accuracy of the multi-modal approach with vision only method. When using the multi-modal fusion the model was able to correctly detect 92.85% of the cracks. Exploring a surface using only tactile required approximately 199 seconds which was reduced to 31 seconds when using both vision and tactile together.

ACKNOWLEDGMENT

The work is supported by the UK EPSRC grant NCNR EP/R02572X/1. Althoefer and Farkhatdinov were supported by the Alan Turing Institute, UK. This work was also supported by the Japan Student Service Organisation with a short stay short visit scholarship.

REFERENCES

- [1] A. Mohan and S. Poobal, "Crack detection using image processing: A critical review and analysis," *Alexandria Engineering Journal*, 2018.
- [2] Y. Yao, S.-T. E. Tung, and B. Glisic, "Crack detection and characterization techniques—an overview," *Structural Control and Health Monitoring*, vol. 21, no. 12, pp. 1387–1413, 2014.
- [3] Y. Chang, Y. Zi, J. Zhao, Z. Yang, W. He, and H. Sun, "An adaptive sparse deconvolution method for distinguishing the overlapping echoes of ultrasonic guided waves for pipeline crack inspection," *Measurement Science and Technology*, vol. 28, no. 3, p. 035002, 2017.

- [4] D. Naragani, M. D. Sangid, P. A. Shade, J. C. Schuren, H. Sharma, J.-S. Park, P. Kenesei, J. V. Bernier, T. J. Turner, and I. Parr, "Investigation of fatigue crack initiation from a non-metallic inclusion via high energy x-ray diffraction microscopy," *Acta Mater.*, 2017.
- [5] R. Adhikari, O. Moselhi, and A. Bagchi, "Image-based retrieval of concrete crack properties for bridge inspection," *Automation in construction*, vol. 39, pp. 180–194, 2014.
- [6] M. R. Jahanshahi, S. F. Masri, C. W. Padgett, and G. S. Sukhatme, "An innovative methodology for detection and quantification of cracks through incorporation of depth perception," *Machine vision and applications*, vol. 24, no. 2, pp. 227–241, 2013.
- [7] H.-N. Nguyen, T.-Y. Kam, and P.-Y. Cheng, "An automatic approach for accurate edge detection of concrete crack utilizing 2d geometric features of crack," *J. Signal Process. Syst.*, vol. 77, no. 3, 2014.
- [8] F. C. Pereira and C. E. Pereira, "Embedded image processing systems for automatic recognition of cracks using uavs," *Ifac-PapersOnline*, vol. 48, no. 10, pp. 16–21, 2015.
- [9] S. Dorafshan, R. J. Thomas, and M. Maguire, "Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete," *Constr Build Mater.*, vol. 186, 2018.
- [10] F.-C. Chen and M. R. Jahanshahi, "Nb-cnn: deep learning-based crack detection using convolutional neural network and naive bayes data fusion," *IEEE Trans. Ind. Electron.*, vol. 65, no. 5, 2017.
- [11] Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, and S. Wang, "Deepcrack: Learning hierarchical convolutional features for crack detection," *IEEE Transactions on Image Processing*, vol. 28, no. 3, 2018.
- [12] Z. Kappasov, J.-A. Corrales, and V. Perdereau, "Tactile sensing in dexterous robot hands," *Rob Auton Syst.*, vol. 74, 2015.
- [13] "Design and characterization of a three-axis hall effect-based soft skin sensor," *Sensors*, vol. 16, no. 4, p. 491, 2016.
- [14] J. Konstantinova, G. Cotugno, A. Stilli, Y. Noh, and K. Althoefer, "Object classification using hybrid fiber optical force/proximity sensor," in *2017 IEEE SENSORS*. IEEE, 2017, pp. 1–3.
- [15] H. Liu, K. C. Nguyen, V. Perdereau, J. Binbo, J. Back, M. Godden, L. D. Seneviratne, and K. Althoefer, "Finger contact sensing and the application in dexterous hand manipulation," *Autonomous Robots*, vol. 39, no. 1, pp. 25–41, 2015.
- [16] F. Palermo, J. Konstantinova, K. Althoefer, S. Poslad, and I. Farkhatdinov, "Implementing tactile and proximity sensing for crack detection," in *IEEE International Conference on Robotics and Automation*, 2020.
- [17] —, "Automatic fracture characterization using tactile and proximity optical sensing," *Frontiers in Robotics and AI*, vol. 7, 2020.
- [18] F. Berghmans, A. F. Fernandez, B. Bricard, F. Vos, M. C. Decreton, A. I. Gusarov, O. Deparis, P. Megret, M. Blondel, S. Caron et al., "Radiation hardness of fiber optic sensors for monitoring and remote handling applications in nuclear environments," in *Process Monitoring with Optical Fibers and Harsh Environment Sensors*, vol. 3538. International Society for Optics and Photonics, 1999, pp. 28–39.
- [19] Y. Morita and T. Seguchi, "Radiation resistance of nylon," *Denki Gakkai Zetsuen Zairyo Kenkyukai Shiryo*, vol. EIM-83, 1983.
- [20] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99.
- [21] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv*, 2015.
- [22] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, "Microsoft coco: Common objects in context," in *European conference on computer vision*. Springer, 2014.
- [23] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, I. Fischer, Z. Wojna, Y. Song, S. Guadarrama et al., "Speed/accuracy trade-offs for modern convolutional object detectors," in *Proceedings of the IEEE CVPR*, 2017, pp. 7310–7311.
- [24] Ç. F. Özgenel and A. G. Sorguç, "Performance comparison of pre-trained convolutional neural networks on crack detection in buildings," in *Proceedings of the International Symposium on Automation and Robotics in Construction*, vol. 35. IAARC Publications, 2018.